CFA Institute®
CFA Program

2022
CFA® PROGRAM
CURRICULUM
LEVEL I
VOLUMES 1–6

About the pagination of this eBook

This eBook contains a multi-volume set.

To navigate this eBook by page number, you will need to use the volume number and the page number, separated by a hyphen.

For example, to go to page 5 of volume 1, type “1-5” in the Go box at the bottom of the screen and click "Go."

To go to page 5 of volume 2, type “2-5”… and so forth.

© 2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006 by CFA Institute. All rights reserved.

This copyright covers material written expressly for this volume by the editor/s as well as the compilation itself. It does not cover the individual selections herein that first appeared elsewhere. Permission to reprint these has been obtained by CFA Institute for this edition only. Further reproductions by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval systems, must be arranged with the individual copyright holders noted.

CFA®, Chartered Financial Analyst®, AIMR-PPS®, and GIPS® are just a few of the trademarks owned by CFA Institute. To view a list of CFA Institute trademarks and the Guide for Use of CFA Institute Marks, please visit our website at www.cfainstitute.org.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional service. If legal advice or other expert assistance is required, the services of a competent professional should be sought.

All trademarks, service marks, registered trademarks, and registered service marks are the property of their respective owners and are used herein for identification purposes only.

ISBN 978-1-950157-60-0 (paper)
ISBN 978-1-119705-02-4 (ebk)

10 9 8 7 6 5 4 3 2 1

Please visit our website at
www.WileyGlobalFinance.com.

WILEY

CFA Institute®
CFA Program

QUANTITATIVE METHODS

CFA® Program Curriculum
2022 • LEVEL I • VOLUME 1

WILEY

CONTENTS

How to Use the CFA Program Curriculum vii

Background on the CBOK vii
Organization of the Curriculum viii
Features of the Curriculum viii
Designing Your Personal Study Program ix
CFA Institute Learning Ecosystem (LES) x
Prep Providers xi
Feedback xii

Quantitative Methods

Study Session 1 **Quantitative Methods (1)** 3

Reading 1 **The Time Value of Money** 5

Introduction 5
Interest Rates 6
Future Value of a Single Cash Flow (Lump Sum) 8
Non-Annual Compounding (Future Value) 13
Continuous Compounding, Stated and Effective Rates 15
  Stated and Effective Rates 16
Future Value of a Series of Cash Flows, Future Value Annuities 17
  Equal Cash Flows—Ordinary Annuity 17
  Unequal Cash Flows 19
Present Value of a Single Cash Flow (Lump Sum) 20
Non-Annual Compounding (Present Value) 22
Present Value of a Series of Equal Cash Flows (Annuities) and Unequal Cash Flows 23
  The Present Value of a Series of Equal Cash Flows 24
  The Present Value of a Series of Unequal Cash Flows 28
Present Value of a Perpetuity and Present Values Indexed at Times other than t=0 29
  Present Values Indexed at Times Other than $t = 0$ 30
Solving for Interest Rates, Growth Rates, and Number of Periods 32
  Solving for Interest Rates and Growth Rates 32
  Solving for the Number of Periods 35
Solving for Size of Annuity Payments (Combining Future Value and Present Value Annuities) 36
Present Value and Future Value Equivalence, Additivity Principle 39
  The Cash Flow Additivity Principle 41
Summary 42
Practice Problems 44
Solutions 49

O indicates an optional segment

Reading 2 **Organizing, Visualizing, and Describing Data** 63
Introduction 64
Data Types 64
  Numerical versus Categorical Data 65
  Cross-Sectional versus Time-Series versus Panel Data 67
  Structured versus Unstructured Data 68
  Data Summarization 72
Organizing Data for Quantitative Analysis 72
Summarizing Data Using Frequency Distributions 75
Summarizing Data Using a Contingency Table 81
Data Visualization 86
  Histogram and Frequency Polygon 86
  Bar Chart 88
  Tree-Map 91
  Word Cloud 92
  Line Chart 93
  Scatter Plot 95
  Heat Map 99
  Guide to Selecting among Visualization Types 100
Measures of Central Tendency 103
  The Arithmetic Mean 103
  The Median 107
  The Mode 109
  Other Concepts of Mean 110
Quantiles 120
  Quartiles, Quintiles, Deciles, and Percentiles 120
  Quantiles in Investment Practice 126
Measures of Dispersion 126
  The Range 126
  The Mean Absolute Deviation 127
  Sample Variance and Sample Standard Deviation 128
Downside Deviation and Coefficient of Variation 131
  Coefficient of Variation 135
The Shape of the Distributions 136
  The Shape of the Distributions: Kurtosis 139
Correlation between Two Variables 142
  Properties of Correlation 143
  Limitations of Correlation Analysis 146
Summary 149
Practice Problems 154
Solutions 166

Reading 3 **Probability Concepts** 175
Introduction, Probability Concepts, and Odds Ratios 176
  Probability, Expected Value, and Variance 176

Conditional and Joint Probability 181
Expected Value (Mean), Variance, and Conditional Measures of Expected Value and Variance 192
Expected Value, Variance, Standard Deviation, Covariances, and Correlations of Portfolio Returns 199
Covariance Given a Joint Probability Function 205
Bayes' Formula 208
    Bayes' Formula 208
Principles of Counting 214
Summary 220
Practice Problems 224
Solutions 230

Study Session 2 **Quantitative Methods (2)** 237

Reading 4 **Common Probability Distributions** 239
Introduction and Discrete Random Variables 240
    Discrete Random Variables 241
Discrete and Continuous Uniform Distribution 244
    Continuous Uniform Distribution 246
Binomial Distribution 250
Normal Distribution 257
    The Normal Distribution 257
    Probabilities Using the Normal Distribution 261
    Standardizing a Random Variable 263
    Probabilities Using the Standard Normal Distribution 263
Applications of the Normal Distribution 265
Lognormal Distribution and Continuous Compounding 269
    The Lognormal Distribution 269
    Continuously Compounded Rates of Return 272
Student's *t*-, Chi-Square, and *F*-Distributions 275
    Student's *t*-Distribution 275
    Chi-Square and *F*-Distribution 277
Monte Carlo Simulation 282
Summary 288
Practice Problems 292
Solutions 299

Reading 5 **Sampling and Estimation** 305
Introduction 306
Sampling Methods 306
    Simple Random Sampling 307
    Stratified Random Sampling 308
    Cluster Sampling 309
    Non-Probability Sampling 310
    Sampling from Different Distributions 315
Distribution of the Sample Mean and the Central Limit Theorem 316
    The Central Limit Theorem 317
    Standard Error of the Sample Mean 319

Point Estimates of the Population Mean 322
    Point Estimators 322
Confidence Intervals for the Population Mean and Selection of Sample Size 326
    Selection of Sample Size 332
Resampling 334
Data Snooping Bias, Sample Selection Bias, Look-Ahead Bias, and Time-Period Bias 338
    Data Snooping Bias 338
    Sample Selection Bias 340
    Look-Ahead Bias 342
    Time-Period Bias 342
Summary 344
Practice Problems 347
Solutions 351

Reading 6 **Hypothesis Testing** 357
Introduction 358
    Why Hypothesis Testing? 358
    Implications from a Sampling Distribution 359
The Process of Hypothesis Testing 360
    Stating the Hypotheses 361
    Two-Sided vs. One-Sided Hypotheses 361
    Selecting the Appropriate Hypotheses 362
Identify the Appropriate Test Statistic 363
    Test Statistics 363
    Identifying the Distribution of the Test Statistic 364
Specify the Level of Significance 364
State the Decision Rule 366
    Determining Critical Values 367
    Decision Rules and Confidence Intervals 368
    Collect the Data and Calculate the Test Statistic 369
Make a Decision 370
    Make a Statistical Decision 370
    Make an Economic Decision 370
    Statistically Significant but Not Economically Significant? 370
The Role of *p*-Values 371
Multiple Tests and Interpreting Significance 374
Tests Concerning a Single Mean 377
Test Concerning Differences between Means with Independent Samples 381
Test Concerning Differences between Means with Dependent Samples 384
Testing Concerning Tests of Variances (Chi-Square Test) 388
    Tests of a Single Variance 388
    Test Concerning the Equality of Two Variances (*F*-Test) 391
Parametric vs. Nonparametric Tests 396
    Uses of Nonparametric Tests 397
    Nonparametric Inference: Summary 397
Tests Concerning Correlation 398
    Parametric Test of a Correlation 399

    Tests Concerning Correlation: The Spearman Rank Correlation Coefficient 401
Test of Independence Using Contingency Table Data 404
Summary 409
Practice Problems 412
Solutions 422

Reading 7 **Introduction to Linear Regression** 431
Simple Linear Regression 431
Estimating the Parameters of a Simple Linear Regression 434
    The Basics of Simple Linear Regression 434
    Estimating the Regression Line 435
    Interpreting the Regression Coefficients 438
    Cross-Sectional vs. Time-Series Regressions 440
Assumptions of the Simple Linear Regression Model 443
    Assumption 1: Linearity 443
    Assumption 2: Homoskedasticity 445
    Assumption 3: Independence 447
    Assumption 4: Normality 448
Analysis of Variance 450
    Breaking down the Sum of Squares Total into Its Components 450
    Measures of Goodness of Fit 451
    ANOVA and Standard Error of Estimate in Simple Linear Regression 453
Hypothesis Testing of Linear Regression Coefficients 455
    Hypothesis Tests of the Slope Coefficient 455
    Hypothesis Tests of the Intercept 459
    Hypothesis Tests of Slope When Independent Variable Is an Indicator Variable 459
    Test of Hypotheses: Level of Significance and *p*-Values 461
Prediction Using Simple Linear Regression and Prediction Intervals 463
Functional Forms for Simple Linear Regression 467
    The Log-Lin Model 468
    The Lin-Log Model 469
    The Log-Log Model 470
    Selecting the Correct Functional Form 472
Summary 474
Practice Problems 477
Solutions 490

Appendices 495
Glossary G-1

O indicates an optional segment

O indicates an optional segment

How to Use the CFA Program Curriculum

Congratulations on your decision to enter the Chartered Financial Analyst (CFA®) Program. This exciting and rewarding program of study reflects your desire to become a serious investment professional. You are embarking on a program noted for its high ethical standards and the breadth of knowledge, skills, and abilities (competencies) it develops. Your commitment should be educationally and professionally rewarding.

The credential you seek is respected around the world as a mark of accomplishment and dedication. Each level of the program represents a distinct achievement in professional development. Successful completion of the program is rewarded with membership in a prestigious global community of investment professionals. CFA charterholders are dedicated to life-long learning and maintaining currency with the ever-changing dynamics of a challenging profession. CFA Program enrollment represents the first step toward a career-long commitment to professional education.

The CFA exam measures your mastery of the core knowledge, skills, and abilities required to succeed as an investment professional. These core competencies are the basis for the Candidate Body of Knowledge (CBOK™). The CBOK consists of four components:

- A broad outline that lists the major CFA Program topic areas (www.cfainstitute.org/programs/cfa/curriculum/cbok);
- Topic area weights that indicate the relative exam weightings of the top-level topic areas (www.cfainstitute.org/programs/cfa/curriculum);
- Learning outcome statements (LOS) that advise candidates about the specific knowledge, skills, and abilities they should acquire from readings covering a topic area (LOS are provided in candidate study sessions and at the beginning of each reading); and
- CFA Program curriculum that candidates receive upon exam registration.

Therefore, the key to your success on the CFA exams is studying and understanding the CBOK. The following sections provide background on the CBOK, the organization of the curriculum, features of the curriculum, and tips for designing an effective personal study program.

BACKGROUND ON THE CBOK

CFA Program is grounded in the practice of the investment profession. CFA Institute performs a continuous practice analysis with investment professionals around the world to determine the competencies that are relevant to the profession, beginning with the Global Body of Investment Knowledge (GBIK®). Regional expert panels and targeted surveys are conducted annually to verify and reinforce the continuous feedback about the GBIK. The practice analysis process ultimately defines the CBOK. The CBOK reflects the competencies that are generally accepted and applied by investment professionals. These competencies are used in practice in a generalist context and are expected to be demonstrated by a recently qualified CFA charterholder.

© 2021 CFA Institute. All rights reserved.

The CFA Institute staff—in conjunction with the Education Advisory Committee and Curriculum Level Advisors, who consist of practicing CFA charterholders—designs the CFA Program curriculum in order to deliver the CBOK to candidates. The exams, also written by CFA charterholders, are designed to allow you to demonstrate your mastery of the CBOK as set forth in the CFA Program curriculum. As you structure your personal study program, you should emphasize mastery of the CBOK and the practical application of that knowledge. For more information on the practice analysis, CBOK, and development of the CFA Program curriculum, please visit www.cfainstitute.org.

ORGANIZATION OF THE CURRICULUM

The Level I CFA Program curriculum is organized into 10 topic areas. Each topic area begins with a brief statement of the material and the depth of knowledge expected. It is then divided into one or more study sessions. These study sessions should form the basic structure of your reading and preparation. Each study session includes a statement of its structure and objective and is further divided into assigned readings. An outline illustrating the organization of these study sessions can be found at the front of each volume of the curriculum.

The readings are commissioned by CFA Institute and written by content experts, including investment professionals and university professors. Each reading includes LOS and the core material to be studied, often a combination of text, exhibits, and in-text examples and questions. End of Reading Questions (EORQs) followed by solutions help you understand and master the material. The LOS indicate what you should be able to accomplish after studying the material. The LOS, the core material, and the EORQs are dependent on each other, with the core material and EORQs providing context for understanding the scope of the LOS and enabling you to apply a principle or concept in a variety of scenarios.

The entire readings, including the EORQs, are the basis for all exam questions and are selected or developed specifically to teach the knowledge, skills, and abilities reflected in the CBOK.

You should use the LOS to guide and focus your study because each exam question is based on one or more LOS and the core material and practice problems associated with the LOS. As a candidate, you are responsible for the entirety of the required material in a study session.

We encourage you to review the information about the LOS on our website (www.cfainstitute.org/programs/cfa/curriculum/study-sessions), including the descriptions of LOS "command words" on the candidate resources page at www.cfainstitute.org.

FEATURES OF THE CURRICULUM

End of Reading Questions/Solutions *All End of Reading Questions (EORQs) as well as their solutions are part of the curriculum and are required material for the exam.* In addition to the in-text examples and questions, these EORQs help demonstrate practical applications and reinforce your understanding of the concepts presented. Some of these EORQs are adapted from past CFA exams and/or may serve as a basis for exam questions.

Glossary For your convenience, each volume includes a comprehensive Glossary. Throughout the curriculum, a **bolded** word in a reading denotes a term defined in the Glossary.

Note that the digital curriculum that is included in your exam registration fee is searchable for key words, including Glossary terms.

LOS Self-Check We have inserted checkboxes next to each LOS that you can use to track your progress in mastering the concepts in each reading.

Source Material The CFA Institute curriculum cites textbooks, journal articles, and other publications that provide additional context or information about topics covered in the readings. As a candidate, you are not responsible for familiarity with the original source materials cited in the curriculum.

Note that some readings may contain a web address or URL. The referenced sites were live at the time the reading was written or updated but may have been deactivated since then.

Some readings in the curriculum cite articles published in the *Financial Analysts Journal®*, which is the flagship publication of CFA Institute. Since its launch in 1945, the *Financial Analysts Journal* has established itself as the leading practitioner-oriented journal in the investment management community. Over the years, it has advanced the knowledge and understanding of the practice of investment management through the publication of peer-reviewed practitioner-relevant research from leading academics and practitioners. It has also featured thought-provoking opinion pieces that advance the common level of discourse within the investment management profession. Some of the most influential research in the area of investment management has appeared in the pages of the *Financial Analysts Journal*, and several Nobel laureates have contributed articles.

Candidates are not responsible for familiarity with *Financial Analysts Journal* articles that are cited in the curriculum. But, as your time and studies allow, we strongly encourage you to begin supplementing your understanding of key investment management issues by reading this, and other, CFA Institute practice-oriented publications through the Research & Analysis webpage (www.cfainstitute.org/en/research).

Errata The curriculum development process is rigorous and includes multiple rounds of reviews by content experts. Despite our efforts to produce a curriculum that is free of errors, there are times when we must make corrections. Curriculum errata are periodically updated and posted by exam level and test date online (www.cfainstitute.org/en/programs/submit-errata). If you believe you have found an error in the curriculum, you can submit your concerns through our curriculum errata reporting process found at the bottom of the Curriculum Errata webpage.

DESIGNING YOUR PERSONAL STUDY PROGRAM

Create a Schedule An orderly, systematic approach to exam preparation is critical. You should dedicate a consistent block of time every week to reading and studying. Complete all assigned readings and the associated problems and solutions in each study session. Review the LOS both before and after you study each reading to ensure that

you have mastered the applicable content and can demonstrate the knowledge, skills, and abilities described by the LOS and the assigned reading. Use the LOS self-check to track your progress and highlight areas of weakness for later review.

Successful candidates report an average of more than 300 hours preparing for each exam. Your preparation time will vary based on your prior education and experience, and you will probably spend more time on some study sessions than on others.

You should allow ample time for both in-depth study of all topic areas and additional concentration on those topic areas for which you feel the least prepared.

CFA INSTITUTE LEARNING ECOSYSTEM (LES)

As you prepare for your exam, we will email you important exam updates, testing policies, and study tips. Be sure to read these carefully.

Your exam registration fee includes access to the CFA Program Learning Ecosystem (LES). This digital learning platform provides access, even offline, to all of the readings and End of Reading Questions found in the print curriculum organized as a series of shorter online lessons with associated EORQs. This tool is your one-stop location for all study materials, including practice questions and mock exams.

The LES provides the following supplemental study tools:

Structured and Adaptive Study Plans The LES offers two ways to plan your study through the curriculum. The first is a structured plan that allows you to move through the material in the way that you feel best suits your learning. The second is an adaptive study plan based on the results of an assessment test that uses actual practice questions.

Regardless of your chosen study path, the LES tracks your level of proficiency in each topic area and presents you with a dashboard of where you stand in terms of proficiency so that you can allocate your study time efficiently.

Flashcards and Game Center The LES offers all the Glossary terms as Flashcards and tracks correct and incorrect answers. Flashcards can be filtered both by curriculum topic area and by action taken—for example, answered correctly, unanswered, and so on. These Flashcards provide a flexible way to study Glossary item definitions.

The Game Center provides several engaging ways to interact with the Flashcards in a game context. Each game tests your knowledge of the Glossary terms a in different way. Your results are scored and presented, along with a summary of candidates with high scores on the game, on your Dashboard.

Discussion Board The Discussion Board within the LES provides a way for you to interact with other candidates as you pursue your study plan. Discussions can happen at the level of individual lessons to raise questions about material in those lessons that you or other candidates can clarify or comment on. Discussions can also be posted at the level of topics or in the initial Welcome section to connect with other candidates in your area.

Practice Question Bank The LES offers access to a question bank of hundreds of practice questions that are in addition to the End of Reading Questions. These practice questions, only available on the LES, are intended to help you assess your mastery of individual topic areas as you progress through your studies. After each practice question, you will receive immediate feedback noting the correct response and indicating the relevant assigned reading so you can identify areas of weakness for further study.

Mock Exams The LES also includes access to three-hour Mock Exams that simulate the morning and afternoon sessions of the actual CFA exam. These Mock Exams are intended to be taken after you complete your study of the full curriculum and take practice questions so you can test your understanding of the curriculum and your readiness for the exam. If you take these Mock Exams within the LES, you will receive feedback afterward that notes the correct responses and indicates the relevant assigned readings so you can assess areas of weakness for further study. We recommend that you take Mock Exams during the final stages of your preparation for the actual CFA exam. For more information on the Mock Exams, please visit www.cfainstitute.org.

PREP PROVIDERS

You may choose to seek study support outside CFA Institute in the form of exam prep providers. After your CFA Program enrollment, you may receive numerous solicitations for exam prep courses and review materials. When considering a prep course, make sure the provider is committed to following the CFA Institute guidelines and high standards in its offerings.

Remember, however, that there are no shortcuts to success on the CFA exams; reading and studying the CFA Program curriculum *is* the key to success on the exam. The CFA Program exams reference only the CFA Institute assigned curriculum; no prep course or review course materials are consulted or referenced.

SUMMARY

Every question on the CFA exam is based on the content contained in the required readings and on one or more LOS. Frequently, an exam question is based on a specific example highlighted within a reading or on a specific practice problem and its solution. To make effective use of the CFA Program curriculum, please remember these key points:

1. All pages of the curriculum are required reading for the exam.
2. All questions, problems, and their solutions are part of the curriculum and are required study material for the exam. These questions are found at the end of the readings in the print versions of the curriculum. In the LES, these questions appear directly after the lesson with which they are associated. The LES provides immediate feedback on your answers and tracks your performance on these questions throughout your study.
3. We strongly encourage you to use the CFA Program Learning Ecosystem. In addition to providing access to all the curriculum material, including EORQs, in the form of shorter, focused lessons, the LES offers structured and adaptive study planning, a Discussion Board to communicate with other candidates, Flashcards, a Game Center for study activities, a test bank of practice questions, and online Mock Exams. Other supplemental study tools, such as eBook and PDF versions of the print curriculum, and additional candidate resources are available at www.cfainstitute.org.
4. Using the study planner, create a schedule and commit sufficient study time to cover the study sessions. You should also plan to review the materials, answer practice questions, and take Mock Exams.
5. Some of the concepts in the study sessions may be superseded by updated rulings and/or pronouncements issued after a reading was published. Candidates are expected to be familiar with the overall analytical framework contained in the assigned readings. Candidates are not responsible for changes that occur after the material was written.

FEEDBACK

At CFA Institute, we are committed to delivering a comprehensive and rigorous curriculum for the development of competent, ethically grounded investment professionals. We rely on candidate and investment professional comments and feedback as we work to improve the curriculum, supplemental study tools, and candidate resources.

Please send any comments or feedback to info@cfainstitute.org. You can be assured that we will review your suggestions carefully. Ongoing improvements in the curriculum will help you prepare for success on the upcoming exams and for a lifetime of learning as a serious investment professional.

Quantitative Methods

STUDY SESSIONS

Study Session 1	Quantitative Methods (1)
Study Session 2	Quantitative Methods (2)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to explain and demonstrate the use of time value of money, data collection and analysis, elementary statistics, probability theory, probability distribution theory, sampling and estimation, hypothesis testing, and simple linear regression in financial decision-making.

The quantitative concepts and applications that follow are fundamental to financial analysis and are used throughout the CFA Program curriculum. Quantitative methods are used widely in securities and risk analysis and in corporate finance to value capital projects and select investments. Descriptive statistics provide the tools to characterize and assess risk and return and other important financial or economic variables. Probability theory, sampling and estimation, and hypothesis testing support investment and risk decision making in the presence of uncertainty. Simple linear regression helps to understand the relationship between two variables and how to make predictions.

© 2021 CFA Institute. All rights reserved.

QUANTITATIVE METHODS

STUDY SESSION

1

Quantitative Methods (1)

This study session introduces quantitative concepts and techniques used in financial analysis and investment decision making. The time value of money and discounted cash flow analysis form the basis for cash flow and security valuation. Methods for organizing and visualizing data are presented; these key skills are required for effectively performing financial analysis. Descriptive statistics used for conveying important data attributes such as central tendency, location, and dispersion are also presented. Characteristics of return distributions such as symmetry, skewness, and kurtosis are also introduced. Finally, all investment forecasts and decisions involve uncertainty: Therefore, probability theory and its application quantifying risk in investment decision making is considered.

READING ASSIGNMENTS

Reading 1	The Time Value of Money by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, DBA, CFA, Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA
Reading 2	Organizing, Visualizing, and Describing Data by Pamela Peterson Drake, PhD, CFA, and Jian Wu, PhD
Reading 3	Probability Concepts by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, DBA, CFA, Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA

© 2021 CFA Institute. All rights reserved.

READING

1

The Time Value of Money

by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, DBA, CFA, Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA

Richard A. DeFusco, PhD, CFA, is at the University of Nebraska-Lincoln (USA). Dennis W. McLeavey, DBA, CFA, is at the University of Rhode Island (USA). Jerald E. Pinto, PhD, CFA, is at CFA Institute (USA). David E. Runkle, PhD, CFA, is at Jacobs Levy Equity Management (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** interpret interest rates as required rates of return, discount rates, or opportunity costs;
☐	**b.** explain an interest rate as the sum of a real risk-free rate and premiums that compensate investors for bearing distinct types of risk;
☐	**c.** calculate and interpret the effective annual rate, given the stated annual interest rate and the frequency of compounding;
☐	**d.** calculate the solution for time value of money problems with different frequencies of compounding;
☐	**e.** calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;
☐	**f.** demonstrate the use of a time line in modeling and solving time value of money problems.

INTRODUCTION

1

As individuals, we often face decisions that involve saving money for a future use, or borrowing money for current consumption. We then need to determine the amount we need to invest, if we are saving, or the cost of borrowing, if we are shopping for a loan. As investment analysts, much of our work also involves evaluating transactions with present and future cash flows. When we place a value on any security, for example, we are attempting to determine the worth of a stream of future cash flows. To carry out all the above tasks accurately, we must understand the mathematics of time value of money problems. Money has time value in that individuals value a given

© 2019 CFA Institute. All rights reserved.

amount of money more highly the earlier it is received. Therefore, a smaller amount of money now may be equivalent in value to a larger amount received at a future date. The **time value of money** as a topic in investment mathematics deals with equivalence relationships between cash flows with different dates. Mastery of time value of money concepts and techniques is essential for investment analysts.

The reading[1] is organized as follows: Section 2 introduces some terminology used throughout the reading and supplies some economic intuition for the variables we will discuss. Sections 3–5 tackle the problem of determining the worth at a future point in time of an amount invested today. Section 6 addresses the future worth of a series of cash flows. These two sections provide the tools for calculating the equivalent value at a future date of a single cash flow or series of cash flows. Sections 7–10 discuss the equivalent value today of a single future cash flow and a series of future cash flows, respectively. In Sections 11–13, we explore how to determine other quantities of interest in time value of money problems.

2

INTEREST RATES

a interpret interest rates as required rates of return, discount rates, or opportunity costs;

b explain an interest rate as the sum of a real risk-free rate and premiums that compensate investors for bearing distinct types of risk;

In this reading, we will continually refer to interest rates. In some cases, we assume a particular value for the interest rate; in other cases, the interest rate will be the unknown quantity we seek to determine. Before turning to the mechanics of time value of money problems, we must illustrate the underlying economic concepts. In this section, we briefly explain the meaning and interpretation of interest rates.

Time value of money concerns equivalence relationships between cash flows occurring on different dates. The idea of equivalence relationships is relatively simple. Consider the following exchange: You pay \$10,000 today and in return receive \$9,500 today. Would you accept this arrangement? Not likely. But what if you received the \$9,500 today and paid the \$10,000 one year from now? Can these amounts be considered equivalent? Possibly, because a payment of \$10,000 a year from now would probably be worth less to you than a payment of \$10,000 today. It would be fair, therefore, to **discount** the \$10,000 received in one year; that is, to cut its value based on how much time passes before the money is paid. An **interest rate**, denoted r, is a rate of return that reflects the relationship between differently dated cash flows. If \$9,500 today and \$10,000 in one year are equivalent in value, then \$10,000 – \$9,500 = \$500 is the required compensation for receiving \$10,000 in one year rather than now. The interest rate—the required compensation stated as a rate of return—is \$500/\$9,500 = 0.0526 or 5.26 percent.

Interest rates can be thought of in three ways. First, they can be considered required rates of return—that is, the minimum rate of return an investor must receive in order to accept the investment. Second, interest rates can be considered discount rates. In the example above, 5.26 percent is that rate at which we discounted the \$10,000 future amount to find its value today. Thus, we use the terms "interest rate" and "discount rate" almost interchangeably. Third, interest rates can be considered opportunity costs. An **opportunity cost** is the value that investors forgo by choosing a particular course of action. In the example, if the party who supplied $9,500 had instead decided to spend it today, he would have forgone earning 5.26 percent on the money. So we can view 5.26 percent as the opportunity cost of current consumption.

---

[1] Examples in this reading and other readings in quantitative methods at Level I were updated in 2018 by Professor Sanjiv Sabherwal of the University of Texas, Arlington.

Economics tells us that interest rates are set in the marketplace by the forces of supply and demand, where investors are suppliers of funds and borrowers are demanders of funds. Taking the perspective of investors in analyzing market-determined interest rates, we can view an interest rate r as being composed of a real risk-free interest rate plus a set of four premiums that are required returns or compensation for bearing distinct types of risk:

$$r = \text{Real risk-free interest rate} + \text{Inflation premium} + \text{Default risk premium} + \text{Liquidity premium} + \text{Maturity premium}$$

- The **real risk-free interest rate** is the single-period interest rate for a completely risk-free security if no inflation were expected. In economic theory, the real risk-free rate reflects the time preferences of individuals for current versus future real consumption.
- The **inflation premium** compensates investors for expected inflation and reflects the average inflation rate expected over the maturity of the debt. Inflation reduces the purchasing power of a unit of currency—the amount of goods and services one can buy with it. The sum of the real risk-free interest rate and the inflation premium is the **nominal risk-free interest rate**.[2] Many countries have governmental short-term debt whose interest rate can be considered to represent the nominal risk-free interest rate in that country. The interest rate on a 90-day US Treasury bill (T-bill), for example, represents the nominal risk-free interest rate over that time horizon.[3] US T-bills can be bought and sold in large quantities with minimal transaction costs and are backed by the full faith and credit of the US government.
- The **default risk premium** compensates investors for the possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount.
- The **liquidity premium** compensates investors for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly. US T-bills, for example, do not bear a liquidity premium because large amounts can be bought and sold without affecting their market price. Many bonds of small issuers, by contrast, trade infrequently after they are issued; the interest rate on such bonds includes a liquidity premium reflecting the relatively high costs (including the impact on price) of selling a position.
- The **maturity premium** compensates investors for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended, in general (holding all else equal). The difference between the interest

---

[2] Technically, 1 plus the nominal rate equals the product of 1 plus the real rate and 1 plus the inflation rate. As a quick approximation, however, the nominal rate is equal to the real rate plus an inflation premium. In this discussion we focus on approximate additive relationships to highlight the underlying concepts.

[3] Other developed countries issue securities similar to US Treasury bills. The French government issues BTFs or negotiable fixed-rate discount Treasury bills (*Bons du Trésor à taux fixe et à intérêts précomptés*) with maturities of up to one year. The Japanese government issues a short-term Treasury bill with maturities of 6 and 12 months. The German government issues at discount both Treasury financing paper (*Finanzierungsschätze des Bundes* or, for short, *Schätze*) and Treasury discount paper (*Bubills*) with maturities up to 24 months. In the United Kingdom, the British government issues gilt-edged Treasury bills with maturities ranging from 1 to 364 days. The Canadian government bond market is closely related to the US market; Canadian Treasury bills have maturities of 3, 6, and 12 months.

rate on longer-maturity, liquid Treasury debt and that on short-term Treasury debt reflects a positive maturity premium for the longer-term debt (and possibly different inflation premiums as well).

Using this insight into the economic meaning of interest rates, we now turn to a discussion of solving time value of money problems, starting with the future value of a single cash flow.

3

FUTURE VALUE OF A SINGLE CASH FLOW (LUMP SUM)

e calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;

f demonstrate the use of a time line in modeling and solving time value of money problems.

In this section, we introduce time value associated with a single cash flow or lump-sum investment. We describe the relationship between an initial investment or **present value (PV)**, which earns a rate of return (the interest rate per period) denoted as r, and its **future value (FV)**, which will be received N years or periods from today.

The following example illustrates this concept. Suppose you invest \$100 (PV = \$100) in an interest-bearing bank account paying 5 percent annually. At the end of the first year, you will have the \$100 plus the interest earned, 0.05 × \$100 = \$5, for a total of \$105. To formalize this one-period example, we define the following terms:

PV = present value of the investment
FV_N = future value of the investment N periods from today
r = rate of interest per period

For $N = 1$, the expression for the future value of amount PV is

$$FV_1 = PV(1 + r) \quad (1)$$

For this example, we calculate the future value one year from today as $FV_1 = \$100(1.05) = \$105$.

Now suppose you decide to invest the initial \$100 for two years with interest earned and credited to your account annually (annual compounding). At the end of the first year (the beginning of the second year), your account will have \$105, which you will leave in the bank for another year. Thus, with a beginning amount of \$105 (PV = \$105), the amount at the end of the second year will be \$105(1.05) = \$110.25. Note that the \$5.25 interest earned during the second year is 5 percent of the amount invested at the beginning of Year 2.

Another way to understand this example is to note that the amount invested at the beginning of Year 2 is composed of the original \$100 that you invested plus the \$5 interest earned during the first year. During the second year, the original principal again earns interest, as does the interest that was earned during Year 1. You can see how the original investment grows:

Original investment	\$100.00
Interest for the first year (\$100 × 0.05)	5.00
Interest for the second year based on original investment (\$100 × 0.05)	5.00
Interest for the second year based on interest earned in the first year (0.05 × \$5.00 interest on interest)	0.25
Total	\$110.25

The \$5 interest that you earned each period on the \$100 original investment is known as **simple interest** (the interest rate times the principal). **Principal** is the amount of funds originally invested. During the two-year period, you earn \$10 of simple interest. The extra \$0.25 that you have at the end of Year 2 is the interest you earned on the Year 1 interest of \$5 that you reinvested.

The interest earned on interest provides the first glimpse of the phenomenon known as **compounding**. Although the interest earned on the initial investment is important, for a given interest rate it is fixed in size from period to period. The compounded interest earned on reinvested interest is a far more powerful force because, for a given interest rate, it grows in size each period. The importance of compounding increases with the magnitude of the interest rate. For example, \$100 invested today would be worth about \$13,150 after 100 years if compounded annually at 5 percent, but worth more than \$20 million if compounded annually over the same time period at a rate of 13 percent.

To verify the \$20 million figure, we need a general formula to handle compounding for any number of periods. The following general formula relates the present value of an initial investment to its future value after N periods:

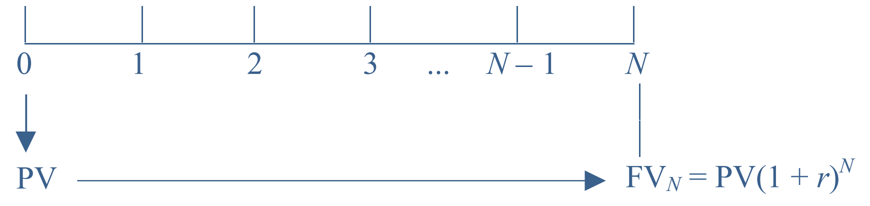
$$FV_N = PV(1 + r)^N \tag{2}$$

where r is the stated interest rate per period and N is the number of compounding periods. In the bank example, $FV_2 = \$100(1 + 0.05)^2 = \$110.25$. In the 13 percent investment example, $FV_{100} = \$100(1.13)^{100} = \$20{,}316{,}287.42$.

The most important point to remember about using the future value equation is that the stated interest rate, r, and the number of compounding periods, N, must be compatible. Both variables must be defined in the same time units. For example, if N is stated in months, then r should be the one-month interest rate, unannualized.

A time line helps us to keep track of the compatibility of time units and the interest rate per time period. In the time line, we use the time index t to represent a point in time a stated number of periods from today. Thus the present value is the amount available for investment today, indexed as $t = 0$. We can now refer to a time N periods from today as $t = N$. The time line in Figure 1 shows this relationship.

Figure 1 The Relationship between an Initial Investment, PV, and Its Future Value, FV



In Figure 1, we have positioned the initial investment, PV, at $t = 0$. Using Equation 2, we move the present value, PV, forward to $t = N$ by the factor $(1 + r)^N$. This factor is called a future value factor. We denote the future value on the time line as FV and position it at $t = N$. Suppose the future value is to be received exactly 10 periods from today's date ($N = 10$). The present value, PV, and the future value, FV, are separated in time through the factor $(1 + r)^{10}$.

The fact that the present value and the future value are separated in time has important consequences:

- We can add amounts of money only if they are indexed at the same point in time.
- For a given interest rate, the future value increases with the number of periods.
- For a given number of periods, the future value increases with the interest rate.

To better understand these concepts, consider three examples that illustrate how to apply the future value formula.

EXAMPLE 1

The Future Value of a Lump Sum with Interim Cash Reinvested at the Same Rate

You are the lucky winner of your state's lottery of \$5 million after taxes. You invest your winnings in a five-year certificate of deposit (CD) at a local financial institution. The CD promises to pay 7 percent per year compounded annually. This institution also lets you reinvest the interest at that rate for the duration of the CD. How much will you have at the end of five years if your money remains invested at 7 percent for five years with no withdrawals?

Solution:

To solve this problem, compute the future value of the \$5 million investment using the following values in Equation 2:

$$\begin{aligned} \text{PV} &= \$5{,}000{,}000 \\ r &= 7\% = 0.07 \\ N &= 5 \\ \text{FV}_N &= \text{PV}(1 + r)^N \\ &= \$5{,}000{,}000(1.07)^5 \\ &= \$5{,}000{,}000(1.402552) \\ &= \$7{,}012{,}758.65 \end{aligned}$$

At the end of five years, you will have \$7,012,758.65 if your money remains invested at 7 percent with no withdrawals.

In this and most examples in this reading, note that the factors are reported at six decimal places but the calculations may actually reflect greater precision. For example, the reported 1.402552 has been rounded up from 1.40255173 (the calculation is actually carried out with more than eight decimal places of precision by the calculator or spreadsheet). Our final result reflects the higher number of decimal places carried by the calculator or spreadsheet.[4]

---

[4] We could also solve time value of money problems using tables of interest rate factors. Solutions using tabled values of interest rate factors are generally less accurate than solutions obtained using calculators or spreadsheets, so practitioners prefer calculators or spreadsheets.

EXAMPLE 2

The Future Value of a Lump Sum with No Interim Cash

An institution offers you the following terms for a contract: For an investment of ¥2,500,000, the institution promises to pay you a lump sum six years from now at an 8 percent annual interest rate. What future amount can you expect?

Solution:

Use the following data in Equation 2 to find the future value:

$$\begin{aligned}
\text{PV} &= ¥2{,}500{,}000\\
r &= 8\% = 0.08\\
N &= 6\\
\text{FV}_N &= \text{PV}(1+r)^N\\
&= ¥2{,}500{,}000(1.08)^6\\
&= ¥2{,}500{,}000(1.586874)\\
&= ¥3{,}967{,}186
\end{aligned}$$

You can expect to receive ¥3,967,186 six years from now.

Our third example is a more complicated future value problem that illustrates the importance of keeping track of actual calendar time.

EXAMPLE 3

The Future Value of a Lump Sum

A pension fund manager estimates that his corporate sponsor will make a \$10 million contribution five years from now. The rate of return on plan assets has been estimated at 9 percent per year. The pension fund manager wants to calculate the future value of this contribution 15 years from now, which is the date at which the funds will be distributed to retirees. What is that future value?

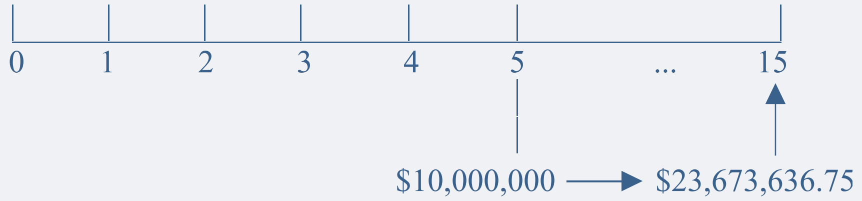
Solution:

By positioning the initial investment, PV, at $t = 5$, we can calculate the future value of the contribution using the following data in Equation 2:

$$\begin{aligned}
\text{PV} &= \$10 \text{ million}\\
r &= 9\% = 0.09\\
N &= 10\\
\text{FV}_N &= \text{PV}(1+r)^N\\
&= \$10{,}000{,}000(1.09)^{10}\\
&= \$10{,}000{,}000(2.367364)\\
&= \$23{,}673{,}636.75
\end{aligned}$$

This problem looks much like the previous two, but it differs in one important respect: its timing. From the standpoint of today ($t = 0$), the future amount of \$23,673,636.75 is 15 years into the future. Although the future value is 10 years from its present value, the present value of \$10 million will not be received for another five years.

Figure 2 The Future Value of a Lump Sum, Initial Investment Not at *t* = 0



As Figure 2 shows, we have followed the convention of indexing today as $t = 0$ and indexing subsequent times by adding 1 for each period. The additional contribution of \$10 million is to be received in five years, so it is indexed as $t = 5$ and appears as such in the figure. The future value of the investment in 10 years is then indexed at $t = 15$; that is, 10 years following the receipt of the \$10 million contribution at $t = 5$. Time lines like this one can be extremely useful when dealing with more-complicated problems, especially those involving more than one cash flow.

In a later section of this reading, we will discuss how to calculate the value today of the \$10 million to be received five years from now. For the moment, we can use Equation 2. Suppose the pension fund manager in Example 3 above were to receive \$6,499,313.86 today from the corporate sponsor. How much will that sum be worth at the end of five years? How much will it be worth at the end of 15 years?

$$\begin{aligned} \text{PV} &= \$6{,}499{,}313.86 \\ r &= 9\% = 0.09 \\ N &= 5 \\ \text{FV}_N &= \text{PV}(1+r)^N \\ &= \$6{,}499{,}313.86(1.09)^5 \\ &= \$6{,}499{,}313.86(1.538624) \\ &= \$10{,}000{,}000 \text{ at the five-year mark} \end{aligned}$$

and

$$\begin{aligned} \text{PV} &= \$6{,}499{,}313.86 \\ r &= 9\% = 0.09 \\ N &= 15 \\ \text{FV}_N &= \text{PV}(1+r)^N \\ &= \$6{,}499{,}313.86(1.09)^{15} \\ &= \$6{,}499{,}313.86(3.642482) \\ &= \$23{,}673{,}636.74 \text{ at the 15-year mark} \end{aligned}$$

These results show that today's present value of about \$6.5 million becomes \$10 million after five years and \$23.67 million after 15 years.

NON-ANNUAL COMPOUNDING (FUTURE VALUE)

4

d calculate the solution for time value of money problems with different frequencies of compounding;

In this section, we examine investments paying interest more than once a year. For instance, many banks offer a monthly interest rate that compounds 12 times a year. In such an arrangement, they pay interest on interest every month. Rather than quote the periodic monthly interest rate, financial institutions often quote an annual interest rate that we refer to as the **stated annual interest rate** or **quoted interest rate**. We denote the stated annual interest rate by r_s. For instance, your bank might state that a particular CD pays 8 percent compounded monthly. The stated annual interest rate equals the monthly interest rate multiplied by 12. In this example, the monthly interest rate is 0.08/12 = 0.0067 or 0.67 percent.[5] This rate is strictly a quoting convention because $(1 + 0.0067)^{12} = 1.083$, not 1.08; the term $(1 + r_s)$ is not meant to be a future value factor when compounding is more frequent than annual.

With more than one compounding period per year, the future value formula can be expressed as

$$\mathrm{FV}_N = \mathrm{PV}\left(1 + \frac{r_s}{m}\right)^{mN} \qquad (3)$$

where r_s is the stated annual interest rate, m is the number of compounding periods per year, and N now stands for the number of years. Note the compatibility here between the interest rate used, r_s/m, and the number of compounding periods, mN. The periodic rate, r_s/m, is the stated annual interest rate divided by the number of compounding periods per year. The number of compounding periods, mN, is the number of compounding periods in one year multiplied by the number of years. The periodic rate, r_s/m, and the number of compounding periods, mN, must be compatible.

EXAMPLE 4

The Future Value of a Lump Sum with Quarterly Compounding

Continuing with the CD example, suppose your bank offers you a CD with a two-year maturity, a stated annual interest rate of 8 percent compounded quarterly, and a feature allowing reinvestment of the interest at the same interest rate. You decide to invest $10,000. What will the CD be worth at maturity?

---

[5] To avoid rounding errors when using a financial calculator, divide 8 by 12 and then press the %*i* key, rather than simply entering 0.67 for %*i*, so we have $(1 + 0.08/12)^{12} = 1.083000$.

Solution:

Compute the future value with Equation 3 as follows:

$$\begin{aligned}
\text{PV} &= \$10{,}000\\
r_s &= 8\% = 0.08\\
m &= 4\\
r_s/m &= 0.08/4 = 0.02\\
N &= 2\\
mN &= 4(2) = 8 \text{ interest periods}\\
\text{FV}_N &= \text{PV}\left(1+\frac{r_s}{m}\right)^{mN}\\
&= \$10{,}000(1.02)^8\\
&= \$10{,}000(1.171659)\\
&= \$11{,}716.59
\end{aligned}$$

At maturity, the CD will be worth \$11,716.59.

The future value formula in Equation 3 does not differ from the one in Equation 2. Simply keep in mind that the interest rate to use is the rate per period and the exponent is the number of interest, or compounding, periods.

EXAMPLE 5

The Future Value of a Lump Sum with Monthly Compounding

An Australian bank offers to pay you 6 percent compounded monthly. You decide to invest A\$1 million for one year. What is the future value of your investment if interest payments are reinvested at 6 percent?

Solution:

Use Equation 3 to find the future value of the one-year investment as follows:

$$\begin{aligned}
\text{PV} &= \text{A}\$1{,}000{,}000\\
r_s &= 6\% = 0.06\\
m &= 12\\
r_s/m &= 0.06/12 = 0.0050\\
N &= 1\\
mN &= 12(1) = 12 \text{ interest periods}\\
\text{FV}_N &= \text{PV}\left(1+\frac{r_s}{m}\right)^{mN}\\
&= \text{A}\$1{,}000{,}000(1.005)^{12}\\
&= \text{A}\$1{,}000{,}000(1.061678)\\
&= \text{A}\$1{,}061{,}677.81
\end{aligned}$$

If you had been paid 6 percent with annual compounding, the future amount would be only A\$1,000,000(1.06) = A\$1,060,000 instead of A\$1,061,677.81 with monthly compounding.

CONTINUOUS COMPOUNDING, STATED AND EFFECTIVE RATES

5

c calculate and interpret the effective annual rate, given the stated annual interest rate and the frequency of compounding;

d calculate the solution for time value of money problems with different frequencies of compounding;

The preceding discussion on compounding periods illustrates discrete compounding, which credits interest after a discrete amount of time has elapsed. If the number of compounding periods per year becomes infinite, then interest is said to compound continuously. If we want to use the future value formula with continuous compounding, we need to find the limiting value of the future value factor for $m \rightarrow \infty$ (infinitely many compounding periods per year) in Equation 3. The expression for the future value of a sum in N years with continuous compounding is

$$\mathrm{FV}_N = \mathrm{PV}e^{r_s N} \tag{4}$$

The term $e^{r_s N}$ is the transcendental number $e \approx 2.7182818$ raised to the power $r_s N$. Most financial calculators have the function e^x.

EXAMPLE 6

The Future Value of a Lump Sum with Continuous Compounding

Suppose a \$10,000 investment will earn 8 percent compounded continuously for two years. We can compute the future value with Equation 4 as follows:

$$\begin{aligned}
\mathrm{PV} &= \$10{,}000 \\
r_s &= 8\% = 0.08 \\
N &= 2 \\
\mathrm{FV}_N &= \mathrm{PV}e^{r_s N} \\
&= \$10{,}000e^{0.08(2)} \\
&= \$10{,}000(1.173511) \\
&= \$11{,}735.11
\end{aligned}$$

With the same interest rate but using continuous compounding, the \$10,000 investment will grow to \$11,735.11 in two years, compared with \$11,716.59 using quarterly compounding as shown in Example 4.

Table 1 shows how a stated annual interest rate of 8 percent generates different ending dollar amounts with annual, semiannual, quarterly, monthly, daily, and continuous compounding for an initial investment of \$1 (carried out to six decimal places).

As Table 1 shows, all six cases have the same stated annual interest rate of 8 percent; they have different ending dollar amounts, however, because of differences in the frequency of compounding. With annual compounding, the ending amount is \$1.08. More frequent compounding results in larger ending amounts. The ending dollar amount with continuous compounding is the maximum amount that can be earned with a stated annual rate of 8 percent.

Table 1 The Effect of Compounding Frequency on Future Value

Frequency	r_s/m	mN	Future Value of $1		
Annual	8%/1 = 8%	1 × 1 = 1	$\$1.00(1.08)$	=	$1.08
Semiannual	8%/2 = 4%	2 × 1 = 2	$\$1.00(1.04)^2$	=	$1.081600
Quarterly	8%/4 = 2%	4 × 1 = 4	$\$1.00(1.02)^4$	=	$1.082432
Monthly	8%/12 = 0.6667%	12 × 1 = 12	$\$1.00(1.006667)^{12}$	=	$1.083000
Daily	8%/365 = 0.0219%	365 × 1 = 365	$\$1.00(1.000219)^{365}$	=	$1.083278
Continuous			$\$1.00e^{0.08(1)}$	=	$1.083287

Table 1 also shows that a $1 investment earning 8.16 percent compounded annually grows to the same future value at the end of one year as a $1 investment earning 8 percent compounded semiannually. This result leads us to a distinction between the stated annual interest rate and the **effective annual rate** (EAR).[6] For an 8 percent stated annual interest rate with semiannual compounding, the EAR is 8.16 percent.

5.1 Stated and Effective Rates

The stated annual interest rate does not give a future value directly, so we need a formula for the EAR. With an annual interest rate of 8 percent compounded semiannually, we receive a periodic rate of 4 percent. During the course of a year, an investment of $1 would grow to $\$1(1.04)^2 = \$1.0816$, as illustrated in Table 1. The interest earned on the $1 investment is $0.0816 and represents an effective annual rate of interest of 8.16 percent. The effective annual rate is calculated as follows:

$$\text{EAR} = (1 + \text{Periodic interest rate})^m - 1 \tag{5}$$

The periodic interest rate is the stated annual interest rate divided by m, where m is the number of compounding periods in one year. Using our previous example, we can solve for EAR as follows: $(1.04)^2 - 1 = 8.16$ percent.

The concept of EAR extends to continuous compounding. Suppose we have a rate of 8 percent compounded continuously. We can find the EAR in the same way as above by finding the appropriate future value factor. In this case, a $1 investment would grow to $\$1e^{0.08(1.0)} = \$1.0833$. The interest earned for one year represents an effective annual rate of 8.33 percent and is larger than the 8.16 percent EAR with semiannual compounding because interest is compounded more frequently. With continuous compounding, we can solve for the effective annual rate as follows:

$$\text{EAR} = e^{r_s} - 1 \tag{6}$$

---

[6] Among the terms used for the effective annual return on interest-bearing bank deposits are annual percentage yield (APY) in the United States and equivalent annual rate (EAR) in the United Kingdom. By contrast, the **annual percentage rate** (APR) measures the cost of borrowing expressed as a yearly rate. In the United States, the APR is calculated as a periodic rate times the number of payment periods per year and, as a result, some writers use APR as a general synonym for the stated annual interest rate. Nevertheless, APR is a term with legal connotations; its calculation follows regulatory standards that vary internationally. Therefore, “stated annual interest rate” is the preferred general term for an annual interest rate that does not account for compounding within the year.

We can reverse the formulas for EAR with discrete and continuous compounding to find a periodic rate that corresponds to a particular effective annual rate. Suppose we want to find the appropriate periodic rate for a given effective annual rate of 8.16 percent with semiannual compounding. We can use Equation 5 to find the periodic rate:

$$\begin{aligned} 0.0816 &= (1 + \text{Periodic rate})^2 - 1 \\ 1.0816 &= (1 + \text{Periodic rate})^2 \\ (1.0816)^{1/2} - 1 &= \text{Periodic rate} \\ (1.04) - 1 &= \text{Periodic rate} \\ 4\% &= \text{Periodic rate} \end{aligned}$$

To calculate the continuously compounded rate (the stated annual interest rate with continuous compounding) corresponding to an effective annual rate of 8.33 percent, we find the interest rate that satisfies Equation 6:

$$\begin{aligned} 0.0833 &= e^{r_s} - 1 \\ 1.0833 &= e^{r_s} \end{aligned}$$

To solve this equation, we take the natural logarithm of both sides. (Recall that the natural log of e^{r_s} is $\ln e^{r_s} = r_s$.) Therefore, $\ln 1.0833 = r_s$, resulting in $r_s = 8$ percent. We see that a stated annual rate of 8 percent with continuous compounding is equivalent to an EAR of 8.33 percent.

FUTURE VALUE OF A SERIES OF CASH FLOWS, FUTURE VALUE ANNUITIES

6

e calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;

f demonstrate the use of a time line in modeling and solving time value of money problems.

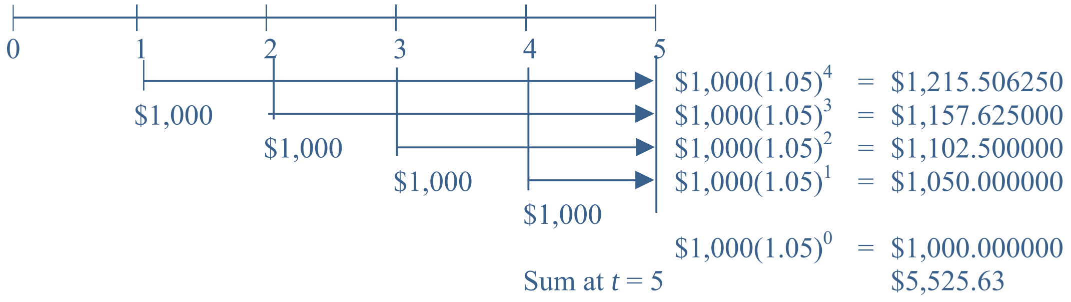
In this section, we consider series of cash flows, both even and uneven. We begin with a list of terms commonly used when valuing cash flows that are distributed over many time periods.

- An **annuity** is a finite set of level sequential cash flows.
- An **ordinary annuity** has a first cash flow that occurs one period from now (indexed at $t = 1$).
- An **annuity due** has a first cash flow that occurs immediately (indexed at $t = 0$).
- A **perpetuity** is a perpetual annuity, or a set of level never-ending sequential cash flows, with the first cash flow occurring one period from now.

6.1 Equal Cash Flows—Ordinary Annuity

Consider an ordinary annuity paying 5 percent annually. Suppose we have five separate deposits of $1,000 occurring at equally spaced intervals of one year, with the first payment occurring at $t = 1$. Our goal is to find the future value of this ordinary annuity after the last deposit at $t = 5$. The increment in the time counter is one year, so the last payment occurs five years from now. As the time line in Figure 3 shows, we find the future value of each \$1,000 deposit as of $t = 5$ with Equation 2, $\mathrm{FV}_N = \mathrm{PV}(1 + r)^N$. The arrows in Figure 3 extend from the payment date to $t = 5$. For instance, the first \$1,000 deposit made at $t = 1$ will compound over four periods. Using Equation 2, we find that the future value of the first deposit at $t = 5$ is $\$1{,}000(1.05)^4 = \$1{,}215.51$. We calculate the future value of all other payments in a similar fashion. (Note that we are finding the future value at $t = 5$, so the last payment does not earn any interest.) With all values now at $t = 5$, we can add the future values to arrive at the future value of the annuity. This amount is \$5,525.63.

Figure 3 The Future Value of a Five-Year Ordinary Annuity



We can arrive at a general annuity formula if we define the annuity amount as A, the number of time periods as N, and the interest rate per period as r. We can then define the future value as

$$\mathrm{FV}_N = A\left[(1 + r)^{N-1} + (1 + r)^{N-2} + (1 + r)^{N-3} + \ldots + (1 + r)^1 + (1 + r)^0\right]$$

which simplifies to

$$\mathrm{FV}_N = A\left[\frac{(1 + r)^N - 1}{r}\right] \tag{7}$$

The term in brackets is the future value annuity factor. This factor gives the future value of an ordinary annuity of \$1 per period. Multiplying the future value annuity factor by the annuity amount gives the future value of an ordinary annuity. For the ordinary annuity in Figure 3, we find the future value annuity factor from Equation 7 as

$$\left[\frac{(1.05)^5 - 1}{0.05}\right] = 5.525631$$

With an annuity amount $A = \$1{,}000$, the future value of the annuity is $\$1{,}000(5.525631) = \$5{,}525.63$, an amount that agrees with our earlier work.

The next example illustrates how to find the future value of an ordinary annuity using the formula in Equation 7.

EXAMPLE 7

The Future Value of an Annuity

Suppose your company's defined contribution retirement plan allows you to invest up to €20,000 per year. You plan to invest €20,000 per year in a stock index fund for the next 30 years. Historically, this fund has earned 9 percent per year on average. Assuming that you actually earn 9 percent a year, how much money will you have available for retirement after making the last payment?

Solution:

Use Equation 7 to find the future amount:

$$A = €20{,}000$$
$$r = 9\% = 0.09$$
$$N = 30$$
$$\text{FV annuity factor} = \frac{(1+r)^N - 1}{r} = \frac{(1.09)^{30} - 1}{0.09} = 136.307539$$
$$\begin{aligned} \text{FV}_N &= €20{,}000(136.307539) \\ &= €2{,}726{,}150.77 \end{aligned}$$

Assuming the fund continues to earn an average of 9 percent per year, you will have €2,726,150.77 available at retirement.

6.2 Unequal Cash Flows

In many cases, cash flow streams are unequal, precluding the simple use of the future value annuity factor. For instance, an individual investor might have a savings plan that involves unequal cash payments depending on the month of the year or lower savings during a planned vacation. One can always find the future value of a series of unequal cash flows by compounding the cash flows one at a time. Suppose you have the five cash flows described in Table 2, indexed relative to the present ($t = 0$).

Table 2 A Series of Unequal Cash Flows and Their Future Values at 5 Percent

Time	Cash Flow ($)	Future Value at Year 5		
$t = 1$	1,000	$\$1{,}000(1.05)^4$	=	$1,215.51
$t = 2$	2,000	$\$2{,}000(1.05)^3$	=	$2,315.25
$t = 3$	4,000	$\$4{,}000(1.05)^2$	=	$4,410.00
$t = 4$	5,000	$\$5{,}000(1.05)^1$	=	$5,250.00
$t = 5$	6,000	$\$6{,}000(1.05)^0$	=	$6,000.00
		Sum	=	$19,190.76

All of the payments shown in Table 2 are different. Therefore, the most direct approach to finding the future value at $t = 5$ is to compute the future value of each payment as of $t = 5$ and then sum the individual future values. The total future value at Year 5 equals $19,190.76, as shown in the third column. Later in this reading, you will learn shortcuts to take when the cash flows are close to even; these shortcuts will allow you to combine annuity and single-period calculations.

7 PRESENT VALUE OF A SINGLE CASH FLOW (LUMP SUM)

e calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;

f demonstrate the use of a time line in modeling and solving time value of money problems.

Just as the future value factor links today's present value with tomorrow's future value, the present value factor allows us to discount future value to present value. For example, with a 5 percent interest rate generating a future payoff of $105 in one year, what current amount invested at 5 percent for one year will grow to $105? The answer is $100; therefore, $100 is the present value of $105 to be received in one year at a discount rate of 5 percent.

Given a future cash flow that is to be received in N periods and an interest rate per period of r, we can use the formula for future value to solve directly for the present value as follows:

$$\mathrm{FV}_N = \mathrm{PV}(1+r)^N$$

$$\mathrm{PV} = \mathrm{FV}_N\left[\frac{1}{(1+r)^N}\right] \qquad (8)$$

$$\mathrm{PV} = \mathrm{FV}_N(1+r)^{-N}$$

We see from Equation 8 that the present value factor, $(1 + r)^{-N}$, is the reciprocal of the future value factor, $(1 + r)^N$.

EXAMPLE 8

The Present Value of a Lump Sum

An insurance company has issued a Guaranteed Investment Contract (GIC) that promises to pay $100,000 in six years with an 8 percent return rate. What amount of money must the insurer invest today at 8 percent for six years to make the promised payment?

Solution:

We can use Equation 8 to find the present value using the following data:

$$\mathrm{FV}_N = \$100{,}000$$

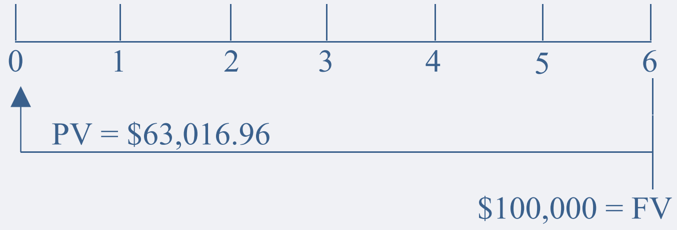
$$r = 8\% = 0.08$$

$$N = 6$$

$$\begin{aligned}\mathrm{PV} &= \mathrm{FV}_N(1+r)^{-N}\\ &= \$100{,}000\left[\frac{1}{(1.08)^6}\right]\\ &= \$100{,}000(0.6301696)\\ &= \$63{,}016.96\end{aligned}$$

We can say that \$63,016.96 today, with an interest rate of 8 percent, is equivalent to \$100,000 to be received in six years. Discounting the \$100,000 makes a future \$100,000 equivalent to \$63,016.96 when allowance is made for the time value of money. As the time line in Figure 4 shows, the \$100,000 has been discounted six full periods.

Figure 4 The Present Value of a Lump Sum to Be Received at Time $t = 6$



EXAMPLE 9

The Projected Present Value of a More Distant Future Lump Sum

Suppose you own a liquid financial asset that will pay you \$100,000 in 10 years from today. Your daughter plans to attend college four years from today, and you want to know what the asset's present value will be at that time. Given an 8 percent discount rate, what will the asset be worth four years from today?

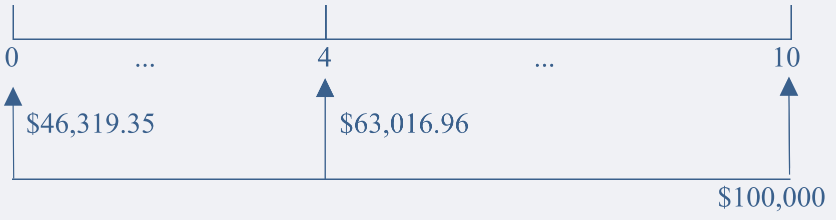
Solution:

The value of the asset is the present value of the asset's promised payment. At $t = 4$, the cash payment will be received six years later. With this information, you can solve for the value four years from today using Equation 8:

$$\begin{aligned} \mathrm{FV}_N &= \$100{,}000 \\ r &= 8\% = 0.08 \\ N &= 6 \end{aligned}$$

$$\begin{aligned} \mathrm{PV} &= \mathrm{FV}_N (1 + r)^{-N} \\ &= \$100{,}000 \frac{1}{(1.08)^6} \\ &= \$100{,}000(0.6301696) \\ &= \$63{,}016.96 \end{aligned}$$

Figure 5 The Relationship between Present Value and Future Value



The time line in Figure 5 shows the future payment of $100,000 that is to be received at $t = 10$. The time line also shows the values at $t = 4$ and at $t = 0$. Relative to the payment at $t = 10$, the amount at $t = 4$ is a projected present value, while the amount at $t = 0$ is the present value (as of today).

Present value problems require an evaluation of the present value factor, $(1 + r)^{-N}$. Present values relate to the discount rate and the number of periods in the following ways:

- For a given discount rate, the farther in the future the amount to be received, the smaller that amount's present value.
- Holding time constant, the larger the discount rate, the smaller the present value of a future amount.

8 NON-ANNUAL COMPOUNDING (PRESENT VALUE)

d calculate the solution for time value of money problems with different frequencies of compounding;

Recall that interest may be paid semiannually, quarterly, monthly, or even daily. To handle interest payments made more than once a year, we can modify the present value formula (Equation 8) as follows. Recall that r_s is the quoted interest rate and equals the periodic interest rate multiplied by the number of compounding periods in each year. In general, with more than one compounding period in a year, we can express the formula for present value as

$$PV = FV_N\left(1 + \frac{r_s}{m}\right)^{-mN} \tag{9}$$

where

m = number of compounding periods per year
r_s = quoted annual interest rate
N = number of years

The formula in Equation 9 is quite similar to that in Equation 8. As we have already noted, present value and future value factors are reciprocals. Changing the frequency of compounding does not alter this result. The only difference is the use of the periodic interest rate and the corresponding number of compounding periods.

The following example illustrates Equation 9.

EXAMPLE 10

The Present Value of a Lump Sum with Monthly Compounding

The manager of a Canadian pension fund knows that the fund must make a lump-sum payment of C$5 million 10 years from now. She wants to invest an amount today in a GIC so that it will grow to the required amount. The current interest rate on GICs is 6 percent a year, compounded monthly. How much should she invest today in the GIC?

Solution:

Use Equation 9 to find the required present value:

$$\begin{aligned}
\text{FV}_N &= \text{C\$}5{,}000{,}000 \\
r_s &= 6\% = 0.06 \\
m &= 12 \\
r_s/m &= 0.06/12 = 0.005 \\
N &= 10 \\
mN &= 12(10) = 120 \\
\text{PV} &= \text{FV}_N\left(1 + \frac{r_s}{m}\right)^{-mN} \\
&= \text{C\$}5{,}000{,}000(1.005)^{-120} \\
&= \text{C\$}5{,}000{,}000(0.549633) \\
&= \text{C\$}2{,}748{,}163.67
\end{aligned}$$

In applying Equation 9, we use the periodic rate (in this case, the monthly rate) and the appropriate number of periods with monthly compounding (in this case, 10 years of monthly compounding, or 120 periods).

PRESENT VALUE OF A SERIES OF EQUAL CASH FLOWS (ANNUITIES) AND UNEQUAL CASH FLOWS

9

- **e** calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;
- **f** demonstrate the use of a time line in modeling and solving time value of money problems.

Many applications in investment management involve assets that offer a series of cash flows over time. The cash flows may be highly uneven, relatively even, or equal. They may occur over relatively short periods of time, longer periods of time, or even stretch on indefinitely. In this section, we discuss how to find the present value of a series of cash flows.

9.1 The Present Value of a Series of Equal Cash Flows

We begin with an ordinary annuity. Recall that an ordinary annuity has equal annuity payments, with the first payment starting one period into the future. In total, the annuity makes N payments, with the first payment at $t = 1$ and the last at $t = N$. We can express the present value of an ordinary annuity as the sum of the present values of each individual annuity payment, as follows:

$$\text{PV} = \frac{A}{(1+r)} + \frac{A}{(1+r)^2} + \frac{A}{(1+r)^3} + \ldots + \frac{A}{(1+r)^{N-1}} + \frac{A}{(1+r)^N} \tag{10}$$

where

A = the annuity amount
r = the interest rate per period corresponding to the frequency of annuity payments (for example, annual, quarterly, or monthly)
N = the number of annuity payments

Because the annuity payment (A) is a constant in this equation, it can be factored out as a common term. Thus the sum of the interest factors has a shortcut expression:

$$\text{PV} = A\left[\frac{1 - \frac{1}{(1+r)^N}}{r}\right] \tag{11}$$

In much the same way that we computed the future value of an ordinary annuity, we find the present value by multiplying the annuity amount by a present value annuity factor (the term in brackets in Equation 11).

EXAMPLE 11

The Present Value of an Ordinary Annuity

Suppose you are considering purchasing a financial asset that promises to pay €1,000 per year for five years, with the first payment one year from now. The required rate of return is 12 percent per year. How much should you pay for this asset?

Solution:

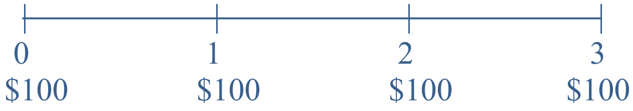
To find the value of the financial asset, use the formula for the present value of an ordinary annuity given in Equation 11 with the following data:

$$A = €1{,}000$$
$$r = 12\% = 0.12$$
$$N = 5$$
$$\text{PV} = A\left[\frac{1 - \dfrac{1}{(1+r)^N}}{r}\right]$$
$$= €1{,}000\left[\frac{1 - \dfrac{1}{(1.12)^5}}{0.12}\right]$$
$$= €1{,}000(3.604776)$$
$$= €3{,}604.78$$

The series of cash flows of €1,000 per year for five years is currently worth €3,604.78 when discounted at 12 percent.

Keeping track of the actual calendar time brings us to a specific type of annuity with level payments: the annuity due. An annuity due has its first payment occurring today ($t = 0$). In total, the annuity due will make N payments. Figure 6 presents the time line for an annuity due that makes four payments of $100.

Figure 6 An Annuity Due of $100 per Period



As Figure 6 shows, we can view the four-period annuity due as the sum of two parts: a $100 lump sum today and an ordinary annuity of $100 per period for three periods. At a 12 percent discount rate, the four $100 cash flows in this annuity due example will be worth $340.18.[7]

Expressing the value of the future series of cash flows in today's dollars gives us a convenient way of comparing annuities. The next example illustrates this approach.

---

[7] There is an alternative way to calculate the present value of an annuity due. Compared to an ordinary annuity, the payments in an annuity due are each discounted one less period. Therefore, we can modify Equation 11 to handle annuities due by multiplying the right-hand side of the equation by $(1 + r)$:

$$\text{PV}(\text{Annuity due}) = A\left\{\left[1 - (1+r)^{-N}\right]/r\right\}(1+r)$$

EXAMPLE 12

An Annuity Due as the Present Value of an Immediate Cash Flow Plus an Ordinary Annuity

You are retiring today and must choose to take your retirement benefits either as a lump sum or as an annuity. Your company's benefits officer presents you with two alternatives: an immediate lump sum of $2 million or an annuity with 20 payments of $200,000 a year with the first payment starting today. The interest rate at your bank is 7 percent per year compounded annually. Which option has the greater present value? (Ignore any tax differences between the two options.)

Solution:

To compare the two options, find the present value of each at time $t = 0$ and choose the one with the larger value. The first option's present value is $2 million, already expressed in today's dollars. The second option is an annuity due. Because the first payment occurs at $t = 0$, you can separate the annuity benefits into two pieces: an immediate $200,000 to be paid today ($t = 0$) and an ordinary annuity of $200,000 per year for 19 years. To value this option, you need to find the present value of the ordinary annuity using Equation 11 and then add $200,000 to it.

$$A = \$200{,}000$$

$$N = 19$$

$$r = 7\% = 0.07$$

$$\begin{aligned} \text{PV} &= A\left[\frac{1 - \dfrac{1}{(1+r)^N}}{r}\right] \\ &= \$200{,}000\left[\frac{1 - \dfrac{1}{(1.07)^{19}}}{0.07}\right] \\ &= \$200{,}000(10.335595) \\ &= \$2{,}067{,}119.05 \end{aligned}$$

The 19 payments of $200,000 have a present value of $2,067,119.05. Adding the initial payment of $200,000 to $2,067,119.05, we find that the total value of the annuity option is $2,267,119.05. The present value of the annuity is greater than the lump sum alternative of $2 million.

We now look at another example reiterating the equivalence of present and future values.

EXAMPLE 13

The Projected Present Value of an Ordinary Annuity

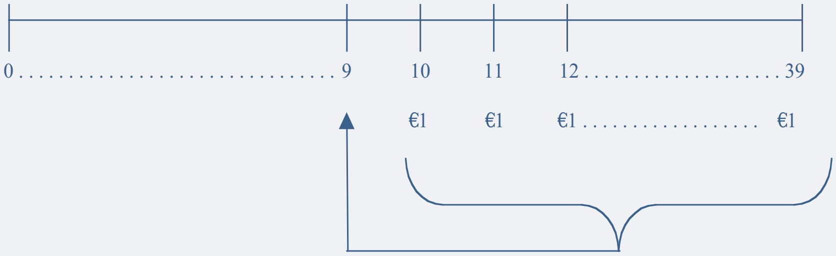
A German pension fund manager anticipates that benefits of €1 million per year must be paid to retirees. Retirements will not occur until 10 years from now at time $t = 10$. Once benefits begin to be paid, they will extend until $t = 39$ for a total of 30 payments. What is the present value of the pension liability if the appropriate annual discount rate for plan liabilities is 5 percent compounded annually?

Solution:

This problem involves an annuity with the first payment at $t = 10$. From the perspective of $t = 9$, we have an ordinary annuity with 30 payments. We can compute the present value of this annuity with Equation 11 and then look at it on a time line.

$$A = €1{,}000{,}000$$
$$r = 5\% = 0.05$$
$$N = 30$$
$$\text{PV} = A\left[\frac{1 - \frac{1}{(1+r)^N}}{r}\right]$$
$$= €1{,}000{,}000\left[\frac{1 - \frac{1}{(1.05)^{30}}}{0.05}\right]$$
$$= €1{,}000{,}000(15.372451)$$
$$= €15{,}372{,}451.03$$

Figure 7 The Present Value of an Ordinary Annuity with First Payment at Time t = 10 (in Millions)



On the time line, we have shown the pension payments of €1 million extending from $t = 10$ to $t = 39$. The bracket and arrow indicate the process of finding the present value of the annuity, discounted back to $t = 9$. The present value of the pension benefits as of $t = 9$ is €15,372,451.03. The problem is to find the present value today (at $t = 0$).

Now we can rely on the equivalence of present value and future value. As Figure 7 shows, we can view the amount at $t = 9$ as a future value from the vantage point of $t = 0$. We compute the present value of the amount at $t = 9$ as follows:

$$\begin{aligned} \text{FV}_N &= \text{€}15{,}372{,}451.03 \text{ (the present value at } t = 9) \\ N &= 9 \\ r &= 5\% = 0.05 \\ \text{PV} &= \text{FV}_N(1 + r)^{-N} \\ &= \text{€}15{,}372{,}451.03(1.05)^{-9} \\ &= \text{€}15{,}372{,}451.03(0.644609) \\ &= \text{€}9{,}909{,}219.00 \end{aligned}$$

The present value of the pension liability is €9,909,219.00.

Example 13 illustrates three procedures emphasized in this reading:

- finding the present or future value of any cash flow series;
- recognizing the equivalence of present value and appropriately discounted future value; and
- keeping track of the actual calendar time in a problem involving the time value of money.

9.2 The Present Value of a Series of Unequal Cash Flows

When we have unequal cash flows, we must first find the present value of each individual cash flow and then sum the respective present values. For a series with many cash flows, we usually use a spreadsheet. Table 3 lists a series of cash flows with the time periods in the first column, cash flows in the second column, and each cash flow's present value in the third column. The last row of Table 3 shows the sum of the five present values.

Table 3 A Series of Unequal Cash Flows and Their Present Values at 5 Percent

Time Period	Cash Flow ($)	Present Value at Year 0		
1	1,000	$\$1{,}000(1.05)^{-1}$	=	$952.38
2	2,000	$\$2{,}000(1.05)^{-2}$	=	$1,814.06
3	4,000	$\$4{,}000(1.05)^{-3}$	=	$3,455.35
4	5,000	$\$5{,}000(1.05)^{-4}$	=	$4,113.51
5	6,000	$\$6{,}000(1.05)^{-5}$	=	$4,701.16
		Sum	=	$15,036.46

We could calculate the future value of these cash flows by computing them one at a time using the single-payment future value formula. We already know the present value of this series, however, so we can easily apply time-value equivalence. The future value of the series of cash flows from Table 2, \$19,190.76, is equal to the single \$15,036.46 amount compounded forward to $t = 5$:

$$\begin{aligned} \mathrm{PV} &= \$15{,}036.46 \\ N &= 5 \\ r &= 5\% = 0.05 \\ \mathrm{FV}_N &= \mathrm{PV}(1+r)^N \\ &= \$15{,}036.46(1.05)^5 \\ &= \$15{,}036.46(1.276282) \\ &= \$19{,}190.76 \end{aligned}$$

PRESENT VALUE OF A PERPETUITY AND PRESENT VALUES INDEXED AT TIMES OTHER THAN T=0

10

e calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;

Consider the case of an ordinary annuity that extends indefinitely. Such an ordinary annuity is called a perpetuity (a perpetual annuity). To derive a formula for the present value of a perpetuity, we can modify Equation 10 to account for an infinite series of cash flows:

$$\mathrm{PV} = A\sum_{t=1}^{\infty}\left[\frac{1}{(1+r)^t}\right] \tag{12}$$

As long as interest rates are positive, the sum of present value factors converges and

$$\mathrm{PV} = \frac{A}{r} \tag{13}$$

To see this, look back at Equation 11, the expression for the present value of an ordinary annuity. As N (the number of periods in the annuity) goes to infinity, the term $1/(1 + r)^N$ approaches 0 and Equation 11 simplifies to Equation 13. This equation will reappear when we value dividends from stocks because stocks have no predefined life span. (A stock paying constant dividends is similar to a perpetuity.) With the first payment a year from now, a perpetuity of \$10 per year with a 20 percent required rate of return has a present value of \$10/0.2 = \$50.

Equation 13 is valid only for a perpetuity with level payments. In our development above, the first payment occurred at $t = 1$; therefore, we compute the present value as of $t = 0$.

Other assets also come close to satisfying the assumptions of a perpetuity. Certain government bonds and preferred stocks are typical examples of financial assets that make level payments for an indefinite period of time.

EXAMPLE 14

The Present Value of a Perpetuity

The British government once issued a type of security called a consol bond, which promised to pay a level cash flow indefinitely. If a consol bond paid £100 per year in perpetuity, what would it be worth today if the required rate of return were 5 percent?

Solution:

To answer this question, we can use Equation 13 with the following data:

$$\begin{aligned} A &= £100 \\ r &= 5\% = 0.05 \\ \text{PV} &= A/r \\ &= £100/0.05 \\ &= £2{,}000 \end{aligned}$$

The bond would be worth £2,000.

10.1 Present Values Indexed at Times Other than $t = 0$

In practice with investments, analysts frequently need to find present values indexed at times other than $t = 0$. Subscripting the present value and evaluating a perpetuity beginning with \$100 payments in Year 2, we find PV_1 = \$100/0.05 = \$2,000 at a 5 percent discount rate. Further, we can calculate today's PV as PV_0 = \$2,000/1.05 = \$1,904.76.

Consider a similar situation in which cash flows of \$6 per year begin at the end of the 4th year and continue at the end of each year thereafter, with the last cash flow at the end of the 10th year. From the perspective of the end of the third year, we are facing a typical seven-year ordinary annuity. We can find the present value of the annuity from the perspective of the end of the third year and then discount that present value back to the present. At an interest rate of 5 percent, the cash flows of \$6 per year starting at the end of the fourth year will be worth \$34.72 at the end of the third year ($t = 3$) and \$29.99 today ($t = 0$).

The next example illustrates the important concept that an annuity or perpetuity beginning sometime in the future can be expressed in present value terms one period prior to the first payment. That present value can then be discounted back to today's present value.

EXAMPLE 15

The Present Value of a Projected Perpetuity

Consider a level perpetuity of £100 per year with its first payment beginning at $t = 5$. What is its present value today (at $t = 0$), given a 5 percent discount rate?

Solution:

First, we find the present value of the perpetuity at $t = 4$ and then discount that amount back to $t = 0$. (Recall that a perpetuity or an ordinary annuity has its first payment one period away, explaining the $t = 4$ index for our present value calculation.)

i. Find the present value of the perpetuity at $t = 4$:

$$\begin{aligned} A &= £100 \\ r &= 5\% = 0.05 \\ \text{PV} &= A/r \\ &= £100/0.05 \\ &= £2{,}000 \end{aligned}$$

ii. Find the present value of the future amount at $t = 4$. From the perspective of $t = 0$, the present value of £2,000 can be considered a future value. Now we need to find the present value of a lump sum:

$$\begin{aligned} \text{FV}_N &= £2{,}000 \text{ (the present value at } t = 4) \\ r &= 5\% = 0.05 \\ N &= 4 \\ \text{PV} &= \text{FV}_N (1 + r)^{-N} \\ &= £2{,}000(1.05)^{-4} \\ &= £2{,}000(0.822702) \\ &= £1{,}645.40 \end{aligned}$$

Today's present value of the perpetuity is £1,645.40.

As discussed earlier, an annuity is a series of payments of a fixed amount for a specified number of periods. Suppose we own a perpetuity. At the same time, we issue a perpetuity obligating us to make payments; these payments are the same size as those of the perpetuity we own. However, the first payment of the perpetuity we issue is at $t = 5$; payments then continue on forever. The payments on this second perpetuity exactly offset the payments received from the perpetuity we own at $t = 5$ and all subsequent dates. We are left with level nonzero net cash flows at $t = 1, 2, 3,$ and 4. This outcome exactly fits the definition of an annuity with four payments. Thus we can construct an annuity as the difference between two perpetuities with equal, level payments but differing starting dates. The next example illustrates this result.

EXAMPLE 16

The Present Value of an Ordinary Annuity as the Present Value of a Current Minus Projected Perpetuity

Given a 5 percent discount rate, find the present value of a four-year ordinary annuity of £100 per year starting in Year 1 as the difference between the following two level perpetuities:

Perpetuity 1	£100 per year starting in Year 1 (first payment at $t = 1$)
Perpetuity 2	£100 per year starting in Year 5 (first payment at $t = 5$)

Solution:

If we subtract Perpetuity 2 from Perpetuity 1, we are left with an ordinary annuity of £100 per period for four years (payments at t = 1, 2, 3, 4). Subtracting the present value of Perpetuity 2 from that of Perpetuity 1, we arrive at the present value of the four-year ordinary annuity:

$$\begin{aligned}
\text{PV}_0(\text{Perpetuity 1}) &= £100 / 0.05 = £2{,}000 \\
\text{PV}_4(\text{Perpetuity 2}) &= £100 / 0.05 = £2{,}000 \\
\text{PV}_0(\text{Perpetuity 2}) &= £2{,}000 / (1.05)^4 = £1{,}645.40 \\
\text{PV}_0(\text{Annuity}) &= \text{PV}_0(\text{Perpetuity 1}) - \text{PV}_0(\text{Perpetuity 2}) \\
&= £2{,}000 - £1{,}645.40 \\
&= £354.60
\end{aligned}$$

The four-year ordinary annuity's present value is equal to £2,000 – £1,645.40 = £354.60.

11 SOLVING FOR INTEREST RATES, GROWTH RATES, AND NUMBER OF PERIODS

e calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;

In the previous examples, certain pieces of information have been made available. For instance, all problems have given the rate of interest, r, the number of time periods, N, the annuity amount, A, and either the present value, PV, or future value, FV. In real-world applications, however, although the present and future values may be given, you may have to solve for either the interest rate, the number of periods, or the annuity amount. In the subsections that follow, we show these types of problems.

11.1 Solving for Interest Rates and Growth Rates

Suppose a bank deposit of €100 is known to generate a payoff of €111 in one year. With this information, we can infer the interest rate that separates the present value of €100 from the future value of €111 by using Equation 2, $\text{FV}_N = \text{PV}(1 + r)^N$, with $N = 1$. With PV, FV, and N known, we can solve for r directly:

$$\begin{aligned}
1 + r &= \text{FV}/\text{PV} \\
1 + r &= €111/€100 = 1.11 \\
r &= 0.11, \text{ or } 11\%
\end{aligned}$$

The interest rate that equates €100 at $t = 0$ to €111 at $t = 1$ is 11 percent. Thus we can state that €100 grows to €111 with a growth rate of 11 percent.

As this example shows, an interest rate can also be considered a growth rate. The particular application will usually dictate whether we use the term "interest rate" or "growth rate." Solving Equation 2 for r and replacing the interest rate r with the growth rate g produces the following expression for determining growth rates:

$$g = (\text{FV}_N/\text{PV})^{1/N} - 1 \tag{14}$$

Below are two examples that use the concept of a growth rate.

EXAMPLE 17

Calculating a Growth Rate (1)

Hyundai Steel, the first Korean steelmaker, was established in 1953. Hyundai Steel's sales increased from ₩14,146.4 billion in 2012 to ₩19,166.0 billion in 2017. However, its net profit declined from ₩796.4 billion in 2012 to ₩727.5 billion in 2017. Calculate the following growth rates for Hyundai Steel for the five-year period from the end of 2012 to the end of 2017:

1. Sales growth rate.
2. Net profit growth rate.

Solution to 1:

To solve this problem, we can use Equation 14, $g = (\mathrm{FV}_N/\mathrm{PV})^{1/N} - 1$. We denote sales in 2012 as PV and sales in 2017 as FV_5. We can then solve for the growth rate as follows:

$$\begin{aligned} g &= \sqrt[5]{₩19{,}166.0/₩14{,}146.4} - 1 \\ &= \sqrt[5]{1.354832} - 1 \\ &= 1.062618 - 1 \\ &= 0.062618 \text{ or about } 6.3\% \end{aligned}$$

The calculated growth rate of about 6.3 percent a year shows that Hyundai Steel's sales grew during the 2012–2017 period.

Solution to 2:

In this case, we can speak of a positive compound rate of decrease or a negative compound growth rate. Using Equation 14, we find

$$\begin{aligned} g &= \sqrt[5]{₩727.5/₩796.4} - 1 \\ &= \sqrt[5]{0.913486} - 1 \\ &= 0.982065 - 1 \\ &= -0.017935 \text{ or about } -1.8\% \end{aligned}$$

In contrast to the positive sales growth, the rate of growth in net profit was approximately −1.8 percent during the 2012–2017 period.

EXAMPLE 18

Calculating a Growth Rate (2)

Toyota Motor Corporation, one of the largest automakers in the world, had consolidated vehicle sales of 8.96 million units in 2018 (fiscal year ending 31 March 2018). This is substantially more than consolidated vehicle sales of 7.35 million units six years earlier in 2012. What was the growth rate in number of vehicles sold by Toyota from 2012 to 2018?

Solution:

Using Equation 14, we find

$$\begin{aligned} g &= \sqrt[6]{8.96/7.35} - 1 \\ &= \sqrt[6]{1.219048} - 1 \\ &= 1.033563 - 1 \\ &= 0.033563 \text{ or about } 3.4\% \end{aligned}$$

The rate of growth in vehicles sold was approximately 3.4 percent during the 2012–2018 period. Note that we can also refer to 3.4 percent as the compound annual growth rate because it is the single number that compounds the number of vehicles sold in 2012 forward to the number of vehicles sold in 2018. Table 4 lists the number of vehicles sold by Toyota from 2012 to 2018.

Table 4 Number of Vehicles Sold, 2012–2018

Year	Number of Vehicles Sold (Millions)	$(1 + g)_t$	t
2012	7.35		0
2013	8.87	8.87/7.35 = 1.206803	1
2014	9.12	9.12/8.87 = 1.028185	2
2015	8.97	8.97/9.12 = 0.983553	3
2016	8.68	8.68/8.97 = 0.967670	4
2017	8.97	8.97/8.68 = 1.033410	5
2018	8.96	8.96/8.97 = 0.998885	6

Source: www.toyota.com.

Table 4 also shows 1 plus the one-year growth rate in number of vehicles sold. We can compute the 1 plus six-year cumulative growth in number of vehicles sold from 2012 to 2018 as the product of quantities (1 + one-year growth rate). We arrive at the same result as when we divide the ending number of vehicles sold, 8.96 million, by the beginning number of vehicles sold, 7.35 million:

$$\begin{aligned} \frac{8.96}{7.35} &= \left(\frac{8.87}{7.35}\right)\left(\frac{9.12}{8.87}\right)\left(\frac{8.97}{9.12}\right)\left(\frac{8.68}{8.97}\right)\left(\frac{8.97}{8.68}\right)\left(\frac{8.96}{8.97}\right) \\ &= (1 + g_1)(1 + g_2)(1 + g_3)(1 + g_4)(1 + g_5)(1 + g_6) \\ 1.219048 &= (1.206803)(1.028185)(0.983553)(0.967670)(1.033410)(0.998885) \end{aligned}$$

The right-hand side of the equation is the product of 1 plus the one-year growth rate in number of vehicles sold for each year. Recall that, using Equation 14, we took the sixth root of 8.96/7.35 = 1.219048. In effect, we were solving for the single value of g which, when compounded over six periods, gives the correct product of 1 plus the one-year growth rates.[8]

In conclusion, we do not need to compute intermediate growth rates as in Table 4 to solve for a compound growth rate g. Sometimes, however, the intermediate growth rates are interesting or informative. For example, most of the 21.9 percent increase in vehicles sold by Toyota from 2012 to 2018 occurred in

---

[8] The compound growth rate that we calculate here is an example of a geometric mean, specifically the geometric mean of the growth rates. We define the geometric mean in the reading on statistical concepts.

2013 as sales increased by 20.7 percent from 2012 to 2013. Elsewhere in Toyota Motor's disclosures, the company noted that all regions except Europe showed a substantial increase in sales in 2013. We can also analyze the variability in growth rates when we conduct an analysis as in Table 4. Sales continued to increase in 2014 but then declined in 2015 and 2016. Sales then increased but the sales in 2017 and 2018 are about the same as in 2015.

The compound growth rate is an excellent summary measure of growth over multiple time periods. In our Toyota Motors example, the compound growth rate of 3.4 percent is the single growth rate that, when added to 1, compounded over six years, and multiplied by the 2012 number of vehicles sold, yields the 2018 number of vehicles sold.

11.2 Solving for the Number of Periods

In this section, we demonstrate how to solve for the number of periods given present value, future value, and interest or growth rates.

EXAMPLE 19

The Number of Annual Compounding Periods Needed for an Investment to Reach a Specific Value

You are interested in determining how long it will take an investment of €10,000,000 to double in value. The current interest rate is 7 percent compounded annually. How many years will it take €10,000,000 to double to €20,000,000?

Solution:

Use Equation 2, $FV_N = PV(1 + r)^N$, to solve for the number of periods, N, as follows:

$$\begin{aligned}(1+r)^N &= FV_N/PV = 2\\ N\ln(1+r) &= \ln(2)\\ N &= \ln(2)/\ln(1+r)\\ &= \ln(2)/\ln(1.07) = 10.24\end{aligned}$$

With an interest rate of 7 percent, it will take approximately 10 years for the initial €10,000,000 investment to grow to €20,000,000. Solving for N in the expression $(1.07)^N = 2.0$ requires taking the natural logarithm of both sides and using the rule that $\ln(x^N) = N\ln(x)$. Generally, we find that $N = [\ln(FV/PV)]/\ln(1 + r)$. Here, $N = \ln(€20{,}000{,}000/€10{,}000{,}000)/\ln(1.07) = \ln(2)/\ln(1.07) = 10.24$.[9]

---

[9] To quickly approximate the number of periods, practitioners sometimes use an ad hoc rule called the **Rule of 72**: Divide 72 by the stated interest rate to get the approximate number of years it would take to double an investment at the interest rate. Here, the approximation gives 72/7 = 10.3 years. The Rule of 72 is loosely based on the observation that it takes 12 years to double an amount at a 6 percent interest rate, giving 6 × 12 = 72. At a 3 percent rate, one would guess it would take twice as many years, 3 × 24 = 72.

12 SOLVING FOR SIZE OF ANNUITY PAYMENTS (COMBINING FUTURE VALUE AND PRESENT VALUE ANNUITIES)

e calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;

f demonstrate the use of a time line in modeling and solving time value of money problems.

In this section, we discuss how to solve for annuity payments. Mortgages, auto loans, and retirement savings plans are classic examples of applications of annuity formulas.

EXAMPLE 20

Calculating the Size of Payments on a Fixed-Rate Mortgage

You are planning to purchase a $120,000 house by making a down payment of $20,000 and borrowing the remainder with a 30-year fixed-rate mortgage with monthly payments. The first payment is due at $t = 1$. Current mortgage interest rates are quoted at 8 percent with monthly compounding. What will your monthly mortgage payments be?

Solution:

The bank will determine the mortgage payments such that at the stated periodic interest rate, the present value of the payments will be equal to the amount borrowed (in this case, $100,000). With this fact in mind, we can use Equation 11, $\text{PV} = A\left[\dfrac{1 - \dfrac{1}{(1+r)^N}}{r}\right]$, to solve for the annuity amount, A, as the present value divided by the present value annuity factor:

$$\begin{aligned}
\text{PV} &= \$100{,}000 \\
r_s &= 8\% = 0.08 \\
m &= 12 \\
r_s/m &= 0.08/12 = 0.006667 \\
N &= 30 \\
mN &= 12 \times 30 = 360
\end{aligned}$$

$$\text{Present value annuity factor} = \frac{1 - \dfrac{1}{\left[1 + (r_s/m)\right]^{mN}}}{r_s/m} = \frac{1 - \dfrac{1}{(1.006667)^{360}}}{0.006667}$$

$$\begin{aligned}
&= 136.283494 \\
A &= \text{PV}/\text{Present value annuity factor} \\
&= \$100{,}000/136.283494 \\
&= \$733.76
\end{aligned}$$

The amount borrowed, \$100,000, is equivalent to 360 monthly payments of \$733.76 with a stated interest rate of 8 percent. The mortgage problem is a relatively straightforward application of finding a level annuity payment.

Next, we turn to a retirement-planning problem. This problem illustrates the complexity of the situation in which an individual wants to retire with a specified retirement income. Over the course of a life cycle, the individual may be able to save only a small amount during the early years but then may have the financial resources to save more during later years. Savings plans often involve uneven cash flows, a topic we will examine in the last part of this reading. When dealing with uneven cash flows, we take maximum advantage of the principle that dollar amounts indexed at the same point in time are additive—the **cash flow additivity principle**.

EXAMPLE 21

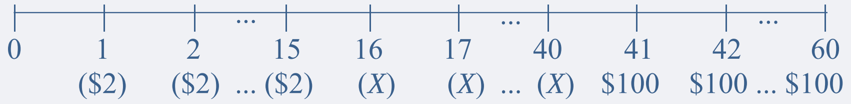
The Projected Annuity Amount Needed to Fund a Future-Annuity Inflow

Jill Grant is 22 years old (at t = 0) and is planning for her retirement at age 63 (at t = 41). She plans to save \$2,000 per year for the next 15 years (t = 1 to t = 15). She wants to have retirement income of \$100,000 per year for 20 years, with the first retirement payment starting at t = 41. How much must Grant save each year from t = 16 to t = 40 in order to achieve her retirement goal? Assume she plans to invest in a diversified stock-and-bond mutual fund that will earn 8 percent per year on average.

Solution:

To help solve this problem, we set up the information on a time line. As Figure 8 shows, Grant will save \$2,000 (an outflow) each year for Years 1 to 15. Starting in Year 41, Grant will start to draw retirement income of \$100,000 per year for 20 years. In the time line, the annual savings is recorded in parentheses (\$2) to show that it is an outflow. The problem is to find the savings, recorded as X, from Year 16 to Year 40.

Figure 8 Solving for Missing Annuity Payments (in Thousands)



Solving this problem involves satisfying the following relationship: the present value of savings (outflows) equals the present value of retirement income (inflows). We could bring all the dollar amounts to t = 40 or to t = 15 and solve for X.

Let us evaluate all dollar amounts at t = 15 (we encourage the reader to repeat the problem by bringing all cash flows to t = 40). As of t = 15, the first payment of X will be one period away (at t = 16). Thus we can value the stream of Xs using the formula for the present value of an ordinary annuity.

This problem involves three series of level cash flows. The basic idea is that the present value of the retirement income must equal the present value of Grant's savings. Our strategy requires the following steps:

1 Find the future value of the savings of \$2,000 per year and index it at t = 15. This value tells us how much Grant will have saved.

2 Find the present value of the retirement income at t = 15. This value tells us how much Grant needs to meet her retirement goals (as of t = 15). Two substeps are necessary. First, calculate the present value of the annuity of \$100,000 per year at t = 40. Use the formula for the present value of an annuity. (Note that the present value is indexed at t = 40 because the first payment is at t = 41.) Next, discount the present value back to t = 15 (a total of 25 periods).

3 Now compute the difference between the amount Grant has saved (Step 1) and the amount she needs to meet her retirement goals (Step 2). Her savings from t = 16 to t = 40 must have a present value equal to the difference between the future value of her savings and the present value of her retirement income.

Our goal is to determine the amount Grant should save in each of the 25 years from t = 16 to t = 40. We start by bringing the \$2,000 savings to t = 15, as follows:

$$
\begin{aligned}
A &= \$2{,}000 \\
r &= 8\% = 0.08 \\
N &= 15 \\
\mathrm{FV} &= A\left[\frac{(1+r)^N - 1}{r}\right] \\
&= \$2{,}000\left[\frac{(1.08)^{15} - 1}{0.08}\right] \\
&= \$2{,}000(27.152114) \\
&= \$54{,}304.23
\end{aligned}
$$

At t = 15, Grant's initial savings will have grown to \$54,304.23.

Now we need to know the value of Grant's retirement income at t = 15. As stated earlier, computing the retirement present value requires two substeps. First, find the present value at t = 40 with the formula in Equation 11; second, discount this present value back to t = 15. Now we can find the retirement income present value at t = 40:

$$
\begin{aligned}
A &= \$100{,}000 \\
r &= 8\% = 0.08 \\
N &= 20 \\
\mathrm{PV} &= A\left[\frac{1 - \dfrac{1}{(1+r)^N}}{r}\right] \\
&= \$100{,}000\left[\frac{1 - \dfrac{1}{(1.08)^{20}}}{0.08}\right] \\
&= \$100{,}000(9.818147) \\
&= \$981{,}814.74
\end{aligned}
$$

The present value amount is as of t = 40, so we must now discount it back as a lump sum to t = 15:

$$\begin{aligned} \mathrm{FV}_N &= \$981{,}814.74 \\ N &= 25 \\ r &= 8\% = 0.08 \\ \mathrm{PV} &= \mathrm{FV}_N (1 + r)^{-N} \\ &= \$981{,}814.74(1.08)^{-25} \\ &= \$981{,}814.74(0.146018) \\ &= \$143{,}362.53 \end{aligned}$$

Now recall that Grant will have saved \$54,304.23 by t = 15. Therefore, in present value terms, the annuity from t = 16 to t = 40 must equal the difference between the amount already saved (\$54,304.23) and the amount required for retirement (\$143,362.53). This amount is equal to \$143,362.53 − \$54,304.23 = \$89,058.30. Therefore, we must now find the annuity payment, A, from t = 16 to t = 40 that has a present value of \$89,058.30. We find the annuity payment as follows:

$$\begin{aligned} \mathrm{PV} &= \$89{,}058.30 \\ r &= 8\% = 0.08 \\ N &= 25 \\ \text{Present value annuity factor} &= \left[\frac{1 - \dfrac{1}{(1 + r)^N}}{r}\right] \\ &= \left[\frac{1 - \dfrac{1}{(1.08)^{25}}}{0.08}\right] \\ &= 10.674776 \\ A &= \mathrm{PV}/\text{Present value annuity factor} \\ &= \$89{,}058.30/10.674776 \\ &= \$8{,}342.87 \end{aligned}$$

Grant will need to increase her savings to \$8,342.87 per year from t = 16 to t = 40 to meet her retirement goal of having a fund equal to \$981,814.74 after making her last payment at t = 40.

PRESENT VALUE AND FUTURE VALUE EQUIVALENCE, ADDITIVITY PRINCIPLE

13

e calculate and interpret the future value (FV) and present value (PV) of a single sum of money, an ordinary annuity, an annuity due, a perpetuity (PV only), and a series of unequal cash flows;

f demonstrate the use of a time line in modeling and solving time value of money problems.

As we have demonstrated, finding present and future values involves moving amounts of money to different points on a time line. These operations are possible because present value and future value are equivalent measures separated in time. Table 5 illustrates this equivalence; it lists the timing of five cash flows, their present values at $t = 0$, and their future values at $t = 5$.

To interpret Table 5, start with the third column, which shows the present values. Note that each \$1,000 cash payment is discounted back the appropriate number of periods to find the present value at $t = 0$. The present value of \$4,329.48 is exactly equivalent to the series of cash flows. This information illustrates an important point: A lump sum can actually generate an annuity. If we place a lump sum in an account that earns the stated interest rate for all periods, we can generate an annuity that is equivalent to the lump sum. Amortized loans, such as mortgages and car loans, are examples of this principle.

Table 5 The Equivalence of Present and Future Values

Time	Cash Flow ($)	Present Value at $t = 0$			Future Value at $t = 5$		
1	1,000	$\$1,000(1.05)^{-1}$	=	\$952.38	$\$1,000(1.05)^{4}$	=	\$1,215.51
2	1,000	$\$1,000(1.05)^{-2}$	=	\$907.03	$\$1,000(1.05)^{3}$	=	\$1,157.63
3	1,000	$\$1,000(1.05)^{-3}$	=	\$863.84	$\$1,000(1.05)^{2}$	=	\$1,102.50
4	1,000	$\$1,000(1.05)^{-4}$	=	\$822.70	$\$1,000(1.05)^{1}$	=	\$1,050.00
5	1,000	$\$1,000(1.05)^{-5}$	=	\$783.53	$\$1,000(1.05)^{0}$	=	\$1,000.00
		Sum:		\$4,329.48	Sum:		\$5,525.64

To see how a lump sum can fund an annuity, assume that we place \$4,329.48 in the bank today at 5 percent interest. We can calculate the size of the annuity payments by using Equation 11. Solving for A, we find

$$A = \frac{\text{PV}}{\dfrac{1 - \left[1/(1+r)^N\right]}{r}}$$

$$= \frac{\$4,329.48}{\dfrac{1 - \left[1/(1.05)^5\right]}{0.05}}$$

$$= \$1,000$$

Table 6 shows how the initial investment of \$4,329.48 can actually generate five \$1,000 withdrawals over the next five years.

To interpret Table 6, start with an initial present value of \$4,329.48 at $t = 0$. From $t = 0$ to $t = 1$, the initial investment earns 5 percent interest, generating a future value of \$4,329.48(1.05) = \$4,545.95. We then withdraw \$1,000 from our account, leaving \$4,545.95 – \$1,000 = \$3,545.95 (the figure reported in the last column for time period 1). In the next period, we earn one year's worth of interest and then make a \$1,000 withdrawal. After the fourth withdrawal, we have \$952.38, which earns 5 percent. This amount then grows to \$1,000 during the year, just enough for us to make the last withdrawal. Thus the initial present value, when invested at 5 percent for five years, generates the \$1,000 five-year ordinary annuity. The present value of the initial investment is exactly equivalent to the annuity.

Now we can look at how future value relates to annuities. In Table 5, we reported that the future value of the annuity was \$5,525.64. We arrived at this figure by compounding the first \$1,000 payment forward four periods, the second \$1,000 forward three periods, and so on. We then added the five future amounts at $t = 5$. The annuity is equivalent to \$5,525.64 at $t = 5$ and \$4,329.48 at $t = 0$. These two dollar measures are thus equivalent. We can verify the equivalence by finding the present value of \$5,525.64, which is $\$5,525.64 \times (1.05)^{-5} = \$4,329.48$. We found this result above when we showed that a lump sum can generate an annuity.

Table 6 How an Initial Present Value Funds an Annuity

Time Period	Amount Available at the Beginning of the Time Period (\$)	Ending Amount before Withdrawal			Withdrawal (\$)	Amount Available after Withdrawal (\$)
1	4,329.48	\$4,329.48(1.05)	=	\$4,545.95	1,000	3,545.95
2	3,545.95	\$3,545.95(1.05)	=	\$3,723.25	1,000	2,723.25
3	2,723.25	\$2,723.25(1.05)	=	\$2,859.41	1,000	1,859.41
4	1,859.41	\$1,859.41(1.05)	=	\$1,952.38	1,000	952.38
5	952.38	\$952.38(1.05)	=	\$1,000	1,000	0

To summarize what we have learned so far: A lump sum can be seen as equivalent to an annuity, and an annuity can be seen as equivalent to its future value. Thus present values, future values, and a series of cash flows can all be considered equivalent as long as they are indexed at the same point in time.

13.1 The Cash Flow Additivity Principle

The cash flow additivity principle—the idea that amounts of money indexed at the same point in time are additive—is one of the most important concepts in time value of money mathematics. We have already mentioned and used this principle; this section provides a reference example for it.

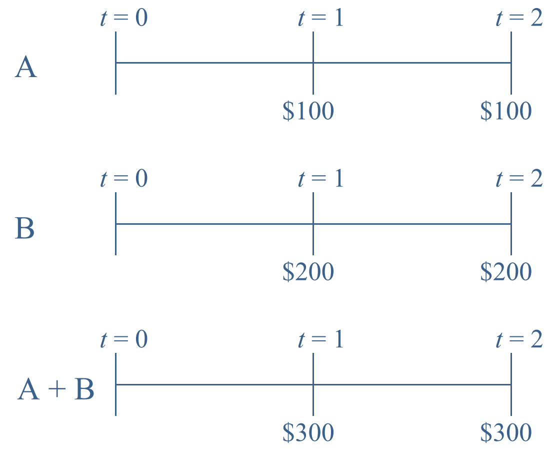
Consider the two series of cash flows shown on the time line in Figure 9. The series are denoted A and B. If we assume that the annual interest rate is 2 percent, we can find the future value of each series of cash flows as follows. Series A's future value is \$100(1.02) + \$100 = \$202. Series B's future value is \$200(1.02) + \$200 = \$404. The future value of (A + B) is \$202 + \$404 = \$606 by the method we have used up to this point. The alternative way to find the future value is to add the cash flows of each series, A and B (call it A + B), and then find the future value of the combined cash flow, as shown in Figure 9.

The third time line in Figure 9 shows the combined series of cash flows. Series A has a cash flow of \$100 at $t = 1$, and Series B has a cash flow of \$200 at $t = 1$. The combined series thus has a cash flow of \$300 at $t = 1$. We can similarly calculate the cash flow of the combined series at $t = 2$. The future value of the combined series (A + B) is \$300(1.02) + \$300 = \$606—the same result we found when we added the future values of each series.

The additivity and equivalence principles also appear in another common situation. Suppose cash flows are \$4 at the end of the first year and \$24 (actually separate payments of \$4 and \$20) at the end of the second year. Rather than finding present values of the first year's \$4 and the second year's \$24, we can treat this situation as a

$4 annuity for two years and a second-year $20 lump sum. If the discount rate were 6 percent, the $4 annuity would have a present value of $7.33 and the $20 lump sum a present value of $17.80, for a total of $25.13.

Figure 9 The Additivity of Two Series of Cash Flows



SUMMARY

In this reading, we have explored a foundation topic in investment mathematics, the time value of money. We have developed and reviewed the following concepts for use in financial applications:

- The interest rate, r, is the required rate of return; r is also called the discount rate or opportunity cost.
- An interest rate can be viewed as the sum of the real risk-free interest rate and a set of premiums that compensate lenders for risk: an inflation premium, a default risk premium, a liquidity premium, and a maturity premium.
- The future value, FV, is the present value, PV, times the future value factor, $(1 + r)^N$.
- The interest rate, r, makes current and future currency amounts equivalent based on their time value.
- The stated annual interest rate is a quoted interest rate that does not account for compounding within the year.
- The periodic rate is the quoted interest rate per period; it equals the stated annual interest rate divided by the number of compounding periods per year.
- The effective annual rate is the amount by which a unit of currency will grow in a year with interest on interest included.
- An annuity is a finite set of level sequential cash flows.
- There are two types of annuities, the annuity due and the ordinary annuity. The annuity due has a first cash flow that occurs immediately; the ordinary annuity has a first cash flow that occurs one period from the present (indexed at $t = 1$).

- On a time line, we can index the present as 0 and then display equally spaced hash marks to represent a number of periods into the future. This representation allows us to index how many periods away each cash flow will be paid.
- Annuities may be handled in a similar approach as single payments if we use annuity factors rather than single-payment factors.
- The present value, PV, is the future value, FV, times the present value factor, $(1 + r)^{-N}$.
- The present value of a perpetuity is A/r, where A is the periodic payment to be received forever.
- It is possible to calculate an unknown variable, given the other relevant variables in time value of money problems.
- The cash flow additivity principle can be used to solve problems with uneven cash flows by combining single payments and annuities.

PRACTICE PROBLEMS

1 The table below gives current information on the interest rates for two two-year and two eight-year maturity investments. The table also gives the maturity, liquidity, and default risk characteristics of a new investment possibility (Investment 3). All investments promise only a single payment (a payment at maturity). Assume that premiums relating to inflation, liquidity, and default risk are constant across all time horizons.

Investment	Maturity (in Years)	Liquidity	Default Risk	Interest Rate (%)
1	2	High	Low	2.0
2	2	Low	Low	2.5
3	7	Low	Low	r_3
4	8	High	Low	4.0
5	8	Low	High	6.5

Based on the information in the above table, address the following:

A Explain the difference between the interest rates on Investment 1 and Investment 2.

B Estimate the default risk premium.

C Calculate upper and lower limits for the interest rate on Investment 3, r_3.

2 A couple plans to set aside $20,000 per year in a conservative portfolio projected to earn 7 percent a year. If they make their first savings contribution one year from now, how much will they have at the end of 20 years?

3 Two years from now, a client will receive the first of three annual payments of $20,000 from a small business project. If she can earn 9 percent annually on her investments and plans to retire in six years, how much will the three business project payments be worth at the time of her retirement?

4 To cover the first year's total college tuition payments for his two children, a father will make a $75,000 payment five years from now. How much will he need to invest today to meet his first tuition goal if the investment earns 6 percent annually?

5 A client can choose between receiving 10 annual $100,000 retirement payments, starting one year from today, or receiving a lump sum today. Knowing that he can invest at a rate of 5 percent annually, he has decided to take the lump sum. What lump sum today will be equivalent to the future annual payments?

6 You are considering investing in two different instruments. The first instrument will pay nothing for three years, but then it will pay $20,000 per year for four years. The second instrument will pay $20,000 for three years and $30,000 in the fourth year. All payments are made at year-end. If your required rate of return on these investments is 8 percent annually, what should you be willing to pay for:

A The first instrument?

B The second instrument (use the formula for a four-year annuity)?

© 2016 CFA Institute. All rights reserved.

7 Suppose you plan to send your daughter to college in three years. You expect her to earn two-thirds of her tuition payment in scholarship money, so you estimate that your payments will be $10,000 a year for four years. To estimate whether you have set aside enough money, you ignore possible inflation in tuition payments and assume that you can earn 8 percent annually on your investments. How much should you set aside now to cover these payments?

8 A client plans to send a child to college for four years starting 18 years from now. Having set aside money for tuition, she decides to plan for room and board also. She estimates these costs at $20,000 per year, payable at the beginning of each year, by the time her child goes to college. If she starts next year and makes 17 payments into a savings account paying 5 percent annually, what annual payments must she make?

9 A couple plans to pay their child's college tuition for 4 years starting 18 years from now. The current annual cost of college is C$7,000, and they expect this cost to rise at an annual rate of 5 percent. In their planning, they assume that they can earn 6 percent annually. How much must they put aside each year, starting next year, if they plan to make 17 equal payments?

10 The nominal risk-free rate is *best* described as the sum of the real risk-free rate and a premium for:

A maturity.

B liquidity.

C expected inflation.

11 Which of the following risk premiums is most relevant in explaining the difference in yields between 30-year bonds issued by the US Treasury and 30-year bonds issued by a small private issuer?

A Inflation

B Maturity

C Liquidity

12 A bank quotes a stated annual interest rate of 4.00%. If that rate is equal to an effective annual rate of 4.08%, then the bank is compounding interest:

A daily.

B quarterly.

C semiannually.

13 The value in six years of $75,000 invested today at a stated annual interest rate of 7% compounded quarterly is *closest* to:

A $112,555.

B $113,330.

C $113,733.

14 A client requires £100,000 one year from now. If the stated annual rate is 2.50% compounded weekly, the deposit needed today is *closest* to:

A £97,500.

B £97,532.

C £97,561.

15 For a lump sum investment of ¥250,000 invested at a stated annual rate of 3% compounded daily, the number of months needed to grow the sum to ¥1,000,000 is *closest* to:

A 555.

B 563.

C 576.

16 Given a €1,000,000 investment for four years with a stated annual rate of 3% compounded continuously, the difference in its interest earnings compared with the same investment compounded daily is *closest* to:

A €1.

B €6.

C €455.

17 An investment pays €300 annually for five years, with the first payment occurring today. The present value (PV) of the investment discounted at a 4% annual rate is *closest* to:

A €1,336.

B €1,389.

C €1,625.

18 A perpetual preferred stock makes its first quarterly dividend payment of $2.00 in five quarters. If the required annual rate of return is 6% compounded quarterly, the stock's present value is *closest* to:

A $31.

B $126.

C $133.

19 A saver deposits the following amounts in an account paying a stated annual rate of 4%, compounded semiannually:

Year	End of Year Deposits ($)
1	4,000
2	8,000
3	7,000
4	10,000

At the end of Year 4, the value of the account is *closest* to:

A $30,432

B $30,447

C $31,677

20 An investment of €500,000 today that grows to €800,000 after six years has a stated annual interest rate *closest* to:

A 7.5% compounded continuously.

B 7.7% compounded daily.

C 8.0% compounded semiannually.

21 A sweepstakes winner may select either a perpetuity of £2,000 a month beginning with the first payment in one month or an immediate lump sum payment of £350,000. If the annual discount rate is 6% compounded monthly, the present value of the perpetuity is:

A less than the lump sum.

B equal to the lump sum.

C greater than the lump sum.

22 At a 5% interest rate per year compounded annually, the present value (PV) of a 10-year ordinary annuity with annual payments of $2,000 is $15,443.47. The PV of a 10-year annuity due with the same interest rate and payments is *closest* to:

A $14,708.

B $16,216.

C $17,443.

23 Grandparents are funding a newborn's future university tuition costs, estimated at $50,000/year for four years, with the first payment due as a lump sum in 18 years. Assuming a 6% effective annual rate, the required deposit today is *closest* to:

A $60,699.

B $64,341.

C $68,201.

24 The present value (PV) of an investment with the following year-end cash flows (CF) and a 12% required annual rate of return is *closest* to:

Year	Cash Flow (€)
1	100,000
2	150,000
5	–10,000

A €201,747.

B €203,191.

C €227,573.

25 A sports car, purchased for £200,000, is financed for five years at an annual rate of 6% compounded monthly. If the first payment is due in one month, the monthly payment is *closest* to:

A £3,847.

B £3,867.

C £3,957.

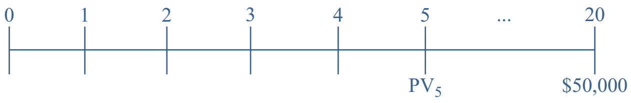
26 Given a stated annual interest rate of 6% compounded quarterly, the level amount that, deposited quarterly, will grow to £25,000 at the end of 10 years is *closest* to:

A £461.

B £474.

C £836.

27 Given the following timeline and a discount rate of 4% a year compounded annually, the present value (PV), as of the end of Year 5 (PV_5), of the cash flow received at the end of Year 20 is *closest* to:



A $22,819.

B $27,763.

C $28,873.

28 A client invests €20,000 in a four-year certificate of deposit (CD) that annually pays interest of 3.5%. The annual CD interest payments are automatically reinvested in a separate savings account at a stated annual interest rate of 2% compounded monthly. At maturity, the value of the combined asset is *closest* to:

A €21,670.

B €22,890.

C €22,950.

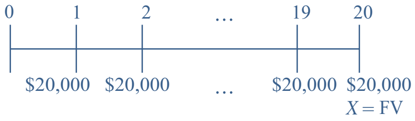
SOLUTIONS

1 **A** Investment 2 is identical to Investment 1 except that Investment 2 has low liquidity. The difference between the interest rate on Investment 2 and Investment 1 is 0.5 percentage point. This amount represents the liquidity premium, which represents compensation for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

B To estimate the default risk premium, find the two investments that have the same maturity but different levels of default risk. Both Investments 4 and 5 have a maturity of eight years. Investment 5, however, has low liquidity and thus bears a liquidity premium. The difference between the interest rates of Investments 5 and 4 is 2.5 percentage points. The liquidity premium is 0.5 percentage point (from Part A). This leaves 2.5 − 0.5 = 2.0 percentage points that must represent a default risk premium reflecting Investment 5's high default risk.

C Investment 3 has liquidity risk and default risk comparable to Investment 2, but with its longer time to maturity, Investment 3 should have a higher maturity premium. The interest rate on Investment 3, r_3, should thus be above 2.5 percent (the interest rate on Investment 2). If the liquidity of Investment 3 were high, Investment 3 would match Investment 4 except for Investment 3's shorter maturity. We would then conclude that Investment 3's interest rate should be less than the interest rate on Investment 4, which is 4 percent. In contrast to Investment 4, however, Investment 3 has low liquidity. It is possible that the interest rate on Investment 3 exceeds that of Investment 4 despite 3's shorter maturity, depending on the relative size of the liquidity and maturity premiums. However, we expect r_3 to be less than 4.5 percent, the expected interest rate on Investment 4 if it had low liquidity. Thus 2.5 percent $< r_3 <$ 4.5 percent.

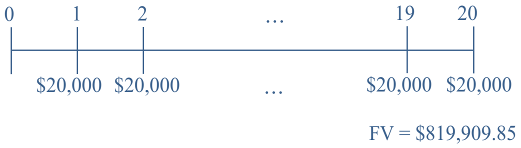
2 **i.** Draw a time line.



ii. Identify the problem as the future value of an annuity.

iii. Use the formula for the future value of an annuity.

$$\begin{aligned}\text{FV}_N &= A\left[\frac{(1+r)^N - 1}{r}\right]\\ &= \$20{,}000\left[\frac{(1+0.07)^{20} - 1}{0.07}\right]\\ &= \$819{,}909.85\end{aligned}$$



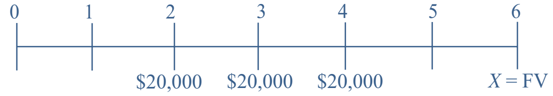
iv. Alternatively, use a financial calculator.

Notation Used on Most Calculators	**Numerical Value for This Problem**
N	20
%i	7
PV	n/a (= 0)
FV **compute**	X
PMT	$20,000

Enter 20 for N, the number of periods. Enter 7 for the interest rate and 20,000 for the payment size. The present value is not needed, so enter 0. Calculate the future value. Verify that you get $819,909.85 to make sure you have mastered your calculator's keystrokes.

In summary, if the couple sets aside $20,000 each year (starting next year), they will have $819,909.85 in 20 years if they earn 7 percent annually.

3 **i.** Draw a time line.



ii. Recognize the problem as the future value of a delayed annuity. Delaying the payments requires two calculations.

iii. Use the formula for the future value of an annuity (Equation 7).

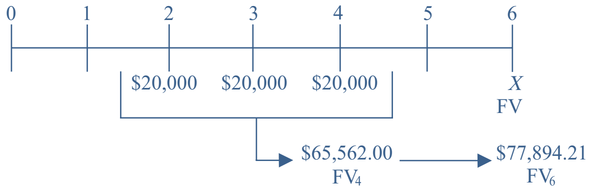
$$\mathrm{FV}_N = A\left[\frac{(1+r)^N - 1}{r}\right]$$

to bring the three $20,000 payments to an equivalent lump sum of $65,562.00 four years from today.

Notation Used on Most Calculators	**Numerical Value for This Problem**
N	3
%i	9
PV	n/a (= 0)
FV **compute**	X
PMT	$20,000

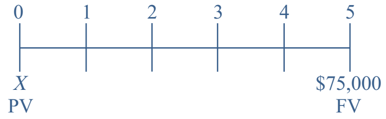
iv. Use the formula for the future value of a lump sum (Equation 2), $\mathrm{FV}_N = \mathrm{PV}(1 + r)^N$, to bring the single lump sum of $65,562.00 to an equivalent lump sum of $77,894.21 six years from today.

Notation Used on Most Calculators	**Numerical Value for This Problem**
N	2
%i	9
PV	$65,562.00
FV **compute**	X
PMT	n/a (= 0)


0
1
2
3
4
5
6
$20,000
$20,000
$20,000
X
FV
$65,562.00
FV4
$77,894.21
FV6

In summary, your client will have $77,894.21 in six years if she receives three yearly payments of $20,000 starting in Year 2 and can earn 9 percent annually on her investments.

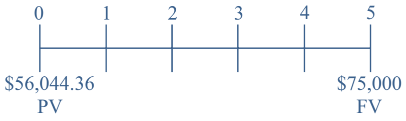
4 **i.** Draw a time line.


0
1
2
3
4
5
X
PV
$75,000
FV

ii. Identify the problem as the present value of a lump sum.

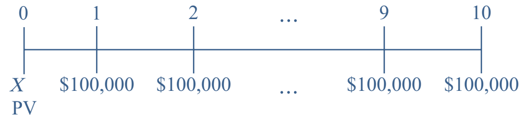
iii. Use the formula for the present value of a lump sum.

$$\begin{aligned}\text{PV} &= \text{FV}_N(1+r)^{-N}\\ &= \$75{,}000(1+0.06)^{-5}\\ &= \$56{,}044.36\end{aligned}$$


0
1
2
3
4
5
$56,044.36
PV
$75,000
FV

In summary, the father will need to invest $56,044.36 today in order to have $75,000 in five years if his investments earn 6 percent annually.

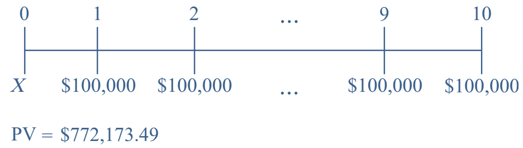
5 **i.** Draw a time line for the 10 annual payments.


0
1
2
…
9
10
X
PV
$100,000
$100,000
…
$100,000
$100,000

ii. Identify the problem as the present value of an annuity.

iii. Use the formula for the present value of an annuity.

$$\begin{aligned}\text{PV} &= A\left[\frac{1-\dfrac{1}{(1+r)^N}}{r}\right]\\ &= \$100{,}000\left[\frac{1-\dfrac{1}{(1+0.05)^{10}}}{0.05}\right]\\ &= \$772{,}173.49\end{aligned}$$


0
1
2
…
9
10
X
$100,000
$100,000
…
$100,000
$100,000
PV = $772,173.49

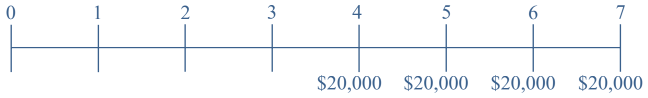
iv. Alternatively, use a financial calculator.

Notation Used on Most Calculators	**Numerical Value for This Problem**
N	10
$\%i$	5
PV **compute**	X
FV	n/a (= 0)
PMT	$100,000

In summary, the present value of 10 payments of $100,000 is $772,173.49 if the first payment is received in one year and the rate is 5 percent compounded annually. Your client should accept no less than this amount for his lump sum payment.

6 **A** To evaluate the first instrument, take the following steps:

i. Draw a time line.



ii.

$$\begin{aligned} \text{PV}_3 &= A\left[\frac{1-\dfrac{1}{(1+r)^N}}{r}\right] \\ &= \$20{,}000\left[\frac{1-\dfrac{1}{(1+0.08)^4}}{0.08}\right] \\ &= \$66{,}242.54 \end{aligned}$$

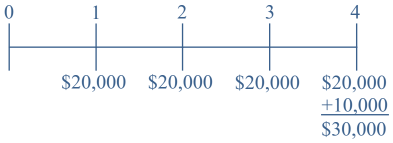
iii.

$$\text{PV}_0 = \frac{\text{PV}_3}{(1+r)^N} = \frac{\$66{,}242.54}{1.08^3} = \$52{,}585.46$$

You should be willing to pay $52,585.46 for this instrument.

B To evaluate the second instrument, take the following steps:

i. Draw a time line.



The time line shows that this instrument can be analyzed as an ordinary annuity of $20,000 with four payments (valued in Step ii below) and a $10,000 payment to be received at t = 4 (valued in Step iii below).

ii.

$$\begin{aligned}\text{PV} &= A\left[\frac{1-\dfrac{1}{(1+r)^N}}{r}\right]\\ &= \$20{,}000\left[\frac{1-\dfrac{1}{(1+0.08)^4}}{0.08}\right]\\ &= \$66{,}242.54\end{aligned}$$

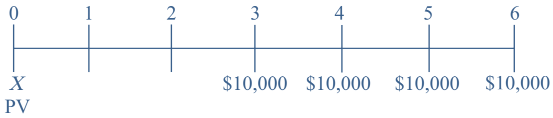
iii.

$$\text{PV} = \frac{\text{FV}_4}{(1+r)^N} = \frac{\$10{,}000}{(1+0.08)^4} = \$7{,}350.30$$

iv. Total = \$66,242.54 + \$7,350.30 = \$73,592.84

You should be willing to pay \$73,592.84 for this instrument.

7 **i.** Draw a time line.



ii. Recognize the problem as a delayed annuity. Delaying the payments requires two calculations.

iii. Use the formula for the present value of an annuity (Equation 11).

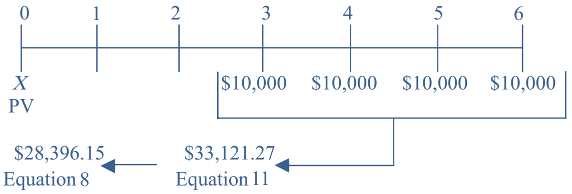
$$\text{PV} = A\left[\frac{1-\dfrac{1}{(1+r)^N}}{r}\right]$$

to bring the four payments of \$10,000 back to a single equivalent lump sum of \$33,121.27 at $t = 2$. Note that we use $t = 2$ because the first annuity payment is then one period away, giving an ordinary annuity.

Notation Used on Most Calculators	Numerical Value for This Problem
N	4
%i	8
PV **compute**	X
PMT	\$10,000

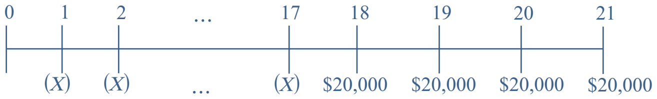
iv. Then use the formula for the present value of a lump sum (Equation 8), $\text{PV} = \text{FV}_N(1 + r)^{-N}$, to bring back the single payment of \$33,121.27 (at $t = 2$) to an equivalent single payment of \$28,396.15 (at $t = 0$).

Notation Used on Most Calculators	Numerical Value for This Problem
N	2
%i	8
PV **compute**	X
FV	\$33,121.27
PMT	n/a (= 0)


0
1
2
3
4
5
6
X
PV
$10,000
$10,000
$10,000
$10,000
$28,396.15
Equation 8
$33,121.27
Equation 11

In summary, you should set aside $28,396.15 today to cover four payments of $10,000 starting in three years if your investments earn a rate of 8 percent annually.

8 **i.** Draw a time line.


0
1
2
…
17
18
19
20
21
(X)
(X)
…
(X)
$20,000
$20,000
$20,000
$20,000

ii. Recognize that you need to equate the values of two annuities.

iii. Equate the value of the four $20,000 payments to a single payment in Period 17 using the formula for the present value of an annuity (Equation 11), with $r = 0.05$. The present value of the college costs as of $t = 17$ is $70,919.

$$\text{PV} = \$20{,}000\left[\frac{1-\dfrac{1}{(1.05)^4}}{0.05}\right] = \$70{,}919$$

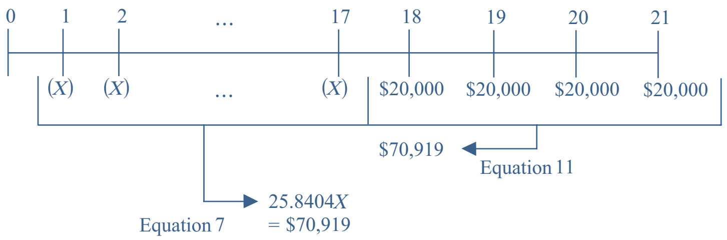
Notation Used on Most Calculators	**Numerical Value for This Problem**
N	4
%*i*	5
PV **compute**	*X*
FV	n/a (= 0)
PMT	$20,000

iv. Equate the value of the 17 investments of *X* to the amount calculated in Step iii, college costs as of $t = 17$, using the formula for the future value of an annuity (Equation 7). Then solve for *X*.

$$\$70{,}919 = \left[\frac{(1.05)^{17}-1}{0.05}\right] = 25.840366X$$

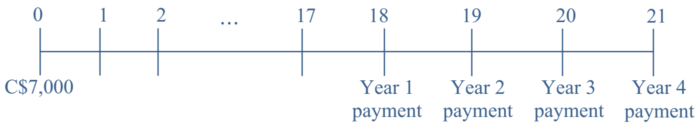
$$X = \$2{,}744.50$$

Notation Used on Most Calculators	**Numerical Value for This Problem**
N	17
%*i*	5
PV	n/a (= 0)
FV	$70,919
PMT **compute**	*X*


0
1
2
…
17
18
19
20
21
(X)
(X)
…
(X)
$20,000
$20,000
$20,000
$20,000
$70,919
Equation 11
25.8404X
= $70,919
Equation 7

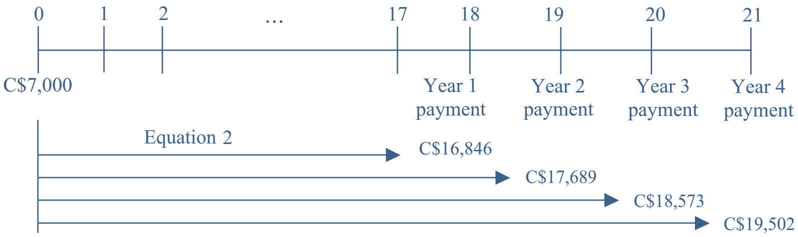
In summary, your client will have to save $2,744.50 each year if she starts next year and makes 17 payments into a savings account paying 5 percent annually.

9 **i.** Draw a time line.


0
1
2
…
17
18
19
20
21
C$7,000
Year 1 payment
Year 2 payment
Year 3 payment
Year 4 payment

ii. Recognize that the payments in Years 18, 19, 20, and 21 are the future values of a lump sum of C$7,000 in Year 0.

iii. With r = 5%, use the formula for the future value of a lump sum (Equation 2), $FV_N = PV\ (1 + r)^N$, four times to find the payments. These future values are shown on the time line below.


0
1
2
…
17
18
19
20
21
C$7,000
Year 1 payment
Year 2 payment
Year 3 payment
Year 4 payment
Equation 2
C$16,846
C$17,689
C$18,573
C$19,502

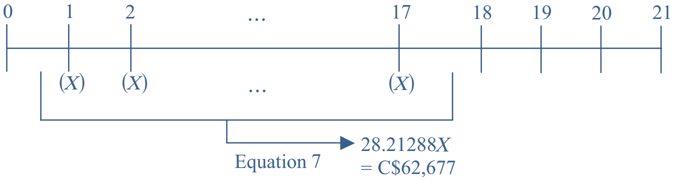
iv. Using the formula for the present value of a lump sum (r = 6%), equate the four college payments to single payments as of t = 17 and add them together. $C\$16{,}846(1.06)^{-1} + C\$17{,}689(1.06)^{-2} + C\$18{,}573(1.06)^{-3} + C\$19{,}502(1.06)^{-4} = C\$62{,}677$

v. Equate the sum of C$62,677 at t = 17 to the 17 payments of X, using the formula for the future value of an annuity (Equation 7). Then solve for X.

$$C\$62{,}677 = X\left[\frac{(1.06)^{17} - 1}{0.06}\right] = 28.21288X$$

$$X = C\$2{,}221.58$$

Notation Used on Most Calculators	**Numerical Value for This Problem**
N	17
$\%i$	6
PV	n/a (= 0)
FV	C$62,677
PMT **compute**	X



In summary, the couple will need to put aside C$2,221.58 each year if they start next year and make 17 equal payments.

10 C is correct. The sum of the real risk-free interest rate and the inflation premium is the nominal risk-free rate.

11 C is correct. US Treasury bonds are highly liquid, whereas the bonds of small issuers trade infrequently and the interest rate includes a liquidity premium. This liquidity premium reflects the relatively high costs (including the impact on price) of selling a position.

12 A is correct. The effective annual rate (EAR) when compounded daily is 4.08%.

$$\text{EAR} = (1 + \text{Periodic interest rate})^m - 1$$

$$\text{EAR} = (1 + 0.04/365)^{365} - 1$$

$$\text{EAR} = (1.0408) - 1 = 0.04081 \approx 4.08\%.$$

13 C is correct, as shown in the following (where FV is future value and PV is present value):

$$\text{FV} = \text{PV}\left(1 + \frac{r_s}{m}\right)^{mN}$$

$$\text{FV}_6 = \$75{,}000\left(1 + \frac{0.07}{4}\right)^{(4\times6)}$$

$$\text{FV}_6 = \$113{,}733.21.$$

14 B is correct because £97,531 represents the present value (PV) of £100,000 received one year from today when today's deposit earns a stated annual rate of 2.50% and interest compounds weekly, as shown in the following equation (where FV is future value):

$$\text{PV} = \text{FV}_N\left(1 + \frac{r_s}{m}\right)^{-mN}$$

$$\text{PV} = £100{,}000\left(1 + \frac{0.025}{52}\right)^{-52}$$

$$\text{PV} = £97{,}531.58.$$

15 A is correct. The effective annual rate (EAR) is calculated as follows:

$$\text{EAR} = (1 + \text{Periodic interest rate})^m - 1$$

$$\text{EAR} = (1 + 0.03/365)^{365} - 1$$

$$\text{EAR} = (1.03045) - 1 = 0.030453 \approx 3.0453\%.$$

Solving for N on a financial calculator results in (where FV is future value and PV is present value):

$(1 + 0.030453)^N = \text{FV}_N/\text{PV} = ¥1{,}000{,}000/¥250{,}000$

= 46.21 years, which multiplied by 12 to convert to months results in 554.5, or ≈ 555 months.

16 B is correct. The difference between continuous compounding and daily compounding is

€127,496.85 – €127,491.29 = €5.56, or ≈ €6, as shown in the following calculations.

With continuous compounding, the investment earns (where PV is present value)

$$\begin{aligned} \text{PV}e^{r_s N} - \text{PV} &= €1{,}000{,}000e^{0.03(4)} - €1{,}000{,}000 \\ &= €1{,}127{,}496.85 - €1{,}000{,}000 \\ &= €127{,}496.85 \end{aligned}$$

With daily compounding, the investment earns:

$€1{,}000{,}000(1 + 0.03/365)^{365(4)} - €1{,}000{,}000 = €1{,}127{,}491.29 - €1{,}000{,}000 = €127{,}491.29$.

17 B is correct, as shown in the following calculation for an annuity (A) due:

$$\text{PV} = A\left[\frac{1 - \dfrac{1}{(1+r)^N}}{r}\right](1 + r)$$

where A = €300, r = 0.04, and N = 5.

$$\text{PV} = €300\left[\frac{1 - \dfrac{1}{(1+.04)^5}}{.04}\right](1.04)$$

PV = €1,388.97, or ≈ €1,389.

18 B is correct. The value of the perpetuity one year from now is calculated as:

PV = A/r, where PV is present value, A is annuity, and r is expressed as a quarterly required rate of return because the payments are quarterly.

PV = \$2.00/(0.06/4)

PV = \$133.33.

The value today is (where FV is future value)

$\text{PV} = \text{FV}_N(1 + r)^{-N}$

$\text{PV} = \$133.33(1 + 0.015)^{-4}$

PV = \$125.62 ≈ \$126.

19 B is correct. To solve for the future value of unequal cash flows, compute the future value of each payment as of Year 4 at the semiannual rate of 2%, and then sum the individual future values, as follows:

Year	End of Year Deposits ($)	Factor	Future Value ($)
1	4,000	$(1.02)^6$	4,504.65
2	8,000	$(1.02)^4$	8,659.46
3	7,000	$(1.02)^2$	7,282.80
4	10,000	$(1.02)^0$	10,000.00
		Sum =	30,446.91

20 C is correct, as shown in the following (where FV is future value and PV is present value):

If:

$$\mathrm{FV}_N = \mathrm{PV}\left(1+\frac{r_s}{m}\right)^{mN}$$

Then:

$$\left(\frac{\mathrm{FV}_N}{\mathrm{PV}}\right)^{\frac{1}{mN}} - 1 = \frac{r_s}{m}$$

$$\left(\frac{800{,}000}{500{,}000}\right)^{\frac{1}{2\times 6}} - 1 = \frac{r_s}{2}$$

$r_s = 0.07988$ (rounded to 8.0%).

21 C is correct. As shown below, the present value (PV) of a £2,000 per month perpetuity is worth approximately £400,000 at a 6% annual rate compounded monthly. Thus, the present value of the annuity (A) is worth more than the lump sum offers.

A = £2,000

r = (6%/12) = 0.005

PV = (A/r)

PV = (£2,000/0.005)

PV = £400,000

22 B is correct.

The present value of a 10-year annuity (A) due with payments of $2,000 at a 5% discount rate is calculated as follows:

$$\mathrm{PV} = A\left[\frac{1-\frac{1}{(1+r)^N}}{r}\right] + \$2{,}000$$

$$PV = \$2{,}000\left[\frac{1-\dfrac{1}{(1+0.05)^9}}{0.05}\right] + \$2{,}000$$

PV = \$16,215.64.

Alternatively, the PV of a 10-year annuity due is simply the PV of the ordinary annuity multiplied by 1.05:

$PV = \$15{,}443.47 \times 1.05$

PV = \$16,215.64.

23 B is correct. First, find the present value (PV) of an ordinary annuity in Year 17 that represents the tuition costs:

$$\$50{,}000\left[\frac{1-\dfrac{1}{(1+0.06)^4}}{0.06}\right]$$

$= \$50{,}000 \times 3.4651$

= \$173,255.28.

Then, find the PV of the annuity in today's dollars (where FV is future value):

$$PV_0 = \frac{FV}{(1+0.06)^{17}}$$

$$PV_0 = \frac{\$173{,}255.28}{(1+0.06)^{17}}$$

$PV_0 = \$64{,}340.85 \approx \$64{,}341.$

24 B is correct, as shown in the following table.

Year	Cash Flow (€)	Formula CF × (1 + *r*)^t	PV at Year 0
1	100,000	$100{,}000(1.12)^{-1}$ =	89,285.71
2	150,000	$150{,}000(1.12)^{-2}$ =	119,579.08
5	–10,000	$-10{,}000(1.12)^{-5}$ =	–5,674.27
			203,190.52

25 B is correct, calculated as follows (where A is annuity and PV is present value):

$$A = (\text{PV of annuity}) \Bigg/ \left[\frac{1 - \dfrac{1}{(1 + r_s/m)^{mN}}}{r_s/m} \right]$$

$$= (£200{,}000) \Bigg/ \left[\frac{1 - \dfrac{1}{(1 + r_s/m)^{mN}}}{r_s/m} \right]$$

$$(£200{,}000) \Bigg/ \left[\frac{1 - \dfrac{1}{(1 + 0.06/12)^{12(5)}}}{0.06/12} \right]$$

$$= (£200{,}000)/51.72556$$

$$= £3{,}866.56$$

26 A is correct. To solve for an annuity (A) payment, when the future value (FV), interest rate, and number of periods is known, use the following equation:

$$\text{FV} = A \left[\frac{\left(1 + \dfrac{r_s}{m}\right)^{mN} - 1}{\dfrac{r}{m}} \right]$$

$$£25{,}000 = A \left[\frac{\left(1 + \dfrac{0.06}{4}\right)^{4 \times 10} - 1}{\dfrac{0.06}{4}} \right]$$

$$A = £460.68$$

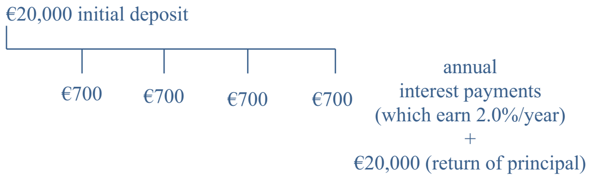
27 B is correct. The PV in Year 5 of a \$50,000 lump sum paid in Year 20 is \$27,763.23 (where FV is future value):

$$\text{PV} = \text{FV}_N(1 + r)^{-N}$$

$$\text{PV} = \$50{,}000(1 + 0.04)^{-15}$$

$$\text{PV} = \$27{,}763.23$$

28 B is correct, as the following cash flows show:


€20,000 initial deposit
€700
€700
€700
€700
annual
interest payments
(which earn 2.0%/year)
+
€20,000 (return of principal)

The four annual interest payments are based on the CD's 3.5% annual rate.

The first payment grows at 2.0% compounded monthly for three years (where FV is future value):

$$\text{FV}_N = €700\left(1 + \frac{0.02}{12}\right)^{3\times12}$$

$$\text{FV}_N = 743.25$$

The second payment grows at 2.0% compounded monthly for two years:

$$\text{FV}_N = €700\left(1 + \frac{0.02}{12}\right)^{2\times12}$$

$$\text{FV}_N = 728.54$$

The third payment grows at 2.0% compounded monthly for one year:

$$\text{FV}_N = €700\left(1 + \frac{0.02}{12}\right)^{1\times12}$$

$$\text{FV}_N = 714.13$$

The fourth payment is paid at the end of Year 4. Its future value is €700.

The sum of all future value payments is as follows:

€20,000.00	CD
€743.25	First payment's *FV*
€728.54	Second payment's *FV*
€714.13	Third payment's *FV*
€700.00	Fourth payment's *FV*
€22,885.92	Total *FV*

READING

2

Organizing, Visualizing, and Describing Data

by Pamela Peterson Drake, PhD, CFA, and Jian Wu, PhD

Pamela Peterson Drake, PhD, CFA, is at James Madison University (USA). Jian Wu, PhD, is at State Street (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** identify and compare data types;
☐	**b.** describe how data are organized for quantitative analysis;
☐	**c.** interpret frequency and related distributions;
☐	**d.** interpret a contingency table;
☐	**e.** describe ways that data may be visualized and evaluate uses of specific visualizations;
☐	**f.** describe how to select among visualization types;
☐	**g.** calculate and interpret measures of central tendency;
☐	**h.** evaluate alternative definitions of mean to address an investment problem;
☐	**i.** calculate quantiles and interpret related visualizations;
☐	**j.** calculate and interpret measures of dispersion;
☐	**k.** calculate and interpret target downside deviation;
☐	**l.** interpret skewness;
☐	**m.** interpret kurtosis;
☐	**n.** interpret correlation between two variables.

© 2020 CFA Institute. All rights reserved.

1 INTRODUCTION

Data have always been a key input for securities analysis and investment management, but the acceleration in the availability and the quantity of data has also been driving the rapid evolution of the investment industry. With the rise of big data and machine learning techniques, investment practitioners are embracing an era featuring large volume, high velocity, and a wide variety of data—allowing them to explore and exploit this abundance of information for their investment strategies.

While this data-rich environment offers potentially tremendous opportunities for investors, turning data into useful information is not so straightforward. Organizing, cleaning, and analyzing data are crucial to the development of successful investment strategies; otherwise, we end up with "garbage in and garbage out" and failed investments. It is often said that 80% of an analyst's time is spent on finding, organizing, cleaning, and analyzing data, while just 20% of her/his time is taken up by model development. So, the importance of having a properly organized, cleansed, and well-analyzed dataset cannot be over-emphasized. With this essential requirement met, an appropriately executed data analysis can detect important relationships within data, uncover underlying structures, identify outliers, and extract potentially valuable insights. Utilizing both visual tools and quantitative methods, like the ones covered in this reading, is the first step in summarizing and understanding data that will be crucial inputs to an investment strategy.

This reading provides a foundation for understanding important concepts that are an indispensable part of the analytical tool kit needed by investment practitioners, from junior analysts to senior portfolio managers. These basic concepts pave the way for more sophisticated tools that will be developed as the quantitative methods topic unfolds and that are integral to gaining competencies in the investment management techniques and asset classes that are presented later in the CFA curriculum.

Section 2 covers core data types, including continuous and discrete numerical data, nominal and ordinal categorical data, and structured versus unstructured data. Organizing data into arrays and data tables and summarizing data in frequency distributions and contingency tables are discussed in Sections 3–5. Section 6 introduces the important topic of data visualization using a range of charts and graphics to summarize, explore, and better understand data. Section 7 covers the key measures of central tendency, including several variants of mean that are especially useful in investments. Quantiles and their investment applications are the focus of Section 8. Key measures of dispersion are discussed in Sections 9 and 10. The shape of data distributions—specifically, skewness and kurtosis—are covered in Section 11. Section 12 provides a graphical introduction to covariance and correlation between two variables. The reading concludes with a Summary.

2 DATA TYPES

a Identify and compare data types

Data can be defined as a collection of numberpanel datas, characters, words, and text—as well as images, audio, and video—in a raw or organized format to represent facts or information. To choose the appropriate statistical methods for summarizing and analyzing data and to select suitable charts for visualizing data, we need to distinguish among different data types. We will discuss data types under three different perspectives of classifications: numerical versus categorical data; cross-sectional vs. time-series vs. panel data; and structured vs. unstructured data.

2.1 Numerical versus Categorical Data

From a statistical perspective, data can be classified into two basic groups: numerical data and categorical data.

2.1.1 *Numerical Data*

Numerical data are values that represent measured or counted quantities as a number and are also called **quantitative data**. Numerical (quantitative) data can be split into two types: continuous data and discrete data.

Continuous data are data that can be measured and can take on any numerical value in a specified range of values. For example, the future value of a lump-sum investment measures the amount of money to be received after a certain period of time bearing an interest rate. The future value could take on a range of values depending on the time period and interest rate. Another common example of continuous data is the price returns of a stock that measures price change over a given period in percentage terms.

Discrete data are numerical values that result from a counting process. So, practically speaking, the data are limited to a finite number of values. For example, the frequency of discrete compounding, m, counts the number of times that interest is accrued and paid out in a given year. The frequency could be monthly ($m = 12$), quarterly ($m = 4$), semi-yearly ($m = 2$), or yearly ($m = 1$).

2.1.2 *Categorical Data*

Categorical data (also called **qualitative data**) are values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize. Usually they can take only a limited number of values that are mutually exclusive. Examples of categorical data for classifying companies include bankrupt vs. not bankrupt and dividends increased vs. no dividend action.

Nominal data are categorical values that are not amenable to being organized in a logical order. An example of nominal data is the classification of publicly listed stocks into 11 sectors, as shown in Exhibit 1, that are defined by the Global Industry Classification Standard (GICS). GICS, developed by Morgan Stanley Capital International (MSCI) and Standard & Poor's (S&P), is a four-tiered, hierarchical industry classification system consisting of 11 sectors, 24 industry groups, 69 industries, and 158 sub-industries. Each sector is defined by a unique text label, as shown in the column named "Sector."

Exhibit 1 Equity Sector Classification by GICS

Sector (Text Label)	Code (Numerical Label)
Energy	10
Materials	15
Industrials	20
Consumer Discretionary	25
Consumer Staples	30
Health Care	35
Financials	40
Information Technology	45
Communication Services	50

(continued)

Exhibit 1 (Continued)

Sector (Text Label)	Code (Numerical Label)
Utilities	55
Real Estate	60

Source: S&P Global Market Intelligence.

Text labels are a common format to represent nominal data, but nominal data can also be coded with numerical labels. As shown below, the column named "Code" contains a corresponding GICS code of each sector as a numerical value. However, the nominal data in numerical format do not indicate ranking, and any arithmetic operations on nominal data are not meaningful. In this example, the energy sector with the code 10 does not represent a lower or higher rank than the real estate sector with the code 60. Often, financial models, such as regression models, require input data to be numerical; so, nominal data in the input dataset must be coded numerically before applying an algorithm (that is, a process for problem solving) for performing the analysis. This would be mainly to identify the category (here, sector) in the model.

Ordinal data are categorical values that can be logically ordered or ranked. For example, the Morningstar and Standard & Poor's star ratings for investment funds are ordinal data in which one star represents a group of funds judged to have had relatively the worst performance, with two, three, four, and five stars representing groups with increasingly better performance or quality as evaluated by those firms.

Ordinal data may also involve numbers to identify categories. For example, in ranking growth-oriented investment funds based on their five-year cumulative returns, we might assign the number 1 to the top performing 10% of funds, the number 2 to next best performing 10% of funds, and so on; the number 10 represents the bottom performing 10% of funds. Despite the fact that categories represented by ordinal data can be ranked higher or lower compared to each other, they do not necessarily establish a numerical difference between each category. Importantly, such investment fund ranking tells us nothing about the difference in performance between funds ranked 1 and 2 compared with the difference in performance between funds ranked 3 and 4 or 9 and 10.

Having discussed different data types from a statistical perspective, it is important to note that at first glance, identifying data types may seem straightforward. In some situations, where categorical data are coded in numerical format, they should be distinguished from numerical data. A sound rule of thumb: Meaningful arithmetic operations can be performed on numerical data but not on categorical data.

EXAMPLE 1

Identifying Data Types (I)

Identify the data type for each of the following kinds of investment-related information:

1 *Number of coupon payments for a corporate bond.* As background, a corporate bond is a contractual obligation between an issuing corporation (i.e., borrower) and bondholders (i.e., lenders) in which the issuer agrees to pay interest—in the form of fixed coupon payments—on specified dates, typically semi-annually, over the life of the bond (i.e., to its maturity date) and to repay principal (i.e., the amount borrowed) at maturity.

2 *Cash dividends per share paid by a public company.* Note that cash dividends are a distribution paid to shareholders based on the number of shares owned.

3 *Credit ratings for corporate bond issues.* As background, credit ratings gauge the bond issuer's ability to meet the promised payments on the bond. Bond rating agencies typically assign bond issues to discrete categories that are in descending order of credit quality (i.e., increasing probability of non-payment or default).

4 *Hedge fund classification types.* Note that hedge funds are investment vehicles that are relatively unconstrained in their use of debt, derivatives, and long and short investment strategies. Hedge fund classification types group hedge funds by the kind of investment strategy they pursue.

Solution to 1

Number of coupon payments are discrete data. For example, a newly-issued 5-year corporate bond paying interest semi-annually (quarterly) will make 10 (20) coupon payments during its life. In this case, coupon payments are limited to a finite number of values; so, they are discrete.

Solution to 2

Cash dividends per share are continuous data since they can take on any non-negative values.

Solution to 3

Credit ratings are ordinal data. A rating places a bond issue in a category, and the categories are ordered with respect to the expected probability of default. But arithmetic operations cannot be done on credit ratings, and the difference in the expected probability of default between categories of highly rated bonds, for example, is not necessarily equal to that between categories of lowly rated bonds.

Solution to 4

Hedge fund classification types are nominal data. Each type groups together hedge funds with similar investment strategies. In contrast to credit ratings for bonds, however, hedge fund classification schemes do not involve a ranking. Thus, such classification schemes are not ordinal data.

2.2 Cross-Sectional versus Time-Series versus Panel Data

Another data classification standard is based on how data are collected, and it categorizes data into three types: cross-sectional, time series, and panel.

Prior to the description of the data types, we need to explain two data-related terminologies: variable and observation. A **variable** is a characteristic or quantity that can be measured, counted, or categorized and is subject to change. A variable can also be called a field, an attribute, or a feature. For example, stock price, market capitalization, dividend and dividend yield, earnings per share (EPS), and price-to-earnings ratio (P/E) are basic data variables for the financial analysis of a public company. An **observation** is the value of a specific variable collected at a point in time or over a specified period of time. For example, last year DEF, Inc. recorded EPS of $7.50. This value represented a 15% annual increase.

Cross-sectional data are a list of the observations of a specific variable from multiple observational units at a given point in time. The observational units can be individuals, groups, companies, trading markets, regions, etc. For example, January inflation rates (i.e., the variable) for each of the euro-area countries (i.e., the observational units) in the European Union for a given year constitute cross-sectional data.

Time-series data are a sequence of observations for a single observational unit of a specific variable collected over time and at discrete and typically equally spaced intervals of time, such as daily, weekly, monthly, annually, or quarterly. For example, the daily closing prices (i.e., the variable) of a particular stock recorded for a given month constitute time-series data.

Panel data are a mix of time-series and cross-sectional data that are frequently used in financial analysis and modeling. Panel data consist of observations through time on one or more variables for multiple observational units. The observations in panel data are usually organized in a matrix format called a data table. Exhibit 2 is an example of panel data showing quarterly earnings per share (i.e., the variable) for three companies (i.e., the observational units) in a given year by quarter. Each column is a time series of data that represents the quarterly EPS observations from Q1 to Q4 of a specific company, and each row is cross-sectional data that represent the EPS of all three companies of a particular quarter.

Exhibit 2 Earnings per Share in Euros of Three Eurozone Companies in a Given Year

Time Period	Company A	Company B	Company C
Q1	13.53	0.84	−0.34
Q2	4.36	0.96	0.08
Q3	13.16	0.79	−2.72
Q4	12.95	0.19	0.09

2.3 Structured versus Unstructured Data

Categorizing data into structured and unstructured types is based on whether or not the data are in a highly organized form.

Structured data are highly organized in a pre-defined manner, usually with repeating patterns. The typical forms of structured data are one-dimensional arrays, such as a time series of a single variable, or two-dimensional data tables, where each column represents a variable or an observation unit and each row contains a set of values for the same columns. Structured data are relatively easy to enter, store, query, and analyze without much manual processing. Typical examples of structured company financial data are:

- Market data: data issued by stock exchanges, such as intra-day and daily closing stock prices and trading volumes.
- Fundamental data: data contained in financial statements, such as earnings per share, price to earnings ratio, dividend yield, and return on equity.
- Analytical data: data derived from analytics, such as cash flow projections or forecasted earnings growth.

Unstructured data, in contrast, are data that do not follow any conventionally organized forms. Some common types of unstructured data are text—such as financial news, posts in social media, and company filings with regulators—and also audio/video, such as managements' earnings calls and presentations to analysts.

Unstructured data are a relatively new classification driven by the rise of alternative data (i.e., data generated from unconventional sources, like electronic devices, social media, sensor networks, and satellites, but also by companies in the normal course of business) and its growing adoption in the financial industry. Unstructured data are typically alternative data as they are usually collected from unconventional sources. By indicating the source from which the data are generated, such data can be classified into three groups:

- Produced by individuals (i.e., via social media posts, web searches, etc.);
- Generated by business processes (i.e., via credit card transactions, corporate regulatory filings, etc.); and
- Generated by sensors (i.e., via satellite imagery, foot traffic by mobile devices, etc.).

Unstructured data may offer new market insights not normally contained in data from traditional sources and may provide potential sources of returns for investment processes. Unlike structured data, however, utilizing unstructured data in investment analysis is challenging. Typically, financial models are able to take only structured data as inputs; therefore, unstructured data must first be transformed into structured data that models can process.

Exhibit 3 shows an excerpt from Form 10-Q (Quarterly Report) filed by Company XYZ with the US Securities and Exchange Commission (SEC) for the fiscal quarter ended 31 March 20XX. The form is an unstructured mix of text and tables, so it cannot be directly used by computers as input to financial models. The SEC has utilized eXtensible Business Reporting Language (XBRL) to structure such data. The data extracted from the XBRL submission can be organized into five tab-delimited TXT format files that contain information about the submission, including taxonomy tags (i.e., financial statement items), dates, units of measure (uom), values (i.e., for the tag items), and more—making it readable by computer. Exhibit 4 shows an excerpt from one of the now structured data tables downloaded from the SEC's EDGAR (Electronic Data Gathering, Analysis, and Retrieval) database.

Exhibit 3 Excerpt from 10-Q of Company XYZ for Fiscal Quarter Ended 31 March 20XX

Company XYZ
Form 10-Q
Fiscal Quarter Ended 31 March 20XX
Table of Contents

Part I		
		Page
Item 1	Financial Statements	1
Item 2	Management's Discussion and Analysis of Financial Condition and Results of Operations	21
Item 3	Quantitative and Qualitative Disclosures About Market Risk	32
Item 4	Controls and Procedures	32
Part II		
Item 1	Legal Proceedings	33
Item 1A	Risk Factors	33
Item 2	Unregistered Sales of Equity Securities and Use of Proceeds	43
Item 3	Defaults Upon Senior Securities	43

(continued)

Exhibit 3 (Continued)

Item 4 Mine Safety Disclosures 43
Item 5 Other Information 43
Item 6 Exhibits 44

Condensed Consolidated Statements of Operations (Unaudited)
(in millions, except number of shares, which are reflected in thousands and per share amounts)

	31 March 20XX
Net sales:	
Products	$46,565
Services	11,450
Total net sales	58,015
Cost of sales:	
Products	32,047
Services	4,147
Total cost of sales	36,194
Gross margin	21,821
Operating expenses:	
Research and development	3,948
Selling, general and administrative	4,458
Total operating expenses	8,406
Operating income	13,415
Other income/(expense), net	378
Income before provision for income taxes	13,793
Provision for income taxes	2,232
Net income	$11,561

Source: EDGAR.

Exhibit 4 Structured Data Extracted from Form 10-Q of Company XYZ for Fiscal Quarter Ended 31 March 20XX

adsh	tag	ddate	uom	value
0000320193-19-000066	RevenueFromContractWithCustomerExcludingAssessedTax	20XX0331	USD	$58,015,000,000
0000320193-19-000066	GrossProfit	20XX0331	USD	$21,821,000,000
0000320193-19-000066	OperatingExpenses	20XX0331	USD	$8,406,000,000

Exhibit 4 (Continued)

adsh	tag	ddate	uom	value
0000320193-19-000066	OperatingIncomeLoss	20XX0331	USD	$13,415,000,000
0000320193-19-000066	NetIncomeLoss	20XX0331	USD	$11,561,000,000

Source: EDGAR.

EXAMPLE 2

Identifying Data Types (II)

1 Which of the following is *most likely* to be structured data?

- **A** Social media posts where consumers are commenting on what they think of a company's new product.
- **B** Daily closing prices during the past month for all companies listed on Japan's Nikkei 225 stock index.
- **C** Audio and video of a CFO explaining her company's latest earnings announcement to securities analysts.

2 Which of the following statements describing panel data is *most accurate*?

- **A** It is a sequence of observations for a single observational unit of a specific variable collected over time at discrete and equally spaced intervals.
- **B** It is a list of observations of a specific variable from multiple observational units at a given point in time.
- **C** It is a mix of time-series and cross-sectional data that are frequently used in financial analysis and modeling.

3 Which of the following data series is *least likely* to be sortable by values?

- **A** Daily trading volumes for stocks listed on the Shanghai Stock Exchange.
- **B** EPS for a given year for technology companies included in the S&P 500 Index.
- **C** Dates of first default on bond payments for a group of bankrupt European manufacturing companies.

4 Which of the following best describes a time series?

- **A** Daily stock prices of the XYZ stock over a 60-month period.
- **B** Returns on four-star rated Morningstar investment funds at the end of the most recent month.
- **C** Stock prices for all stocks in the FTSE100 on 31 December of the most recent calendar year.

Solution to 1

B is correct as daily closing prices constitute structured data. A is incorrect as social media posts are unstructured data. C is incorrect as audio and video are unstructured data.

Solution to 2

C is correct as it most accurately describes panel data. A is incorrect as it describes time-series data. B is incorrect as it describes cross-sectional data.

Solution to 3

C is correct as dates are ordinal data that can be sorted by chronological order but not by value. A and B are incorrect as both daily trading volumes and earnings per share (EPS) are numerical data, so they can be sorted by values.

Solution to 4

A is correct since a time series is a sequence of observations of a specific variable (XYZ stock price) collected over time (60 months) and at discrete intervals of time (daily). B and C are both incorrect as they are cross-sectional data.

2.4 Data Summarization

b Describe how data are organized for quantitative analysis

Given the wide variety of possible formats of **raw data**, which are data available in their original form as collected, such data typically cannot be used by humans or computers to directly extract information and insights. Organizing data into a one-dimensional array or a two-dimensional array is typically the first step in data analytics and modeling. In this section, we will illustrate the construction of these typical data organization formats. We will also introduce two useful tools that can efficiently summarize one-variable and two-variable data: frequency distributions and contingency tables, respectively. Both of them can give us a quick snapshot of the data and allow us to find patterns in the data and associations between variables.

3 ORGANIZING DATA FOR QUANTITATIVE ANALYSIS

b Describe how data are organized for quantitative analysis

Quantitative analysis and modeling typically require input data to be in a clean and formatted form, so raw data are usually not suitable for use directly by analysts. Depending upon the number of variables, raw data can be organized into two typical formats for quantitative analysis: one-dimensional arrays and two-dimensional rectangular arrays.

A **one-dimensional array** is the simplest format for representing a collection of data of the same data type, so it is suitable for representing a single variable. Exhibit 5 is an example of a one-dimensional array that shows the closing price for the first 10 trading days for ABC Inc. stock after the company went public. Closing prices are time-series data collected at daily intervals, so it is natural to organize them into a time-ordered sequence. The time-series format also facilitates future data updates to the existing dataset. In this case, closing prices for future trading sessions can be easily added to the end of the array with no alteration of previously formatted data.

More importantly, in contrast to compiling the data randomly in an unorganized manner, organizing such data by its time-series nature preserves valuable information beyond the basic **descriptive statistics** that summarize central tendency and spread variation in the data's distribution. For example, by simply plotting the data against time, we can learn whether the data demonstrate any increasing or decreasing trends over time or whether the time series repeats certain patterns in a systematic way over time.

Exhibit 5 One-Dimensional Array: Daily Closing Price of ABC Inc. Stock

Observation by Day	Stock Price ($)
1	57.21
2	58.26
3	58.64
4	56.19
5	54.78
6	54.26
7	56.88
8	54.74
9	52.42
10	50.14

A **two-dimensional rectangular array** (also called a **data table**) is one of the most popular forms for organizing data for processing by computers or for presenting data visually for consumption by humans. Similar to the structure in an Excel spreadsheet, a data table is comprised of columns and rows to hold multiple variables and multiple observations, respectively. When a data table is used to organize the data of one single observational unit (i.e., a single company), each column represents a different variable (feature or attribute) of that observational unit, and each row holds an observation for the different variables; successive rows represent the observations for successive time periods. In other words, observations of each variable are a time-series sequence that is sorted in either ascending or descending time order. Consequently, observations of different variables must be sorted and aligned to the same time scale. Example 3 shows how to organize a raw dataset for a company collected online into a machine-readable data table.

EXAMPLE 3

Organizing a Company's Raw Data into a Data Table

Suppose you are conducting a valuation analysis of ABC Inc., which has been listed on the stock exchange for two years. The metrics to be used in your valuation include revenue, earnings per share (EPS), and dividends paid per share (DPS). You have retrieved the last two years of ABC's quarterly data from the exchange's website, which is shown in Exhibit 6. The data available online are pre-organized into a tabular format, where each column represents a fiscal year and each row represents a particular quarter with values of the three measures clustered together.

Exhibit 6 Metrics of ABC Inc. Retrieved Online

Fiscal Quarter	Year 1 (Fiscal Year)	Year 2 (Fiscal Year)
March		
Revenue	$3,784(M)	$4,097(M)
EPS	1.37	−0.34
DPS	N/A	N/A
June		
Revenue	$4,236(M)	$5,905(M)
EPS	1.78	3.89
DPS	N/A	0.25
September		
Revenue	$4,187(M)	$4,997(M)
EPS	−3.38	−2.88
DPS	N/A	0.25
December		
Revenue	$3,889(M)	$4,389(M)
EPS	−8.66	−3.98
DPS	N/A	0.25

Use the data to construct a two-dimensional rectangular array (i.e., data table) with the columns representing the metrics for valuation and the observations arranged in a time-series sequence.

Solution:

To construct a two-dimensional rectangular array, we first need to determine the data table structure. The columns have been specified to represent the three valuation metrics (i.e., variables): revenue, EPS and DPS. The rows should be the observations for each variable in a time ordered sequence. In this example, the data for the valuation measures will be organized in the same quarterly intervals as the raw data retrieved online, starting from Q1 Year 1 to Q4 Year 2. Then, the observations from the original table can be placed accordingly into the data table by variable name and by filing quarter. Exhibit 7 shows the raw data reorganized in the two-dimensional rectangular array (by date and associated valuation metric), which can now be used in financial analysis and is readable by a computer.

It is worth pointing out that in case of missing values while organizing data, how to handle them depends largely on why the data are missing. In this example, dividends (DPS) in the first five quarters are missing because ABC Inc. did not authorize (and pay) any dividends. So, filling the dividend column with zeros is appropriate. If revenue, EPS, and DPS of a given quarter are missing due to particular data source issues, however, these missing values cannot be simply replaced with zeros; this action would result in incorrect interpretation. Instead, the missing values might be replaced with the latest available data or with interpolated values, depending on how the data will be consumed or modeled.

Exhibit 7 Data Table for ABC Inc.

	Revenue ($ Million)	EPS ($)	DPS ($)
Q1 Year 1	3,784	1.37	0
Q2 Year 1	4,236	1.78	0
Q3 Year 1	4,187	−3.38	0
Q4 Year 1	3,889	−8.66	0
Q1 Year 2	4,097	−0.34	0
Q2 Year 2	5,905	3.89	0.25
Q3 Year 2	4,997	−2.88	0.25
Q4 Year 2	4,389	−3.98	0.25

SUMMARIZING DATA USING FREQUENCY DISTRIBUTIONS

4

c **Interpret frequency and related distributions**

We now discuss various tabular formats for describing data based on the count of observations. These tables are a necessary step toward building a true visualization of a dataset. Later, we shall see how bar charts, tree-maps, and heat maps, among other graphic tools, are used to visualize important properties of a dataset.

A **frequency distribution** (also called a one-way table) is a tabular display of data constructed either by counting the observations of a variable by distinct values or groups or by tallying the values of a numerical variable into a set of numerically ordered bins. It is an important tool for initially summarizing data by groups or bins for easier interpretation.

Constructing a frequency distribution of a categorical variable is relatively straightforward and can be stated in the following two basic steps:

1 Count the number of observations for each unique value of the variable.

2 Construct a table listing each unique value and the corresponding counts, and then sort the records by number of counts in descending or ascending order to facilitate the display.

Exhibit 8 shows a frequency distribution of a portfolio's stock holdings by sectors (the variables), which are defined by GICS. The portfolio contains a total of 479 stocks that have been individually classified into 11 GICS sectors (first column). The stocks are counted by sector and are summarized in the second column, absolute frequency. The **absolute frequency**, or simply the raw frequency, is the actual number of observations counted for each unique value of the variable (i.e., each sector). Often it is desirable to express the frequencies in terms of percentages, so we also show the **relative frequency** (in the third column), which is calculated as the absolute frequency of each unique value of the variable divided by the total number of observations. The relative frequency provides a normalized measure of the distribution of the data, allowing comparisons between datasets with different numbers of total observations.

Exhibit 8 Frequency Distribution for a Portfolio by Sector

Sector (Variable)	Absolute Frequency	Relative Frequency
Industrials	73	15.2%
Information Technology	69	14.4%
Financials	67	14.0%
Consumer Discretionary	62	12.9%
Health Care	54	11.3%
Consumer Staples	33	6.9%
Real Estate	30	6.3%
Energy	29	6.1%
Utilities	26	5.4%
Materials	26	5.4%
Communication Services	10	2.1%
Total	**479**	**100.0%**

A frequency distribution table provides a snapshot of the data, and it facilitates finding patterns. Examining the distribution of absolute frequency in Exhibit 8, we see that the largest number of stocks (73), accounting for 15.2% of the stocks in the portfolio, are held in companies in the industrials sector. The sector with the least number of stocks (10) is communication services, which represents just 2.1% of the stocks in the portfolio.

It is also easy to see that the top four sectors (i.e., industrials, information technology, financials, and consumer discretionary) have very similar relative frequencies, between 15.2% and 12.9%. Similar relative frequencies, between 6.9% and 5.4%, are also seen among several other sectors. Note that the absolute frequencies add up to the total number of stocks in the portfolio (479), and the sum of the relative frequencies should be equal to 100%.

Frequency distributions also help in the analysis of large amounts of numerical data. The procedure for summarizing numerical data is a bit more involved than that for summarizing categorical data because it requires creating non-overlapping bins (also called **intervals** or buckets) and then counting the observations falling into each bin. One procedure for constructing a frequency distribution for numerical data can be stated as follows:

1. Sort the data in ascending order.
2. Calculate the range of the data, defined as Range = Maximum value − Minimum value.
3. Decide on the number of bins (k) in the frequency distribution.
4. Determine bin width as Range/k.
5. Determine the first bin by adding the bin width to the minimum value. Then, determine the remaining bins by successively adding the bin width to the prior bin's end point and stopping after reaching a bin that includes the maximum value.

6. Determine the number of observations falling into each bin by counting the number of observations whose values are equal to or exceed the bin minimum value yet are less than the bin's maximum value. The exception is in the last bin, where the maximum value is equal to the last bin's maximum, and therefore, the observation with the maximum value is included in this bin's count.
7. Construct a table of the bins listed from smallest to largest that shows the number of observations falling into each bin.

In Step 4, when rounding the bin width, round up (rather than down) to ensure that the final bin includes the maximum value of the data.

These seven steps are basic guidelines for constructing frequency distributions. In practice, however, we may want to refine the above basic procedure. For example, we may want the bins to begin and end with whole numbers for ease of interpretation. Another practical refinement that promotes interpretation is to start the first bin at the nearest whole number below the minimum value.

As this procedure implies, a frequency distribution groups data into a set of bins, where each bin is defined by a unique set of values (i.e., beginning and ending points). Each observation falls into only one bin, and the total number of bins covers all the values represented in the data. The frequency distribution is the list of the bins together with the corresponding measures of frequency.

To illustrate the basic procedure, suppose we have 12 observations sorted in ascending order (*Step 1*):

−4.57, −4.04, −1.64, 0.28, 1.34, 2.35, 2.38, 4.28, 4.42, 4.68, 7.16, and 11.43.

The minimum observation is −4.57, and the maximum observation is +11.43. So, the range is +11.43 − (−4.57) = 16 (*Step 2*).

If we set $k = 4$ (*Step 3*), then the bin width is 16/4 = 4 (*Step 4*).

Exhibit 9 shows the repeated addition of the bin width of 4 to determine the endpoint for each of the bins (*Step 5*).

Exhibit 9 Determining Endpoints of the Bins

−4.57	+	4.0	=	−0.57
−0.57	+	4.0	=	3.43
3.43	+	4.0	=	7.43
7.40	+	4.0	=	11.43

Thus, the bins are [−4.57 to −0.57), [−0.57 to 3.43), [3.43 to 7.43), and [7.43 to 11.43], where the notation [−4.57 to −0.57) indicates −4.57 ≤ observation < −0.57. The parentheses indicate that the endpoints are not included in the bins, and the square brackets indicate that the beginning points and the last endpoint are included in the bin. Exhibit 10 summarizes Steps 5 through 7.

Exhibit 10 Frequency Distribution

Bin				Absolute Frequency
A	−4.57	≤ observation <	−0.57	3
B	−0.57	≤ observation <	3.43	4
C	3.43	≤ observation <	7.43	4
D	7.43	≤ observation ≤	11.43	1

Note that the bins do not overlap, so each observation can be placed uniquely into one bin, and the last bin includes the maximum value.

We turn to these issues in discussing the construction of frequency distributions for daily returns of the fictitious Euro-Asia-Africa (EAA) Equity Index. The dataset of daily returns of the EAA Equity Index spans a five-year period and consists of 1,258 observations with a minimum value of −4.1% and a maximum value of 5.0%. Thus, the range of the data is 5% − (−4.1%) = 9.1%, approximately. [The mean daily return—mean as a measure of central tendency will be discussed shortly—is 0.04%.]

The decision on the number of bins (*k*) into which we should group the observations often involves inspecting the data and exercising judgment. How much detail should we include? If we use too few bins, we will summarize too much and may lose pertinent characteristics. Conversely, if we use too many bins, we may not summarize enough and may introduce unnecessary noise.

We can establish an appropriate value for *k* by evaluating the usefulness of the resulting bin width. A large number of empty bins may indicate that we are attempting to over-organize the data to present too much detail. Starting with a relatively small bin width, we can see whether or not the bins are mostly empty and whether or not the value of *k* associated with that bin width is too large. If the bins are mostly empty, implying that *k* is too large, we can consider increasingly larger bins (i.e., smaller values of *k*) until we have a frequency distribution that effectively summarizes the distribution.

Suppose that for ease of interpretation we want to use a bin width stated in whole rather than fractional percentages. In the case of the daily EAA Equity Index returns, a 1% bin width would be associated with 9.1/1 = 9.1 bins, which can be rounded up to *k* = 10 bins. That number of bins will cover a range of 1% × 10 = 10%. By constructing the frequency distribution in this manner, we will also have bins that end and begin at a value of 0%, thereby allowing us to count the negative and positive returns in the data. Without too much work, we have found an effective way to summarize the data.

Exhibit 11 shows the frequency distribution for the daily returns of the EAA Equity Index using return bins of 1%, where the first bin includes returns from −5.0% to −4.0% (exclusive, meaning < −4%) and the last bin includes daily returns from 4.0% to 5.0% (inclusive, meaning ≤ 5%). Note that to facilitate interpretation, the first bin starts at the nearest whole number below the minimum value (so, at −5.0%).

Exhibit 11 includes two other useful ways to present the data (which can be computed in a straightforward manner once we have established the absolute and relative frequency distributions): the cumulative absolute frequency and the cumulative relative frequency. The **cumulative absolute frequency** cumulates (meaning, adds up) the absolute frequencies as we move from the first bin to the last bin. Similarly, the **cumulative relative frequency** is a sequence of partial sums of the relative frequencies. For the last bin, the cumulative absolute frequency will equal the number observations in the dataset (1,258), and the cumulative relative frequency will equal 100%.

Exhibit 11 Frequency Distribution for Daily Returns of EAA Equity Index

Return Bin (%)	Absolute Frequency	Relative Frequency (%)	Cumulative Absolute Frequency	Cumulative Relative Frequency (%)
−5.0 to −4.0	1	0.08	1	0.08
−4.0 to −3.0	7	0.56	8	0.64
−3.0 to −2.0	23	1.83	31	2.46
−2.0 to −1.0	77	6.12	108	8.59
−1.0 to 0.0	470	37.36	578	45.95
0.0 to 1.0	555	44.12	1,133	90.06

Exhibit 11 (Continued)

Return Bin (%)	Absolute Frequency	Relative Frequency (%)	Cumulative Absolute Frequency	Cumulative Relative Frequency (%)
1.0 to 2.0	110	8.74	1,243	98.81
2.0 to 3.0	13	1.03	1,256	99.84
3.0 to 4.0	1	0.08	1,257	99.92
4.0 to 5.0	1	0.08	1,258	100.00

As Exhibit 11 shows, the absolute frequencies vary widely, ranging from 1 to 555. The bin encompassing returns between 0% and 1% has the most observations (555), and the corresponding relative frequency tells us these observations account for 44.12% of the total number of observations. The frequency distribution gives us a sense of not only where most of the observations lie but also whether the distribution is evenly spread. It is easy to see that the vast majority of observations (37.36% + 44.12% = 81.48%) lie in the middle two bins spanning −1% to 1%. We can also see that not many observations are greater than 3% or less than −4%. Moreover, as there are bins with 0% as ending or beginning points, we are able to count positive and negative returns in the data. Looking at the cumulative relative frequency in the last column, we see that the bin of −1% to 0% shows a cumulative relative frequency of 45.95%. This indicates that 45.95% of the observations lie below the daily return of 0% and that 54.05% of the observations are positive daily returns.

It is worth noting that other than being summarized in tables, frequency distributions also can be effectively represented in visuals, which will be discussed shortly in the section on data visualization.

EXAMPLE 4

Constructing a Frequency Distribution of Country Index Returns

Suppose we have the annual equity index returns of a given year for 18 different countries, as shown in Exhibit 12, and we are asked to summarize the data.

Exhibit 12 Annual Equity Index Returns for 18 Countries

Market	Index Return (%)
Country A	7.7
Country B	8.5
Country C	9.1
Country D	5.5
Country E	7.1
Country F	9.9
Country G	6.2
Country H	6.8
Country I	7.5

(continued)

Exhibit 12 (Continued)

Market	Index Return (%)
Country J	8.9
Country K	7.4
Country L	8.6
Country M	9.6
Country N	7.7
Country O	6.8
Country P	6.1
Country Q	8.8
Country R	7.9

Construct a frequency distribution table from these data and state some key findings from the summarized data.

Solution:

The first step in constructing a frequency distribution table is to sort the return data in ascending order:

Market	Index Return (%)
Country D	5.5
Country P	6.1
Country G	6.2
Country H	6.8
Country O	6.8
Country E	7.1
Country K	7.4
Country I	7.5
Country A	7.7
Country N	7.7
Country R	7.9
Country B	8.5
Country L	8.6
Country Q	8.8
Country J	8.9
Country C	9.1
Country M	9.6
Country F	9.9

The second step is to calculate the range of the data, which is 9.9% – 5.5% = 4.4%.
The third step is to decide on the number of bins. Here, we will use $k = 5$.
The fourth step is to determine the bin width. Here, it is 4.4%/5 = 0.88%, which we will round up to 1.0%.
The fifth step is to determine the bins, which are as follows:

5.0% + 1.0% = 6.0%
6.0% + 1.0% = 7.0%
7.0% + 1.0% = 8.0%

8.0% + 1.0% = 9.0%
9.0% + 1.0% = 10.0%

For ease of interpretation, the first bin is set to begin with the nearest whole number (5.0%) below the minimum value (5.5%) of the data series.

The sixth step requires counting the return observations falling into each bin, and the seventh (last) step is use these results to construct the final frequency distribution table.

Exhibit 13 presents the frequency distribution table, which summarizes the data in Exhibit 12 into five bins spanning 5% to 10%. Note that with 18 countries, the relative frequency for one observation is calculated as 1/18 = 5.56%.

Exhibit 13 Frequency Distribution of Equity Index Returns

Return Bin (%)	Absolute Frequency	Relative Frequency (%)	Cumulative Absolute Frequency	Cumulative Relative Frequency (%)
5.0 to 6.0	1	5.56	1	5.56
6.0 to 7.0	4	22.22	5	27.78
7.0 to 8.0	6	33.33	11	61.11
8.0 to 9.0	4	22.22	15	83.33
9.0 to 10.0	3	16.67	18	100.00

As Exhibit 13 shows, there is substantial variation in these equity index returns. One-third of the observations fall in the 7.0 to 8.0% bin, making it the bin with the most observations. Both the 6.0 to 7.0% bin and the 8.0 to 9.0% bin hold four observations each, accounting for 22.22% of the total number of the observations, respectively. The two remaining bins have fewer observations, one or three observations, respectively.

SUMMARIZING DATA USING A CONTINGENCY TABLE

5

d Interpret a contingency table

We have shown that the frequency distribution table is a powerful tool to summarize data for one variable. How can we summarize data for two variables simultaneously? A contingency table provides a solution to this question.

A **contingency table** is a tabular format that displays the frequency distributions of two or more categorical variables simultaneously and is used for finding patterns between the variables. A contingency table for two categorical variables is also known as a two-way table. Contingency tables are constructed by listing all the levels (i.e., categories) of one variable as rows and all the levels of the other variable as columns in the table. A contingency table having R levels of one variable in rows and C levels of the other variable in columns is referred to as an $R \times C$ table. Note that each variable in a contingency table must have a finite number of levels, which can be either ordered (ordinal data) or unordered (nominal data). Importantly, the data displayed in the cells of the contingency table can be either a frequency (count) or a relative frequency (percentage) based on either overall total, row totals, or column totals.

Exhibit 14 presents a 5 × 3 contingency table that summarizes the number of stocks (i.e., frequency) in a particular portfolio of 1,000 stocks by two variables, sector and company market capitalization. Sector has five levels, with each one being a GICS-defined sector. Market capitalization (commonly referred to as "market cap") is defined for a company as the number of shares outstanding times the price per share. The stocks in this portfolio are categorized by three levels of market capitalization: large cap, more than $10 billion; mid cap, $10 billion to $2 billion; and small cap, less than $2 billion.

Exhibit 14 Portfolio Frequencies by Sector and Market Capitalization

	Market Capitalization Variable (3 Levels)			
Sector Variable (5 Levels)	**Small**	**Mid**	**Large**	**Total**
Communication Services	55	35	20	110
Consumer Staples	50	30	30	110
Energy	175	95	20	290
Health Care	275	105	55	435
Utilities	20	25	10	55
Total	**575**	**290**	**135**	**1,000**

The entries in the cells of the contingency table show the number of stocks of each sector with a given level of market cap. For example, there are 275 small-cap health care stocks, making it the portfolio's largest subgroup in terms of frequency. These data are also called **joint frequencies** because you are joining one variable from the row (i.e., sector) and the other variable from the column (i.e., market cap) to count observations. The joint frequencies are then added across rows and across columns, and these corresponding sums are called **marginal frequencies**. For example, the marginal frequency of health care stocks in the portfolio is the sum of the joint frequencies across all three levels of market cap, so 435 (= 275 + 105 + 55). Similarly, adding the joint frequencies of small-cap stocks across all five sectors gives the marginal frequency of small-cap stocks of 575 (= 55 + 50 + 175 + 275 + 20).

Clearly, health care stocks and small-cap stocks have the largest marginal frequencies among sector and market cap, respectively, in this portfolio. Note the marginal frequencies represent the frequency distribution for each variable. Finally, the marginal frequencies for each variable must sum to the total number of stocks (overall total) in the portfolio—here, 1,000 (shown in the lower right cell).

Similar to the one-way frequency distribution table, we can express frequency in percentage terms as relative frequency by using one of three options. We can divide the joint frequencies by: a) the total count; b) the marginal frequency on a row; or c) the marginal frequency on a column.

Exhibit 15 shows the contingency table using relative frequencies based on total count. It is readily apparent that small-cap health care and energy stocks comprise the largest portions of the total portfolio, at 27.5% (= 275/1,000) and 17.5% (= 175/1,000), respectively, followed by mid-cap health care and energy stocks, at 10.5% and 9.5%, respectively. Together, these two sectors make up nearly three-quarters of the portfolio (43.5% + 29.0% = 72.5%).

Exhibit 15 Relative Frequencies as Percentage of Total

	Market Capitalization Variable (3 Levels)			
Sector Variable (5 Levels)	**Small**	**Mid**	**Large**	**Total**
Communication Services	5.5%	3.5%	2.0%	11.0%
Consumer Staples	5.0%	3.0%	3.0%	11.0%
Energy	17.5%	9.5%	2.0%	29.0%
Health Care	27.5%	10.5%	5.5%	43.5%
Utilities	2.0%	2.5%	1.0%	5.5%
Total	**57.5%**	**29.0%**	**13.5%**	**100%**

Exhibit 16 shows relative frequencies based on marginal frequencies of market cap (i.e., columns). From this perspective, it is clear that the health care and energy sectors dominate the other sectors at each level of market capitalization: 78.3% (= 275/575 + 175/575), 69.0% (= 105/290 + 95/290), and 55.6% (= 55/135 + 20/135), for small, mid, and large caps, respectively. Note that there may be a small rounding error difference between these results and the numbers shown in Exhibit 15.

Exhibit 16 Relative Frequencies: Sector as Percentage of Market Cap

	Market Capitalization Variable (3 Levels)			
Sector Variable (5 Levels)	**Small**	**Mid**	**Large**	**Total**
Communication Services	9.6%	12.1%	14.8%	11.0%
Consumer Staples	8.7%	10.3%	22.2%	11.0%
Energy	30.4%	32.8%	14.8%	29.0%
Health Care	47.8%	36.2%	40.7%	43.5%
Utilities	3.5%	8.6%	7.4%	5.5%
Total	100.0%	100.0%	100.0%	**100.0%**

In conclusion, the findings from these contingency tables using frequencies and relative frequencies indicate that in terms of the number of stocks, the portfolio can be generally described as a small- to mid-cap-oriented health care and energy sector portfolio that also includes stocks of several other defensive sectors.

As an analytical tool, contingency tables can be used in different applications. One application is for evaluating the performance of a classification model (in this case, the contingency table is called a **confusion matrix**). Suppose we have a model for classifying companies into two groups: those that default on their bond payments and those that do not default. The confusion matrix for displaying the model's results will be a 2 × 2 table showing the frequency of actual defaults versus the model's predicted frequency of defaults. Exhibit 17 shows such a confusion matrix for a sample of 2,000 non-investment-grade bonds. Using company characteristics and other inputs, the model correctly predicts 300 cases of bond defaults and 1,650 cases of no defaults.

Exhibit 17 Confusion Matrix for Bond Default Prediction Model

Predicted Default	Actual Default Yes	Actual Default No	Total
Yes	300	40	**340**
No	10	1,650	**1,660**
Total	**310**	**1,690**	**2,000**

We can also observe that this classification model incorrectly predicts default in 40 cases where no default actually occurred and also incorrectly predicts no default in 10 cases where default actually did occur. Later in the CFA Program curriculum you will learn how to construct a confusion matrix, how to calculate related model performance metrics, and how to use them to evaluate and tune a classification model.

Another application of contingency tables is to investigate potential association between two categorical variables. For example, revisiting Exhibit 14, one may ask whether the distribution of stocks by sectors is independent of the levels of market capitalization? Given the dominance of small-cap and mid-cap health care and energy stocks, the answer is likely, no.

One way to test for a potential association between categorical variables is to perform a **chi-square test of independence**. Essentially, the procedure involves using the marginal frequencies in the contingency table to construct a table with expected values of the observations. The actual values and expected values are used to derive the chi-square test statistic. This test statistic is then compared to a value from the chi-square distribution for a given level of significance. If the test statistic is greater than the chi-square distribution value, then there is evidence to reject the claim of independence, implying a significant association exists between the categorical variables. The following example describes how a contingency table is used to set up this test of independence.

EXAMPLE 5

Contingency Tables and Association between Two Categorical Variables

Suppose we randomly pick 315 investment funds and classify them two ways: by fund style, either a growth fund or a value fund; and by risk level, either low risk or high risk. Growth funds primarily invest in stocks whose earnings are expected to grow at a faster rate than earnings for the broad stock market. Value funds primarily invest in stocks that appear to be undervalued relative to their fundamental values. Risk here refers to volatility in the return of a given investment fund, so low (high) volatility implies low (high) risk. The data are summarized in a 2 × 2 contingency table shown in Exhibit 18.

Exhibit 18 Contingency Table by Investment Fund Style and Risk Level

	Low Risk	High Risk
Growth	73	26
Value	183	33

1 Calculate the number of growth funds and number of value funds out of the total funds.

2 Calculate the number of low-risk and high-risk funds out of the total funds.

3 Describe how the contingency table is used to set up a test for independence between fund style and risk level.

Solution to 1

The task is to calculate the marginal frequencies by fund style, which is done by adding joint frequencies across the rows. Therefore, the marginal frequency for growth is 73 + 26 = 99, and the marginal frequency for value is 183 + 33 = 216.

Solution to 2

The task is to calculate the marginal frequencies by fund risk, which is done by adding joint frequencies down the columns. Therefore, the marginal frequency for low risk is 73 + 183 = 256, and the marginal frequency for high risk is 26 + 33 = 59.

Solution to 3

Based on the procedure mentioned for conducting a chi-square test of independence, we would perform the following three steps.

Step 1: Add the marginal frequencies and overall total to the contingency table. We have also included the relative frequency table for observed values.

Exhibit 19a Observed Marginal Frequencies and Relative Frequencies

	Observed Values		
	Low Risk	**High Risk**	
Growth	73	26	**99**
Value	183	33	**216**
	256	**59**	**315**

	Observed Values		
	Low Risk	**High Risk**	
Growth	74%	26%	**100%**
Value	85%	15%	**100%**

Step 2: Use the marginal frequencies in the contingency table to construct a table with expected values of the observations. To determine expected values for each cell, multiply the respective row total by the respective column total, then divide by the overall total. So, for $cell_{i,j}$ (in *ith* row and *jth* column):

$$\text{Expected Value}_{i,j} = (\text{Total Row } i \times \text{Total Column } j)/\text{Overall Total} \tag{1}$$

For example,

Expected value for Growth/Low Risk is: (99 × 256)/ 315 = 80.46; and

Expected value for Value/High Risk is: (216 × 59) / 315 = 40.46.

The table of expected values (and accompanying relative frequency table) are:

Exhibit 19b Expected Marginal Frequencies and Relative Frequencies

Observed Values	Low Risk	High Risk		Observed Values	Low Risk	High Risk	
Growth	80.457	18.543	**99**	Growth	81%	19%	**100%**
Value	175.543	40.457	**216**	Value	81%	19%	**100%**
	256	**59**	**315**				

Step 3: Use the actual values and the expected values of observation counts to derive the chi-square test statistic, which is then compared to a value from the chi-square distribution for a given level of significance. If the test statistic is greater than the chi-square distribution value, then there is evidence of a significant association between the categorical variables.

6 DATA VISUALIZATION

e Describe ways that data may be visualized and evaluate uses of specific visualizations

Visualization is the presentation of data in a pictorial or graphical format for the purpose of increasing understanding and for gaining insights into the data. As has been said, "a picture is worth a thousand words." In this section, we discuss a variety of charts that are useful for understanding distributions, making comparisons, and exploring potential relationships among data. Specifically, we will cover visualizing frequency distributions of numerical and categorical data by using plots that represent multi-dimensional data for discovering relationships and by interpreting visuals that display unstructured data.

6.1 Histogram and Frequency Polygon

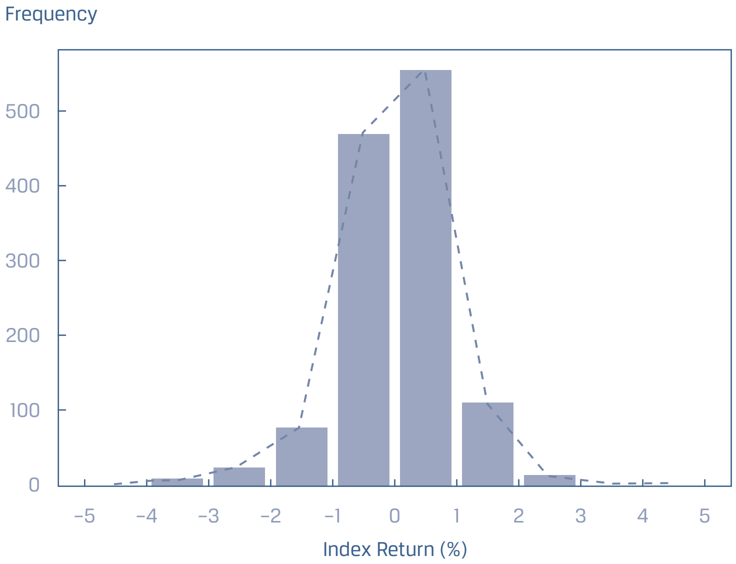
A **histogram** is a chart that presents the distribution of numerical data by using the height of a bar or column to represent the absolute frequency of each bin or interval in the distribution.

To construct a histogram from a continuous variable, we first need to split the data into bins and summarize the data into a frequency distribution table, such as the one we constructed in Exhibit 11. In a histogram, the y-axis generally represents the absolute frequency or the relative frequency in percentage terms, while the x-axis usually represents the bins of the variable. Using the frequency distribution table in Exhibit 11, we plot the histogram of daily returns of the EAA Equity Index, as shown in Exhibit 20. The bars are of equal width, representing the bin width of 1% for each return interval. The bars are usually drawn with no spaces in between, but small gaps can also be added between adjacent bars to increase readability, as in this exhibit. In this case, the height of each bar represents the absolute frequency for each return bin. A quick glance can tell us that the return bin 0% to 1% (exclusive) has the highest frequency, with more than 500 observations (555, to be exact), and it is represented by the tallest bar in the histogram.

An advantage of the histogram is that it can effectively present a large amount of numerical data that has been grouped into a frequency distribution and can allow a quick inspection of the shape, center, and spread of the distribution to better understand it. For example, in Exhibit 20, despite the histogram of daily EAA Equity Index returns appearing bell-shaped and roughly symmetrical, most bars to the right side of the origin (i.e., zero) are taller than those on the left side, indicating that more observations lie in the bins in positive territory. Remember that in the earlier discussion of this return distribution, it was noted that 54.1% of the observations are positive daily returns.

As mentioned, histograms can also be created with relative frequencies—the choice of using absolute versus relative frequency depends on the question being answered. An absolute frequency histogram best answers the question of how many items are in each bin, while a relative frequency histogram gives the proportion or percentage of the total observations in each bin.

Exhibit 20 Histogram Overlaid with Frequency Polygon for Daily Returns of EAA Equity Index

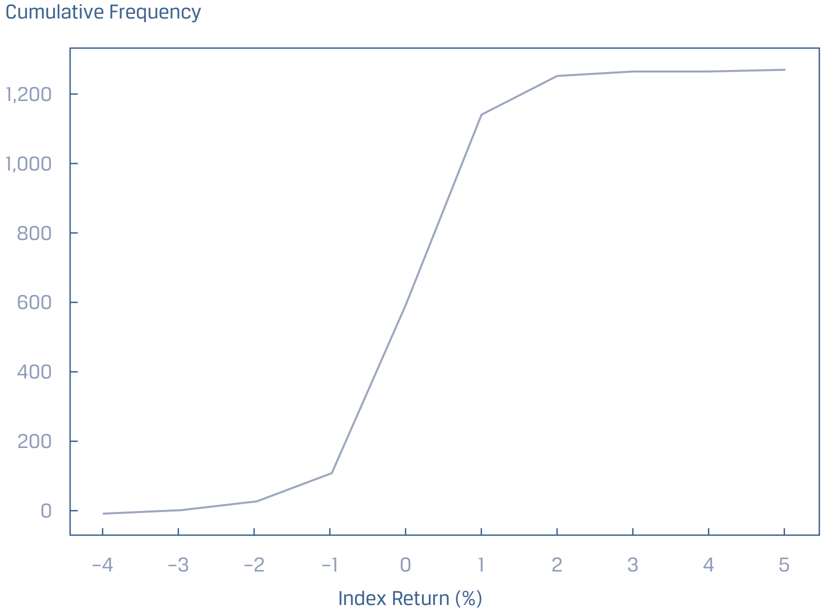


Another graphical tool for displaying frequency distributions is the frequency polygon. To construct a **frequency polygon**, we plot the midpoint of each return bin on the x-axis and the absolute frequency for that bin on the y-axis. We then connect neighboring points with a straight line. Exhibit 20 shows the frequency polygon that overlays the histogram. In the graph, for example, the return interval 1% to 2% (exclusive) has a frequency of 110, so we plot the return-interval midpoint of 0.5% (which is 1.50% on the x-axis) and a frequency of 110 (on the y-axis). Importantly, the frequency polygon can quickly convey a visual understanding of the distribution since it displays frequency as an area under the curve.

Another form for visualizing frequency distributions is the **cumulative frequency distribution chart**. Such a chart can plot either the cumulative absolute frequency or the cumulative relative frequency on the y-axis against the upper limit of the interval. The cumulative frequency distribution chart allows us to see the number or the percentage of the observations that lie below a certain value. To construct the cumulative frequency distribution, we graph the returns in the fourth (i.e., Cumulative Absolute Frequency) or fifth (i.e., Cumulative Relative Frequency) column of Exhibit 11 against the upper limit of each return interval.

Exhibit 21 presents the graph of the cumulative absolute frequency distribution for the daily returns on the EAA Equity Index. Notice that the cumulative distribution tends to flatten out when returns are extremely negative or extremely positive because the frequencies in these bins are quite small. The steep slope in the middle of Exhibit 21 reflects the fact that most of the observations—[(470 + 555)/1,258], or 81.5%—lie in the neighborhood of −1.0% to 1.0%.

Exhibit 21 Cumulative Absolute Frequency Distribution of Daily Returns of EAA Equity Index



6.2 Bar Chart

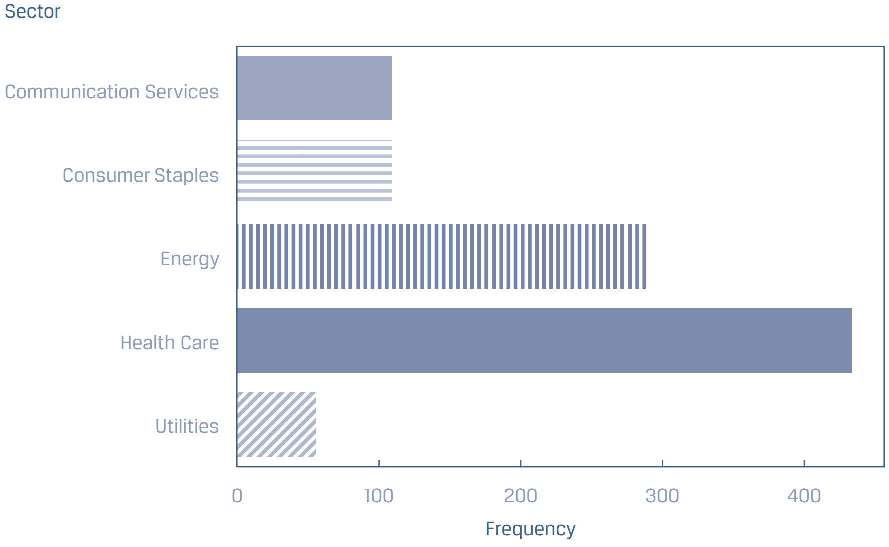
As we have demonstrated, the histogram is an efficient graphical tool to present the frequency distribution of numerical data. The frequency distribution of categorical data can be plotted in a similar type of graph called a **bar chart**. In a bar chart, each bar represents a distinct category, with the bar's height proportional to the frequency of the corresponding category.

Similar to plotting a histogram, the construction of a bar chart with one categorical variable first requires a frequency distribution table summarized from the variable. Note that the bars can be plotted vertically or horizontally. In a vertical bar chart, the y-axis still represents the absolute frequency or the relative frequency. Different from the histogram, however, is that the x-axis in a bar chart represents the mutually exclusive categories to be *compared* rather than bins that group numerical data.

For example, using the marginal frequencies for the five GICS sectors shown in the last column in Exhibit 14, we plot a horizontal bar chart in Exhibit 22 to show the frequency of stocks by sector in the portfolio. The bars are of equal width to represent each sector, and sufficient space should be between adjacent bars to separate them from each other. Because this is a horizontal bar chart—in this case, the x-axis shows the absolute frequency and the y-axis represents the sectors—the length of each bar represents the absolute frequency of each sector. Since sectors are nominal data with no logical ordering, the bars representing sectors may be arranged in any order. However, in the particular case where the categories in a bar chart are ordered by frequency in descending order and the chart includes a line displaying cumulative relative frequency, then it is called a Pareto Chart. The chart is often used to highlight dominant categories or the most important groups.

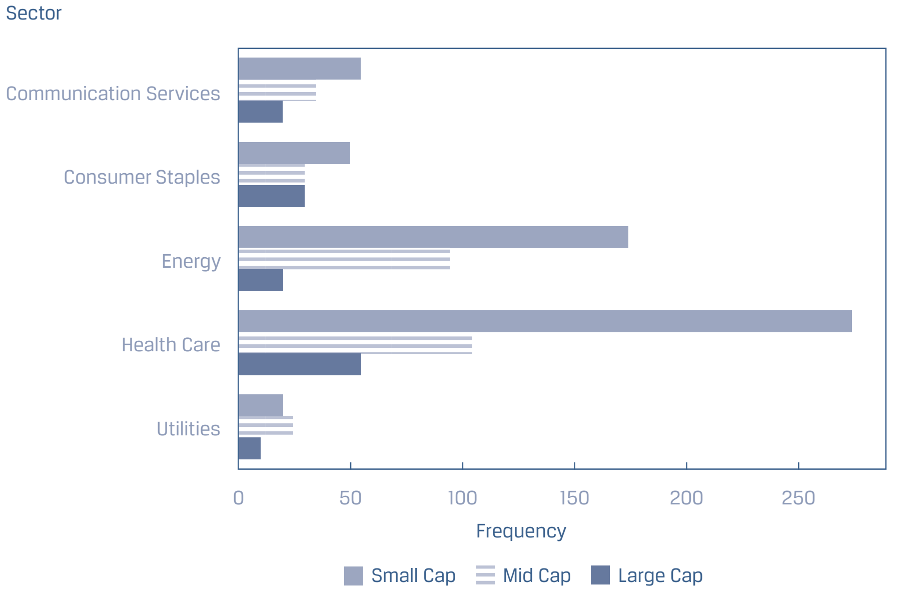
Bar charts provide a snapshot to show the comparison between categories of data. As shown in Exhibit 22, the sector in which the portfolio holds most stocks is the health care sector, with 435 stocks, followed by the energy sector, with 290 stocks. The sector in which the portfolio has the least number of stocks is utilities, with 55 stocks. To compare categories more accurately, in some cases we may add the frequency count to the right end of each bar (or the top end of each bar in the case of a vertical bar chart).

Exhibit 22 Frequency by Sector for Stocks in a Portfolio



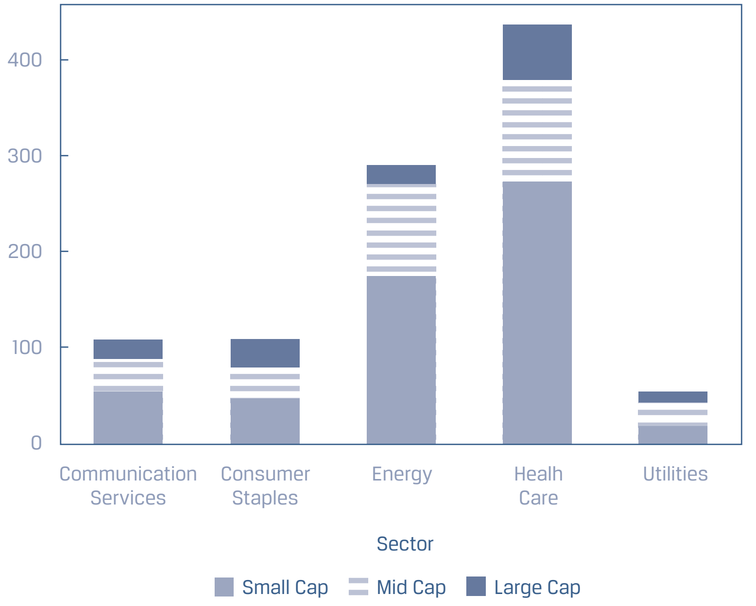
The bar chart shown in Exhibit 22 can present the frequency distribution of only one categorical variable. In the case of two categorical variables, we need an enhanced version of the bar chart, called a **grouped bar chart** (also known as a **clustered bar chart**), to show joint frequencies. Using the joint frequencies by sector and by level of market capitalization given in Exhibit 14, for example, we show how a grouped bar chart is constructed in Exhibit 23. While the *y*-axis still represents the same categorical variable (the distinct GICS sectors as in Exhibit 22), in Exhibit 23 three bars are clustered side-by-side within the same sector to represent the three respective levels of market capitalization. The bars within each cluster should be colored differently to distinguish between them, but the color schemes for the sub-groups must be identical across the sector clusters, as shown by the legend at the upper right of Exhibit 23. Additionally, the bars in each sector cluster must always be placed in the same order throughout the chart. It is easy to see that the small-cap heath care stocks are the sub-group with the highest frequency (275), and we can also see that small-cap stocks are the largest sub-group within each sector—except for utilities, where mid cap is the largest.

Exhibit 23 **Frequency by Sector and Level of Market Capitalization for Stocks in a Portfolio**



An alternative form for presenting the joint frequency distribution of two categorical variables is a **stacked bar chart**. In the vertical version of a stacked bar chart, the bars representing the sub-groups are placed on top of each other to form a single bar. Each subsection of the bar is shown in a different color to represent the contribution of each sub-group, and the overall height of the stacked bar represents the marginal frequency for the category. Exhibit 23 can be replotted in a stacked bar chart, as shown in Exhibit 24.

Exhibit 24 **Frequency by Sector and Level of Market Capitalization in a Stacked Bar Chart**

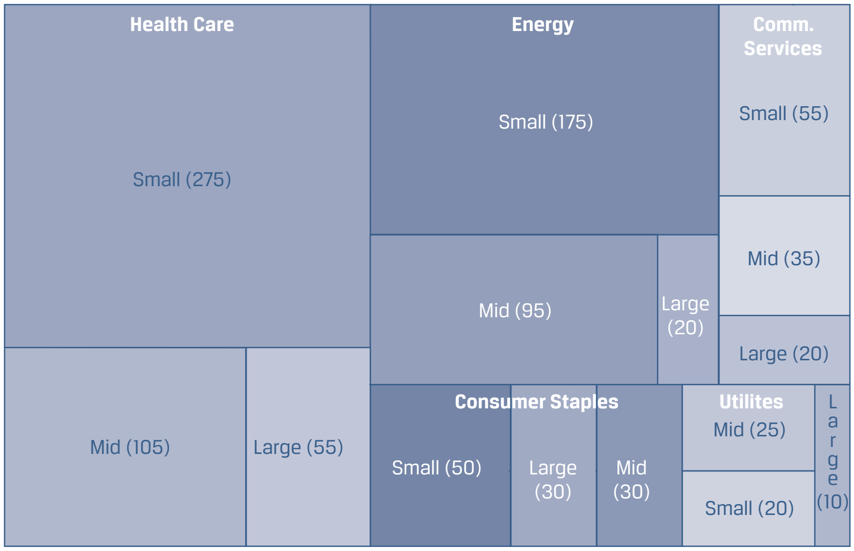


We have shown that the frequency distribution of categorical data can be clearly and efficiently presented by using a bar chart. However, it is worth noting that applications of bar charts may be extended to more general cases when categorical data are associated with numerical data. For example, suppose we want to show a company's quarterly profits over the past one year. In this case, we can plot a vertical bar chart where each bar represents one of the four quarters in a time order and its height indicates the value of profits for that quarter.

6.3 Tree-Map

In addition to bar charts and grouped bar charts, another graphical tool for displaying categorical data is a **tree-map**. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group. For example, referring back to the marginal frequencies by GICS sector in Exhibit 14, we plot a tree-map in Exhibit 25 to represent the frequency distribution by sector for stocks in the portfolio. The tree-map clearly shows that health care is the sector with the largest number of stocks in the portfolio, which is represented by the rectangle with the largest area.

Exhibit 25 Tree-Map for Frequency Distribution by Sector in a Portfolio



Note that this example also depicts one more categorical variable (i.e., level of market capitalization). The tree-map can represent data with additional dimensions by displaying a set of nested rectangles. To show the joint frequencies of sub-groups by sector and level of market capitalization, as given in Exhibit 14, we can split each existing rectangle for sector into three sub-rectangles to represent small-cap, mid-cap, and large-cap stocks, respectively. In this case, the area of each nested rectangle would be proportional to the number of stocks in each market capitalization sub-group. The exhibit clearly shows that small-cap health care is the sub-group with the largest number of stocks. It is worth noting a caveat for using tree-maps: Tree-maps become difficult to read if the hierarchy involves more than three levels.

6.4 Word Cloud

So far, we have shown how to visualize the frequency distribution of numerical data or categorical data. However, can we find a chart to depict the frequency of unstructured data—particularly, textual data? A **word cloud** (also known as **tag cloud**) is a visual device for representing textual data. A word cloud consists of words extracted from a source of textual data, with the size of each distinct word being proportional to the frequency with which it appears in the given text. Note that common words (e.g., "a," "it," "the") are generally stripped out to focus on key words that convey the most meaningful information. This format allows us to quickly perceive the most frequent terms among the given text to provide information about the nature of the text, including topic and whether or not the text conveys positive or negative news. Moreover, words conveying different sentiment may be displayed in different colors. For example, "profit" typically indicates positive sentiment so might be displayed in green, while "loss" typically indicates negative sentiment and may be shown in red.

Exhibit 26 is an excerpt from the Management's Discussion and Analysis (MDA) section of the 10-Q filing for QXR Inc. for the quarter ended 31 March 20XX. Taking this text, we can create a word cloud, as shown in Exhibit 27. A quick glance at the word cloud tells us that the following words stand out (i.e., they were used most frequently in the MDA text): "billion," "revenue," "year," "income," "growth," and "financial." Note that specific words, such as "income" and "growth," typically convey positive sentiment, as contrasted with such words as "loss" and "decline," which typically convey negative sentiment. In conclusion, word clouds are a useful tool for visualizing textual data that can facilitate understanding the topic of the text as well as the sentiment it may convey.

Exhibit 26 Excerpt of MDA Section in Form 10-Q of QXR Inc. for Quarter Ended 31 March 20XX

MANAGEMENT'S DISCUSSION AND ANALYSIS OF FINANCIAL CONDITION AND RESULTS OF OPERATIONS

Please read the following discussion and analysis of our financial condition and results of operations together with our consolidated financial statements and related notes included under Part I, Item 1 of this Quarterly Report on Form 10-Q

Executive Overview of Results

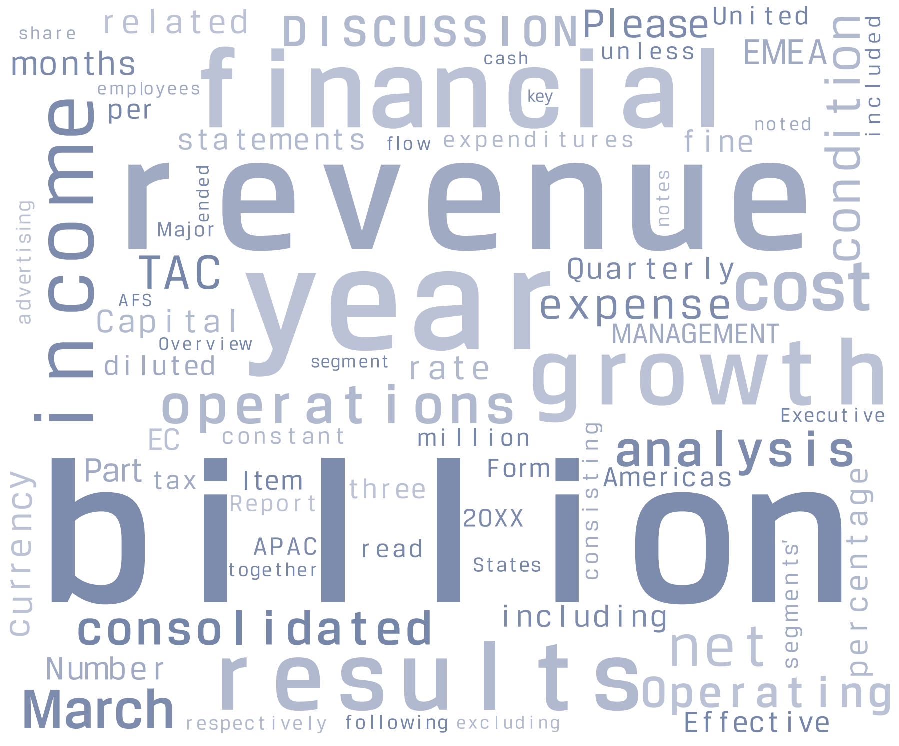
Below are our key financial results for the three months ended March 31, 20XX (consolidated unless otherwise noted):

- Revenues of $36.3 billion and revenue growth of 17% year over year, constant currency revenue growth of 19% year over year.
- Major segment revenues of $36.2 billion with revenue growth of 17% year over year and other segments' revenues of $170 million with revenue growth of 13% year over year.
- Revenues from the United States, EMEA, APAC, and Other Americas were $16.5 billion, $11.8 billion, $6.1 billion, and $1.9 billion, respectively.
- Cost of revenues was $16.0 billion, consisting of TAC of $6.9 billion and other cost of revenues of $9.2 billion. Our TAC as a percentage of advertising revenues were 22%.
- Operating expenses (excluding cost of revenues) were $13.7 billion, including the EC AFS fine of $1.7 billion.
- Income from operations was $6.6 billion
- Other income (expense), net, was $1.5 billion.

Exhibit 26 (Continued)

- Effective tax rate was 18%
- Net income was $6.7 billion with diluted net income per share of $9.50.
- Operating cash flow was $12.0 billion.
- Capital expenditures were $4.6 billion.

Exhibit 27 Word Cloud Visualizing Excerpted Text in MDA Section in Form 10-Q of QXR Inc.



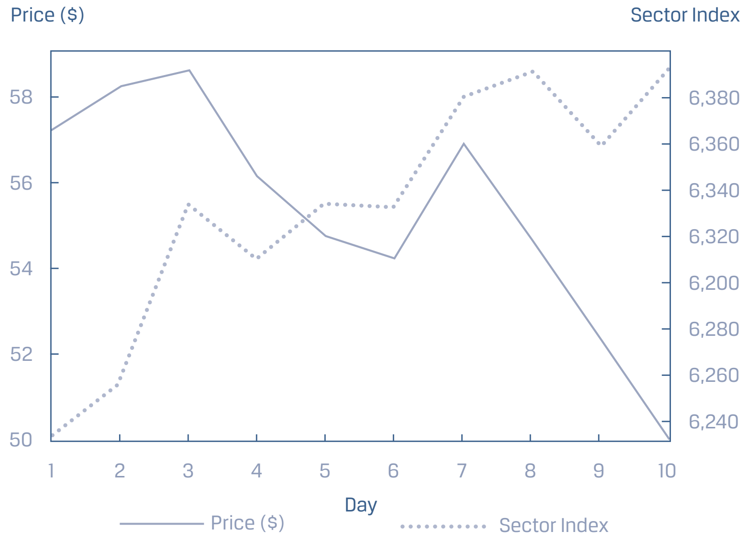
6.5 Line Chart

A **line chart** is a type of graph used to visualize ordered observations. Often a line chart is used to display the change of data series over time. Note that the frequency polygon in Exhibit 20 and the cumulative frequency distribution chart in Exhibit 21 are also line charts but used particularly in those instances for representing data frequency distributions.

Constructing a line chart is relatively straightforward: We first plot all the data points against horizontal and vertical axes and then connect the points by straight line segments. For example, to show the 10-day daily closing prices of ABC Inc. stock presented in Exhibit 5, we first construct a chart with the *x*-axis representing time (in days) and the *y*-axis representing stock price (in dollars). Next, plot each closing price as points against both axes, and then use straight line segments to join the points together, as shown in Exhibit 28.

An important benefit of a line chart is that it facilitates showing changes in the data and underlying trends in a clear and concise way. This helps to understand the current data and also helps with forecasting the data series. In Exhibit 28, for example, it is easy to spot the price changes over the first 10 trading days since ABC's initial public offering (IPO). We see that the stock price peaked on Day 3 and then traded lower. Following a partial recovery on Day 7, it declined steeply to around $50 on Day 10. In contrast, although the one-dimensional data array table in Exhibit 5 displays the same values as the line chart, the data table by itself does not provide a quick snapshot of changes in the data or facilitate understanding underlying trends. This is why line charts are helpful for visualization, particularly in cases of large amounts of data (i.e., hundreds, or even thousands, of data points).

Exhibit 28 Daily Closing Prices of ABC Inc.'s Stock and Its Sector Index



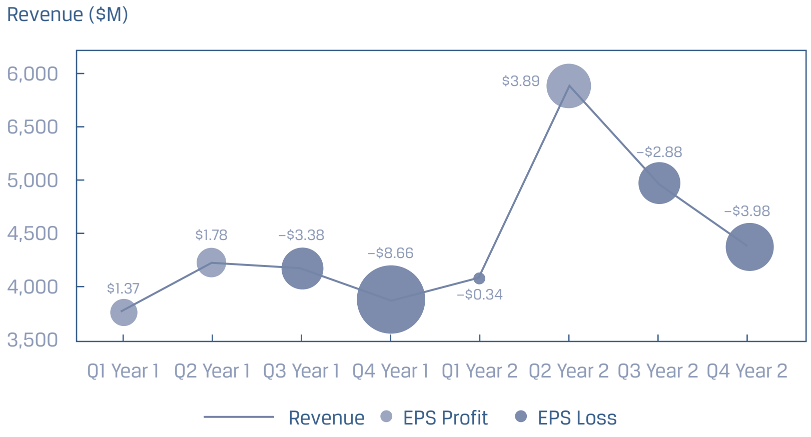
A line chart is also capable of accommodating more than one set of data points, which is especially helpful for making comparisons. We can add a line to represent each group of data (e.g., a competitor's stock price or a sector index), and each line would have a distinct color or line pattern identified in a legend. For example, Exhibit 28 also includes a plot of ABC's sector index (i.e., the sector index for which ABC stock is a member, like health care or energy) over the same period. The sector index is displayed with its own distinct color to facilitate comparison. Note also that because the sector index has a different range (approximately 6,230 to 6,390) than ABCs' stock ($50 to $59 per share), we need a secondary y-axis to correctly display the sector index, which is on the right-hand side of the exhibit.

This comparison can help us understand whether ABC's stock price movement over the period is due to potential mispricing of its share issuance or instead due to industry-specific factors that also affect its competitors' stock prices. The comparison shows that over the period, the sector index moved in a nearly opposite trend versus ABC's stock price movement. This indicates that the steep decline in ABC's stock price is less likely attributable to sector-specific factors and more likely due to potential over-pricing of its IPO or to other company-specific factors.

When an observational unit (here, ABC Inc.) has more than two features (or variables) of interest, it would be useful to show the multi-dimensional data all in one chart to gain insights from a more holistic view. How can we add an additional dimension to a two-dimensional line chart? We can replace the data points with varying-sized bubbles to represent a third dimension of the data. Moreover, these bubbles may even be color-coded to present additional information. This version of a line chart is called a **bubble line chart**.

Exhibit 7, for example, presented three types of quarterly data for ABC Inc. for use in a valuation analysis. We would like to plot two of them, revenue and earnings per share (EPS), over the two-year period. As shown in Exhibit 29, with the x-axis representing time (i.e., quarters) and the y-axis representing revenue in millions of dollars, we can plot the revenue data points against both axes to form a typical line chart. Next, each marker representing a revenue data point is replaced by a circular bubble with its size proportional to the magnitude of the EPS in the corresponding quarter. Moreover, the bubbles are colored in a binary scheme with green representing profits and red representing losses. In this way, the bubble line chart reflects the changes for both revenue and EPS simultaneously, and it also shows whether the EPS represents a profit or a loss.

Exhibit 29 Quarterly Revenue and EPS of ABC Incorporated



As depicted, ABC's earning were quite volatile during its initial two years as a public company. Earnings started off as a profit of $1.37/share but finished the first year with a big loss of –$8.66/share, during which time revenue experienced only small fluctuations. Furthermore, while revenues and earnings both subsequently recovered sharply—peaking in Q2 of Year 2—revenues then declined, and the company returned to significant losses (–3.98/share) by the end of Year 2.

6.6 Scatter Plot

A **scatter plot** is a type of graph for visualizing the joint variation in two numerical variables. It is a useful tool for displaying and understanding potential relationships between the variables.

A scatter plot is constructed with the x-axis representing one variable and the y-axis representing the other variable. It uses dots to indicate the values of the two variables for a particular point in time, which are plotted against the corresponding axes. Suppose an analyst is investigating potential relationships between sector index returns and returns for the broad market, such as the S&P 500 Index. Specifically, he or she is interested in the relative performance of two sectors, information technology (IT) and utilities, compared to the market index over a specific five-year period. The analyst has obtained the sector and market index returns for each month over the five years under investigation and plotted the data points in the scatter plots, shown in Exhibit 30 for IT versus the S&P 500 returns and in Exhibit 31 for utilities versus the S&P 500 returns.

Despite their relatively straightforward construction, scatter plots convey lots of valuable information. First, it is important to inspect for any potential association between the two variables. The pattern of the scatter plot may indicate no apparent relationship, a linear association, or a non-linear relationship. A scatter plot with randomly distributed data points would indicate no clear association between the two variables. However, if the data points seem to align along a straight line, then there may exist a significant relationship among the variables. A positive (negative) slope for the line of data points indicates a positive (negative) association, *meaning the variables move in the same (opposite) direction*. Furthermore, the strength of the association can be determined by how closely the data points are clustered around the line. Tight (loose) clustering signals a potentially stronger (weaker) relationship.

Exhibit 30 Scatter Plot of Information Technology Sector Index Return vs. S&P 500 Index Return

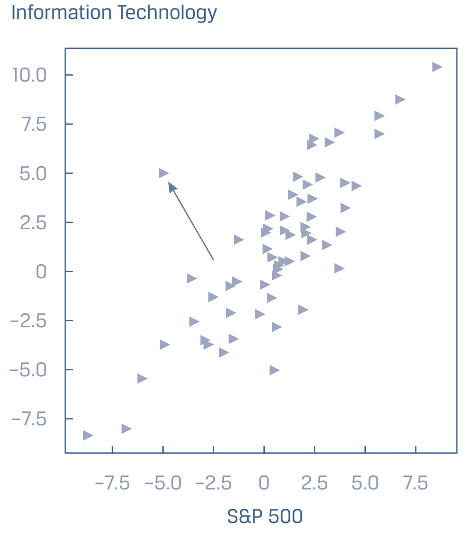
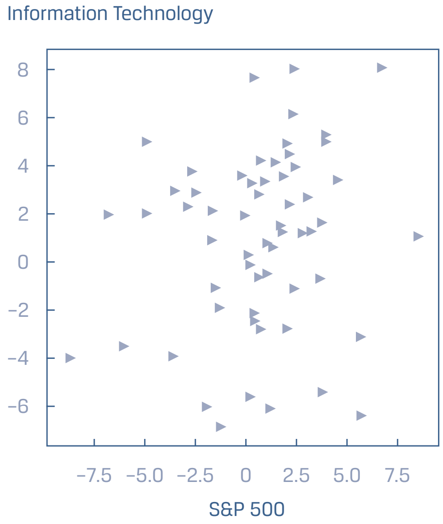


Exhibit 31 Scatter Plot of Utilities Sector Index Return vs. S&P 500 Index Return

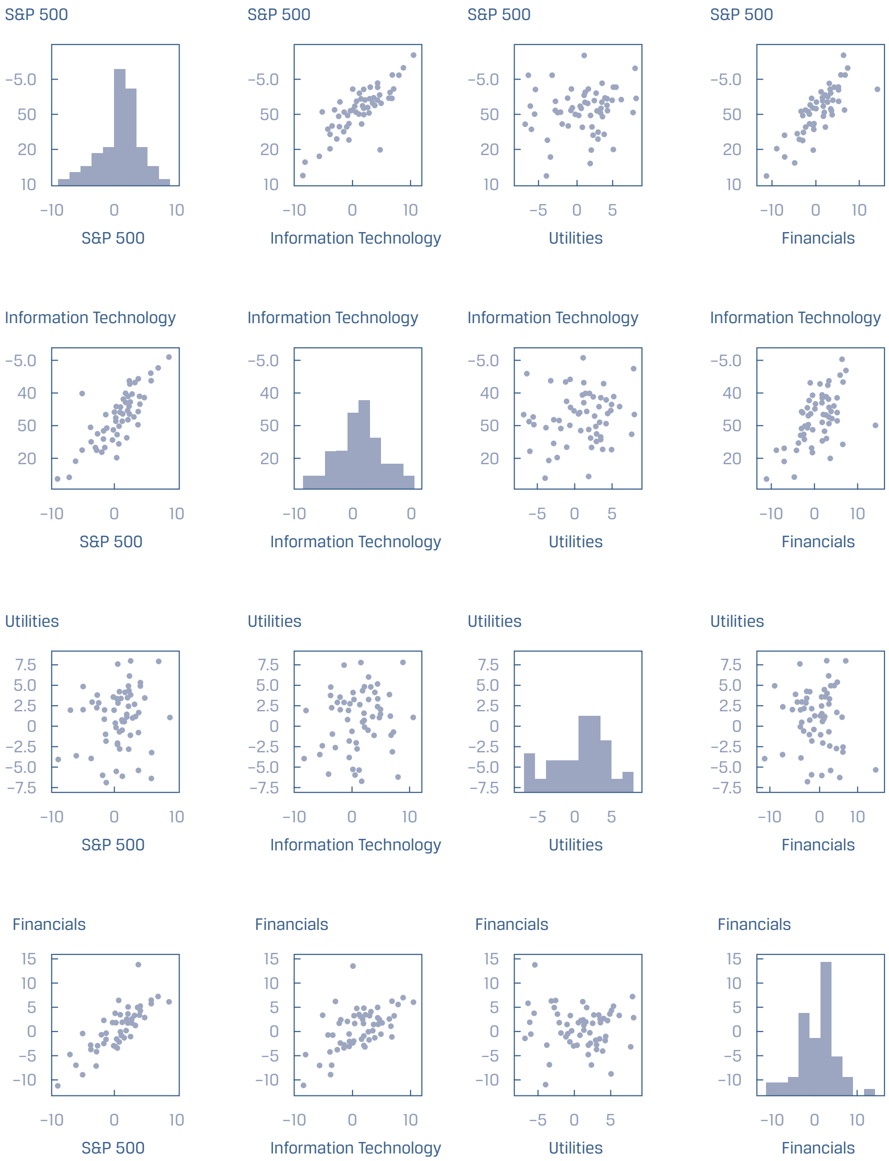


Examining Exhibit 30, we can see the returns of the IT sector are highly positively associated with S&P 500 Index returns because the data points are tightly clustered along a positively sloped line. Exhibit 31 tells a different story for relative performance of the utilities sector and S&P 500 index returns: The data points appear to be distributed in no discernable pattern, indicating no clear relationship among these variables. Second, observing the data points located toward the ends of each axis, which represent the maximum or minimum values, provides a quick sense of the data range. Third, assuming that a relationship among the variables is apparent, inspecting the scatter plot can help to spot extreme values (i.e., outliers). For example, an outlier data point is readily detected in Exhibit 30, as indicated by the arrow. As you will learn later in the CFA Program curriculum, finding these extreme values and handling them with appropriate measures is an important part of the financial modeling process.

Scatter plots are a powerful tool for finding patterns between two variables, for assessing data range, and for spotting extreme values. In practice, however, there are situations where we need to inspect for pairwise associations among many variables—for example, when conducting feature selection from dozens of variables to build a predictive model.

A **scatter plot matrix** is a useful tool for organizing scatter plots between pairs of variables, making it easy to inspect all pairwise relationships in one combined visual. For example, suppose the analyst would like to extend his or her investigation by adding another sector index. He or she can use a scatter plot matrix, as shown in Exhibit 32, which now incorporates four variables, including index returns for the S&P 500 and for three sectors: IT, utilities, and financials.

Exhibit 32 Pairwise Scatter Plot Matrix



The scatter plot matrix contains each combination of bivariate scatter plot (i.e., S&P 500 vs. each sector, IT vs. utilities, IT vs. financials, and financials vs. utilities) as well as univariate frequency distribution histograms for each variable plotted along the diagonal. In this way, the scatter plot matrix provides a concise visual summary of each variable and of potential relationships among them. Importantly, the construction of the scatter plot matrix is typically a built-in function in most major statistical software packages, so it is relatively easy to implement. It is worth pointing out that the upper triangle of the matrix is the mirror image of the lower triangle, so the compact form of the scatter plot matrix that uses only the lower triangle is also appropriate.

With the addition of the financial sector, the bottom panel of Exhibit 32 reveals the following additional information, which can support sector allocation in the portfolio construction process:

- Strong positive relationship between returns of financial and S&P 500;

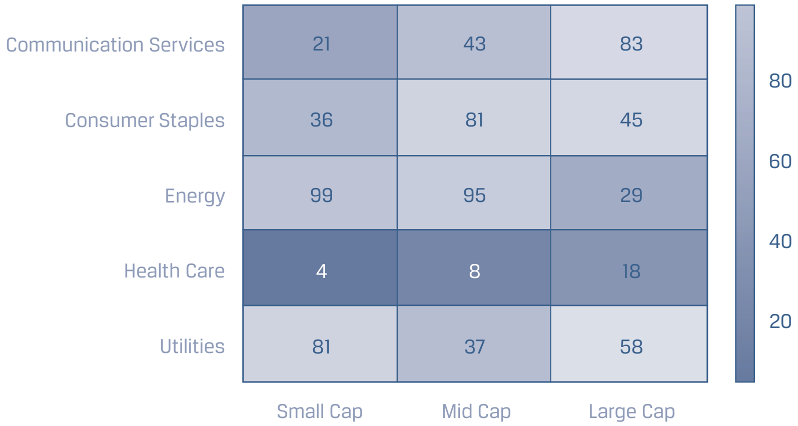
- Positive relationship between returns of financial and IT; and
- No clear relationship between returns of financial and utilities.

It is important to note that despite their usefulness, scatter plots and scatter plot matrixes should not be considered as a substitute for robust statistical tests; rather, they should be used alongside such tests for best results.

6.7 Heat Map

A **heat map** is a type of graphic that organizes and summarizes data in a tabular format and represents them using a color spectrum. For example, given a portfolio, we can create a contingency table that summarizes the joint frequencies of the stock holdings by sector and by level of market capitalization, as in Exhibit 33.

Exhibit 33 Frequencies by Sector and Market Capitalization in Heat Map



	Small Cap	Mid Cap	Large Cap
Communication Services	21	43	83
Consumer Staples	36	81	45
Energy	99	95	29
Health Care	4	8	18
Utilities	81	37	58

Cells in the chart are color-coded to differentiate high values from low values by using the color scheme defined in the color spectrum on the right side of the chart. As shown by the heat map, this portfolio has the largest exposure (in terms of number of stocks) to small- and mid-cap energy stocks. It has substantial exposures to large-cap communications services, mid-cap consumer staples, and small-cap utilities; however, exposure to the health care sector is limited. In sum, the heat map reveals this portfolio to be relatively well-diversified among sectors and market-cap levels. Besides their use in displaying frequency distributions, heat maps are commonly used for visualizing the degree of correlation among different variables.

EXAMPLE 6

Evaluating Data Visuals

1 You have a cumulative absolute frequency distribution graph (similar to the one in Exhibit 21) of daily returns over a five-year period for an index of Asian equity markets.

Interpret the meaning of the slope of such a graph.

2 You are creating a word cloud for a visual representation of text on a company's quarterly earnings announcements over the past three years. The word cloud uses font size to indicate word frequency. This particular company has experienced both quarterly profits and losses during the period under investigation.

Describe how the word cloud might be used to convey information besides word frequency.

3 You are examining a scatter plot of monthly stock returns, similar to the one in Exhibit 30, for two technology companies: one is a hardware manufacturer, and the other is a software developer. The scatter plot shows a strong positive association among their returns.

Describe what other information the scatter plot can provide.

4 You are reading a vertical bar chart displaying the sales of a company over the past five years. The sales of the first four years seem nearly flat as the corresponding bars are nearly the same height, but the bar representing the sales of the most recent year is approximately three times as high as the other bars.

Explain whether we can conclude that the sales of the fifth year tripled compared to sales in the earlier years.

Solution 1

The slope of the graph of a cumulative absolute frequency distribution reflects the change in the number of observations between two adjacent return bins. A steep (flat) slope indicates a large (small) change in the frequency of observations between adjacent return bins.

Solution 2

Color can add an additional dimension to the information conveyed in the word cloud. For example, red can be used for "losses" and other words conveying negative sentiment, and green can be used for "profit" and other words indicative of positive sentiment.

Solution 3

Besides the sign and degree of association of the stocks' returns, the scatter plot can provide a visual representation of whether the association is linear or non-linear, the maximum and minimum values for the return observations, and an indication of which observations may have extreme values (i.e., are potential outliers).

Solution 4

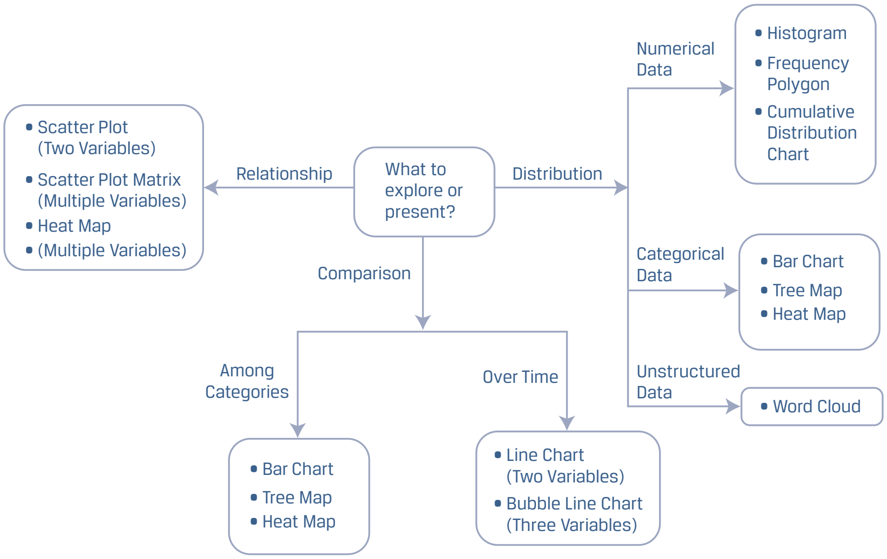
Typically, the heights of bars in a vertical bar chart are proportional to the values that they represent. However, if the graph is using a truncated y-axis (i.e., one that does not start at zero), then values are not accurately represented by the height of bars. Therefore, we need to examine the y-axis of the bar chart before concluding that sales in the fifth year were triple the sales of the prior years.

6.8 Guide to Selecting among Visualization Types

f Describe How to Select among Visualization Types

We have introduced and discussed a variety of different visualization types that are regularly used in investment practice. When it comes to selecting a chart for visualizing data, the intended purpose is the key consideration: Is it for exploring and/or presenting distributions or relationships, or is it for making comparisons? Given your intended purpose, the best selection is typically the simplest visual that conveys the message or achieves the specific goal. Exhibit 34 presents a flow chart for facilitating selection among the visualization types we have discussed. Finally, note that some visualization types, such as bar chart and heat map, may be suitable for several different purposes.

Exhibit 34 Flow Chart of Selecting Visualization Types


What to explore or present?
Relationship
Distribution
Comparison
Scatter Plot (Two Variables)
Scatter Plot Matrix (Multiple Variables)
Heat Map
(Multiple Variables)
Numerical Data
Histogram
Frequency Polygon
Cumulative Distribution Chart
Categorical Data
Bar Chart
Tree Map
Heat Map
Unstructured Data
Word Cloud
Among Categories
Bar Chart
Tree Map
Heat Map
Over Time
Line Chart (Two Variables)
Bubble Line Chart (Three Variables)

Data visualization is a powerful tool to show data and gain insights into data. However, we need to be cautious that a graph could be misleading if data are misrepresented or the graph is poorly constructed. There are numerous different ways that may lead to a misleading graph. We list four typical pitfalls here that analysts should avoid.

First, an improper chart type is selected to present data, which would hinder the accurate interpretation of data. For example, to investigate the correlation between two data series, we can construct a scatter plot to visualize the joint variation between two variables. In contrast, plotting the two data series separately in a line chart would make it rather difficult to examine the relationship.

Second, data are selectively plotted in favor of the conclusion an analyst intends to draw. For example, data presented for an overly short time period may appear to show a trend that is actually noise—that is, variation within the data's normal range if examining the data over a longer time period. So, presenting data for too short a time window may mistakenly point to a non-existing trend.

Third, data are improperly plotted in a truncated graph that has a y-axis that does not start at zero. In some situations, the truncated graph can create the false impression of significant differences when there is actually only a small difference. For example, suppose a vertical bar chart is used to compare annual revenues of two companies, one with \$9 billion and the other with \$10 billion. If the y-axis starts at \$8 billion, then the bar heights would inaccurately imply that the latter company's revenue is twice the former company's revenue.

Last, but not least, is the improper scaling of axes. For example, given a line chart, setting a higher than necessary maximum on the y-axis tends to compress the graph into an area close to the x-axis. This causes the graph to appear to be less steep and less volatile than if it was properly plotted. In sum, analysts need to avoid these misuses of visualization when charting data and must ensure the ethical use of data visuals.

EXAMPLE 7

Selecting Visualization Types

1 A portfolio manager plans to buy several stocks traded on a small emerging market exchange but is concerned whether the market can provide sufficient liquidity to support her purchase order size. As the first step, she wants to analyze the daily trading volumes of one of these stocks over the past five years.

Explain which type of chart can best provide a quick view of trading volume for the given period.

2 An analyst is building a model to predict stock market downturns. According to the academic literature and his practitioner knowledge and expertise, he has selected 10 variables as potential predictors. Before continuing to construct the model, the analyst would like to get a sense of how closely these variables are associated with the broad stock market index and whether any pair of variables are associated with each other.

Describe the most appropriate visual to select for this purpose.

3 Central Bank members meet regularly to assess the economy and decide on any interest rate changes. Minutes of their meetings are published on the Central Bank's website. A quantitative researcher wants to analyze the meeting minutes for use in building a model to predict future economic growth.

Explain which type of chart is most appropriate for creating an overview of the meeting minutes.

4 A private investor wants to add a stock to her portfolio, so she asks her financial adviser to compare the three-year financial performances (by quarter) of two companies. One company experienced consistent revenue and earnings growth, while the other experienced volatile revenue and earnings growth, including quarterly losses.

Describe the chart the adviser should use to best show these performance differences.

Solution to 1

The five-year history of daily trading volumes contains a large amount of numerical data. Therefore, a histogram is the best chart for grouping these data into frequency distribution bins and for showing a quick snapshot of the shape, center, and spread of the data's distribution.

Solution to 2

To inspect for a potential relationship between two variables, a scatter plot is a good choice. But with 10 variables, plotting individual scatter plots is not an efficient approach. Instead, utilizing a scatter plot matrix would give the analyst a good overview in one comprehensive visual of all the pairwise associations between the variables.

Solution to 3

Since the meeting minutes consist of textual data, a word cloud would be the most suitable tool to visualize the textual data and facilitate the researcher's understanding of the topic of the text as well as the sentiment, positive or negative, it may convey.

Solution to 4

The best chart for making this comparison would be a bubble line chart using two different color lines to represent the quarterly revenues for each company. The bubble sizes would then indicate the magnitude of each company's quarterly earnings, with green bubbles signifying profits and red bubbles indicating losses.

MEASURES OF CENTRAL TENDENCY

7

g Calculate and interpret measures of central tendency

So far, we have discussed methods we can use to organize and present data so that they are more understandable. The frequency distribution of an asset return series, for example, reveals much about the nature of the risks that investors may encounter in a particular asset. Although frequency distributions, histograms, and contingency tables provide a convenient way to summarize a series of observations, these methods are just a first step toward describing the data. In this section, we discuss the use of quantitative measures that explain characteristics of data. Our focus is on measures of central tendency and other measures of location. A **measure of central tendency** specifies where the data are centered. Measures of central tendency are probably more widely used than any other statistical measure because they can be computed and applied relatively easily. **Measures of location** include not only measures of central tendency but other measures that illustrate the location or distribution of data.

In the following subsections, we explain the common measures of central tendency—the arithmetic mean, the median, the mode, the weighted mean, the geometric mean, and the harmonic mean. We also explain other useful measures of location, including quartiles, quintiles, deciles, and percentiles.

A **statistic** is a summary measure of a set of observations, and descriptive statistics summarize the central tendency and spread variation in the distribution of data. If the statistic summarizes the set of all possible observations of a **population**, we refer to the statistic as a parameter. If the statistic summarizes a set of observations that is a subset of the population, we refer to the statistic as a **sample statistic**, often leav8ing off the word "sample" and simply referring to it as a statistic. While measures of central tendency and location can be calculated for populations and **samples**, our focus is on sample measures (i.e., sample statistics) as it is rare that an investment manager would be dealing with an entire population of data.

7.1 The Arithmetic Mean

Analysts and portfolio managers often want one number that describes a representative possible outcome of an investment decision. The arithmetic mean is one of the most frequently used measures of the center of data.

> **Definition of Arithmetic Mean.** The **arithmetic mean** is the sum of the values of the observations divided by the number of observations.

7.1.1 *The Sample Mean*

The sample mean is the arithmetic mean or arithmetic average computed for a sample. As you will see, we use the terms "mean" and "average" interchangeably. Often, we cannot observe every member of a population; instead, we observe a subset or sample of the population.

Sample Mean Formula. The **sample mean** or average, $\overline{X}$ (read “X-bar”), is the arithmetic mean value of a sample:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad (2)$$

where n is the number of observations in the sample.

Equation 2 tells us to sum the values of the observations (X_i) and divide the sum by the number of observations. For example, if a sample of market capitalizations for six publicly traded Australian companies contains the values (in AUD billions) 35, 30, 22, 18, 15, and 12, the sample mean market cap is 132/6 = A$22 billion. As previously noted, the sample mean is a statistic (that is, a descriptive measure of a sample).

Means can be computed for individual units or over time. For instance, the sample might be the return on equity (ROE) in a given year for a sample of 25 companies in the FTSE Eurotop 100, an index of Europe’s 100 largest companies. In this case, we calculate the mean ROE in that year as an average across 25 individual units. When we examine the characteristics of some units at a specific point in time (such as ROE for the FTSE Eurotop 100), we are examining cross-sectional data; the mean of these observations is the cross-sectional mean. If the sample consists of the historical monthly returns on the FTSE Eurotop 100 for the past five years, however, then we have time-series data; the mean of these observations is the time-series mean. We will examine specialized statistical methods related to the behavior of time series in the reading on time-series analysis.

Except in cases of large datasets with many observations, we should not expect any of the actual observations to equal the mean; sample means provide only a summary of the data being analyzed. Also, although in some cases the number of values below the mean is quite close to the number of values above the mean, this need not be the case. As an analyst, you will often need to find a few numbers that describe the characteristics of the distribution, and we will consider more later. The mean is generally the statistic that you use as a measure of the typical outcome for a distribution. You can then use the mean to compare the performance of two different markets. For example, you might be interested in comparing the stock market performance of investments in Asia Pacific with investments in Europe. You can use the mean returns in these markets to compare investment results.

EXAMPLE 8

Calculating a Cross-Sectional Mean

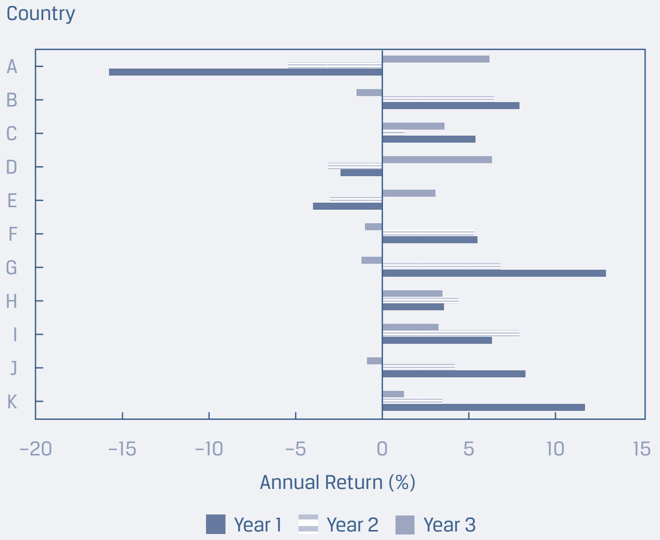
Suppose we want to examine the performance of a sample of selected stock indexes from 11 different countries. The 52-week percentage change is reported in Exhibit 35 for Year 1, Year 2, and Year 3 for the sample of indexes.

Exhibit 35 Annual Returns for Years 1 to 3 for Selected Countries’ Stock Indexes

	52-Week Return (%)		
Index	**Year 1**	**Year 2**	**Year 3**
Country A	−15.6	−5.4	6.1
Country B	7.8	6.3	−1.5

Exhibit 35 (Continued)

	52-Week Return (%)		
Index	**Year 1**	**Year 2**	**Year 3**
Country C	5.3	1.2	3.5
Country D	−2.4	−3.1	6.2
Country E	−4.0	−3.0	3.0
Country F	5.4	5.2	−1.0
Country G	12.7	6.7	−1.2
Country H	3.5	4.3	3.4
Country I	6.2	7.8	3.2
Country J	8.1	4.1	−0.9
Country K	11.5	3.4	1.2



Using the data provided, calculate the sample mean return for the 11 indexes for each year.

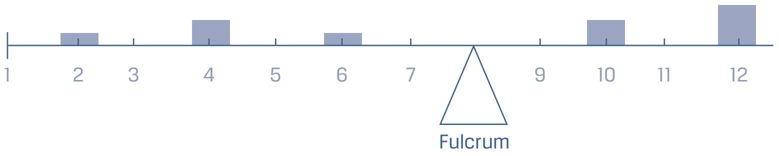
Solution:

For Year 3, the calculation applies Equation 2 to the returns for Year 3: (6.1 – 1.5 + 3.5 + 6.2 + 3.0 – 1.0 – 1.2 + 3.4 + 3.2 – 0.9 + 1.2)/11 = 22.0/11 = 2.0%. Using a similar calculation, the sample mean is 3.5% for Year 1 and 2.5% for Year 2.

7.1.2 *Properties of the Arithmetic Mean*

The arithmetic mean can be likened to the center of gravity of an object. Exhibit 36 expresses this analogy graphically by plotting nine hypothetical observations on a bar. The nine observations are 2, 4, 4, 6, 10, 10, 12, 12, and 12; the arithmetic mean is 72/9 = 8. The observations are plotted on the bar with various heights based on their frequency (that is, 2 is one unit high, 4 is two units high, and so on). When the bar is placed on a fulcrum, it balances only when the fulcrum is located at the point on the scale that corresponds to the arithmetic mean.

Exhibit 36 Center of Gravity Analogy for the Arithmetic Mean



As analysts, we often use the mean return as a measure of the typical outcome for an asset. As in Example 8, however, some outcomes are above the mean and some are below it. We can calculate the distance between the mean and each outcome, which is the deviation. Mathematically, it is always true that the sum of the deviations around the mean equals 0. We can see this by using the definition of the arithmetic mean shown in Equation 2, multiplying both sides of the equation by n: $n\overline{X} = \sum_{i=1}^{n} X_i$. The sum of the deviations from the mean is calculated as follows:

$$\sum_{i=1}^{n}\left(X_i - \overline{X}\right) = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \overline{X} = \sum_{i=1}^{n} X_i - n\overline{X} = 0$$

Deviations from the arithmetic mean are important information because they indicate risk. The concept of deviations around the mean forms the foundation for the more complex concepts of variance, skewness, and kurtosis, which we will discuss later.

A property and potential drawback of the arithmetic mean is its sensitivity to extreme values, or outliers. Because all observations are used to compute the mean and are given equal weight (i.e., importance), the arithmetic mean can be pulled sharply upward or downward by extremely large or small observations, respectively. For example, suppose we compute the arithmetic mean of the following seven numbers: 1, 2, 3, 4, 5, 6, and 1,000. The mean is 1,021/7 = 145.86, or approximately 146. Because the magnitude of the mean, 146, is so much larger than most of the observations (the first six), we might question how well it represents the location of the data. Perhaps the most common approach in such cases is to report the median, or middle value, in place of or in addition to the mean.

7.1.3 *Outliers*

In practice, although an extreme value or outlier in a financial dataset may just represent a rare value in the population, it may also reflect an error in recording the value of an observation or an observation generated from a different population from that producing the other observations in the sample. In the latter two cases, in particular, the arithmetic mean could be misleading. So, what do we do? The first step is to examine the data, either by inspecting the sample observations if the sample is not too large or by using visualization approaches. Once we are comfortable that we have identified and eliminated errors (that is, we have "cleaned" the data), we can then address what to do with extreme values in the sample. When dealing with a sample that has extreme values, there may be a possibility of transforming the variable (e.g., a log transformation) or of selecting another variable that achieves the same purpose. However, if alternative model specifications or variable transformations are not possible, then here are three options for dealing with extreme values:

Option 1	Do nothing; use the data without any adjustment.
Option 2	Delete all the outliers.
Option 3	Replace the outliers with another value.

The first option is appropriate if the values are legitimate, correct observations, and it is important to reflect the whole of the sample distribution. Outliers may contain meaningful information, so excluding or altering these values may reduce valuable information. Further, because identifying a data point as extreme leaves it up to the judgment of the analyst, leaving in all observations eliminates that need to judge a value as extreme.

The second option excludes the extreme observations. One measure of central tendency in this case is the **trimmed mean**, which is computed by excluding a stated small percentage of the lowest and highest values and then computing an arithmetic mean of the remaining values. For example, a 5% trimmed mean discards the lowest 2.5% and the highest 2.5% of values and computes the mean of the remaining 95% of values. A trimmed mean is used in sports competitions when judges' lowest and highest scores are discarded in computing a contestant's score.

The third option involves substituting values for the extreme values. A measure of central tendency in this case is the **winsorized mean**. It is calculated by assigning a stated percentage of the lowest values equal to one specified low value and a stated percentage of the highest values equal to one specified high value, and then it computes a mean from the restated data. For example, a 95% winsorized mean sets the bottom 2.5% of values equal to the value at or below which 2.5% of all the values lie (as will be seen shortly, this is called the "2.5th percentile" value) and the top 2.5% of values equal to the value at or below which 97.5% of all the values lie (the "97.5th percentile" value).

In Exhibit 37, we show the differences among these options for handling outliers using daily returns for the fictitious Euro-Asia-Africa (EAA) Equity Index in Exhibit 11.

Exhibit 37 Handling Outliers: Daily Returns to an Index

Consider the fictitious EAA Equity Index. Using daily returns on the EAA Equity Index for the period of five years, consisting of 1,258 trading days, we can see the effect of trimming and winsorizing the data:

	Arithmetic Mean (%)	Trimmed Mean [Trimmed 5%] (%)	Winsorized Mean [95%] (%)
Mean	0.035	0.048	0.038
Number of Observations	1,258	1,194	1,258

The trimmed mean eliminates the lowest 2.5% of returns, which in this sample is any daily return less than −1.934%, and it eliminates the highest 2.5%, which in this sample is any daily return greater than 1.671%. The result of this trimming is that the mean is calculated using 1,194 observations instead of the original sample's 1,258 observations.

The winsorized mean substitutes −1.934% for any return below −1.934 and substitutes 1.671% for any return above 1.671. The result in this case is that the trimmed and winsorized means are above the arithmetic mean.

7.2 The Median

A second important measure of central tendency is the median.

Definition of Median. The **median** is the value of the middle item of a set of items that has been sorted into ascending or descending order. In an odd-numbered sample of n items, the median is the value of the item that occupies the $(n + 1)/2$ position. In an even-numbered sample, we define the median as the mean of the values of items occupying the $n/2$ and $(n + 2)/2$ positions (the two middle items).

Suppose we have a return on assets (in %) for each of three companies: 0.0, 2.0, and 2.1. With an odd number of observations ($n = 3$), the median occupies the $(n + 1)/2 = 4/2 =$ 2nd position. The median is 2.0%. The value of 2.0% is the "middlemost" observation: One lies above it, and one lies below it. Whether we use the calculation for an even- or odd-numbered sample, an equal number of observations lie above and below the median. A distribution has only one median.

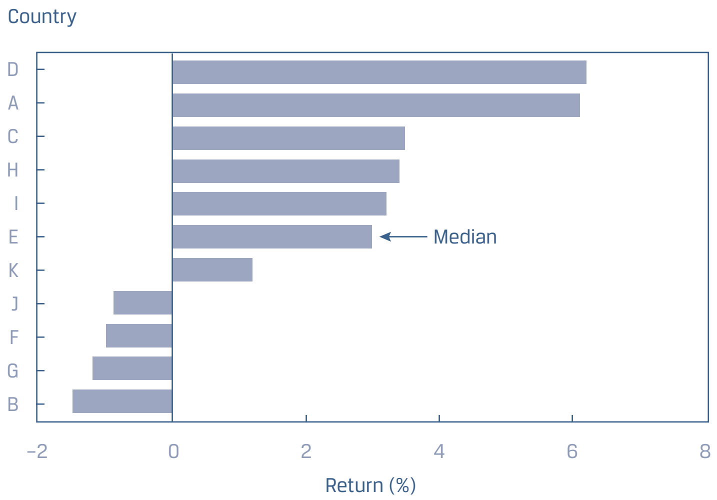
A potential advantage of the median is that, unlike the mean, extreme values do not affect it. For example, if a sample consists of the observations of 1, 2, 3, 4, 5, 6 and 1,000, the median is 4. The median is not influenced by the extremely large outcome of 1,000. In other words, the median is affected less by outliers than the mean and therefore is useful in describing data that follow a distribution that is not symmetric, such as revenues.

The median, however, does not use all the information about the size of the observations; it focuses only on the relative position of the ranked observations. Calculating the median may also be more complex. To do so, we need to order the observations from smallest to largest, determine whether the sample size is even or odd, and then on that basis, apply one of two calculations. Mathematicians express this disadvantage by saying that the median is less mathematically tractable than the mean.

We use the data from Exhibit 35 to demonstrate finding the median, reproduced in Exhibit 38 in ascending order of the return for Year 3, with the ranked position from 1 (lowest) to 11 (highest) indicated. Because this sample has 11 observations, the median is the value in the sorted array that occupies the $(11 + 1)/2 =$ 6th position. Country E's index occupies the sixth position and is the median. The arithmetic mean for Year 3 for this sample of indexes is 2.0%, whereas the median is 3.0.

Exhibit 38 Returns on Selected Country Stock Indexes for Year 3 in Ascending Order

Index	Year 3 Return (%)		Position
Country B	−1.5		1
Country G	−1.2		2
Country F	−1.0		3
Country J	−0.9		4
Country K	1.2		5
Country E	3.0	←	6
Country I	3.2		7
Country H	3.4		8
Country C	3.5		9
Country A	6.1		10
Country D	6.2		11



If a sample has an even number of observations, the median is the mean of the two values in the middle. For example, if our sample in Exhibit 38 had 12 indexes instead of 11, the median would be the mean of the values in the sorted array that occupy the sixth and the seventh positions.

7.3 The Mode

The third important measure of central tendency is the mode.

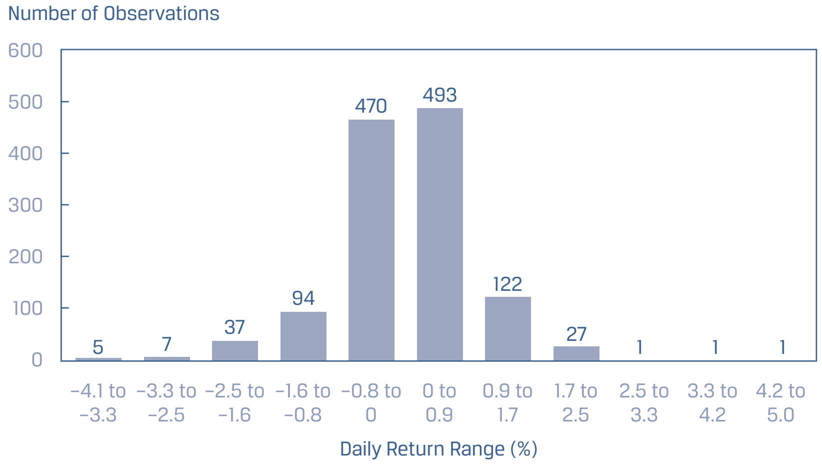
> **Definition of Mode.** The **mode** is the most frequently occurring value in a distribution.

A distribution can have more than one mode, or even no mode. When a distribution has a single value that is most frequently occurring, the distribution is said to be **unimodal**. If a distribution has two most frequently occurring values, then it has two modes and is called **bimodal**. If the distribution has three most frequently occurring values, then it is **trimodal**. When all the values in a dataset are different, the distribution has no mode because no value occurs more frequently than any other value.

Stock return data and other data from continuous distributions may not have a modal outcome. When such data are grouped into bins, however, we often find an interval (possibly more than one) with the highest frequency: the **modal interval** (or intervals). Consider the frequency distribution of the daily returns for the EAA Equity Index over five years that we looked at in Exhibit 11. A histogram for the frequency distribution of these daily returns is shown in Exhibit 39. The modal interval always has the highest bar in the histogram; in this case, the modal interval is 0.0 to 0.9%, and this interval has 493 observations out of a total of 1,258 observations.

Notice that this histogram in Exhibit 39 looks slightly different from the one in Exhibit 11, since this one has 11 bins and follows the seven-step procedure exactly. Thus, the bin width is 0.828 [= (5.00 – −4.11)/11], and the first bin begins at the minimum value of −4.11%. It was noted previously that for ease of interpretation, in practice bin width is often rounded up to the nearest whole number; the first bin can start at the nearest whole number below the minimum value. These refinements and the use of 10 bins were incorporated into the histogram in Exhibit 11, which has a modal interval of 0.0% to 1.0%.

Exhibit 39 Histogram of Daily Returns on the EAA Equity Index



The mode is the only measure of central tendency that can be used with nominal data. For example, when we categorize investment funds into different styles and assign a number to each style, the mode of these categorized data is the most frequent investment fund style.

7.4 Other Concepts of Mean

Earlier we explained the arithmetic mean, which is a fundamental concept for describing the central tendency of data. An advantage of the arithmetic mean over two other measures of central tendency, the median and mode, is that the mean uses all the information about the size of the observations. The mean is also relatively easy to work with mathematically.

However, other concepts of mean are very important in investments. In the following sections, we discuss such concepts.

7.4.1 *The Weighted Mean*

The concept of weighted mean arises repeatedly in portfolio analysis. In the arithmetic mean, all sample observations are equally weighted by the factor $1/n$. In working with portfolios, we often need the more general concept of weighted mean to allow for different (i.e., unequal) weights on different observations.

To illustrate the weighted mean concept, an investment manager with $100 million to invest might allocate $70 million to equities and $30 million to bonds. The portfolio, therefore, has a weight of 0.70 on stocks and 0.30 on bonds. How do we calculate the return on this portfolio? The portfolio's return clearly involves an averaging of the returns on the stock and bond investments. The mean that we compute, however, must reflect the fact that stocks have a 70% weight in the portfolio and bonds have a 30% weight. The way to reflect this weighting is to multiply the return on the stock investment by 0.70 and the return on the bond investment by 0.30, then sum the two results. This sum is an example of a weighted mean. It would be incorrect to take an arithmetic mean of the return on the stock and bond investments, equally weighting the returns on the two asset classes.

Weighted Mean Formula. The **weighted mean** $\overline{X}_w$ (read "X-bar sub-w"), for a set of observations $X_1, X_2, \ldots, X_n$ with corresponding weights of $w_1, w_2, \ldots, w_n$, is computed as:

$$\overline{X}_w = \sum_{i=1}^{n} w_i X_i \tag{3}$$

where the sum of the weights equals 1; that is, $\sum_i w_i = 1$.

In the context of portfolios, a positive weight represents an asset held long and a negative weight represents an asset held short.

The formula for the weighted mean can be compared to the formula for the arithmetic mean. For a set of observations $X_1, X_2, \ldots, X_n$, let the weights $w_1, w_2, \ldots, w_n$ all equal $1/n$. Under this assumption, the formula for the weighted mean is $(1/n)\sum_{i=1}^{n} X_i$. This is the formula for the arithmetic mean. Therefore, the arithmetic mean is a special case of the weighted mean in which all the weights are equal.

EXAMPLE 9

Calculating a Weighted Mean

Using the country index data shown in Exhibit 35, consider a portfolio that consists of three funds that track three countries' indexes: County C, Country G, and Country K. The portfolio weights and index returns are as follows:

Index Tracked by Fund	Allocation (%)	Annual Return (%) Year 1	Year 2	Year 3
Country C	25%	5.3	1.2	3.5
Country G	45%	12.7	6.7	−1.2
Country K	30%	11.5	3.4	1.2

Using the information provided, calculate the returns on the portfolio for each year.

Solution

Converting the percentage asset allocation to decimal form, we find the mean return as the weighted average of the funds' returns. We have:

Mean portfolio return for Year 1 $= 0.25\,(5.3) + 0.45\,(12.7) + 0.30(11.5)$
$= 10.50\%$

Mean portfolio return for Year 2 $= 0.25\,(1.2) + 0.45\,(6.7) + 0.30\,(3.4)$
$= 4.34\%$

Mean portfolio return for Year 3 $= 0.25\,(3.5) + 0.45\,(-1.2) + 0.30\,(1.2)$
$= 0.70\%$

This example illustrates the general principle that a portfolio return is a weighted sum. Specifically, a portfolio's return is the weighted average of the returns on the assets in the portfolio; the weight applied to each asset's return is the fraction of the portfolio invested in that asset.

Market indexes are computed as weighted averages. For market-capitalization weighted indexes, such as the CAC-40 in France, the TOPIX in Japan, or the S&P 500 in the United States, each included stock receives a weight corresponding to its market value divided by the total market value of all stocks in the index.

Our illustrations of weighted mean use past data, but they might just as well use forward-looking data. When we take a weighted average of forward-looking data, the weighted mean is the expected value. Suppose we make one forecast for the year-end level of the S&P 500 assuming economic expansion and another forecast for the year-end level of the S&P 500 assuming economic contraction. If we multiply the first forecast by the probability of expansion and the second forecast by the probability of contraction and then add these weighted forecasts, we are calculating the expected value of the S&P 500 at year-end. If we take a weighted average of possible future returns on the S&P 500, where the weights are the probabilities, we are computing the S&P 500's expected return. The probabilities must sum to 1, satisfying the condition on the weights in the expression for weighted mean, Equation 3.

7.4.2 *The Geometric Mean*

The geometric mean is most frequently used to average rates of change over time or to compute the growth rate of a variable. In investments, we frequently use the geometric mean to either average a time series of rates of return on an asset or a portfolio or to compute the growth rate of a financial variable, such as earnings or sales. The geometric mean is defined by the following formula.

> **Geometric Mean Formula.** The **geometric mean**, $\bar{X}_G$, of a set of observations $X_1, X_2, \ldots, X_n$ is:

$$\bar{X}_G = \sqrt[n]{X_1 X_2 X_3 \ldots X_n} \text{ with } X_i \geq 0 \text{ for } i = 1, 2, \ldots, n. \quad (4)$$

Equation 4 has a solution, and the geometric mean exists only if the product under the square root sign is non-negative. Therefore, we must impose the restriction that all the observations X_i are greater than or equal to zero. We can solve for the geometric mean directly with any calculator that has an exponentiation key (on most calculators, y^x). We can also solve for the geometric mean using natural logarithms. Equation 4 can also be stated as

$$ln\bar{X}_G = \frac{1}{n} ln\left(X_1 X_2 X_3 \ldots X_n\right)$$

or, because the logarithm of a product of terms is equal to the sum of the logarithms of each of the terms, as

$$ln\bar{X}_G = \frac{\sum_{i=1}^{n} lnX_i}{n}$$

When we have computed $ln\bar{X}_G$, then $\bar{X}_G = e^{ln\bar{X}_G}$ (on most calculators, the key for this step is e^x).

Risky assets can have negative returns up to −100% (if their price falls to zero), so we must take some care in defining the relevant variables to average in computing a geometric mean. We cannot just use the product of the returns for the sample and then take the nth root because the returns for any period could be negative. We must recast the returns to make them positive. We do this by adding 1.0 to the returns expressed as decimals, where R_t represents the return in period t. The term $(1 + R_t)$ represents the year-ending value relative to an initial unit of investment at the beginning of the year. As long as we use $(1 + R_t)$, the observations will never be negative because the biggest negative return is –100%. The result is the geometric mean of $1 + R_t$; by then subtracting 1.0 from this result, we obtain the geometric mean of the individual returns R_t.

An equation that summarizes the calculation of the geometric mean return, R_G, is a slightly modified version of Equation 4 in which X_i represents "1 + return in decimal form." Because geometric mean returns use time series, we use a subscript t indexing time as well. We calculate one plus the geometric mean return as:

$$1 + R_G = \sqrt[T]{(1 + R_1)(1 + R_2)\ldots(1 + R_T)}$$

We can represent this more compactly as:

$$1 + R_G = \left[\prod_{t=1}^{T}(1 + R_t)\right]^{\frac{1}{T}}$$

where the capital Greek letter 'pi,' Π, denotes the arithmetical operation of multiplication of the T terms. Once we subtract one, this becomes the formula for the geometric mean return.

For example, the returns on Country B's index are given in Exhibit 35 as 7.8, 6.3, and –1.5%. Putting the returns into decimal form and adding 1.0 produces 1.078, 1.063, and 0.985. Using Equation 4, we have $\sqrt[3]{(1.078)(1.063)(0.985)} = \sqrt[3]{1.128725} =$ 1.041189. This number is 1 plus the geometric mean rate of return. Subtracting 1.0 from this result, we have 1.041189 – 1.0 = 0.041189, or approximately 4.12%. This is lower than the arithmetic mean for County B's index of 4.2%.

Geometric Mean Return Formula. Given a time series of holding period returns R_t, $t = 1, 2, \ldots, T$, the geometric mean return over the time period spanned by the returns R_1 through R_T is:

$$R_G = \left[\prod_{t=1}^{T}(1 + R_t)\right]^{\frac{1}{T}} - 1 \qquad (5)$$

We can use Equation 5 to solve for the geometric mean return for any return data series. Geometric mean returns are also referred to as compound returns. If the returns being averaged in Equation 5 have a monthly frequency, for example, we may call the geometric mean monthly return the compound monthly return. The next example illustrates the computation of the geometric mean while contrasting the geometric and arithmetic means.

EXAMPLE 10

Geometric and Arithmetic Mean Returns

Using the data in Exhibit 35, calculate the arithmetic mean and the geometric mean returns over the three years for each of the three stock indexes: those of Country D, Country E, and Country F.

Solution

The arithmetic mean returns calculations are:

	Annual Return (%)			**Sum**	
	Year 1	**Year 2**	**Year 3**	$\sum_{i=1}^{3} R_i$	**Arithmetic Mean**
Country D	−2.4	−3.1	6.2	0.7	0.233
Country E	−4.0	−3.0	3.0	−4.0	−1.333
Country F	5.4	5.2	−1.0	9.6	3.200

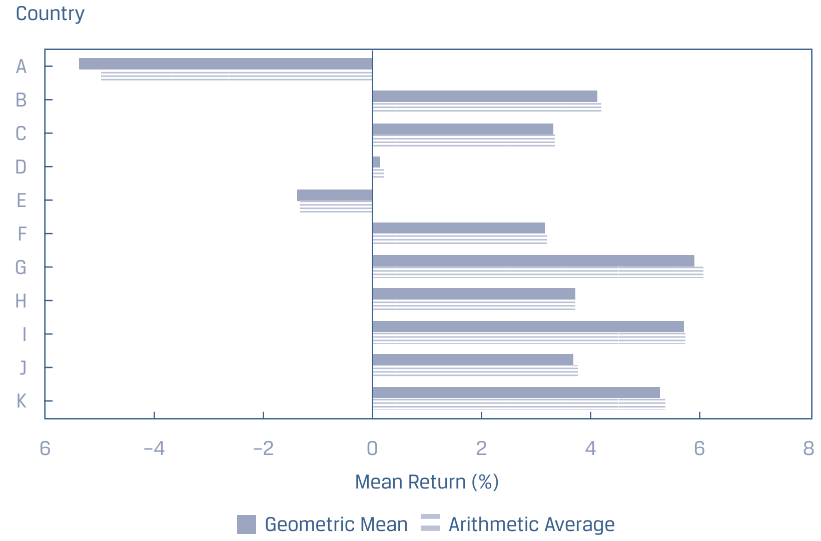
Geometric mean returns calculations are:

	1 + Return in Decimal Form (1 + R_t)			**Product**	**3rd root**	
	Year 1	**Year 2**	**Year 3**	$\prod_{t}^{T}(1+R_t)$	$\left[\prod_{t}^{3}(1+R_t)\right]^{\frac{1}{3}}$	**Geometric mean return (%)**
Country D	0.976	0.969	1.062	1.00438	1.00146	0.146
Country E	0.960	0.970	1.030	0.95914	0.98619	−1.381
Country F	1.054	1.052	0.990	1.09772	1.03157	3.157

In Example 10, the geometric mean return is less than the arithmetic mean return for each country's index returns. In fact, the geometric mean is always less than or equal to the arithmetic mean. The only time that the two means will be equal is when there is no variability in the observations—that is, when all the observations in the series are the same.

In general, the difference between the arithmetic and geometric means increases with the variability within the sample; the more disperse the observations, the greater the difference between the arithmetic and geometric means. Casual inspection of the returns in Exhibit 35 and the associated graph of means suggests a greater variability for Country A's index relative to the other indexes, and this is confirmed with the greater deviation of the geometric mean return (−5.38%) from the arithmetic mean return (−4.97%), as we show in Exhibit 40. How should the analyst interpret these results?

Exhibit 40 Arithmetic and Geometric Mean Returns for Country Stock Indexes: Years 1 to 3



The geometric mean return represents the growth rate or compound rate of return on an investment. One unit of currency invested in a fund tracking the Country B index at the beginning of Year 1 would have grown to $(1.078)(1.063)(0.985) = 1.128725$ units of currency, which is equal to 1 plus the geometric mean return compounded over three periods: $[1 + 0.041189]^3 = 1.128725$, confirming that the geometric mean is the compound rate of return. With its focus on the profitability of an investment over a multi-period horizon, the geometric mean is of key interest to investors. The arithmetic mean return, focusing on average single-period performance, is also of interest. Both arithmetic and geometric means have a role to play in investment management, and both are often reported for return series.

For reporting historical returns, the geometric mean has considerable appeal because it is the rate of growth or return we would have to earn each year to match the actual, cumulative investment performance. Suppose we purchased a stock for €100 and two years later it was worth €100, with an intervening year at €200. The geometric mean of 0% is clearly the compound rate of growth during the two years, which we can confirm by compounding the returns: $[(1 + 1.00)(1 - 0.50)]^{1/2} - 1 = 0\%$. Specifically, the ending amount is the beginning amount times $(1 + R_G)^2$. The geometric mean is an excellent measure of past performance.

The arithmetic mean, which is [100% + −50%]/2 = 25% in the above example, can distort our assessment of historical performance. As we noted previously, the arithmetic mean is always greater than or equal to the geometric mean. If we want to estimate the average return over a one-period horizon, we should use the arithmetic mean because the arithmetic mean is the average of one-period returns. If we want to estimate the average returns over more than one period, however, we should use the geometric mean of returns because the geometric mean captures how the total returns are linked over time. In a forward-looking context, a financial analyst calculating expected risk premiums may find that the weighted mean is appropriate, with the probabilities of the possible outcomes used as the weights.

Dispersion in cash flows or returns causes the arithmetic mean to be larger than the geometric mean. The more dispersion in the sample of returns, the more divergence exists between the arithmetic and geometric means. If there is zero variance in a sample of observations, the geometric and arithmetic return are equal.

7.4.3 *The Harmonic Mean*

The arithmetic mean, the weighted mean, and the geometric mean are the most frequently used concepts of mean in investments. A fourth concept, the **harmonic mean**, $\bar{X}_H$, is another measure of central tendency. The harmonic mean is appropriate in cases in which the variable is a rate or a ratio. The terminology "harmonic" arises from its use of a type of series involving reciprocals known as a harmonic series.

Harmonic Mean Formula. The harmonic mean of a set of observations X_1, X_2, …, X_n is:

$$\bar{X}_H = \frac{n}{\sum_{i=1}^{n}(1/X_i)} \text{ with } X_i > 0 \text{ for } i = 1, 2, \ldots, n. \tag{6}$$

The harmonic mean is the value obtained by summing the reciprocals of the observations—terms of the form $1/X_i$—then averaging that sum by dividing it by the number of observations n, and, finally, taking the reciprocal of the average.

The harmonic mean may be viewed as a special type of weighted mean in which an observation's weight is inversely proportional to its magnitude. For example, if there is a sample of observations of 1, 2, 3, 4, 5, 6, and 1,000, the harmonic mean is 2.8560. Compared to the arithmetic mean of 145.8571, we see the influence of the outlier (the 1,000) to be much less than in the case of the arithmetic mean. So, the harmonic mean is quite useful as a measure of central tendency in the presence of outliers.

The harmonic mean is used most often when the data consist of rates and ratios, such as P/Es. Suppose three peer companies have P/Es of 45, 15, and 15. The arithmetic mean is 25, but the harmonic mean, which gives less weight to the P/E of 45, is 19.3.

EXAMPLE 11

Harmonic Mean Returns and the Returns on Selected Country Stock Indexes

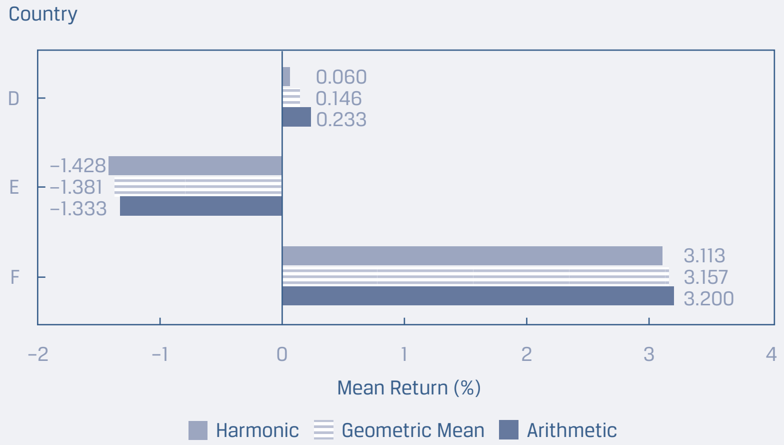
Using data in Exhibit 35, calculate the harmonic mean return over the 2016–2018 period for three stock indexes: Country D, Country E, and Country F.

Calculating the Harmonic Mean for the Indexes

	Inverse of 1 + Return, or $\frac{1}{(1+X_i)}$ where X_i is the return in decimal form					
Index	**Year 1**	**Year 2**	**Year 3**	$\sum_{i}^{n} 1/X_i$	$\frac{n}{\sum_{i}^{n} 1/X_i}$	**Harmonic Mean (%)**
Country D	1.02459	1.03199	0.94162	2.99820	1.00060	0.05999
Country E	1.04167	1.03093	0.97087	3.04347	0.98572	−1.42825
Country F	0.94877	0.95057	1.01010	2.90944	1.03113	3.11270

Comparing the three types of means, we see the arithmetic mean is higher than the geometric mean return, and the geometric mean return is higher than the harmonic mean return. We can see the differences in these means in the following graph:

Harmonic, Geometric, and Arithmetic Means of Selected Country Indexes



The harmonic mean is a relatively specialized concept of the mean that is appropriate for averaging ratios ("amount per unit") when the ratios are repeatedly applied to a fixed quantity to yield a variable number of units. The concept is best explained through an illustration. A well-known application arises in the investment strategy known as **cost averaging**, which involves the periodic investment of a fixed amount of money. In this application, the ratios we are averaging are prices per share at different purchase dates, and we are applying those prices to a constant amount of money to yield a variable number of shares. An illustration of the harmonic mean to cost averaging is provided in Example 12.

EXAMPLE 12

Cost Averaging and the Harmonic Mean

Suppose an investor purchases €1,000 of a security each month for $n = 2$ months. The share prices are €10 and €15 at the two purchase dates. What is the average price paid for the security?

Purchase in the first month = €1,000/€10 = 100 shares

Purchase in the second month = €1,000/€15 = 66.67 shares

The purchases are 166.67 shares in total, and the price paid per share is €2,000/166.67 = €12.

The average price paid is in fact the harmonic mean of the asset's prices at the purchase dates. Using Equation 6, the harmonic mean price is 2/[(1/10) + (1/15)] = €12. The value €12 is less than the arithmetic mean purchase price (€10 + €15)/2 = €12.5.

However, we could find the correct value of €12 using the weighted mean formula, where the weights on the purchase prices equal the shares purchased at a given price as a proportion of the total shares purchased. In our example, the calculation would be (100/166.67)€10.00 + (66.67/166.67)€15.00 = €12. If we had invested varying amounts of money at each date, we could not use the harmonic mean formula. We could, however, still use the weighted mean formula.

h Evaluate alternative definitions of mean to address an investment problem

Since they use the same data but involve different progressions in their respective calculations (that is, arithmetic, geometric, and harmonic progressions) the arithmetic, geometric, and harmonic means are mathematically related to one another. While we will not go into the proof of this relationship, the basic result follows:

Arithmetic mean × Harmonic mean = Geometric mean^2.

However, the key question is: Which mean to use in what circumstances?

EXAMPLE 13

Calculating the Arithmetic, Geometric, and Harmonic Means for P/Es

Each year in December, a securities analyst selects her 10 favorite stocks for the next year. Exhibit 41 gives the P/E, the ratio of share price to projected earnings per share (EPS), for her top-10 stock picks for the next year.

Exhibit 41 Analyst's 10 Favorite Stocks for Next Year

Stock	P/E
Stock 1	22.29
Stock 2	15.54
Stock 3	9.38
Stock 4	15.12
Stock 5	10.72
Stock 6	14.57
Stock 7	7.20
Stock 8	7.97
Stock 9	10.34
Stock 10	8.35

For these 10 stocks,

1 Calculate the arithmetic mean P/E.

2 Calculate the geometric mean P/E.

3 Calculate the harmonic mean P/E.

Solution

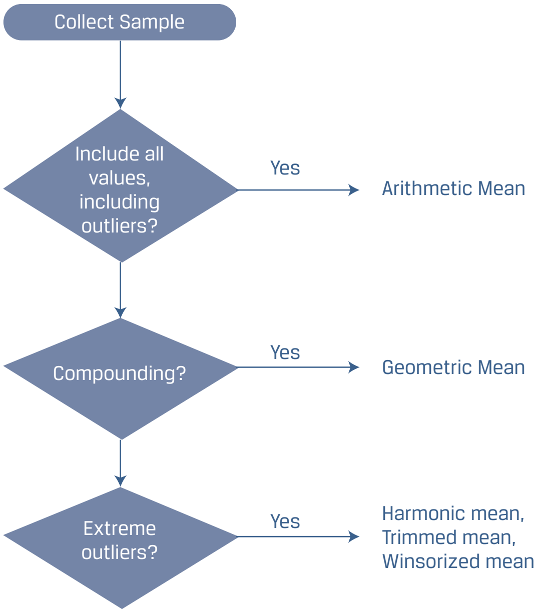
Stock	P/E	Natural Log of the P/E $ln(X_i)$	Inverse of the P/E $1/X_i$
Stock 1	22.29	3.1041	0.0449
Stock 2	15.54	2.7434	0.0644
Stock 3	9.38	2.2386	0.1066
Stock 4	15.12	2.7160	0.0661
Stock 5	10.72	2.3721	0.0933
Stock 6	14.57	2.6790	0.0686
Stock 7	7.20	1.9741	0.1389
Stock 8	7.97	2.0757	0.1255
Stock 9	10.34	2.3360	0.0967
Stock 10	8.35	2.1223	0.1198
Sum	121.48	24.3613	0.9247

1. The arithmetic mean is 121.48/10 = 12.1480.
2. The geometric mean is $e^{24.3613/10} = 11.4287$.
3. The harmonic mean is 10/0.9247 = 10.8142.

A mathematical fact concerning the harmonic, geometric, and arithmetic means is that unless all the observations in a dataset have the same value, the harmonic mean is less than the geometric mean, which, in turn, is less than the arithmetic mean. The choice of which mean to use depends on many factors, as we describe in Exhibit 42:

- Are there outliers that we want to include?
- Is the distribution symmetric?
- Is there compounding?
- Are there extreme outliers?

Exhibit 42 Deciding Which Central Tendency Measure to Use



8 QUANTILES

i. Calculate quantiles and interpret related visualizations

Having discussed measures of central tendency, we now examine an approach to describing the location of data that involves identifying values at or below which specified proportions of the data lie. For example, establishing that 25, 50, and 75% of the annual returns on a portfolio are at or below the values −0.05, 0.16, and 0.25, respectively, provides concise information about the distribution of portfolio returns. Statisticians use the word **quantile** (or **fractile**) as the most general term for a value at or below which a stated fraction of the data lies. In the following section, we describe the most commonly used quantiles—quartiles, quintiles, deciles, and percentiles—and their application in investments.

8.1 Quartiles, Quintiles, Deciles, and Percentiles

We know that the median divides a distribution of data in half. We can define other dividing lines that split the distribution into smaller sizes. **Quartiles** divide the distribution into quarters, **quintiles** into fifths, **deciles** into tenths, and **percentiles** into hundredths. Given a set of observations, the yth percentile is the value at or below which y% of observations lie. Percentiles are used frequently, and the other measures can be defined with respect to them. For example, the first quartile (Q_1) divides a distribution such that 25% of the observations lie at or below it; therefore, the first quartile is also the 25th percentile. The second quartile (Q_2) represents the 50th percentile, and the third quartile (Q_3) represents the 75th percentile (i.e., 75%of the observations lie at or below it). The **interquartile range** (IQR) is the difference between the third quartile and the first quartile, or IQR = $Q_3 - Q_1$.

When dealing with actual data, we often find that we need to approximate the value of a percentile. For example, if we are interested in the value of the 75th percentile, we may find that no observation divides the sample such that exactly 75% of the observations lie at or below that value. The following procedure, however, can help us determine or estimate a percentile. The procedure involves first locating the position of the percentile within the set of observations and then determining (or estimating) the value associated with that position.

Let P_y be the value at or below which y% of the distribution lies, or the yth percentile. (For example, P_{18} is the point at or below which 18% of the observations lie; this implies that 100 − 18 = 82% of the observations are greater than P_{18}.) The formula for the position (or location) of a percentile in an array with n entries sorted in ascending order is:

$$L_y = (n + 1)\frac{y}{100} \tag{7}$$

where y is the percentage point at which we are dividing the distribution, and L_y is the location (L) of the percentile (P_y) in the array sorted in ascending order. The value of L_y may or may not be a whole number. In general, as the sample size increases, the percentile location calculation becomes more accurate; in small samples it may be quite approximate.

To summarize:

- When the location, L_y, is a whole number, the location corresponds to an actual observation. For example, if we are determining the third quartile (Q_3) in a sample of size n = 11, then L_y would be L_{75} = (11 + 1)(75/100) = 9, and the third quartile would be $P_{75} = X_9$, where X_i is defined as the value of the observation in the ith ($i = L_{75}$, so 9th), position of the data sorted in ascending order.
- When L_y is not a whole number or integer, L_y lies between the two closest integer numbers (one above and one below), and we use **linear interpolation** between those two places to determine P_y. Interpolation means estimating an unknown value on the basis of two known values that surround it (i.e., lie above and below it); the term "linear" refers to a straight-line estimate.

Example 14 illustrates the calculation of various quantiles for the daily return on the EAA Equity Index.

EXAMPLE 14

Percentiles, Quintiles, and Quartiles for the EAA Equity Index

Using the daily returns on the fictitious EAA Equity Index over five years and ranking them by return, from lowest to highest daily return, we show the return bins from 1 (the lowest 5%) to 20 (the highest 5%) as follows:

Exhibit 43 EAA Equity Index Daily Returns Grouped by Size of Return

Bin	Cumulative Percentage of Sample Trading Days (%)	Daily Return (%) Between*		Number of Observations
		Lower Bound	Upper Bound	
1	5	−4.108	−1.416	63
2	10	−1.416	−0.876	63
3	15	−0.876	−0.629	63
4	20	−0.629	−0.432	63
5	25	−0.432	−0.293	63
6	30	−0.293	−0.193	63
7	35	−0.193	−0.124	62
8	40	−0.124	−0.070	63
9	45	−0.070	−0.007	63
10	50	−0.007	0.044	63
11	55	0.044	0.108	63
12	60	0.108	0.173	63
13	65	0.173	0.247	63
14	70	0.247	0.343	62
15	75	0.343	0.460	63
16	80	0.460	0.575	63
17	85	0.575	0.738	63
18	90	0.738	0.991	63
19	95	0.991	1.304	63
20	100	1.304	5.001	63

Note that because of the continuous nature of returns, it is not likely for a return to fall on the boundary for any bin other than the minimum (Bin = 1) and maximum (Bin = 20).

1. Identify the 10th and 90th percentiles.
2. Identify the first, second, and third quintiles.
3. Identify the first and third quartiles.
4. Identify the median.
5. Calculate the interquartile range.

Solution to 1

The 10th and 90th percentiles correspond to the bins or ranked returns that include 10% and 90% of the daily returns, respectively. The 10th percentile corresponds to the return of −0.876% (and includes returns of that much and lower), and the 90th percentile corresponds to the return of 0.991% (and lower).

Solution to 2

The first quintile corresponds to the lowest 20% of the ranked data, or −0.432% (and lower).

The second quintile corresponds to the lowest 40% of the ranked data, or −0.070% (and lower).

The third quintile corresponds to the lowest 60% of the ranked data, or 0.173% (and lower).

Solution to 3

The first quartile corresponds to the lowest 25% of the ranked data, or −0.293% (and lower).

The third quartile corresponds to the lowest 75% of the ranked data, or 0.460% (and lower).

Solution to 4

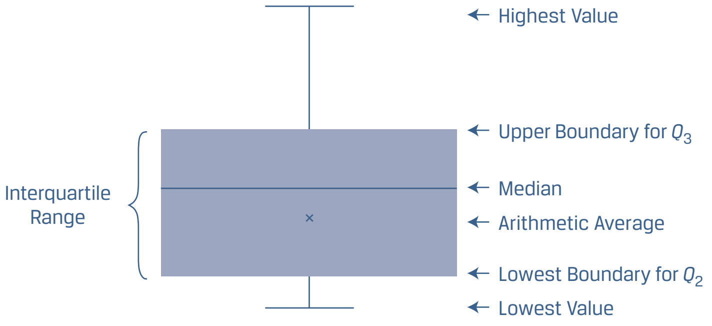
The median is the return for which 50% of the data lies on either side, which is 0.044%, the highest daily return in the 10th bin out of 20.

Solution to 5

The interquartile range is the difference between the third and first quartiles, 0.460% and −0.293%, or 0.753%.

One way to visualize the dispersion of data across quartiles is to use a diagram, such as a box and whisker chart. A **box and whisker plot** consists of a "box" with "whiskers" connected to the box, as shown in Exhibit 44. The "box" represents the lower bound of the second quartile and the upper bound of the third quartile, with the median or arithmetic average noted as a measure of central tendency of the entire distribution. The whiskers are the lines that run from the box and are bounded by the "fences," which represent the lowest and highest values of the distribution.

Exhibit 44 Box and Whisker Plot



There are several variations for box and whisker displays. For example, for ease in detecting potential outliers, the fences of the whiskers may be a function of the interquartile range instead of the highest and lowest values like that in Exhibit 44.

In Exhibit 44, visually, the interquartile range is the height of the box and the fences are set at extremes. But another form of box and whisker plot typically uses 1.5 times the interquartile range for the fences. Thus, the upper fence is 1.5 times the interquartile range added to the upper bound of Q_3, and the lower fence is 1.5 times the interquartile range subtracted from the lower bound of Q_2. Observations beyond the fences (i.e., outliers) may also be displayed.

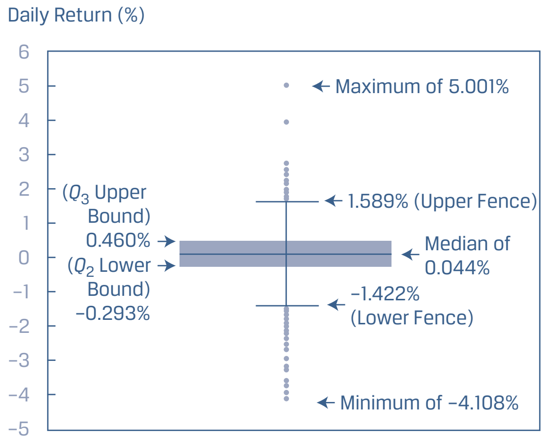
We can see the role of outliers in such a box and whisker plot using the EAA Equity Index daily returns, as shown in Exhibit 45. Referring back to Exhibit 43 (Example 13), we know:

- The maximum and minimum values of the distribution are 5.001 and −4.108, respectively, while the median (50th percentile) value is 0.044.

- The interquartile range is 0.753 [= 0.460 − (−0.293)], and when multiplied by 1.5 and added to the Q_3 upper bound of 0.460 gives an upper fence of 1.589 [= (1.5 × 0.753) + 0.460].
- The lower fence is determined in a similar manner, using the Q_2 lower bound, to be −1.422 [= −(1.5 × 0.753) + (−0.293)].

As noted, any observation above (below) the upper (lower) fence is deemed to be an outlier.

Exhibit 45 Box and Whisker Chart for EAA Equity Index Daily Returns



EXAMPLE 15

Quantiles

Consider the results of an analysis focusing on the market capitalizations of a sample of 100 firms:

Bin	Cumulative Percentage of Sample (%)	Market Capitalization (in billions of €) Lower Bound	Market Capitalization (in billions of €) Upper Bound	Number of Observations
1	5	0.28	15.45	5
2	10	15.45	21.22	5
3	15	21.22	29.37	5
4	20	29.37	32.57	5
5	25	32.57	34.72	5
6	30	34.72	37.58	5
7	35	37.58	39.90	5
8	40	39.90	41.57	5
9	45	41.57	44.86	5
10	50	44.86	46.88	5
11	55	46.88	49.40	5
12	60	49.40	51.27	5
13	65	51.27	53.58	5
14	70	53.58	56.66	5

Bin	Cumulative Percentage of Sample (%)	Market Capitalization (in billions of €) Lower Bound	Market Capitalization (in billions of €) Upper Bound	Number of Observations
15	75	56.66	58.34	5
16	80	58.34	63.10	5
17	85	63.10	67.06	5
18	90	67.06	73.00	5
19	95	73.00	81.62	5
20	100	81.62	96.85	5

Using this information, answer the following five questions.

1 The tenth percentile corresponds to observations in bins:

A 2.

B 1 and 2.

C 19 and 20.

2 The second quintile corresponds to observations in bins:

A 8

B 5, 6, 7, and 8.

C 6, 7, 8, 9, and 10.

3 The fourth quartile corresponds to observations in bins:

A 17.

B 17, 18, 19, and 20.

C 16, 17, 18, 19, and 20.

4 The median is *closest* to:

A 44.86.

B 46.88.

C 49.40.

5 The interquartile range is *closest* to:

A 20.76.

B 23.62.

C 25.52.

Solution to 1

B is correct because the tenth percentile corresponds to the lowest 10% of the observations in the sample, which are in bins 1 and 2.

Solution to 2

B is correct because the second quintile corresponds to the second 20% of observations. The first 20% consists of bins 1 through 4. The second 20% of observations consists of bins 5 through 8.

Solution to 3

C is correct because a quartile consists of 25% of the data, and the last 25% of the 20 bins are 16 through 20.

Solution to 4

B is correct because this is the center of the 20 bins. The market capitalization of 46.88 is the highest value of the 10th bin and the lowest value of the 11th bin.

Solution to 5

B is correct because the interquartile range is the difference between the lowest value in the second quartile and the highest value in the third quartile. The lowest value of the second quartile is 34.72, and the highest value of the third quartile is 58.34. Therefore, the interquartile range is 58.34 − 34.72 = 23.62.

8.2 Quantiles in Investment Practice

In this section, we briefly discuss the use of quantiles in investments. Quantiles are used in portfolio performance evaluation as well as in investment strategy development and research.

Investment analysts use quantiles every day to rank performance—for example, the performance of portfolios. The performance of investment managers is often characterized in terms of the percentile or quartile in which they fall relative to the performance of their peer group of managers. The Morningstar investment fund star rankings, for example, associate the number of stars with percentiles of performance relative to similar-style investment funds.

Another key use of quantiles is in investment research. For example, analysts often refer to the set of companies with returns falling below the 10th percentile cutoff point as the bottom return decile. Dividing data into quantiles based on some characteristic allows analysts to evaluate the impact of that characteristic on a quantity of interest. For instance, empirical finance studies commonly rank companies based on the market value of their equity and then sort them into deciles. The first decile contains the portfolio of those companies with the smallest market values, and the tenth decile contains those companies with the largest market values. Ranking companies by decile allows analysts to compare the performance of small companies with large ones.

9

MEASURES OF DISPERSION

j Calculate and interpret measures of dispersion

Few would disagree with the importance of expected return or mean return in investments: The mean return tells us where returns, and investment results, are centered. To more completely understand an investment, however, we also need to know how returns are dispersed around the mean. **Dispersion** is the variability around the central tendency. If mean return addresses reward, then dispersion addresses risk.

In this section, we examine the most common measures of dispersion: range, mean absolute deviation, variance, and standard deviation. These are all measures of **absolute dispersion**. Absolute dispersion is the amount of variability present without comparison to any reference point or benchmark.

These measures are used throughout investment practice. The variance or standard deviation of return is often used as a measure of risk pioneered by Nobel laureate Harry Markowitz. Other measures of dispersion, mean absolute deviation and range, are also useful in analyzing data.

9.1 The Range

We encountered range earlier when we discussed the construction of frequency distributions. It is the simplest of all the measures of dispersion.

Definition of Range. The **range** is the difference between the maximum and minimum values in a dataset:

$$\text{Range} = \text{Maximum value} - \text{Minimum value}. \tag{8}$$

As an illustration of range, consider Exhibit 35, our example of annual returns for countries' stock indexes. The range of returns for Year 1 is the difference between the returns of Country G's index and Country A's index, or 12.7 − (−15.6) = 28.3%. The range of returns for Year 3 is the difference between the returns for the County D index and the Country B index, or 6.2 − (−1.5) = 7.7%.

An alternative definition of range specifically reports the maximum and minimum values. This alternative definition provides more information than does the range as defined in Equation 8. In other words, in the above-mentioned case for Year 1, the range is reported as "from 12.7% to −15.6%."

One advantage of the range is ease of computation. A disadvantage is that the range uses only two pieces of information from the distribution. It cannot tell us how the data are distributed (that is, the shape of the distribution). Because the range is the difference between the maximum and minimum returns, it can reflect extremely large or small outcomes that may not be representative of the distribution.

9.2 The Mean Absolute Deviation

Measures of dispersion can be computed using all the observations in the distribution rather than just the highest and lowest. But how should we measure dispersion? Our previous discussion on properties of the arithmetic mean introduced the notion of distance or deviation from the mean $\left(X_i - \overline{X}\right)$ as a fundamental piece of information used in statistics. We could compute measures of dispersion as the arithmetic average of the deviations around the mean, but we would encounter a problem: The deviations around the mean always sum to 0. If we computed the mean of the deviations, the result would also equal 0. Therefore, we need to find a way to address the problem of negative deviations canceling out positive deviations.

One solution is to examine the absolute deviations around the mean as in the **mean absolute deviation**. This is also known as the average absolute deviation.

Mean Absolute Deviation Formula. The mean absolute deviation (MAD) for a sample is:

$$\text{MAD} = \frac{\sum_{i=1}^{n}\left|X_i - \overline{X}\right|}{n} \tag{9}$$

where $\overline{X}$ is the sample mean, n is the number of observations in the sample, and the | | indicate the absolute value of what is contained within these bars.

In calculating MAD, we ignore the signs of the deviations around the mean. For example, if $X_i = -11.0$ and $\overline{X} = 4.5$, the absolute value of the difference is $|-11.0 - 4.5| = |-15.5| = 15.5$. The mean absolute deviation uses all of the observations in the sample and is thus superior to the range as a measure of dispersion. One technical drawback of MAD is that it is difficult to manipulate mathematically compared with the next measure we will introduce, sample variance. Example 16 illustrates the use of the range and the mean absolute deviation in evaluating risk.

EXAMPLE 16

Mean Absolute Deviation for Selected Countries' Stock Index Returns

Using the country stock index returns in Exhibit 35, calculate the mean absolute deviation of the index returns for each year. Note the sample mean returns ($\bar{X}$) are 3.5%, 2.5%, and 2.0% for Years 1, 2, and 3, respectively.

Solution

	Absolute Value of Deviation from the Mean $\lvert X_i - \bar{X} \rvert$		
	Year 1	**Year 2**	**Year 3**
Country A	19.1	7.9	4.1
Country B	4.3	3.8	3.5
Country C	1.8	1.3	1.5
Country D	5.9	5.6	4.2
Country E	7.5	5.5	1.0
Country F	1.9	2.7	3.0
Country G	9.2	4.2	3.2
Country H	0.0	1.8	1.4
Country I	2.7	5.3	1.2
Country J	4.6	1.6	2.9
Country K	8.0	0.9	0.8
Sum	65.0	40.6	26.8
MAD	5.91	3.69	2.44

For Year 3, for example, the sum of the absolute deviations from the arithmetic mean ($\bar{X} = 2.0$) is 26.8. We divide this by 11, with the resulting MAD of 2.44.

9.3 Sample Variance and Sample Standard Deviation

The mean absolute deviation addressed the issue that the sum of deviations from the mean equals zero by taking the absolute value of the deviations. A second approach to the treatment of deviations is to square them. The variance and standard deviation, which are based on squared deviations, are the two most widely used measures of dispersion. **Variance** is defined as the average of the squared deviations around the mean. **Standard deviation** is the positive square root of the variance. The following discussion addresses the calculation and use of variance and standard deviation.

9.3.1 *Sample Variance*

In investments, we often do not know the mean of a population of interest, usually because we cannot practically identify or take measurements from each member of the population. We then estimate the population mean using the mean from a sample drawn from the population, and we calculate a sample variance or standard deviation.

Sample Variance Formula. The **sample variance**, s^2, is:

$$s^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1} \tag{10}$$

where $\bar{X}$ is the sample mean and n is the number of observations in the sample.

Given knowledge of the sample mean, we can use Equation 10 to calculate the sum of the squared differences from the mean, taking account of all n items in the sample, and then to find the mean squared difference by dividing the sum by $n - 1$. Whether a difference from the mean is positive or negative, squaring that difference results in a positive number. Thus, variance takes care of the problem of negative deviations from the mean canceling out positive deviations by the operation of squaring those deviations.

For the sample variance, by dividing by the sample size minus 1 (or $n - 1$) rather than n, we improve the statistical properties of the sample variance. In statistical terms, the sample variance defined in Equation 10 is an unbiased estimator of the population variance (a concept covered later in the curriculum on sampling). The quantity $n - 1$ is also known as the number of degrees of freedom in estimating the population variance. To estimate the population variance with s^2, we must first calculate the sample mean, which itself is an estimated parameter. Therefore, once we have computed the sample mean, there are only $n - 1$ independent pieces of information from the sample; that is, if you know the sample mean and $n - 1$ of the observations, you could calculate the missing sample observation.

9.3.2 *Sample Standard Deviation*

Because the variance is measured in squared units, we need a way to return to the original units. We can solve this problem by using standard deviation, the square root of the variance. Standard deviation is more easily interpreted than the variance because standard deviation is expressed in the same unit of measurement as the observations. By taking the square root, we return the values to the original unit of measurement. Suppose we have a sample with values in euros. Interpreting the standard deviation in euros is easier than interpreting the variance in squared euros.

Sample Standard Deviation Formula. The **sample standard deviation**, s, is:

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}} \tag{11}$$

where $\bar{X}$ is the sample mean and n is the number of observations in the sample.

To calculate the sample standard deviation, we first compute the sample variance. We then take the square root of the sample variance. The steps for computing the sample variance and the standard deviation are provided in Exhibit 46.

Exhibit 46 Steps to Calculate Sample Standard Deviation and Variance

Step	Description	Notation
1	Calculate the sample mean	$\bar{X}$
2	Calculate the deviations from the sample mean	$\left(X_i - \bar{X}\right)$

(continued)

Exhibit 46 (Continued)

Step	Description	Notation
3	Calculate each observation's squared deviation from the sample mean	$\left(X_i - \bar{X}\right)^2$
4	Sum the squared deviations from the mean	$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$
5	Divide the sum of squared deviations from the mean by $n - 1$. This is the variance (s^2).	$\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}$
6	Take the square root of the sum of the squared deviations divided by $n - 1$. This is the standard deviation (s).	$\sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}}$

We illustrate the process of calculating the sample variance and standard deviation in Example 17 using the returns of the selected country stock indexes presented in Exhibit 35.

EXAMPLE 17

Calculating Sample Variance and Standard Deviation for Returns on Selected Country Stock Indexes

Using the sample information on country stock indexes in Exhibit 35, calculate the sample variance and standard deviation of the sample of index returns for Year 3.

Solution

Index	Sample Observation	Deviation from the Sample Mean	Squared Deviation
Country A	6.1	4.1	16.810
Country B	−1.5	−3.5	12.250
Country C	3.5	1.5	2.250
Country D	6.2	4.2	17.640
Country E	3.0	1.0	1.000
Country F	−1.0	−3.0	9.000
Country G	−1.2	−3.2	10.240
Country H	3.4	1.4	1.960
Country I	3.2	1.2	1.440
Country J	−0.9	−2.9	8.410
Country K	1.2	−0.8	0.640
Sum	**22.0**	**0.0**	**81.640**

Sample variance = 81.640/10 = 8.164

Sample standard deviation = $\sqrt{8.164}$ = 2.857

In addition to looking at the cross-sectional standard deviation as we did in Example 17, we could also calculate the standard deviation of a given country's returns across time (that is, the three years). Consider Country F, which has an arithmetic mean return of 3.2%. The sample standard deviation is calculated as:

$$\sqrt{\frac{(0.054-0.032)^2+(0.052-0.032)^2+(-0.01-0.032)^2}{2}}$$

$$=\sqrt{\frac{0.000484+0.000400+0.001764}{2}}$$

$$=\sqrt{0.001324}$$

$$=3.6387\%.$$

Because the standard deviation is a measure of dispersion about the arithmetic mean, we usually present the arithmetic mean and standard deviation together when summarizing data. When we are dealing with data that represent a time series of percentage changes, presenting the geometric mean—representing the compound rate of growth—is also very helpful.

9.3.3 *Dispersion and the Relationship between the Arithmetic and the Geometric Means*

We can use the sample standard deviation to help us understand the gap between the arithmetic mean and the geometric mean. The relation between the arithmetic mean $(\bar{X})$ and geometric mean $(\bar{X}_G)$ is:

$$\bar{X}_G \approx \bar{X}-\frac{s^2}{2}$$

In other words, the larger the variance of the sample, the wider the difference between the geometric mean and the arithmetic mean.

Using the data for Country F from Example 8, the geometric mean return is 3.1566%, the arithmetic mean return is 3.2%, and the factor $s^2/2$ is 0.001324/2 = 0.0662%:

$$3.1566\% \approx 3.2\% - 0.0662\%$$

$$3.1566\% \approx 3.1338\%.$$

This relation informs us that the more disperse or volatile the returns, the larger the gap between the geometric mean return and the arithmetic mean return.

DOWNSIDE DEVIATION AND COEFFICIENT OF VARIATION

10

k Calculate and interpret target downside deviation

An asset's variance or standard deviation of returns is often interpreted as a measure of the asset's risk. Variance and standard deviation of returns take account of returns above and below the mean, or upside and downside risks, respectively. However, investors are typically concerned only with **downside risk**—for example, returns below the mean or below some specified minimum target return. As a result, analysts have developed measures of downside risk.

In practice, we may be concerned with values of return (or another variable) below some level other than the mean. For example, if our return objective is 6.0% annually (our minimum acceptable return), then we may be concerned particularly with returns below 6.0% a year. The 6.0% is the target. The target downside deviation, also referred to as the **target semideviation**, is a measure of dispersion of the observations (here, returns) below the target. To calculate a sample target semideviation, we first specify the target. After identifying observations below the target, we find the sum of the squared negative deviations from the target, divide that sum by the total number of observations in the sample minus 1, and, finally, take the square root.

Sample Target Semideviation Formula. The target semideviation, s_{Target}, is:

$$s_{\text{Target}} = \sqrt{\sum_{\text{for all} X_i \leq B}^{n} \frac{(X_i - B)^2}{n - 1}}, \tag{12}$$

where B is the target and n is the total number of sample observations. We illustrate this in Example 18.

EXAMPLE 18

Calculating Target Downside Deviation

Suppose the monthly returns on a portfolio are as shown:

Monthly Portfolio Returns

Month	Return (%)
January	5
February	3
March	–1
April	–4
May	4
June	2
July	0
August	4
September	3
October	0
November	6
December	5

1 Calculate the target downside deviation when the target return is 3%.

2 If the target return were 4%, would your answer be different from that for question 1? Without using calculations, explain how would it be different?

Solution to 1

Month	Observation	Deviation from the 3% Target	Deviations below the Target	Squared Deviations below the Target
January	5	2	—	—
February	3	0	—	—
March	–1	–4	–4	16
April	–4	–7	–7	49
May	4	1	—	—
June	2	–1	–1	1
July	0	–3	–3	9
August	4	1	—	—
September	3	0	—	—
October	0	–3	–3	9
November	6	3	—	—
December	5	2	—	—
Sum				**84**

$$\textbf{Target semideviation} = \sqrt{\frac{84}{11}} = 2.7634\%$$

Solution to 2

If the target return is higher, then the existing deviations would be larger and there would be several more values in the deviations and squared deviations below the target; so, the target semideviation would be larger.

How does the target downside deviation relate to the sample standard deviation? We illustrate the differences between the target downside deviation and the standard deviation in Example 19, using the data in Example 18.

EXAMPLE 19

Comparing the Target Downside Deviation with the Standard Deviation

1 Given the data in Example 18, calculate the sample standard deviation.

2 Given the data in Example 18, calculate the target downside deviation if the target is 2%.

3 Compare the standard deviation, the target downside deviation if the target is 2%, and the target downside deviation if the target is 3%.

Solution to 1

Month	Observation	Deviation from the mean	Squared deviation
January	5	2.75	7.5625
February	3	0.75	0.5625

(continued)

Month	Observation	Deviation from the mean	Squared deviation
March	−1	−3.25	10.5625
April	−4	−6.25	39.0625
May	4	1.75	3.0625
June	2	−0.25	0.0625
July	0	−2.25	5.0625
August	4	1.75	3.0625
September	3	0.75	0.5625
October	0	−2.25	5.0625
November	6	3.75	14.0625
December	5	2.75	7.5625
Sum	27		**96.2500**

The sample standard deviation is $\sqrt{\frac{96.2500}{11}} = 2.958\%$.

Solution to 2

Month	Observation	Deviation from the 2% Target	Deviations below the Target	Squared Deviations below the Target
January	5	3	—	—
February	3	1	—	—
March	−1	−3	−3	9
April	−4	−6	−6	36
May	4	2	—	—
June	2	0	—	—
July	0	−2	−2	4
August	4	2	—	—
September	3	1	—	—
October	0	−2	−2	4
November	6	4	—	—
December	5	3	—	—
Sum				**53**

The target semideviation with 2% target = $\sqrt{\frac{53}{11}} = 2.195\%$.

Solution to 3

The standard deviation is based on the deviation from the mean, which is 2.25%. The standard deviation includes all deviations from the mean, not just those below it. This results in a sample standard deviation of 2.958%.

Considering just the four observations below the 2% target, the target semideviation is 2.195%. It is less than the sample standard deviation since target semideviation captures only the downside risk (i.e., deviations below the target). Considering target semideviation with a 3% target, there are now five observations below 3%, so the target semideviation is higher, at 2.763%.

10.1 Coefficient of Variation

We noted earlier that the standard deviation is more easily interpreted than variance because standard deviation uses the same units of measurement as the observations. We may sometimes find it difficult to interpret what standard deviation means in terms of the relative degree of variability of different sets of data, however, either because the datasets have markedly different means or because the datasets have different units of measurement. In this section, we explain a measure of relative dispersion, the coefficient of variation that can be useful in such situations. **Relative dispersion** is the amount of dispersion relative to a reference value or benchmark.

The coefficient of variation is helpful in such situations as that just described (i.e., datasets with markedly different means or different units of measurement).

Coefficient of Variation Formula. The **coefficient of variation**, CV, is the ratio of the standard deviation of a set of observations to their mean value:

$$\text{CV} = s/\overline{X} \tag{13}$$

where s is the sample standard deviation and $\overline{X}$ is the sample mean.

When the observations are returns, for example, the coefficient of variation measures the amount of risk (standard deviation) per unit of reward (mean return). An issue that may arise, especially when dealing with returns, is that if $\overline{X}$ is negative, the statistic is meaningless.

The CV may be stated as a multiple (e.g., 2 times) or as a percentage (e.g., 200%). Expressing the magnitude of variation among observations relative to their average size, the coefficient of variation permits direct comparisons of dispersion across different datasets. Reflecting the correction for scale, the coefficient of variation is a scale-free measure (that is, it has no units of measurement).

We illustrate the usefulness of coefficient of variation for comparing datasets with markedly different standard deviations using two hypothetical samples of companies in Example 20.

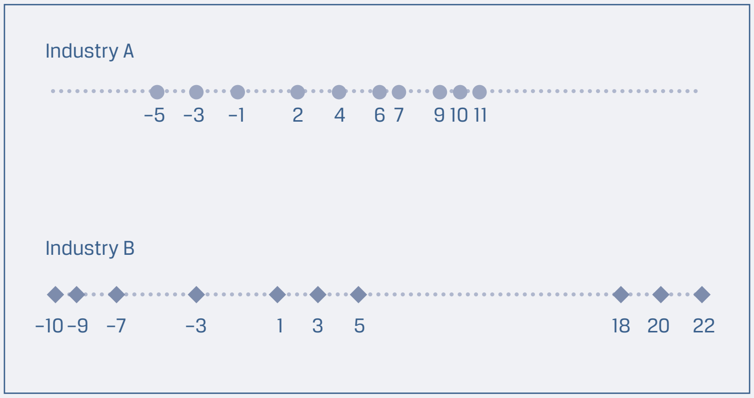
EXAMPLE 20

Coefficient of Variation of Returns on Assets

Suppose an analyst collects the return on assets (in percentage terms) for ten companies for each of two industries:

Company	Industry A	Industry B
1	−5	−10
2	−3	−9
3	−1	−7
4	2	−3
5	4	1
6	6	3
7	7	5
8	9	18
9	10	20
10	11	22

These data can be represented graphically as the following:



1 Calculate the average return on assets (ROA) for each industry.
2 Calculate the standard deviation of ROA for each industry.
3 Calculate the coefficient of variation of ROA for each industry.

Solution to 1

The arithmetic mean for both industries is the sum divided by 10, or 40/10 = 4%.

Solution to 2

The standard deviation using Equation 11 for Industry A is 5.60, and for Industry B the standard deviation is 12.12.

Solution to 3

The coefficient of variation for Industry A = 5.60/4 = 1.40.

The coefficient of variation for Industry B = 12.12/4 = 3.03.

Though the two industries have the same arithmetic mean ROA, the dispersion is different—with Industry B's returns on assets being much more disperse than those of Industry A. The coefficients of variation for these two industries reflects this, with Industry B having a larger coefficient of variation. The interpretation is that the risk per unit of mean return is more than two times (2.16 = 3.03/1.40) greater for Industry B compared to Industry A.

11 THE SHAPE OF THE DISTRIBUTIONS

l. Interpret skewness

Mean and variance may not adequately describe an investment's distribution of returns. In calculations of variance, for example, the deviations around the mean are squared, so we do not know whether large deviations are likely to be positive or negative. We need to go beyond measures of central tendency and dispersion to reveal other important characteristics of the distribution. One important characteristic of interest to analysts is the degree of symmetry in return distributions.

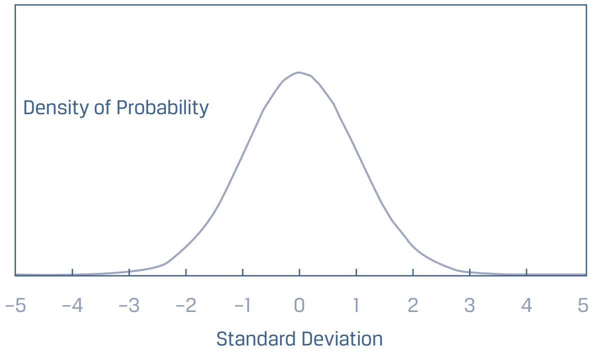
If a return distribution is symmetrical about its mean, each side of the distribution is a mirror image of the other. Thus, equal loss and gain intervals exhibit the same frequencies. If the mean is zero, for example, then losses from −5% to −3% occur with about the same frequency as gains from 3% to 5%.

One of the most important distributions is the normal distribution, depicted in Exhibit 47. This symmetrical, bell-shaped distribution plays a central role in the mean–variance model of portfolio selection; it is also used extensively in financial risk management. The normal distribution has the following characteristics:

- Its mean, median, and mode are equal.
- It is completely described by two parameters—its mean and variance (or standard deviation).

But with any distribution other than a normal distribution, more information than the mean and variance is needed to characterize its shape.

Exhibit 47 The Normal Distribution

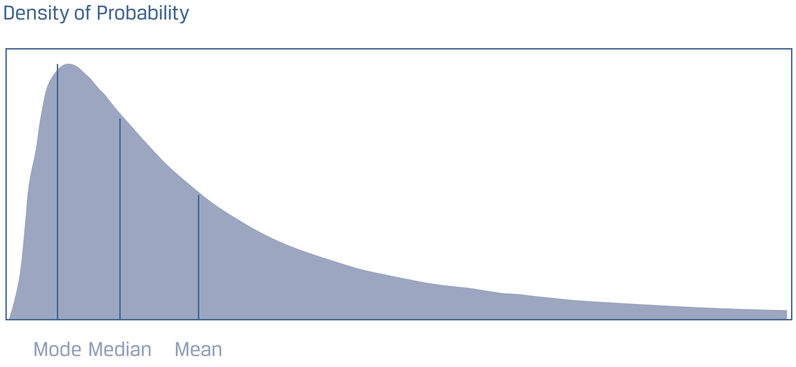


A distribution that is not symmetrical is **skewed**. A return distribution with positive skew has frequent small losses and a few extreme gains. A return distribution with negative skew has frequent small gains and a few extreme losses. Exhibit 48 shows continuous positively and negatively skewed distributions. The continuous positively skewed distribution shown has a long tail on its right side; the continuous negatively skewed distribution shown has a long tail on its left side.

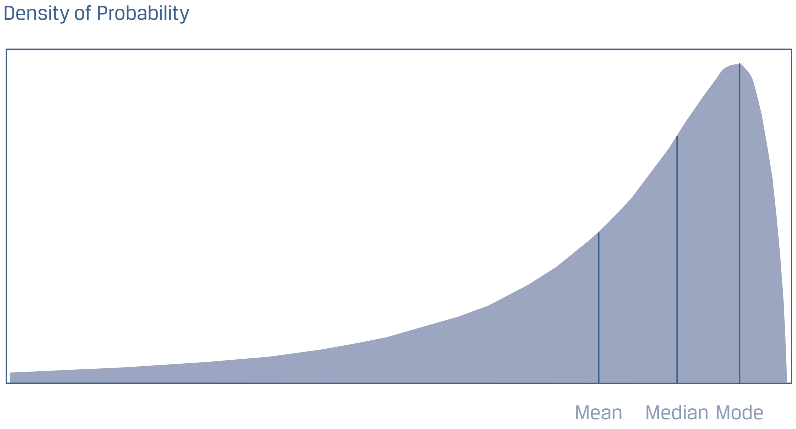
For a continuous positively skewed unimodal distribution, the mode is less than the median, which is less than the mean. For the continuous negatively skewed unimodal distribution, the mean is less than the median, which is less than the mode. For a given expected return and standard deviation, investors should be attracted by a positive skew because the mean return lies above the median. Relative to the mean return, positive skew amounts to limited, though frequent, downside returns compared with somewhat unlimited, but less frequent, upside returns.

Exhibit 48 Properties of Skewed Distributions

A. Positively Skewed



B. Negatively Skewed



Skewness is the name given to a statistical measure of skew. (The word "skewness" is also sometimes used interchangeably for "skew.") Like variance, skewness is computed using each observation's deviation from its mean. **Skewness** (sometimes referred to as relative skewness) is computed as the average cubed deviation from the mean standardized by dividing by the standard deviation cubed to make the measure free of scale. A symmetric distribution has skewness of 0, a positively skewed distribution has positive skewness, and a negatively skewed distribution has negative skewness, as given by this measure.

We can illustrate the principle behind the measure by focusing on the numerator. Cubing, unlike squaring, preserves the sign of the deviations from the mean. If a distribution is positively skewed with a mean greater than its median, then more than half of the deviations from the mean are negative and less than half are positive. However, for the sum of the cubed deviations to be positive, the losses must be small and likely and the gains less likely but more extreme. Therefore, if skewness is positive, the average magnitude of positive deviations is larger than the average magnitude of negative deviations.

The approximation for computing **sample skewness** when n is large (100 or more) is:

$$\text{Skewness} \approx \left(\frac{1}{n}\right)\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^3}{s^3}$$

A simple example illustrates that a symmetrical distribution has a skewness measure equal to 0. Suppose we have the following data: 1, 2, 3, 4, 5, 6, 7, 8, and 9. The mean outcome is 5, and the deviations are −4, −3, −2, −1, 0, 1, 2, 3, and 4. Cubing the deviations yields −64, −27, −8, −1, 0, 1, 8, 27, and 64, with a sum of 0. The numerator of skewness (and so skewness itself) is thus equal to 0, supporting our claim.

As you will learn as the CFA Program curriculum unfolds, different investment strategies may tend to introduce different types and amounts of skewness into returns.

11.1 The Shape of the Distributions: Kurtosis

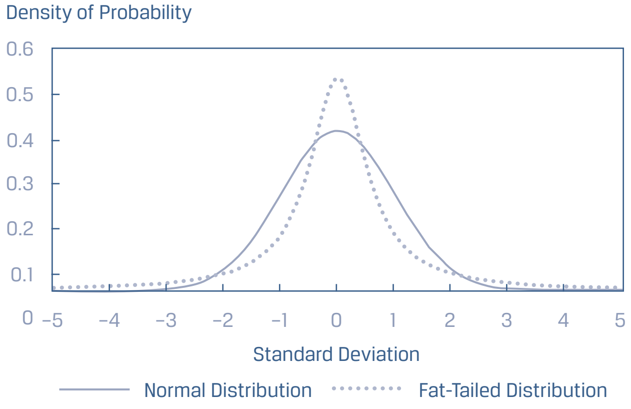
m Interpret kurtosis

In the previous section, we discussed how to determine whether a return distribution deviates from a normal distribution because of skewness. Another way in which a return distribution might differ from a normal distribution is its relative tendency to generate large deviations from the mean. Most investors would perceive a greater chance of extremely large deviations from the mean as increasing risk.

Kurtosis is a measure of the combined weight of the tails of a distribution relative to the rest of the distribution—that is, the proportion of the total probability that is outside of, say, 2.5 standard deviations of the mean. A distribution that has fatter tails than the normal distribution is referred to as **leptokurtic** or **fat-tailed**; a distribution that has thinner tails than the normal distribution is referred to as being **platykurtic** or **thin-tailed**; and a distribution similar to the normal distribution as concerns relative weight in the tails is called **mesokurtic**. A fat-tailed (thin-tailed) distribution tends to generate more-frequent (less-frequent) extremely large deviations from the mean than the normal distribution.

Exhibit 49 illustrates a fat-tailed distribution. It has fatter tails than the normal distribution. By construction, the fat-tailed and normal distributions in this exhibit have the same mean, standard deviation, and skewness. Note that this fat-tailed distribution is more likely than the normal distribution to generate observations in the tail regions defined by the intersection of graphs near a standard deviation of about ±2.5. This fat-tailed distribution is also more likely to generate observations that are near the mean, defined here as the region ±1 standard deviation around the mean. In compensation, to have probabilities sum to 1, this distribution generates fewer observations in the regions between the central region and the two tail regions.

Exhibit 49 Fat-Tailed Distribution Compared to the Normal Distribution



The calculation for kurtosis involves finding the average of deviations from the mean raised to the fourth power and then standardizing that average by dividing by the standard deviation raised to the fourth power. A normal distribution has kurtosis of 3.0, so a fat-tailed distribution has a kurtosis of above 3 and a thin-tailed distribution of below 3.0.

Excess kurtosis is the kurtosis relative to the normal distribution. For a large sample size (n = 100 or more), **sample excess kurtosis** (K_E) is approximately as follows:

$$K_E \approx \left[\left(\frac{1}{n} \right) \frac{\sum_{i=1}^{n} \left(X_i - \bar{X} \right)^4}{s^4} \right] - 3$$

As with skewness, this measure is free of scale. Many statistical packages report estimates of sample **excess kurtosis**, labeling this as simply "kurtosis."

Excess kurtosis thus characterizes kurtosis relative to the normal distribution. A normal distribution has excess kurtosis equal to 0. A fat-tailed distribution has excess kurtosis greater than 0, and a thin-tailed distribution has excess kurtosis less than 0. A return distribution with positive excess kurtosis—a fat-tailed return distribution—has more frequent extremely large deviations from the mean than a normal distribution.

Summarizing:

If kurtosis is …	then excess kurtosis is …	Therefore, the distribution is …	And we refer to the distribution as being …
above 3.0	above 0.	fatter-tailed than the normal distribution.	fat-tailed (leptokurtic).
equal to 3.0	equal to 0.	similar in tails to the normal distribution.	mesokurtic.
less than 3.0	less than 0.	thinner-tailed than the normal distribution.	thin-tailed (platykurtic).

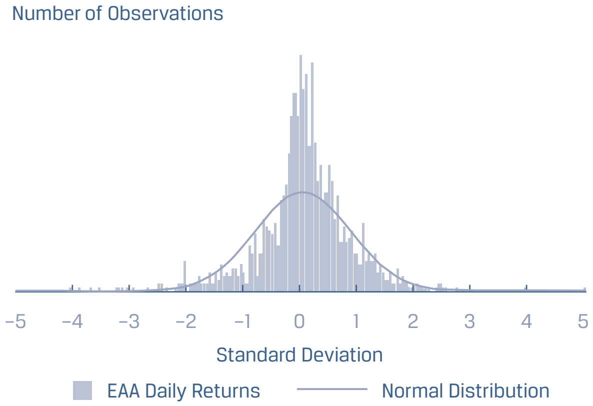
Most equity return series have been found to be fat-tailed. If a return distribution is fat-tailed and we use statistical models that do not account for the distribution, then we will underestimate the likelihood of very bad or very good outcomes. Using the data on the daily returns of the fictitious EAA Equity Index, we see the skewness and kurtosis of these returns in Exhibit 50.

Exhibit 50 Skewness and Kurtosis of EAA Equity Index Daily Returns

	Daily Return (%)
Arithmetic mean	0.0347
Standard deviation	0.8341
	Measure of Symmetry
Skewness	−0.4260
Excess kurtosis	3.7962

We can see this graphically, comparing the distribution of the daily returns with a normal distribution with the same mean and standard deviation:

Exhibit 50 (Continued)



Using both the statistics and the graph, we see the following:

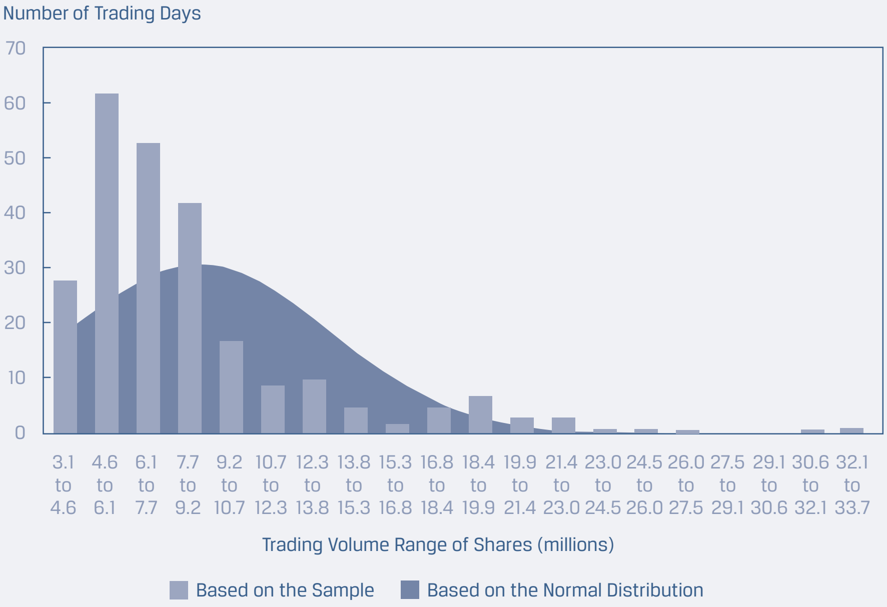
- The distribution is negatively skewed, as indicated by the negative calculated skewness of −0.4260 and the influence of observations below the mean of 0.0347%.
- The highest frequency of returns occurs within the −0.5 to 0.0 standard deviations from the mean (i.e., negatively skewed).
- The distribution is fat-tailed, as indicated by the positive excess kurtosis of 3.7962. We can see fat tails, a concentration of returns around the mean, and fewer observations in the regions between the central region and the two-tail regions.

EXAMPLE 21

Interpreting Skewness and Kurtosis

Consider the daily trading volume for a stock for one year, as shown in the graph below. In addition to the count of observations within each bin or interval, the number of observations anticipated based on a normal distribution (given the sample arithmetic average and standard deviation) is provided in the chart as well. The average trading volume per day for this stock in this year is 8.6 million shares, and the standard deviation is 4.9 million shares.

Histogram of Daily Trading Volume for a Stock for One Year



1. Describe whether or not this distribution is skewed. If so, what could account for this situation?
2. Describe whether or not this distribution displays kurtosis. How would you make this determination?

Solution to 1

The distribution appears to be skewed to the right, or positively skewed. This is likely due to: (1) no possible negative trading volume on a given trading day, so the distribution is truncated at zero; and (2) greater-than-typical trading occurring relatively infrequently, such as when there are company-specific announcements.

The actual skewness for this distribution is 2.1090, which supports this interpretation.

Solution to 2

The distribution appears to have excess kurtosis, with a right-side fat tail and with maximum shares traded in the 4.6 to 6.1 million range, exceeding what is expected if the distribution was normally distributed. There are also fewer observations than expected between the central region and the tail.

The actual excess kurtosis for this distribution is 5.2151, which supports this interpretation.

12 CORRELATION BETWEEN TWO VARIABLES

- **Interpret correlation between two variables**

Now that we have some understanding of sample variance and standard deviation, we can more formally consider the concept of correlation between two random variables that we previously explored visually in the scatter plots in Section 6. **Correlation** is a measure of the linear relationship between two random variables.

The first step is to consider how two variables vary together, their covariance.

Definition of Sample Covariance. The **sample covariance** (s_{XY}) is a measure of how two variables in a sample move together:

$$s_{XY} = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{n-1} \tag{14}$$

Equation 14 indicates that the sample covariance is the average value of the product of the deviations of observations on two random variables (X_i and Yi) from their sample means. If the random variables are returns, the units would be returns squared. Also, note the use of $n - 1$ in the denominator, which ensures that the sample covariance is an unbiased estimate of population covariance.

Stated simply, covariance is a measure of the joint variability of two random variables. If the random variables vary in the same direction—for example, X tends to be above its mean when Y is above its mean, and X tends to be below its mean when Y is below its mean—then their covariance is positive. If the variables vary in the opposite direction relative to their respective means, then their covariance is negative.

By itself, the size of the covariance measure is difficult to interpret as it is not normalized and so depends on the magnitude of the variables. This brings us to the normalized version of covariance, which is the correlation coefficient.

Definition of Sample Correlation Coefficient. The **sample correlation coefficient** is a standardized measure of how two variables in a sample move together. The sample correlation coefficient (r_{XY}) is the ratio of the sample covariance to the product of the two variables' standard deviations:

$$r_{XY} = \frac{s_{XY}}{s_X s_Y} \tag{15}$$

Importantly, the correlation coefficient expresses the strength of the linear relationship between the two random variables.

12.1 Properties of Correlation

We now discuss the correlation coefficient, or simply correlation, and its properties in more detail, as follows:

1. Correlation ranges from −1 and +1 for two random variables, X and Y:

   $-1 \le r_{XY} \le +1.$

2. A correlation of 0 (uncorrelated variables) indicates an absence of any linear (that is, straight-line) relationship between the variables.
3. A positive correlation close to +1 indicates a strong positive linear relationship. A correlation of 1 indicates a perfect linear relationship.
4. A negative correlation close to −1 indicates a strong negative (that is, inverse) linear relationship. A correlation of −1 indicates a perfect inverse linear relationship.

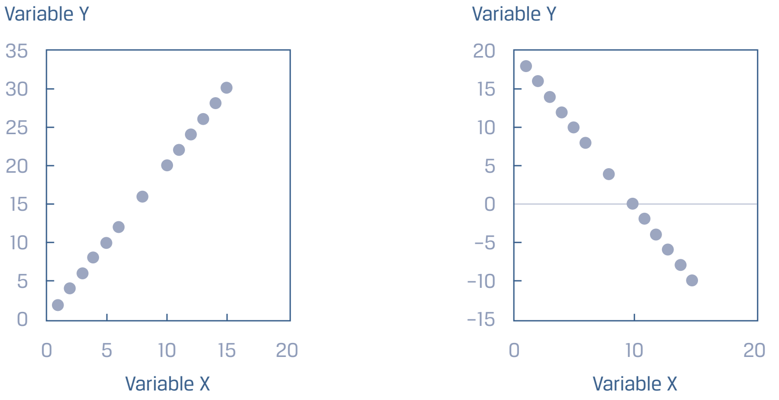
We will make use of scatter plots, similar to those used previously in our discussion of data visualization, to illustrate correlation. In contrast to the correlation coefficient, which expresses the relationship between two data series using a single number, a scatter plot depicts the relationship graphically. Therefore, scatter plots are a very useful tool for the sensible interpretation of a correlation coefficient.

Exhibit 51 shows examples of scatter plots. Panel A shows the scatter plot of two variables with a correlation of +1. Note that all the points on the scatter plot in Panel A lie on a straight line with a positive slope. Whenever variable *X* increases by one unit, variable *Y* increases by two units. Because all of the points in the graph lie on a straight line, an increase of one unit in *X* is associated with exactly a two-unit increase in *Y*, regardless of the level of *X*. Even if the slope of the line were different (but positive), the correlation between the two variables would still be +1 as long as all the points lie on that straight line. Panel B shows a scatter plot for two variables with a correlation coefficient of −1. Once again, the plotted observations all fall on a straight line. In this graph, however, the line has a negative slope. As *X* increases by one unit, *Y* decreases by two units, regardless of the initial value of *X*.

Exhibit 51 Scatter Plots Showing Various Degrees of Correlation

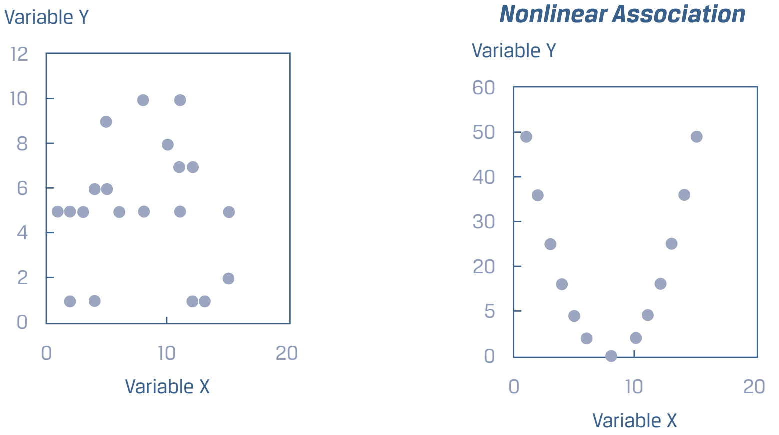
A. Variables With a Correlation of +1

B. Variables With a Correlation of −1



C. Variables With a Correlation of 0

D. Variables With a Strong Nonlinear Association



Panel C shows a scatter plot of two variables with a correlation of 0; they have no linear relation. This graph shows that the value of variable *X* tells us nothing about the value of variable *Y*. Panel D shows a scatter plot of two variables that have a non-linear relationship. Because the correlation coefficient is a measure of the linear association between two variables, it would not be appropriate to use the correlation coefficient in this case.

Example 22 is meant to reinforce your understanding of how to interpret covariance and correlation.

EXAMPLE 22

Interpreting the Correlation Coefficient

Consider the statistics for the returns over twelve months for three funds, A, B, and C, shown in Exhibit 52.

Exhibit 52

	Fund A	Fund B	Fund C
Arithmetic average	2.9333	3.2250	2.6250
Standard deviation	2.4945	2.4091	3.6668

The covariances are represented in the upper-triangle (shaded area) of the matrix shown in Exhibit 53.

Exhibit 53

	Fund A	Fund B	Fund C
Fund A	6.2224	5.7318	−3.6682
Fund B		5.8039	−2.3125
Fund C			13.4457

The covariance of Fund A and Fund B returns, for example, is 5.7318.

Why show just the upper-triangle of this matrix? Because the covariance of Fund A and Fund B returns is the same as the covariance of Fund B and Fund A returns.

The diagonal of the matrix in Exhibit 53 is the variance of each fund's return. For example, the variance of Fund A returns is 6.2224, but the covariance of Fund A and Fund B returns is 5.7138.

The correlations among the funds' returns are given in Exhibit 54, where the correlations are reported in the upper-triangle (shaded area) of the matrix. Note that the correlation of a fund's returns with itself is +1, so the diagonal in the correlation matrix consists of 1.000.

Exhibit 54

	Fund A	Fund B	Fund C
Fund A	1.0000	0.9538	−0.4010
Fund B		1.0000	−0.2618
Fund C			1.0000

1 Interpret the correlation between Fund A's returns and Fund B's returns.

2 Interpret the correlation between Fund A's returns and Fund C's returns.

3 Describe the relationship of the covariance of these returns and the correlation of returns.

Solutions

1 The correlation of Fund A and Fund B returns is 0.9538, which is positive and close to 1.0. This means that when returns of Fund A tend to be above their mean, Fund B's returns also tend to be above their mean. Graphically, we would observe a positive, but not perfect, linear relationship between the returns for the two funds.

2 The correlation of Fund A's returns and Fund C's returns is −0.4010, which indicates that when Fund A's returns are above their mean, Fund B's returns tend to be below their mean. This implies a negative slope when graphing the returns of these two funds, but it would not be a perfect inverse relationship.

3 There are two negative correlations: Fund A returns with Fund C returns, and Fund B returns with Fund C returns. What determines the sign of the correlation is the sign of the covariance, which in each of these cases is negative. When the covariance between fund returns is positive, such as between Fund A and Fund B returns, the correlation is positive. This follows from the fact that the correlation coefficient is the ratio of the covariance of the two funds' returns to the product of their standard deviations.

12.2 Limitations of Correlation Analysis

Exhibit 51 illustrates that correlation measures the linear association between two variables, but it may not always be reliable. Two variables can have a strong *nonlinear* relation and still have a very low correlation. For example, the relation $Y = (X - 4)^2$ is a nonlinear relation contrasted to the linear relation $Y = 2X - 4$. The nonlinear relation between variables X and Y is shown in Panel D. Below a level of 4 for X, Y increases with decreasing values of X. When X is 4 or greater, however, Y increases whenever X increases. Even though these two variables are perfectly associated, there is no linear association between them (hence, no meaningful correlation).

Correlation may also be an unreliable measure when outliers are present in one or both of the variables. As we have seen, outliers are small numbers of observations at either extreme (small or large) of a sample. The correlation may be quite sensitive to outliers. In such a situation, we should consider whether it makes sense to exclude those outlier observations and whether they are noise or news. As a general rule, we must determine whether a computed sample correlation changes greatly by removing outliers. We must also use judgment to determine whether those outliers contain information about the two variables' relationship (and should thus be included in the correlation analysis) or contain no information (and should thus be excluded). If they are to be excluded from the correlation analysis, as we have seen previously, outlier observations can be handled by trimming or winsorizing the dataset.

Importantly, keep in mind that correlation does not imply causation. Even if two variables are highly correlated, one does not necessarily cause the other in the sense that certain values of one variable bring about the occurrence of certain values of the other.

Moreover, with visualizations too, including scatter plots, we must be on guard against unconsciously making judgments about causal relationships that may or may not be supported by the data.

The term **spurious correlation** has been used to refer to: 1) correlation between two variables that reflects chance relationships in a particular dataset; 2) correlation induced by a calculation that mixes each of two variables with a third variable; and 3) correlation between two variables arising not from a direct relation between them but from their relation to a third variable.

As an example of the chance relationship, consider the monthly US retail sales of beer, wine, and liquor and the atmospheric carbon dioxide levels from 2000–2018. The correlation is 0.824, indicating that there is a positive relation between the two. However, there is no reason to suspect that the levels of atmospheric carbon dioxide are related to the retail sales of beer, wine, and liquor.

As an example of the second kind of spurious correlation, two variables that are uncorrelated may be correlated if divided by a third variable. For example, consider a cross-sectional sample of companies' dividends and total assets. While there may be a low correlation between these two variables, dividing each by market capitalization may increase the correlation.

As an example of the third kind of spurious correlation, height may be positively correlated with the extent of a person's vocabulary, but the underlying relationships are between age and height and between age and vocabulary.

Investment professionals must be cautious in basing investment strategies on high correlations. Spurious correlations may suggest investment strategies that appear profitable but actually would not be, if implemented.

A further issue is that correlation does not tell the whole story about the data. Consider Anscombe's Quartet, discussed in Exhibit 55, where very dissimilar graphs can be developed with variables that have the same mean, same standard deviation, and same correlation.

Exhibit 55 Anscombe's Quartet

Francis Anscombe, a British statistician, developed datasets that illustrate why just looking at summary statistics (that is, mean, standard deviation, and correlation) does not fully describe the data. He created four datasets (designated I, II, III, and IV), each with two variables, X and Y, such that:

- The Xs in each dataset have the same mean and standard deviation, 9.00 and 3.32, respectively.
- The Ys in each dataset have the same mean and standard deviation, 7.50 and 2.03, respectively.
- The Xs and Ys in each dataset have the same correlation of 0.82.

(continued)

Exhibit 55 (Continued)

	I		II		III		IV	
Observation	***X***	***Y***	***X***	***Y***	***X***	***Y***	***X***	***Y***
1	10	8.04	10	9.14	10	7.46	8	6.6
2	8	6.95	8	8.14	8	6.77	8	5.8
3	13	7.58	13	8.74	13	12.74	8	7.7
4	9	8.81	9	8.77	9	7.11	8	8.8
5	11	8.33	11	9.26	11	7.81	8	8.5
6	14	9.96	14	8.1	14	8.84	8	7
7	6	7.24	6	6.13	6	6.08	8	5.3
8	4	4.26	4	3.1	4	5.39	19	13
9	12	10.8	12	9.13	12	8.15	8	5.6
10	7	4.82	7	7.26	7	6.42	8	7.9
11	5	5.68	5	4.74	5	5.73	8	6.9
N	11	11	11	11	11	11	11	11
Mean	9.00	7.50	9.00	7.50	9.00	7.50	9.00	7.50
Standard deviation	3.32	2.03	3.32	2.03	3.32	2.03	3.32	2.03
Correlation	0.82		0.82		0.82		0.82	

While the X variable has the same values for I, II, and III in the quartet of datasets, the Y variables are quite different, creating different relationships. The four datasets are:

I An approximate linear relationship between X and Y.

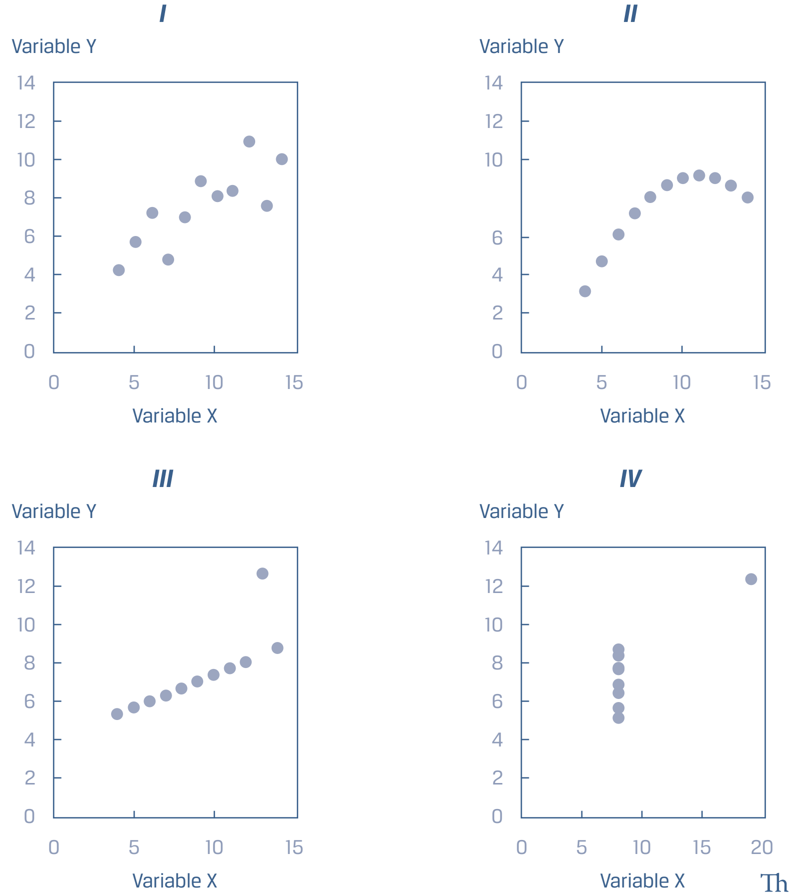
II A curvilinear relationship between X and Y.

III A linear relationship except for one outlier.

IV A constant X with the exception of one outlier.

Depicting the quartet visually,

Exhibit 55 (Continued)



The bottom line? Knowing the means and standard deviations of the two variables, as well as the correlation between them, does not tell the entire story.

Source: Francis John Anscombe, "Graphs in Statistical Analysis," *The American Statistician* 27 (February 1973): 17–21.

SUMMARY

In this reading, we have presented tools and techniques for organizing, visualizing, and describing data that permit us to convert raw data into useful information for investment analysis.

- Data can be defined as a collection of numbers, characters, words, and text—as well as images, audio, and video—in a raw or organized format to represent facts or information.
- From a statistical perspective, data can be classified as numerical data and categorical data. Numerical data (also called quantitative data) are values that represent measured or counted quantities as a number. Categorical data (also called qualitative data) are values that describe a quality or characteristic of a group of observations and usually take only a limited number of values that are mutually exclusive.

- Numerical data can be further split into two types: continuous data and discrete data. Continuous data can be measured and can take on any numerical value in a specified range of values. Discrete data are numerical values that result from a counting process and therefore are limited to a finite number of values.
- Categorical data can be further classified into two types: nominal data and ordinal data. Nominal data are categorical values that are not amenable to being organized in a logical order, while ordinal data are categorical values that can be logically ordered or ranked.
- Based on how they are collected, data can be categorized into three types: cross-sectional, time series, and panel. Time-series data are a sequence of observations for a single observational unit on a specific variable collected over time and at discrete and typically equally spaced intervals of time. Cross-sectional data are a list of the observations of a specific variable from multiple observational units at a given point in time. Panel data are a mix of time-series and cross-sectional data that consists of observations through time on one or more variables for multiple observational units.
- Based on whether or not data are in a highly organized form, they can be classified into structured and unstructured types. Structured data are highly organized in a pre-defined manner, usually with repeating patterns. Unstructured data do not follow any conventionally organized forms; they are typically alternative data as they are usually collected from unconventional sources.
- Raw data are typically organized into either a one-dimensional array or a two-dimensional rectangular array (also called a data table) for quantitative analysis.
- A frequency distribution is a tabular display of data constructed either by counting the observations of a variable by distinct values or groups or by tallying the values of a numerical variable into a set of numerically ordered bins. Frequency distributions permit us to evaluate how data are distributed.
- The relative frequency of observations in a bin (interval or bucket) is the number of observations in the bin divided by the total number of observations. The cumulative relative frequency cumulates (adds up) the relative frequencies as we move from the first bin to the last, thus giving the fraction of the observations that are less than the upper limit of each bin.
- A contingency table is a tabular format that displays the frequency distributions of two or more categorical variables simultaneously. One application of contingency tables is for evaluating the performance of a classification model (using a confusion matrix). Another application of contingency tables is to investigate a potential association between two categorical variables by performing a chi-square test of independence.
- Visualization is the presentation of data in a pictorial or graphical format for the purpose of increasing understanding and for gaining insights into the data.
- A histogram is a bar chart of data that have been grouped into a frequency distribution. A frequency polygon is a graph of frequency distributions obtained by drawing straight lines joining successive midpoints of bars representing the class frequencies.
- A bar chart is used to plot the frequency distribution of categorical data, with each bar representing a distinct category and the bar's height (or length) proportional to the frequency of the corresponding category. Grouped bar charts or stacked bar charts can present the frequency distribution of multiple categorical variables simultaneously.

- A tree-map is a graphical tool to display categorical data. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group. Additional dimensions of categorical data can be displayed by nested rectangles.
- A word cloud is a visual device for representing textual data, with the size of each distinct word being proportional to the frequency with which it appears in the given text.
- A line chart is a type of graph used to visualize ordered observations and often to display the change of data series over time. A bubble line chart is a special type of line chart that uses varying-sized bubbles as data points to represent an additional dimension of data.
- A scatter plot is a type of graph for visualizing the joint variation in two numerical variables. It is constructed by drawing dots to indicate the values of the two variables plotted against the corresponding axes. A scatter plot matrix organizes scatter plots between pairs of variables into a matrix format to inspect all pairwise relationships between more than two variables in one combined visual.
- A heat map is a type of graphic that organizes and summarizes data in a tabular format and represents it using a color spectrum. It is often used in displaying frequency distributions or visualizing the degree of correlation among different variables.
- The key consideration when selecting among chart types is the intended purpose of visualizing data (i.e., whether it is for exploring/presenting distributions or relationships or for making comparisons).
- A population is defined as all members of a specified group. A sample is a subset of a population.
- A parameter is any descriptive measure of a population. A sample statistic (statistic, for short) is a quantity computed from or used to describe a sample.
- Sample statistics—such as measures of central tendency, measures of dispersion, skewness, and kurtosis—help with investment analysis, particularly in making probabilistic statements about returns.
- Measures of central tendency specify where data are centered and include the mean, median, and mode (i.e., the most frequently occurring value).
- The arithmetic mean is the sum of the observations divided by the number of observations. It is the most frequently used measure of central tendency.
- The median is the value of the middle item (or the mean of the values of the two middle items) when the items in a set are sorted into ascending or descending order. The median is not influenced by extreme values and is most useful in the case of skewed distributions.
- The mode is the most frequently observed value and is the only measure of central tendency that can be used with nominal data. A distribution may be unimodal (one mode), bimodal (two modes), trimodal (three modes), or have even more modes.
- A portfolio's return is a weighted mean return computed from the returns on the individual assets, where the weight applied to each asset's return is the fraction of the portfolio invested in that asset.
- The geometric mean, $\bar{X}_G$, of a set of observations $X_1, X_2, \ldots, X_n$, is $\bar{X}_G = \sqrt[n]{X_1 X_2 X_3 \ldots X_n}$, with $X_i \geq 0$ for $i = 1, 2, \ldots, n$. The geometric mean is especially important in reporting compound growth rates for time-series data. The geometric mean will always be less than an arithmetic mean whenever there is variance in the observations.

- The harmonic mean, $\bar{X}_H$, is a type of weighted mean in which an observation's weight is inversely proportional to its magnitude.
- Quantiles—such as the median, quartiles, quintiles, deciles, and percentiles—are location parameters that divide a distribution into halves, quarters, fifths, tenths, and hundredths, respectively.
- A box and whiskers plot illustrates the interquartile range (the "box") as well as a range outside of the box that is based on the interquartile range, indicated by the "whiskers."
- Dispersion measures—such as the range, mean absolute deviation (MAD), variance, standard deviation, target downside deviation, and coefficient of variation—describe the variability of outcomes around the arithmetic mean.
- The range is the difference between the maximum value and the minimum value of the dataset. The range has only a limited usefulness because it uses information from only two observations.
- The MAD for a sample is the average of the absolute deviations of observations from the mean, $\frac{\sum_{i=1}^{n}\left|X_i - \bar{X}\right|}{n}$, where $\bar{X}$ is the sample mean and n is the number of observations in the sample.
- The variance is the average of the squared deviations around the mean, and the standard deviation is the positive square root of variance. In computing sample variance (s^2) and sample standard deviation (s), the average squared deviation is computed using a divisor equal to the sample size minus 1.
- The target downside deviation, or target semideviation, is a measure of the risk of being below a given target. It is calculated as the square root of the average squared deviations from the target, but it includes only those observations below the target (B), or $\sqrt{\sum_{\text{for all} X_i \le B}^{n} \frac{\left(X_i - B\right)^2}{n - 1}}$.
- The coefficient of variation, CV, is the ratio of the standard deviation of a set of observations to their mean value. By expressing the magnitude of variation among observations relative to their average size, the CV permits direct comparisons of dispersion across different datasets. Reflecting the correction for scale, the CV is a scale-free measure (i.e., it has no units of measurement).
- Skew or skewness describes the degree to which a distribution is asymmetric about its mean. A return distribution with positive skewness has frequent small losses and a few extreme gains compared to a normal distribution. A return distribution with negative skewness has frequent small gains and a few extreme losses compared to a normal distribution. Zero skewness indicates a symmetric distribution of returns.
- Kurtosis measures the combined weight of the tails of a distribution relative to the rest of the distribution. A distribution with fatter tails than the normal distribution is referred to as fat-tailed (leptokurtic); a distribution with thinner tails than the normal distribution is referred to as thin-tailed (platykurtic). Excess kurtosis is kurtosis minus 3, since 3 is the value of kurtosis for all normal distributions.
- The correlation coefficient is a statistic that measures the association between two variables. It is the ratio of covariance to the product of the two variables' standard deviations. A positive correlation coefficient indicates that the two variables tend to move together, whereas a negative coefficient indicates that

the two variables tend to move in opposite directions. Correlation does not imply causation, simply association. Issues that arise in evaluating correlation include the presence of outliers and spurious correlation.

PRACTICE PROBLEMS

1 Published ratings on stocks ranging from 1 (strong sell) to 5 (strong buy) are examples of which measurement scale?

- **A** Ordinal
- **B** Continuous
- **C** Nominal

2 Data values that are categorical and not amenable to being organized in a logical order are *most likely* to be characterized as:

- **A** ordinal data.
- **B** discrete data.
- **C** nominal data.

3 Which of the following data types would be classified as being categorical?

- **A** Discrete
- **B** Nominal
- **C** Continuous

4 A fixed-income analyst uses a proprietary model to estimate bankruptcy probabilities for a group of firms. The model generates probabilities that can take any value between 0 and 1. The resulting set of estimated probabilities would *most likely* be characterized as:

- **A** ordinal data.
- **B** discrete data.
- **C** continuous data.

5 An analyst uses a software program to analyze unstructured data—specifically, management's earnings call transcript for one of the companies in her research coverage. The program scans the words in each sentence of the transcript and then classifies the sentences as having negative, neutral, or positive sentiment. The resulting set of sentiment data would *most likely* be characterized as:

- **A** ordinal data.
- **B** discrete data.
- **C** nominal data.

Use the following information to answer Questions 6 and 7

An equity analyst gathers total returns for three country equity indexes over the past four years. The data are presented below.

© 2020 CFA Institute. All rights reserved.

Time Period	Index A	Index B	Index C
Year *t*–3	15.56%	11.84%	–4.34%
Year t–2	–4.12%	–6.96%	9.32%
Year *t*–1	11.19%	10.29%	–12.72%
Year *t*	8.98%	6.32%	21.44%

6 Each individual column of data in the table can be *best* characterized as:

A panel data.

B time-series data.

C cross-sectional data.

7 Each individual row of data in the table can be *best* characterized as:

A panel data.

B time-series data.

C cross-sectional data.

---

8 A two-dimensional rectangular array would be most suitable for organizing a collection of raw:

A panel data.

B time-series data.

C cross-sectional data.

9 In a frequency distribution, the absolute frequency measure:

A represents the percentages of each unique value of the variable.

B represents the actual number of observations counted for each unique value of the variable.

C allows for comparisons between datasets with different numbers of total observations.

10 An investment fund has the return frequency distribution shown in the following exhibit.

Return Interval (%)	Absolute Frequency
–10.0 to –7.0	3
–7.0 to –4.0	7
–4.0 to –1.0	10
–1.0 to +2.0	12
+2.0 to +5.0	23
+5.0 to +8.0	5

Which of the following statements is correct?

A The relative frequency of the bin "–1.0 to +2.0" is 20%.

B The relative frequency of the bin "+2.0 to +5.0" is 23%.

C The cumulative relative frequency of the bin "+5.0 to +8.0" is 91.7%.

11 An analyst is using the data in the following exhibit to prepare a statistical report.

Portfolio's Deviations from Benchmark Return for a 12-Year Period (%)			
Year 1	2.48	Year 7	–9.19
Year 2	–2.59	Year 8	–5.11
Year 3	9.47	Year 9	1.33
Year 4	–0.55	Year 10	6.84
Year 5	–1.69	Year 11	3.04
Year 6	–0.89	Year 12	4.72

The cumulative relative frequency for the bin $-1.71\% \le x < 2.03\%$ is *closest* to:

- **A** 0.250.
- **B** 0.333.
- **C** 0.583.

Use the following information to answer questions 12 and 13

A fixed-income portfolio manager creates a contingency table of the number of bonds held in her portfolio by sector and bond rating. The contingency table is presented here:

	Bond Rating		
Sector	**A**	**AA**	**AAA**
Communication Services	25	32	27
Consumer Staples	30	25	25
Energy	100	85	30
Health Care	200	100	63
Utilities	22	28	14

12 The marginal frequency of energy sector bonds is *closest* to:

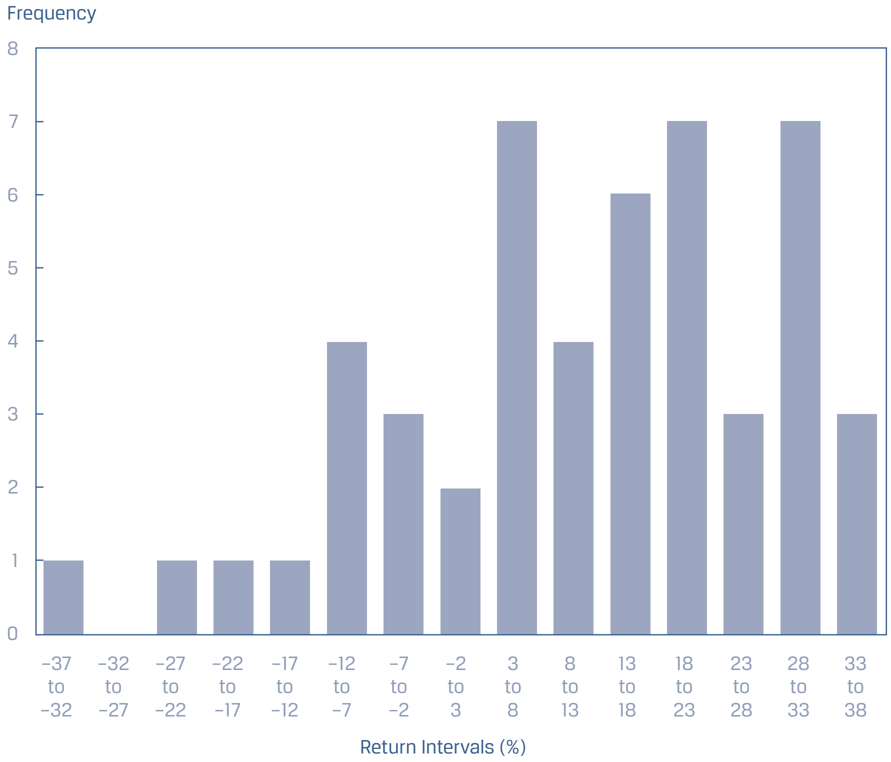
- **A** 27.
- **B** 85.
- **C** 215.

13 The relative frequency of AA rated energy bonds, based on the total count, is *closest* to:

- **A** 10.5%.
- **B** 31.5%.
- **C** 39.5%.

The following information relates to Questions 14–15

The following histogram shows a distribution of the S&P 500 Index annual returns for a 50-year period:



14 The bin containing the median return is:

A 3% to 8%.

B 8% to 13%.

C 13% to 18%.

15 Based on the previous histogram, the distribution is *best* described as being:

A unimodal.

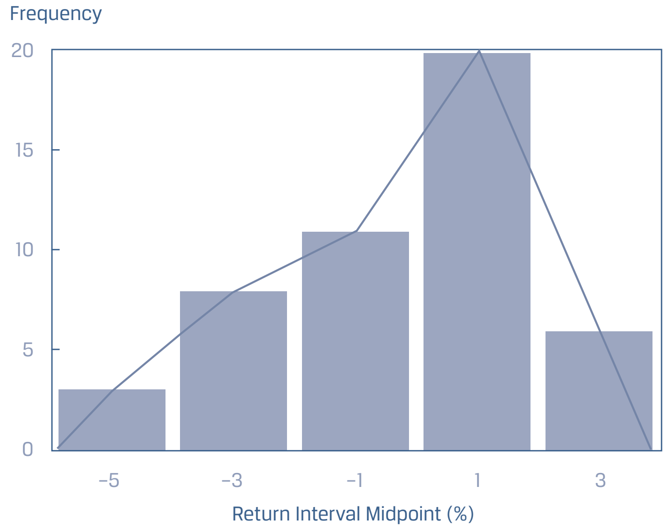
B bimodal.

C trimodal.

---

16 The following is a frequency polygon of monthly exchange rate changes in the US dollar/Japanese yen spot exchange rate for a four-year period. A positive change represents yen appreciation (the yen buys more dollars), and a negative change represents yen depreciation (the yen buys fewer dollars).

Monthly Changes in the US Dollar/Japanese Yen Spot Exchange Rate



Based on the chart, yen appreciation:

A occurred more than 50% of the time.

B was less frequent than yen depreciation.

C in the 0.0 to 2.0 interval occurred 20% of the time.

17 A bar chart that orders categories by frequency in descending order and includes a line displaying cumulative relative frequency is referred to as a:

A Pareto Chart.

B grouped bar chart.

C frequency polygon.

18 Which visualization tool works *best* to represent unstructured, textual data?

A Tree-Map

B Scatter plot

C Word cloud

19 A tree-map is best suited to illustrate:

A underlying trends over time.

B joint variations in two variables.

C value differences of categorical groups.

20 A line chart with two variables—for example, revenues and earnings per share—is best suited for visualizing:

A the joint variation in the variables.

B underlying trends in the variables over time.

C the degree of correlation between the variables.

21 A heat map is best suited for visualizing the:

A frequency of textual data.

B degree of correlation between different variables.

C shape, center, and spread of the distribution of numerical data.

22 Which valuation tool is recommended to be used if the goal is to make comparisons of three or more variables over time?

A Heat map

B Bubble line chart

C Scatter plot matrix

23 The annual returns for three portfolios are shown in the following exhibit. Portfolios P and R were created in Year 1, Portfolio Q in Year 2.

	Annual Portfolio Returns (%)				
	Year 1	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Portfolio P	−3.0	4.0	5.0	3.0	7.0
Portfolio Q		−3.0	6.0	4.0	8.0
Portfolio R	1.0	−1.0	4.0	4.0	3.0

The median annual return from portfolio creation to Year 5 for:

A Portfolio P is 4.5%.

B Portfolio Q is 4.0%.

C Portfolio R is higher than its arithmetic mean annual return.

24 At the beginning of Year X, an investor allocated his retirement savings in the asset classes shown in the following exhibit and earned a return for Year X as also shown.

Asset Class	**Asset Allocation (%)**	**Asset Class Return for Year X (%)**
Large-cap US equities	20.0	8.0
Small-cap US equities	40.0	12.0
Emerging market equities	25.0	−3.0
High-yield bonds	15.0	4.0

The portfolio return for Year X is *closest to*:

A 5.1%.

B 5.3%.

C 6.3%.

25 The following exhibit shows the annual returns for Fund Y.

	Fund Y (%)
Year 1	19.5
Year 2	−1.9
Year 3	19.7
Year 4	35.0
Year 5	5.7

The geometric mean return for Fund Y is *closest* to:

A 14.9%.

B 15.6%.

C 19.5%.

26 A portfolio manager invests €5,000 annually in a security for four years at the prices shown in the following exhibit.

	Purchase Price of Security (€ per unit)
Year 1	62.00
Year 2	76.00
Year 3	84.00
Year 4	90.00

The average price is *best* represented as the:

A harmonic mean of €76.48.

B geometric mean of €77.26.

C arithmetic average of €78.00.

The following information relates to Questions 27–28

The following exhibit shows the annual MSCI World Index total returns for a 10-year period.

Year 1	15.25%	Year 6	30.79%
Year 2	10.02%	Year 7	12.34%
Year 3	20.65%	Year 8	–5.02%
Year 4	9.57%	Year 9	16.54%
Year 5	–40.33%	Year 10	27.37%

27 The fourth quintile return for the MSCI World Index is *closest* to:

A 20.65%.

B 26.03%.

C 27.37%.

28 For Year 6–Year 10, the mean absolute deviation of the MSCI World Index total returns is *closest* to:

A 10.20%.

B 12.74%.

C 16.40%.

---

29 Annual returns and summary statistics for three funds are listed in the following exhibit:

	Annual Returns (%)		
Year	**Fund ABC**	**Fund XYZ**	**Fund PQR**
Year 1	–20.0	–33.0	–14.0
Year 2	23.0	–12.0	–18.0
Year 3	–14.0	–12.0	6.0
Year 4	5.0	–8.0	–2.0
Year 5	–14.0	11.0	3.0

Year	Annual Returns (%) Fund ABC	Fund XYZ	Fund PQR
Mean	−4.0	−10.8	−5.0
Standard deviation	17.8	15.6	10.5

The fund with the highest absolute dispersion is:

A Fund PQR if the measure of dispersion is the range.

B Fund XYZ if the measure of dispersion is the variance.

C Fund ABC if the measure of dispersion is the mean absolute deviation.

30 The mean monthly return and the standard deviation for three industry sectors are shown in the following exhibit.

Sector	Mean Monthly Return (%)	Standard Deviation of Return (%)
Utilities (UTIL)	2.10	1.23
Materials (MATR)	1.25	1.35
Industrials (INDU)	3.01	1.52

Based on the coefficient of variation, the riskiest sector is:

A utilities.

B materials.

C industrials.

31 The average return for Portfolio A over the past twelve months is 3%, with a standard deviation of 4%. The average return for Portfolio B over this same period is also 3%, but with a standard deviation of 6%. The geometric mean return of Portfolio A is 2.85%. The geometric mean return of Portfolio B is:

A less than 2.85%.

B equal to 2.85%.

C greater than 2.85%.

32 An analyst calculated the excess kurtosis of a stock's returns as −0.75. From this information, we conclude that the distribution of returns is:

A normally distributed.

B thin-tailed compared to the normal distribution.

C fat-tailed compared to the normal distribution.

33 When analyzing investment returns, which of the following statements is correct?

A The geometric mean will exceed the arithmetic mean for a series with non-zero variance.

B The geometric mean measures an investment's compound rate of growth over multiple periods.

C The arithmetic mean measures an investment's terminal value over multiple periods.

The following information relates to Questions 34–38

A fund had the following experience over the past 10 years:

Year	Return
1	4.5%
2	6.0%
3	1.5%
4	−2.0%
5	0.0%
6	4.5%
7	3.5%
8	2.5%
9	5.5%
10	4.0%

34 The arithmetic mean return over the 10 years is *closest* to:

A 2.97%.

B 3.00%.

C 3.33%.

35 The geometric mean return over the 10 years is *closest* to:

A 2.94%.

B 2.97%.

C 3.00%.

36 The harmonic mean return over the 10 years is *closest* to:

A 2.94%.

B 2.97%.

C 3.00%.

37 The standard deviation of the 10 years of returns is *closest* to:

A 2.40%.

B 2.53%.

C 7.58%.

38 The target semideviation of the returns over the 10 years if the target is 2% is *closest* to:

A 1.42%.

B 1.50%.

C 2.01%.

---

39 A correlation of 0.34 between two variables, X and Y, is *best* described as:

A changes in X causing changes in Y.

B a positive association between X and Y.

C a curvilinear relationship between X and Y.

40 Which of the following is a potential problem with interpreting a correlation coefficient?

A Outliers

B Spurious correlation

C Both outliers and spurious correlation

The following relates to questions 41 and 42

An analyst is evaluating the tendency of returns on the portfolio of stocks she manages to move along with bond and real estate indexes. She gathered monthly data on returns and the indexes:

	Returns (%)		
	Portfolio Returns	Bond Index Returns	Real Estate Index Returns
Arithmetic average	5.5	3.2	7.8
Standard deviation	8.2	3.4	10.3

	Portfolio Returns and Bond Index Returns	Portfolio Returns and Real Estate Index Returns
Covariance	**18.9**	**–55.9**

41 Without calculating the correlation coefficient, the correlation of the portfolio returns and the bond index returns is:

A negative.

B zero.

C positive.

42 Without calculating the correlation coefficient, the correlation of the portfolio returns and the real estate index returns is:

A negative.

B zero.

C positive.

---

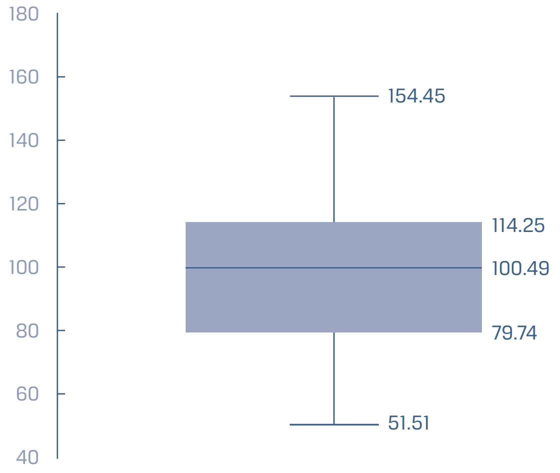
43 Consider two variables, A and B. If variable A has a mean of –0.56, variable B has a mean of 0.23, and the covariance between the two variables is positive, the correlation between these two variables is:

A negative.

B zero.

C positive.

The following information relates to Questions 44–45



44 The median is *closest* to:

- **A** 34.51.
- **B** 100.49.
- **C** 102.98.

45 The interquartile range is *closest* to:

- **A** 13.76.
- **B** 25.74.
- **C** 34.51.

The following information relates to Questions 46–48

An analyst examined a cross-section of annual returns for 252 stocks and calculated the following statistics:

Arithmetic Average	9.986%
Geometric Mean	9.909%
Variance	0.001723
Skewness	0.704
Excess Kurtosis	0.503

46 The coefficient of variation is closest to:

- **A** 0.02.
- **B** 0.42.
- **C** 2.41.

47 This distribution is best described as:

- **A** negatively skewed.
- **B** having no skewness.
- **C** positively skewed.

48 Compared to the normal distribution, this sample's distribution is best described as having tails of the distribution with:

- **A** less probability than the normal distribution.
- **B** the same probability as the normal distribution.
- **C** more probability than the normal distribution.

SOLUTIONS

1 A is correct. Ordinal scales sort data into categories that are ordered with respect to some characteristic and may involve numbers to identify categories but do not assure that the differences between scale values are equal. The buy rating scale indicates that a stock ranked 5 is expected to perform better than a stock ranked 4, but it tells us nothing about the performance difference between stocks ranked 4 and 5 compared with the performance difference between stocks ranked 1 and 2, and so on.

2 C is correct. Nominal data are categorical values that are not amenable to being organized in a logical order. A is incorrect because ordinal data are categorical data that can be logically ordered or ranked. B is incorrect because discrete data are numerical values that result from a counting process; thus, they can be ordered in various ways, such as from highest to lowest value.

3 B is correct. Categorical data (or qualitative data) are values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize. The two types of categorical data are nominal data and ordinal data. Nominal data are categorical values that are not amenable to being organized in a logical order, while ordinal data are categorical values that can be logically ordered or ranked. A is incorrect because discrete data would be classified as numerical data (not categorical data). C is incorrect because continuous data would be classified as numerical data (not categorical data).

4 C is correct. Continuous data are data that can be measured and can take on any numerical value in a specified range of values. In this case, the analyst is estimating bankruptcy probabilities, which can take on any value between 0 and 1. Therefore, the set of bankruptcy probabilities estimated by the analyst would likely be characterized as continuous data. A is incorrect because ordinal data are categorical values that can be logically ordered or ranked. Therefore, the set of bankruptcy probabilities would not be characterized as ordinal data. B is incorrect because discrete data are numerical values that result from a counting process, and therefore the data are limited to a finite number of values. The proprietary model used can generate probabilities that can take any value between 0 and 1; therefore, the set of bankruptcy probabilities would not be characterized as discrete data.

5 A is correct. Ordinal data are categorical values that can be logically ordered or ranked. In this case, the classification of sentences in the earnings call transcript into three categories (negative, neutral, or positive) describes ordinal data, as the data can be logically ordered from positive to negative. B is incorrect because discrete data are numerical values that result from a counting process. In this case, the analyst is categorizing sentences (i.e., unstructured data) from the earnings call transcript as having negative, neutral, or positive sentiment. Thus, these categorical data do not represent discrete data. C is incorrect because nominal data are categorical values that are not amenable to being organized in a logical order. In this case, the classification of unstructured data (i.e., sentences from the earnings call transcript) into three categories (negative, neutral, or positive) describes ordinal (not nominal) data, as the data can be logically ordered from positive to negative.

6 B is correct. Time-series data are a sequence of observations of a specific variable collected over time and at discrete and typically equally spaced intervals of time, such as daily, weekly, monthly, annually, and quarterly. In this case, each

column is a time series of data that represents annual total return (the specific variable) for a given country index, and it is measured annually (the discrete interval of time). A is incorrect because panel data consist of observations through time on one or more variables for multiple observational units. The entire table of data is an example of panel data showing annual total returns (the variable) for three country indexes (the observational units) by year. C is incorrect because cross-sectional data are a list of the observations of a specific variable from multiple observational units at a given point in time. Each row (not column) of data in the table represents cross-sectional data.

7 C is correct. Cross-sectional data are observations of a specific variable from multiple observational units at a given point in time. Each row of data in the table represents cross-sectional data. The specific variable is annual total return, the multiple observational units are the three countries' indexes, and the given point in time is the time period indicated by the particular row. A is incorrect because panel data consist of observations through time on one or more variables for multiple observational units. The entire table of data is an example of panel data showing annual total returns (the variable) for three country indexes (the observational units) by year. B is incorrect because time-series data are a sequence of observations of a specific variable collected over time and at discrete and typically equally spaced intervals of time, such as daily, weekly, monthly, annually, and quarterly. In this case, each column (not row) is a time series of data that represents annual total return (the specific variable) for a given country index, and it is measured annually (the discrete interval of time).

8 A is correct. Panel data consist of observations through time on one or more variables for multiple observational units. A two-dimensional rectangular array, or data table, would be suitable here as it is comprised of columns to hold the variable(s) for the observational units and rows to hold the observations through time. B is incorrect because a one-dimensional (not a two-dimensional rectangular) array would be most suitable for organizing a collection of data of the same data type, such as the time-series data from a single variable. C is incorrect because a one-dimensional (not a two-dimensional rectangular) array would be most suitable for organizing a collection of data of the same data type, such as the same variable for multiple observational units at a given point in time (cross-sectional data).

9 B is correct. In a frequency distribution, the absolute frequency, or simply the raw frequency, is the actual number of observations counted for each unique value of the variable. A is incorrect because the relative frequency, which is calculated as the absolute frequency of each unique value of the variable divided by the total number of observations, presents the absolute frequencies in terms of percentages. C is incorrect because the relative (not absolute) frequency provides a normalized measure of the distribution of the data, allowing comparisons between datasets with different numbers of total observations.

10 A is correct. The relative frequency is the absolute frequency of each bin divided by the total number of observations. Here, the relative frequency is calculated as: (12/60) × 100 = 20%. B is incorrect because the relative frequency of this bin is (23/60) × 100 = 38.33%. C is incorrect because the cumulative relative frequency of the last bin must equal 100%.

11 C is correct. The cumulative relative frequency of a bin identifies the fraction of observations that are less than the upper limit of the given bin. It is determined by summing the relative frequencies from the lowest bin up to and including the given bin. The following exhibit shows the relative frequencies for all the bins of the data from the previous exhibit:

Lower Limit (%)	Upper Limit (%)	Absolute Frequency	Relative Frequency	Cumulative Relative Frequency
−9.19 ≤	< −5.45	1	0.083	0.083
−5.45 ≤	< −1.71	2	0.167	0.250
−1.71 ≤	< 2.03	4	0.333	0.583
2.03 ≤	< 5.77	3	0.250	0.833
5.77 ≤	≤ 9.47	2	0.167	1.000

The bin −1.71% ≤ x < 2.03% has a cumulative relative frequency of 0.583.

12 C is correct. The marginal frequency of energy sector bonds in the portfolio is the sum of the joint frequencies across all three levels of bond rating, so 100 + 85 + 30 = 215. A is incorrect because 27 is the relative frequency for energy sector bonds based on the total count of 806 bonds, so 215/806 = 26.7%, not the marginal frequency. B is incorrect because 85 is the joint frequency for AA rated energy sector bonds, not the marginal frequency.

13 A is correct. The relative frequency for any value in the table based on the total count is calculated by dividing that value by the total count. Therefore, the relative frequency for AA rated energy bonds is calculated as 85/806 = 10.5%.

B is incorrect because 31.5% is the relative frequency for AA rated energy bonds, calculated based on the marginal frequency for all AA rated bonds, so 85/(32 + 25 + 85 + 100 + 28), not based on total bond counts. C is incorrect because 39.5% is the relative frequency for AA rated energy bonds, calculated based on the marginal frequency for all energy bonds, so 85/(100 + 85 + 30), not based on total bond counts.

14 C is correct. Because 50 data points are in the histogram, the median return would be the mean of the 50/2 = 25th and (50 + 2)/2 = 26th positions. The sum of the return bin frequencies to the left of the 13% to 18% interval is 24. As a result, the 25th and 26th returns will fall in the 13% to 18% interval.

15 C is correct. The mode of a distribution with data grouped in intervals is the interval with the highest frequency. The three intervals of 3% to 8%, 18% to 23%, and 28% to 33% all have a high frequency of 7.

16 A is correct. Twenty observations lie in the interval "0.0 to 2.0," and six observations lie in the "2.0 to 4.0" interval. Together, they represent 26/48, or 54.17%, of all observations, which is more than 50%.

17 A is correct. A bar chart that orders categories by frequency in descending order and includes a line displaying cumulative relative frequency is called a Pareto Chart. A Pareto Chart is used to highlight dominant categories or the most important groups. B is incorrect because a grouped bar chart or clustered bar chart is used to present the frequency distribution of two categorical variables. C is incorrect because a frequency polygon is used to display frequency distributions.

18 C is correct. A word cloud, or tag cloud, is a visual device for representing unstructured, textual data. It consists of words extracted from text with the size of each word being proportional to the frequency with which it appears in the given text. A is incorrect because a tree-map is a graphical tool for displaying and comparing categorical data, not for visualizing unstructured, textual data. B is incorrect because a scatter plot is used to visualize the joint variation in two numerical variables, not for visualizing unstructured, textual data.

19 C is correct. A tree-map is a graphical tool used to display and compare categorical data. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group. A is incorrect because a line chart, not a tree-map, is used to display the change in a data series over time. B is incorrect because a scatter plot, not a tree-map, is used to visualize the joint variation in two numerical variables.

20 B is correct. An important benefit of a line chart is that it facilitates showing changes in the data and underlying trends in a clear and concise way. Often a line chart is used to display the changes in data series over time. A is incorrect because a scatter plot, not a line chart, is used to visualize the joint variation in two numerical variables. C is incorrect because a heat map, not a line chart, is used to visualize the values of joint frequencies among categorical variables.

21 B is correct. A heat map is commonly used for visualizing the degree of correlation between different variables. A is incorrect because a word cloud, or tag cloud, not a heat map, is a visual device for representing textual data with the size of each distinct word being proportional to the frequency with which it appears in the given text. C is incorrect because a histogram, not a heat map, depicts the shape, center, and spread of the distribution of numerical data.

22 B is correct. A bubble line chart is a version of a line chart where data points are replaced with varying-sized bubbles to represent a third dimension of the data. A line chart is very effective at visualizing trends in three or more variables over time. A is incorrect because a heat map differentiates high values from low values and reflects the correlation between variables but does not help in making comparisons of variables over time. C is incorrect because a scatterplot matrix is a useful tool for organizing scatterplots between pairs of variables, making it easy to inspect all pairwise relationships in one combined visual. However, it does not help in making comparisons of these variables over time.

23 C is correct. The median of Portfolio R is 0.8% higher than the mean for Portfolio R.

24 C is correct. The portfolio return must be calculated as the weighted mean return, where the weights are the allocations in each asset class:

$$(0.20 \times 8\%) + (0.40 \times 12\%) + (0.25 \times -3\%) + (0.15 \times 4\%) = 6.25\%, \text{ or } \approx 6.3\%.$$

25 A is correct. The geometric mean return for Fund Y is found as follows:

$$\begin{aligned}\text{Fund Y} &= [(1 + 0.195) \times (1 - 0.019) \times (1 + 0.197) \times (1 + 0.350) \times (1 + 0.057)]^{(1/5)} - 1\\ &= 14.9\%.\end{aligned}$$

26 A is correct. The harmonic mean is appropriate for determining the average price per unit. It is calculated by summing the reciprocals of the prices, then averaging that sum by dividing by the number of prices, then taking the reciprocal of the average:

$$4/[(1/62.00) + (1/76.00) + (1/84.00) + (1/90.00)] = €76.48.$$

27 B is correct. Quintiles divide a distribution into fifths, with the fourth quintile occurring at the point at which 80% of the observations lie below it. The fourth quintile is equivalent to the 80th percentile. To find the yth percentile (P_y), we first must determine its location. The formula for the location (L_y) of a yth percentile in an array with n entries sorted in ascending order is $L_y = (n + 1) \times (y/100)$. In this case, $n = 10$ and $y = 80\%$, so

$$L_{80} = (10 + 1) \times (80/100) = 11 \times 0.8 = 8.8.$$

With the data arranged in ascending order (−40.33%, −5.02%, 9.57%, 10.02%, 12.34%, 15.25%, 16.54%, 20.65%, 27.37%, and 30.79%), the 8.8th position would be between the 8th and 9th entries, 20.65% and 27.37%, respectively. Using linear interpolation, $P_{80} = X_8 + (L_y - 8) \times (X_9 - X_8)$,

$$\begin{aligned} P_{80} &= 20.65 + (8.8 - 8) \times (27.37 - 20.65) \\ &= 20.65 + (0.8 \times 6.72) = 20.65 + 5.38 \\ &= 26.03\%. \end{aligned}$$

28 A is correct. The formula for mean absolute deviation (MAD) is

$$\text{MAD} = \frac{\sum_{i=1}^{n} \left| X_i - \bar{X} \right|}{n}$$

Column 1: Sum annual returns and divide by n to find the arithmetic mean $\left(\bar{X}\right)$ of 16.40%.

Column 2: Calculate the absolute value of the difference between each year's return and the mean from Column 1. Sum the results and divide by n to find the MAD.

These calculations are shown in the following exhibit:

	Column 1		Column 2
Year	**Return**		$\left\|X_i - \bar{X}\right\|$
Year 6	30.79%		14.39%
Year 7	12.34%		4.06%
Year 8	−5.02%		21.42%
Year 9	16.54%		0.14%
Year 10	27.37%		10.97%
Sum:	82.02%	Sum:	50.98%
n:	5	n:	5
$\bar{X}$:	16.40%	MAD:	10.20%

29 C is correct. The mean absolute deviation (MAD) of Fund ABC's returns is greater than the MAD of both of the other funds.

$$\text{MAD} = \frac{\sum_{i=1}^{n} \left| X_i - \bar{X} \right|}{n}, \text{ where } \bar{X} \text{ is the arithmetic mean of the series.}$$

MAD for Fund ABC =

$$\frac{\left|-20-(-4)\right| + \left|23-(-4)\right| + \left|-14-(-4)\right| + \left|5-(-4)\right| + \left|-14-(-4)\right|}{5} = 14.4\%$$

MAD for Fund XYZ =

$$\frac{|-33-(-10.8)|+|-12-(-10.8)|+|-12-(-10.8)|+|-8-(-10.8)|+|11-(-10.8)|}{5}=9.8\%$$

MAD for Fund PQR =

$$\frac{|-14-(-5)|+|-18-(-5)|+|6-(-5)|+|-2-(-5)|+|3-(-5)|}{5}=8.8\%$$

A and B are incorrect because the range and variance of the three funds are as follows:

	Fund ABC	Fund XYZ	Fund PQR
Range	43%	44%	24%
Variance	317	243	110

The numbers shown for variance are understood to be in "percent squared" terms so that when taking the square root, the result is standard deviation in percentage terms. Alternatively, by expressing standard deviation and variance in decimal form, one can avoid the issue of units. In decimal form, the variances for Fund ABC, Fund XYZ, and Fund PQR are 0.0317, 0.0243, and 0.0110, respectively.

30 B is correct. The coefficient of variation (CV) is the ratio of the standard deviation to the mean, where a higher CV implies greater risk per unit of return.

$$CV_{UTIL}=\frac{s}{\overline{X}}=\frac{1.23\%}{2.10\%}=0.59$$

$$CV_{MATR}=\frac{s}{\overline{X}}=\frac{1.35\%}{1.25\%}=1.08$$

$$CV_{INDU}=\frac{s}{\overline{X}}=\frac{1.52\%}{3.01\%}=0.51$$

31 A is correct. The more disperse a distribution, the greater the difference between the arithmetic mean and the geometric mean.

32 B is correct. The distribution is thin-tailed relative to the normal distribution because the excess kurtosis is less than zero.

33 B is correct. The geometric mean compounds the periodic returns of every period, giving the investor a more accurate measure of the terminal value of an investment.

34 B is correct. The sum of the returns is 30.0%, so the arithmetic mean is 30.0%/10 = 3.0%.

35 B is correct.

Year	Return	1+ Return
1	4.5%	1.045
2	6.0%	1.060
3	1.5%	1.015
4	−2.0%	0.980
5	0.0%	1.000
6	4.5%	1.045

(continued)

Year	Return	1+ Return
7	3.5%	1.035
8	2.5%	1.025
9	5.5%	1.055
10	4.0%	1.040

The product of the 1 + Return is 1.3402338.

Therefore, $\bar{X}_G = \sqrt[10]{1.3402338} - 1 = 2.9717\%$.

36 A is correct.

Year	Return	1+ Return	1/(1+Return)
1	4.5%	1.045	0.957
2	6.0%	1.060	0.943
3	1.5%	1.015	0.985
4	−2.0%	0.980	1.020
5	0.0%	1.000	1.000
6	4.5%	1.045	0.957
7	3.5%	1.035	0.966
8	2.5%	1.025	0.976
9	5.5%	1.055	0.948
10	4.0%	1.040	0.962
Sum			**9.714**

The harmonic mean return = (n/Sum of reciprocals) − 1 = (10 / 9.714) − 1.

The harmonic mean return = 2.9442%.

37 B is correct.

Year	Return	Deviation	Deviation Squared
1	4.5%	0.0150	0.000225
2	6.0%	0.0300	0.000900
3	1.5%	−0.0150	0.000225
4	−2.0%	−0.0500	0.002500
5	0.0%	−0.0300	0.000900
6	4.5%	0.0150	0.000225
7	3.5%	0.0050	0.000025
8	2.5%	−0.0050	0.000025
9	5.5%	0.0250	0.000625
10	4.0%	0.0100	0.000100
Sum		0.0000	**0.005750**

The standard deviation is the square root of the sum of the squared deviations divided by $n - 1$:

$$s = \sqrt{\frac{0.005750}{9}} = 2.5276\%.$$

38 B is correct.

Year	Return	Deviation Squared below Target of 2%
1	4.5%	
2	6.0%	
3	1.5%	0.000025
4	−2.0%	0.001600
5	0.0%	0.000400
6	4.5%	
7	3.5%	
8	2.5%	
9	5.5%	
10	4.0%	
Sum		0.002025

The target semi-deviation is the square root of the sum of the squared deviations from the target, divided by $n - 1$:

$$s_{\text{Target}} = \sqrt{\frac{0.002025}{9}} = 1.5\%.$$

39 B is correct. The correlation coefficient is positive, indicating that the two series move together.

40 C is correct. Both outliers and spurious correlation are potential problems with interpreting correlation coefficients.

41 C is correct. The correlation coefficient is positive because the covariation is positive.

42 A is correct. The correlation coefficient is negative because the covariation is negative.

43 C is correct. The correlation coefficient is positive because the covariance is positive. The fact that one or both variables have a negative mean does not affect the sign of the correlation coefficient.

44 B is correct. The median is indicated within the box, which is the 100.49 in this diagram.

45 C is correct. The interquartile range is the difference between 114.25 and 79.74, which is 34.51.

46 B is correct. The coefficient of variation is the ratio of the standard deviation to the arithmetic average, or $\sqrt{0.001723}/0.09986 = 0.416$.

47 C is correct. The skewness is positive, so it is right-skewed (positively skewed).

48 C is correct. The excess kurtosis is positive, indicating that the distribution is “fat-tailed”; therefore, there is more probability in the tails of the distribution relative to the normal distribution.

READING

3

Probability Concepts

**by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, CFA,
Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA**

Richard A. DeFusco, PhD, CFA, is at the University of Nebraska-Lincoln (USA). Dennis W. McLeavey, DBA, CFA, is at the University of Rhode Island (USA). Jerald E. Pinto, PhD, CFA, is at CFA Institute (USA). David E. Runkle, PhD, CFA, is at Jacobs Levy Equity Management (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** define a random variable, an outcome, and an event;
☐	**b.** identify the two defining properties of probability, including mutually exclusive and exhaustive events, and compare and contrast empirical, subjective, and a priori probabilities;
☐	**c.** describe the probability of an event in terms of odds for and against the event;
☐	**d.** calculate and interpret conditional probabilities;
☐	**e.** demonstrate the application of the multiplication and addition rules for probability;
☐	**f.** compare and contrast dependent and independent events;
☐	**g.** calculate and interpret an unconditional probability using the total probability rule;
☐	**h.** calculate and interpret the expected value, variance, and standard deviation of random variables;
☐	**i.** explain the use of conditional expectation in investment applications;
☐	**j.** interpret a probability tree and demonstrate its application to investment problems;
☐	**k.** calculate and interpret the expected value, variance, standard deviation, covariances, and correlations of portfolio returns;

(continued)

CFA Institute would like to thank John Stowe, PhD, CFA, for his significant contributions in revising this reading.

© 2021 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	The candidate should be able to:
☐	**l.** calculate and interpret the covariances of portfolio returns using the joint probability function;
☐	**m.** calculate and interpret an updated probability using Bayes' formula;
☐	**n.** identify the most appropriate method to solve a particular counting problem and analyze counting problems using factorial, combination, and permutation concepts.

1 INTRODUCTION, PROBABILITY CONCEPTS, AND ODDS RATIOS

a define a random variable, an outcome, and an event

Investment decisions are made in a risky environment. The tools that allow us to make decisions with consistency and logic in this setting are based on probability concepts. This reading presents the essential probability tools needed to frame and address many real-world problems involving risk. These tools apply to a variety of issues, such as predicting investment manager performance, forecasting financial variables, and pricing bonds so that they fairly compensate bondholders for default risk. Our focus is practical. We explore the concepts that are most important to investment research and practice. Among these are independence, as it relates to the predictability of returns and financial variables; expectation, as analysts continually look to the future in their analyses and decisions; and variability, variance or dispersion around expectation, as a risk concept important in investments. The reader will acquire specific skills and competencies in using these probability concepts to understand risks and returns on investments.

1.1 Probability, Expected Value, and Variance

The probability concepts and tools necessary for most of an analyst's work are relatively few and straightforward but require thought to apply. This section presents the essentials for working with probability, expectation, and variance, drawing on examples from equity and fixed income analysis.

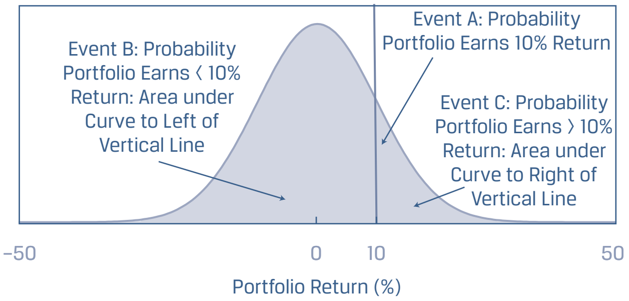
An investor's concerns center on returns. The return on a risky asset is an example of a random variable.

- Definition of Random Variable. A **random variable** is a quantity whose future outcomes are uncertain.
- Definition of Outcome. An **outcome** is a possible value of a random variable.

Using Exhibit 1 as an example, a portfolio manager may have a return objective of 10% a year. The portfolio manager's focus at the moment may be on the likelihood of earning a return that is less than 10% over the next year. Ten percent is a particular value or outcome of the random variable "portfolio return." Although we may be concerned about a single outcome, frequently our interest may be in a set of outcomes. The concept of "event" covers both.

- Definition of Event. An **event** is a specified set of outcomes.

Exhibit 1 Visualizing Probability



An event can be a single outcome—for example, the portfolio earns a return of (exactly) 10%. We can capture the portfolio manager's concerns by defining another event as the portfolio earns a return below 10%. This second event, referring as it does to all possible returns greater than or equal to –100% (the worst possible return, losing all the money in the portfolio) but less than 10%, contains an infinite number of outcomes. To save words, it is common to use a capital letter in italics to represent a defined event. We could define A = *the portfolio earns a return of 10%* and B = *the portfolio earns a return below 10%.*

To return to the portfolio manager's concern, how likely is it that the portfolio will earn a return below 10%? The answer to this question is a **probability**: a number between 0 and 1 that measures the chance that a stated event will occur. If the probability is 0.65 that the portfolio earns a return below 10%, there is a 65% chance of that event happening. If an event is impossible, it has a probability of 0. If an event is certain to happen, it has a probability of 1. If an event is impossible or a sure thing, it is not random at all. So, 0 and 1 bracket all the possible values of a probability.

To reiterate, a probability can be thought of as the likelihood that something will happen. If it has a probability of 1, it is likely to happen 100% of the time, and if it has a probably of 0, it is likely to never happen. Some people think of probabilities as akin to relative frequencies. If something is expected to happen 30 times out of 100, the probability is 0.30. The probability is the number of ways that an (equally likely) event can happen divided by the total number of possible outcomes.

b identify the two defining properties of probability, including mutually exclusive and exhaustive events, and compare and contrast empirical, subjective, and a priori probabilities

Probability has two properties, which together constitute its definition.

- Definition of Probability. The two defining properties of a probability are:
  1. The probability of any event E is a number between 0 and 1: $0 \le P(E) \le 1$.
  2. The sum of the probabilities of any set of mutually exclusive and exhaustive events equals 1.

P followed by parentheses stands for "the probability of (the event in parentheses)," as in $P(E)$ for "the probability of event E." We can also think of P as a rule or function that assigns numerical values to events consistent with Properties 1 and 2.

In the above definition, the term **mutually exclusive** means that only one event can occur at a time; **exhaustive** means that the events cover all possible outcomes. Referring back to Exhibit 1, the events A = *the portfolio earns a return of 10%* and B = *the portfolio earns a return below 10%* are mutually exclusive because A and B cannot both occur at the same time. For example, a return of 8.1% means that B has occurred and A has not occurred. Although events A and B are mutually exclusive, they are not exhaustive because they do not cover outcomes such as a return of 11%.

Suppose we define a third event: C = *the portfolio earns a return above 10%.* Clearly, A, B, and C are mutually exclusive and exhaustive events. Each of $P(A)$, $P(B)$, and $P(C)$ is a number between 0 and 1, and $P(A) + P(B) + P(C) = 1$.

Earlier, to illustrate a concept, we assumed a probability of 0.65 for a portfolio earning less than 10%, without justifying the particular assumption. We also talked about using assigned probabilities of outcomes to calculate the probability of events, without explaining how such a probability distribution might be estimated. Making actual financial decisions using inaccurate probabilities could have grave consequences. How, in practice, do we estimate probabilities? This topic is a field of study in itself, but there are three broad approaches to estimating probabilities. In investments, we often estimate the probability of an event as a relative frequency of occurrence based on historical data. This method produces an **empirical probability**. For example, suppose you noted that 51 of the 60 stocks in a particular large-cap equity index pay dividends. The empirical probability of the stocks in the index paying a dividend is $P(\textit{stock is dividend paying}) = 51 / 60 = 0.85$.

Relationships must be stable through time for empirical probabilities to be accurate. We cannot calculate an empirical probability of an event not in the historical record or a reliable empirical probability for a very rare event. In some cases, then, we may adjust an empirical probability to account for perceptions of changing relationships. In other cases, we have no empirical probability to use at all. We may also make a personal assessment of probability without reference to any particular data. Another type of probability is a **subjective probability**, one drawing on personal or subjective judgment. Subjective probabilities are of great importance in investments. Investors, in making buy and sell decisions that determine asset prices, often draw on subjective probabilities.

For many well-defined problems, we can deduce probabilities by reasoning about the problem. The resulting probability is an **a priori probability**, one based on logical analysis rather than on observation or personal judgment. Because a priori and empirical probabilities generally do not vary from person to person, they are often grouped as **objective probabilities**.

For examples of the three types of probabilities, suppose you want to estimate the probability of flipping a coin and getting exactly two heads out of five flips. For the empirical probability, you do the experiment 100 times (five flips each time) and find that you get two heads 33 times. The empirical probability would be 33/100 = 0.33. For a subjective judgement, you think the probability is somewhere between 0.25 and 0.50, so you split the difference and choose 0.375. For the a priori probability, you assume that the binomial probability function (discussed later in the curriculum) applies, and the mathematical probability of two heads out of five flips is 0.3125.

c describe the probability of an event in terms of odds for and against the event

Another way of stating probabilities often encountered in investments is in terms of odds—for instance, "the odds for *E*" or the "odds against *E*." A probability is the fraction of the time you expect an event to occur, and the odds for an event is the probability that an event will occur divided by the probability that the event will not occur. Consider a football team that has a 0.25 probability of winning the World Cup, and a 0.75 probability of losing. The odds for winning are 0.25/0.75 = 0.33 (and the odds for losing are 0.75/0.25 = 3.0). If another team has a 0.80 probability of winning, the odds for winning would be 0.80/0.20 = 4.0. If, for a third team, the probability of winning was 0.50, the odds are even: odds = 0.50/0.50 = 1. If the probability is low, the odds are very close to the probability. For example, if the probability of winning is 0.05, the odds for winning are 0.05/0.95 = 0.0526.

EXAMPLE 1

Odds of Passing a Quantitative Methods Investment Course

Two of your colleagues are taking a quantitative methods investment course.

1. If your first colleague has a 0.40 probability of passing, what are his odds for passing?
2. If your second colleague has odds of passing of 4 to 1, what is the probability of her passing?

Solution for 1:

The odds are the probability of passing divided by the probability of not passing. The odds are 0.40 / 0.60 = 2/3 ≈ 0.667.

Solution for 2:

The odds = *Probability* (*passing*) / *Probability* (*not passing*). If Y = Probability of passing, then 4 = Y / (1 − Y). Solving for Y, we get 0.80 as the probability of passing.

We interpret probabilities stated in terms of odds as follows:

- Probability Stated as Odds. Given a probability $P(E)$,
  1. **Odds for *E*** = $P(E)/[1 - P(E)]$. The odds for E are the probability of E divided by 1 minus the probability of E. Given odds for E of "a to b," the implied probability of E is $a/(a + b)$.

  In the example, the statement that your second colleague's odds of passing the exam are 4 to 1 means that the probability of the event is 4/(4 + 1) = 4/5 = 0.80.

  2. **Odds against *E*** = $[1 - P(E)]/P(E)$, the reciprocal of odds for E. Given odds against E of "a to b," the implied probability of E is $b/(a + b)$.

  In the example, if the odds against your second colleague passing the exam are 1 to 4, this means that the probability of the event is 1/(4 + 1) = 1/5 = 0.20.

To further explain odds for an event, if $P(E) = 1/8$, the odds for E are $(1/8)/(7/8) = (1/8)(8/7) = 1/7$, or "1 to 7." For each occurrence of E, we expect seven cases of nonoccurrence; out of eight cases in total, therefore, we expect E to happen once, and the probability of E is 1/8. In wagering, it is common to speak in terms of the odds against something, as in Statement 2. For odds of "15 to 1" against E (an implied probability of E of 1/16), a $1 wager on E, if successful, returns $15 in profits plus the $1 staked in the wager. We can calculate the bet's anticipated profit as follows:

Win:	Probability = 1/16; Profit =$15
Loss:	Probability = 15/16; Profit = −$1

Anticipated profit = (1/16)($15) + (15/16)(−$1) = $0

Weighting each of the wager's two outcomes by the respective probability of the outcome, if the odds (probabilities) are accurate, the anticipated profit of the bet is $0.

EXAMPLE 2

Profiting from Inconsistent Probabilities

You are examining the common stock of two companies in the same industry in which an important antitrust decision will be announced next week. The first company, SmithCo Corporation, will benefit from a governmental decision that there is no antitrust obstacle related to a merger in which it is involved. You believe that SmithCo's share price reflects a 0.85 probability of such a decision. A second company, Selbert Corporation, will equally benefit from a "go ahead" ruling. Surprisingly, you believe Selbert stock reflects only a 0.50 probability of a favorable decision. Assuming your analysis is correct, what investment strategy would profit from this pricing discrepancy?

Consider the logical possibilities. One is that the probability of 0.50 reflected in Selbert's share price is accurate. In that case, Selbert is fairly valued, but SmithCo is overvalued, because its current share price overestimates the probability of a "go ahead" decision. The second possibility is that the probability of 0.85 is accurate. In that case, SmithCo shares are fairly valued, but Selbert shares, which build in a lower probability of a favorable decision, are undervalued. You diagram the situation as shown in Exhibit 2.

Exhibit 2 Worksheet for Investment Problem

	True Probability of a "Go Ahead" Decision	
	0.50	**0.85**
SmithCo	Shares Overvalued	Shares Fairly Valued
Selbert	Shares Fairly Valued	Shares Undervalued
Strategy	Short-Sell Smith / Buy Selbert	Sell Smith / Buy Selbert

The 0.50 probability column shows that Selbert shares are a better value than SmithCo shares. Selbert shares are also a better value if a 0.85 probability is accurate. Thus, SmithCo shares are overvalued relative to Selbert shares.

Your investment actions depend on your confidence in your analysis and on any investment constraints you face (such as constraints on selling stock short). Selling short or shorting stock means selling borrowed shares in the hope of repurchasing them later at a lower price. A conservative strategy would be to buy Selbert shares and reduce or eliminate any current position in SmithCo. The most aggressive strategy is to short SmithCo stock (relatively overvalued) and simultaneously buy the stock of Selbert (relatively undervalued). The prices of SmithCo and Selbert shares reflect probabilities that are not consistent. According to one of the most important probability results for investments, the **Dutch Book Theorem**, inconsistent probabilities create profit opportunities. In our example, investors' buy and sell decisions exploit the inconsistent probabilities to eliminate the profit opportunity and inconsistency.

CONDITIONAL AND JOINT PROBABILITY

2

d calculate and interpret conditional probabilities

To understand the meaning of a probability in investment contexts, we need to distinguish between two types of probability: unconditional and conditional. Both unconditional and conditional probabilities satisfy the definition of probability stated earlier, but they are calculated or estimated differently and have different interpretations. They provide answers to different questions.

The probability in answer to the straightforward question "What is the probability of this event *A*?" is an **unconditional probability**, denoted $P(A)$. Suppose the question is "What is the probability that *the stock earns a return above the risk-free rate* (event *A*)?" The answer is an unconditional probability that can be viewed as the ratio of two quantities. The numerator is the sum of the probabilities of stock returns above the risk-free rate. Suppose that sum is 0.70. The denominator is 1, the sum of the probabilities of all possible returns. The answer to the question is $P(A) = 0.70$.

Contrast the question "What is the probability of *A*?" with the question "What is the probability of *A*, given that *B* has occurred?" The probability in answer to this last question is a **conditional probability**, denoted $P(A \mid B)$ (read: "the probability of *A* given *B*").

Suppose we want to know the probability that *the stock earns a return above the risk-free rate* (event *A*), given that *the stock earns a positive return* (event *B*). With the words "given that," we are restricting returns to those larger than 0%—a new element in contrast to the question that brought forth an unconditional probability. The conditional probability is calculated as the ratio of two quantities. The numerator is the sum of the probabilities of stock returns above the risk-free rate; in this particular case, the numerator is the same as it was in the unconditional case, which we gave as 0.70. The denominator, however, changes from 1 to the sum of the probabilities for all outcomes (returns) above 0%. Suppose that number is 0.80, a larger number than 0.70 because returns between 0 and the risk-free rate have some positive probability of occurring. Then $P(A \mid B) = 0.70/0.80 = 0.875$. If we observe that the stock earns a positive return, the probability of a return above the risk-free rate is greater than the unconditional probability, which is the probability of the event given no other information. To review, an unconditional probability is the probability of an event without any restriction (i.e., a standalone probability). A conditional probability, in contrast, is a probability of an event given that another event has occurred.

To state an exact definition of conditional probability, we first need to introduce the concept of joint probability. Suppose we ask the question "What is the probability of both *A* and *B* happening?" The answer to this question is a **joint probability**, denoted $P(AB)$ (read: "the probability of *A* and *B*"). If we think of the probability of *A* and the probability of *B* as sets built of the outcomes of one or more random variables, the joint probability of *A* and *B* is the sum of the probabilities of the outcomes they have in common. For example, consider two events: *the stock earns a return above the risk-free rate* (*A*) and *the stock earns a positive return* (*B*). The outcomes of *A* are contained within (a subset of) the outcomes of *B*, so $P(AB)$ equals $P(A)$. We can now state a formal definition of conditional probability that provides a formula for calculating it.

- Definition of Conditional Probability. The conditional probability of *A* given that *B* has occurred is equal to the joint probability of *A* and *B* divided by the probability of *B* (assumed not to equal 0).

$$P(A \mid B) = P(AB)/P(B),\ P(B) \neq 0 \tag{1}$$

For example, suppose B happens half the time, $P(B) = 0.50$, and A and B both happen 10% of the time, $P(AB) = 0.10$. What is the probability that A happens, given that B happens? That is $P(A \mid B) = P(AB)/P(B) = 0.10 / 0.50 = 0.20$.

e demonstrate the application of the multiplication and addition rules for probability

Sometimes we know the conditional probability $P(A \mid B)$ and we want to know the joint probability $P(AB)$. We can obtain the joint probability from the following multiplication rule for probabilities, Equation 1 rearranged.

- **Multiplication Rule for Probability**. The joint probability of A and B can be expressed as

$$P(AB) = P(A \mid B)P(B) \tag{2}$$

With the same numbers above, if B happens 50% of the time, and the probability of A given that B happens is 20%, the joint probability of A and B happening is $P(AB) = P(A \mid B)P(B) = 0.20 \times 0.50 = 0.10$.

EXAMPLE 3

Conditional Probabilities and Predictability of Mutual Fund Performance (1)

An analyst conducts a study of the returns of 200 mutual funds over a two-year period. For each year, the total returns for the funds were ranked, and the top 50% of funds were labeled winners; the bottom 50% were labeled losers. Exhibit 3 shows the percentage of those funds that were winners in two consecutive years, winners in one year and then losers in the next year, losers then winners, and finally losers in both years. The winner–winner entry, for example, shows that 66% of the first-year winner funds were also winners in the second year. The four entries in the table can be viewed as conditional probabilities.

Exhibit 3 Persistence of Returns: Conditional Probability for Year 2 Performance Given Year 1 Performance

	Year 2 Winner	Year 2 Loser
Year 1 Winner	66%	34%
Year 1 Loser	34%	66%

Based on the data in Exhibit 3, answer the following questions:

1 State the four events needed to define the four conditional probabilities.

2 State the four entries of the table as conditional probabilities using the form $P(\textit{this event} \mid \textit{that event})$ = number.

3 Are the conditional probabilities in Question 2 empirical, a priori, or subjective probabilities?

4 Using information in the table, calculate the probability of the event a *fund is a loser in both Year 1 and Year 2.* (Note that because 50% of funds are categorized as losers in each year, the unconditional probability that a fund is labeled a loser in either year is 0.5.)

Solution to 1:

The four events needed to define the conditional probabilities are as follows:

Fund is a Year 1 winner

Fund is a Year 1 loser

Fund is a Year 2 loser

Fund is a Year 2 winner

Solution to 2:

From Row 1:

$P(\text{fund is a Year 2 winner} \mid \text{fund is a Year 1 winner}) = 0.66$

$P(\text{fund is a Year 2 loser} \mid \text{fund is a Year 1 winner}) = 0.34$

From Row 2:

$P(\text{fund is a Year 2 winner} \mid \text{fund is a Year 1 loser}) = 0.34$

$P(\text{fund is a Year 2 loser} \mid \text{fund is a Year 1 loser}) = 0.66$

Solution to 3:

These probabilities are calculated from data, so they are empirical probabilities.

Solution to 4:

The estimated probability is 0.33. Let A represent the event that a *fund is a Year 2 loser*, and let B represent the event that the *fund is a Year 1 loser*. Therefore, the event AB is the event that a *fund is a loser in both Year 1 and Year 2.* From Exhibit 3, $P(A \mid B) = 0.66$ and $P(B) = 0.50$. Thus, using Equation 2, we find that

$$P(AB) = P(A \mid B)P(B) = 0.66(0.50) = 0.33$$

or a probability of 0.33. Note that Equation 2 states that the joint probability of A and B equals the probability of A given B times the probability of B. Because $P(AB) = P(BA)$, the expression $P(AB) = P(BA) = P(B \mid A)P(A)$ is equivalent to Equation 2.

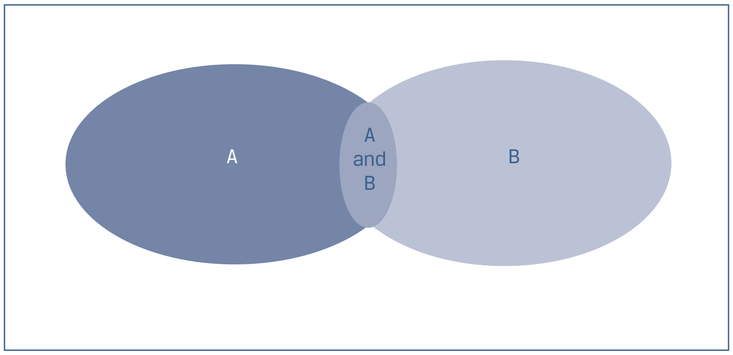
When we have two events, A and B, that we are interested in, we often want to know the probability that either A or B occurs. Here the word "or" is inclusive, meaning that either A or B occurs or that both A and B occur. Put another way, the probability of A or B is the probability that at least one of the two events occurs. Such probabilities are calculated using the **addition rule for probabilities**.

- **Addition Rule for Probabilities.** Given events A and B, the probability that A or B occurs, or both occur, is equal to the probability that A occurs, plus the probability that B occurs, minus the probability that both A and B occur.

$$P(A \text{ or } B) = P(A) + P(B) - P(AB) \quad (3)$$

If we think of the individual probabilities of A and B as sets built of outcomes of one or more random variables, the first step in calculating the probability of A or B is to sum the probabilities of the outcomes in A to obtain $P(A)$. If A and B share any outcomes, then if we now added $P(B)$ to $P(A)$, we would count twice the probabilities of those shared outcomes. So we add to $P(A)$ the quantity $[P(B) - P(AB)]$, which is the probability of outcomes in B net of the probability of any outcomes already counted when we computed $P(A)$. Exhibit 4 illustrates this process; we avoid double-counting the outcomes in the intersection of A and B by subtracting $P(AB)$. As an example of the calculation, if $P(A) = 0.50$, $P(B) = 0.40$, and $P(AB) = 0.20$, then $P(A \text{ or } B) = 0.50 + 0.40 - 0.20 = 0.70$. Only if the two events A and B were mutually exclusive, so that $P(AB) = 0$, would it be correct to state that $P(A \text{ or } B) = P(A) + P(B)$.

Exhibit 4 Addition Rule for Probabilities



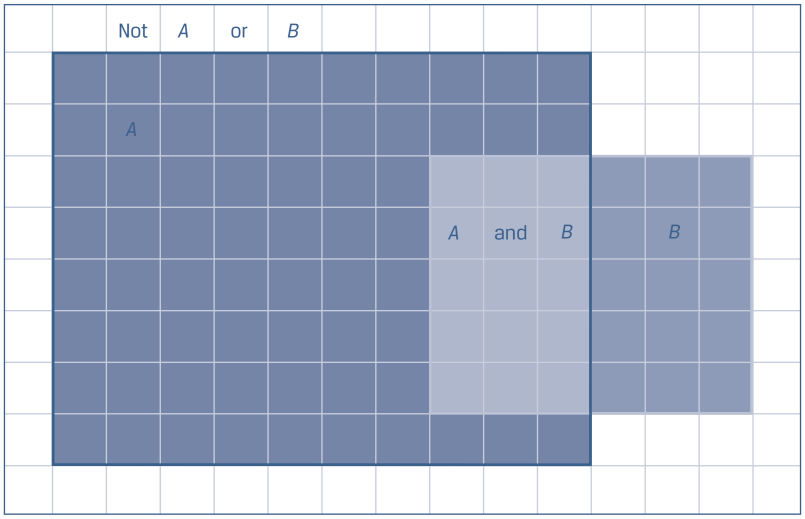
Example 4 illustrates the relation between empirical frequencies and unconditional, conditional, and joint probabilities as well as the multiplication and addition rules for probability.

EXAMPLE 4

Frequencies and Probability Concepts

Analysts often discuss the frequencies of events as well as their probabilities. In Exhibit 5, there are 150 cells, each representing one trading day. Outcome A, one of the 80 trading days when the stock market index increased, is represented by the dark-shaded rectangle with 80 cells. Outcome B, one the 30 trading days when interest rates decreased, is represented by the light -bordered rectangle with 30 cells. The overlap between these two rectangles, when both events A and B occurred—the stock market index increased, and interest rates decreased—happened 15 times and is represented by the intermediate-shaded rectangle.

Exhibit 5 Frequencies for Two Events



- The frequency of *A* (stock market index increased) is 80 and has an unconditional probability *P*(*A*) = 80/150 = 0.533.
- The frequency of *B* (interest rates decreased) is 30 and has an unconditional probability *P*(*B*) = 30/150 = 0.20.
- The frequency of *A* and *B* (stock market index increased, and interest rates decreased) is 15 and has a joint probability, *P*(*AB*) = 15/150 = 0.10.

The frequency of *A* or *B* (stock market index increased or interest rates decreased is 95, and *P*(*A* or *B*) is 95/150 = 0.633. Using the addition rule for probabilities, the probability of *A* or *B* is *P*(*A* or *B*) = *P*(*A*) + *P*(*B*) – *P*(*AB*) = 80/150 + 30/150 – 15/150 = 95/150 = 0.633. The probability of not *A* or *B* (stock market index did not increase or interest rates did not decrease) = 1 – 95/150 = 55/150 = 0.367.

The conditional probability of *A* given *B*, *P*(*A*|*B*), stock market index increased given that interest rates decreased, was 15/30 = 0.50, which is also *P*(*A*|*B*) = *P*(*AB*) / *P*(*B*) = (15/150) / (30/150) = 0.10 / 0.20 = 0.50

The next example shows how much useful information can be obtained using the probability rules presented to this point.

EXAMPLE 5

Probability of a Limit Order Executing

You have two buy limit orders outstanding on the same stock. A limit order to buy stock at a stated price is an order to buy at that price or lower. A number of vendors, including an internet service that you use, supply the estimated probability that a limit order will be filled within a stated time horizon, given the current stock price and the price limit. One buy order (Order 1) was placed at a price limit of \$10. The probability that it will execute within one hour is 0.35. The second buy order (Order 2) was placed at a price limit of \$9.75; it has a 0.25 probability of executing within the same one-hour time frame.

1 What is the probability that either Order 1 or Order 2 will execute?
2 What is the probability that Order 2 executes, given that Order 1 executes?

Solution to 1:

The probability is 0.35. The two probabilities that are given are $P(\textit{Order 1 executes}) = 0.35$ and $P(\textit{Order 2 executes}) = 0.25$. Note that if Order 2 executes, it is certain that Order 1 also executes because the price must pass through \$10 to reach \$9.75. Thus,

$$P(\textit{Order 1 executes} \mid \textit{Order 2 executes}) = 1$$

and using the multiplication rule for probabilities,

$$P(\textit{Order 1 executes} \text{ and } \textit{Order 2 executes}) = P(\textit{Order 1 executes} \mid \textit{Order 2 executes})P(\textit{Order 2 executes}) = 1(0.25) = 0.25$$

To answer the question, we use the addition rule for probabilities:

$$\begin{aligned}&P(\textit{Order 1 executes} \text{ or } \textit{Order 2 executes}) = P(\textit{Order 1 executes})\\&+ P(\textit{Order 2 executes}) - P(\textit{Order 1 executes} \text{ and } \textit{Order 2 executes})\\&= 0.35 + 0.25 - 0.25 = 0.35\end{aligned}$$

Note that the outcomes for which Order 2 executes are a subset of the outcomes for which Order 1 executes. After you count the probability that Order 1 executes, you have counted the probability of the outcomes for which Order 2 also executes. Therefore, the answer to the question is the probability that Order 1 executes, 0.35.

Solution to 2:

If the first order executes, the probability that the second order executes is 0.714. In the solution to Part 1, you found that $P(\textit{Order 1 executes} \text{ and } \textit{Order 2 executes}) = P(\textit{Order 1 executes} \mid \textit{Order 2 executes})P(\textit{Order 2 executes}) = 1(0.25) = 0.25$. An equivalent way to state this joint probability is useful here:

$$\begin{aligned}&P(\textit{Order 1 executes} \text{ and } \textit{Order 2 executes}) = 0.25\\&= P(\textit{Order 2 executes} \mid \textit{Order 1 executes})P(\textit{Order 1 executes})\end{aligned}$$

Because $P(\textit{Order 1 executes}) = 0.35$ was a given, you have one equation with one unknown:

$$0.25 = P(\textit{Order 2 executes} \mid \textit{Order 1 executes})(0.35)$$

You conclude that $P(\textit{Order 2 executes} \mid \textit{Order 1 executes}) = 0.25/0.35 = 0.714$. You can also use Equation 1 to obtain this answer.

f compare and contrast dependent and independent events

The concepts of independence and dependence are of great interest to investment analysts. These concepts bear on such basic investment questions as which financial variables are useful for investment analysis, whether asset returns can be predicted, and whether superior investment managers can be selected based on their past records.

Two events are independent if the occurrence of one event does not affect the probability of occurrence of the other event.

- Definition of Independent Events. Two events A and B are **independent** if and only if $P(A \mid B) = P(A)$ or, equivalently, $P(B \mid A) = P(B)$.

The logic of independence is clear: A and B are independent if the conditional probability of A given B, $P(A \mid B)$, is the same as the unconditional probability of A, $P(A)$. Independence means that knowing B tells you nothing about A.

For an example of independent events, suppose that event A is the bankruptcy of Company A, and event B is the bankruptcy of Company B. If the probability of bankruptcy of Company A is $P(A) = 0.20$, and the probability of bankruptcy of Company A given that Company B goes bankrupt is the same, $P(A \mid B) = 0.20$, then event A is independent of event B.

When two events are not independent, they are **dependent**: The probability of occurrence of one is related to the occurrence of the other. If we are trying to forecast one event, information about a dependent event may be useful, but information about an independent event will not be useful. For example, suppose an announcement is released that a biotech company will be acquired at an attractive price by another company. If the prices of pharmaceutical companies increase as a result of this news, the companies' stock prices are not independent of the biotech takeover announcement event. For a different example, if two events are mutually exclusive, then knowledge that one event has occurred gives us information that the other (mutually exclusive) event cannot occur.

When two events are independent, the multiplication rule for probabilities, Equation 2, simplifies because $P(A \mid B)$ in that equation then equals $P(A)$.

- **Multiplication Rule for Independent Events**. When two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B.

$$P(AB) = P(A)P(B) \tag{4}$$

Therefore, if we are interested in two independent events with probabilities of 0.75 and 0.50, respectively, the probability that both will occur is $0.375 = 0.75(0.50)$. The multiplication rule for independent events generalizes to more than two events; for example, if A, B, and C are independent events, then $P(ABC) = P(A)P(B)P(C)$.

EXAMPLE 6

BankCorp's Earnings per Share (1)

As part of your work as a banking industry analyst, you build models for forecasting earnings per share of the banks you cover. Today you are studying BankCorp. The historical record shows that in 55% of recent quarters, BankCorp's EPS has increased sequentially, and in 45% of quarters, EPS has decreased or remained unchanged sequentially. At this point in your analysis, you are assuming that changes in sequential EPS are independent.

Earnings per share for 2Q:Year 1 (that is, EPS for the second quarter of Year 1) were larger than EPS for 1Q:Year 1.

1 What is the probability that *3Q:Year 1 EPS will be larger than 2Q:Year 1 EPS* (a positive change in sequential EPS)?

2 What is the probability that EPS decreases or remains unchanged in the next two quarters?

Solution to 1:

Under the assumption of independence, the probability that *3Q:Year 1 EPS will be larger than 2Q:Year 1 EPS* is the unconditional probability of positive change, 0.55. The fact that 2Q:Year 1 EPS was larger than 1Q:Year 1 EPS is not useful information, because the next change in EPS is independent of the prior change.

Solution to 2:

Assuming independence, the probability is 0.2025 = 0.45(0.45).

The following example illustrates how difficult it is to satisfy a set of independent criteria even when each criterion by itself is not necessarily stringent.

EXAMPLE 7

Screening Stocks for Investment

You have developed a stock screen—a set of criteria for selecting stocks. Your investment universe (the set of securities from which you make your choices) is 905 large- and medium-cap US equities, specifically all stocks that are members of the S&P 500 and S&P 400 Indexes. Your criteria capture different aspects of the stock selection problem; you believe that the criteria are independent of each other, to a close approximation.

Criterion	Number of stocks meeting criterion	Fraction of stocks meeting criterion
First valuation criterion	556	0.614
Second valuation criterion	489	0.540
Analyst coverage criterion	600	0.663
Profitability criterion	490	0.541
Financial strength criterion	313	0.346

How many stocks do you expect to pass your screen?

Only 37 stocks out of 905 should pass through your screen. If you define five events—*the stock passes the first valuation criterion, the stock passes the second valuation criterion, the stock passes the analyst coverage criterion, the company passes the profitability criterion, the company passes the financial strength criterion* (say events A, B, C, D, and E, respectively)—then the probability that a stock will pass all five criteria, under independence, is

$$P(ABCDE) = P(A)P(B)P(C)P(D)P(E) = (0.614)(0.540)(0.663)(0.541)(0.346) = 0.0411$$

Although only one of the five criteria is even moderately strict (the strictest lets 34.6% of stocks through), the probability that a stock can pass all five criteria is only 0.0411, or about 4%. If the criteria are independent, the size of the list of candidate investments is expected to be 0.0411(905) = 37 stocks.

An area of intense interest to investment managers and their clients is whether records of past performance are useful in identifying repeat winners and losers. The following example shows how this issue relates to the concept of independence.

EXAMPLE 8

Conditional Probabilities and Predictability of Mutual Fund Performance (2)

The purpose of the mutual fund study introduced in Example 3 was to address the question of repeat mutual fund winners and losers. If the status of a fund as a winner or a loser in one year is independent of whether it is a winner in the next year, the practical value of performance ranking is questionable. Using the four events defined in Example 3 as building blocks, we can define the following events to address the issue of predictability of mutual fund performance:

Fund is a Year 1 winner and *fund is a Year 2 winner*

Fund is a Year 1 winner and *fund is a Year 2 loser*

Fund is a Year 1 loser and *fund is a Year 2 winner*

Fund is a Year 1 loser and *fund is a Year 2 loser*

In Part 4 of Example 3, you calculated that

P(*fund is a Year 2 loser* and *fund is a Year 1 loser*) = 0.33

If the ranking in one year is independent of the ranking in the next year, what will you expect P(*fund is a Year 2 loser* and *fund is a Year 1 loser*) to be? Interpret the empirical probability 0.33.

By the multiplication rule for independent events, P(*fund is a Year 2 loser* and *fund is a Year 1 loser*) = P(*fund is a Year 2 loser*)P(*fund is a Year 1 loser*). Because 50% of funds are categorized as losers in each year, the unconditional probability that a fund is labeled a loser in either year is 0.50. Thus P(*fund is a Year 2 loser*)P(*fund is a Year 1 loser*) = 0.50(0.50) = 0.25. If the status of a fund as a loser in one year is independent of whether it is a loser in the prior year, we conclude that P(*fund is a Year 2 loser* and *fund is a Year 1 loser*) = 0.25. This probability is a priori because it is obtained from reasoning about the problem. You could also reason that the four events described above define categories and that if funds are randomly assigned to the four categories, there is a 1/4 probability of *fund is a Year 1 loser* and *fund is a Year 2 loser.* If the classifications in Year 1 and Year 2 were dependent, then the assignment of funds to categories would not be random. The empirical probability of 0.33 is above 0.25. Is this apparent predictability the result of chance? Further analysis would be necessary to determine whether these results would allow you to reject the hypothesis that investment returns are independent between Year 1 and Year 2.

g calculate and interpret an unconditional probability using the total probability rule

In many practical problems, we logically analyze a problem as follows: We formulate scenarios that we think affect the likelihood of an event that interests us. We then estimate the probability of the event, given the scenario. When the scenarios (conditioning events) are mutually exclusive and exhaustive, no possible outcomes are left out. We can then analyze the event using the **total probability rule**. This rule explains the unconditional probability of the event in terms of probabilities conditional on the scenarios.

The total probability rule is stated below for two cases. Equation 5 gives the simplest case, in which we have two scenarios. One new notation is introduced: If we have an event or scenario S, the event not-S, called the **complement** of S, is written S^C. Note that $P(S) + P(S^C) = 1$, as either S or not-S must occur. Equation 6 states the rule for the general case of n mutually exclusive and exhaustive events or scenarios.

- **Total Probability Rule.**

$$\begin{aligned} P(A) &= P(AS) + P\left(AS^C\right) \\ &= P(A \mid S)P(S) + P\left(A \mid S^C\right)P\left(S^C\right) \end{aligned} \tag{5}$$

$$\begin{aligned} P(A) &= P(AS_1) + P(AS_2) + \ldots + P(AS_n) \\ &= P(A \mid S_1)P(S_1) + P(A \mid S_2)P(S_2) + \ldots + P(A \mid S_n)P(S_n) \end{aligned} \tag{6}$$

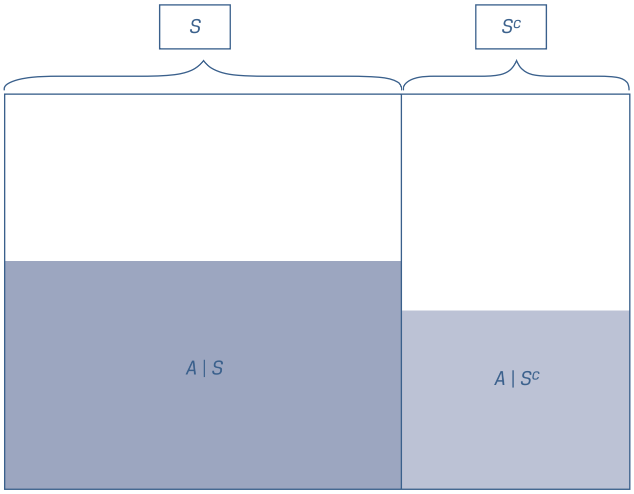
where S_1, S_2, …, S_n are mutually exclusive and exhaustive scenarios or events.

Equation 6 states the following: The probability of any event [$P(A)$] can be expressed as a weighted average of the probabilities of the event, given scenarios [terms such $P(A \mid S_1)$]; the weights applied to these conditional probabilities are the respective probabilities of the scenarios [terms such as $P(S_1)$ multiplying $P(A \mid S_1)$], and the scenarios must be mutually exclusive and exhaustive. Among other applications, this rule is needed to understand Bayes' formula, which we discuss later.

Exhibit 6 is a visual representation of the total probability rule. Panel A illustrates Equation 5 for the total probability rule when there are two scenarios (S and its complement S^C). For two scenarios, the probabilities of S and S^C sum to 1, and the probability of A is a weighted average where the probability of A in each scenario is weighted by the probability of each scenario. Panel B of Exhibit 6 illustrates Equation 6 for the total probability rule when there are n scenarios. The scenarios are mutually exclusive and exhaustive, and the sum of the probabilities for the scenarios is 1. Like the two-scenario case, the probability of A given the n-scenarios is a weighted average of the conditional probabilities of A in each scenario, using as weights the probability of each scenario.

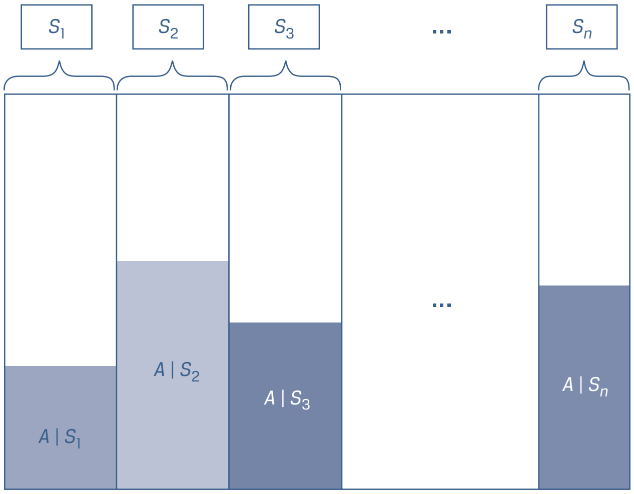
Exhibit 6 The Total Probability Rule for Two Scenarios and for *n* Scenarios

A. Total Probability Rule for Two Scenarios (S and S^C)



$P(S) + P(S^C) = 1\,;\ \ P(A) = P(AS) + P(AS^C) = P(A \mid S)\,P(S) + P(A \mid S^C)\,P(S^C)$

B. Total Probability Rule for n Scenarios



$S_1, S_2, \ldots S_n$, are mutually exclusive and exhaustive scenarios, such that

$$\sum_{i=1}^{n} P(S_i) = 1$$

$$\begin{aligned} P(A) &= P(AS_1) + P(AS_2) + \ldots + P(AS_n) \\ &= P(A \mid S_1)\,P(S_1) + P(A \mid S_2)\,P(S_2) + \ldots + P(A \mid S_n)\,P(S_n) \end{aligned}$$

In the next example, we use the total probability rule to develop a consistent set of views about BankCorp's earnings per share.

EXAMPLE 9

BankCorp's Earnings per Share (2)

You are continuing your investigation into whether you can predict the direction of changes in BankCorp's quarterly EPS. You define four events:

Event	Probability
A = *Change in sequential EPS is positive next quarter*	0.55
A^C = *Change in sequential EPS is 0 or negative next quarter*	0.45
S = *Change in sequential EPS is positive in the prior quarter*	0.55
S^C = *Change in sequential EPS is 0 or negative in the prior quarter*	0.45

On inspecting the data, you observe some persistence in EPS changes: Increases tend to be followed by increases, and decreases by decreases. The first probability estimate you develop is P(*change in sequential EPS is positive next quarter* | *change in sequential EPS is 0 or negative in the prior quarter*) = $P(A \mid S^C)$ = 0.40. The most recent quarter's EPS (2Q:Year 1) is announced, and the change is a positive sequential change (the event S). You are interested in forecasting EPS for 3Q:Year 1.

1 Write this statement in probability notation: "the probability that the change in sequential EPS is positive next quarter, given that the change in sequential EPS is positive the prior quarter."

2 Calculate the probability in Part 1. (Calculate the probability that is consistent with your other probabilities or beliefs.)

Solution to 1:

In probability notation, this statement is written $P(A \mid S)$.

Solution to 2:

The probability is 0.673 that the change in sequential EPS is positive for 3Q:Year 1, given the positive change in sequential EPS for 2Q:Year 1, as shown below.

According to Equation 5, $P(A) = P(A \mid S)P(S) + P(A \mid S^C)P(S^C)$. The values of the probabilities needed to calculate $P(A \mid S)$ are already known: $P(A) = 0.55$, $P(S) = 0.55$, $P(S^C) = 0.45$, and $P(A \mid S^C) = 0.40$. Substituting into Equation 5,

$$0.55 = P(A \mid S)(0.55) + 0.40(0.45)$$

Solving for the unknown, $P(A \mid S) = [0.55 - 0.40(0.45)]/0.55 = 0.672727$, or 0.673.

You conclude that P(*change in sequential EPS is positive next quarter* | *change in sequential EPS is positive the prior quarter*) = 0.673. Any other probability is not consistent with your other estimated probabilities. Reflecting the persistence in EPS changes, this conditional probability of a positive EPS change, 0.673, is greater than the unconditional probability of an EPS increase, 0.55.

3 EXPECTED VALUE (MEAN), VARIANCE, AND CONDITIONAL MEASURES OF EXPECTED VALUE AND VARIANCE

h calculate and interpret the expected value, variance, and standard deviation of random variables

The expected value of a random variable is an essential quantitative concept in investments. Investors continually make use of expected values—in estimating the rewards of alternative investments, in forecasting EPS and other corporate financial variables and ratios, and in assessing any other factor that may affect their financial position. The expected value of a random variable is defined as follows:

- Definition of Expected Value. The **expected value** of a random variable is the probability-weighted average of the possible outcomes of the random variable. For a random variable X, the expected value of X is denoted $E(X)$.

Expected value (for example, expected stock return) looks either to the future, as a forecast, or to the "true" value of the mean (the population mean). We should distinguish expected value from the concepts of historical or sample mean. The sample mean also summarizes in a single number a central value. However, the sample mean presents a central value for a particular set of observations as an equally weighted average of those observations. In sum, the contrast is forecast versus historical, or population versus sample.

EXAMPLE 10

BankCorp's Earnings per Share (3)

You continue with your analysis of BankCorp's EPS. In Exhibit 7, you have recorded a probability distribution for BankCorp's EPS for the current fiscal year.

Exhibit 7 Probability Distribution for BankCorp's EPS

Probability	EPS ($)
0.15	2.60
0.45	2.45
0.24	2.20
0.16	2.00
1.00	

What is the expected value of BankCorp's EPS for the current fiscal year?

Following the definition of expected value, list each outcome, weight it by its probability, and sum the terms.

$$E(\text{EPS}) = 0.15(\$2.60) + 0.45(\$2.45) + 0.24(\$2.20) + 0.16(\$2.00) = \$2.3405$$

The expected value of EPS is $2.34.

An equation that summarizes your calculation in Example 10 is

$$E(X) = P(X_1)X_1 + P(X_2)X_2 + \ldots + P(X_n)X_n = \sum_{i=1}^{n} P(X_i)X_i \quad (7)$$

where X_i is one of n possible outcomes of the random variable X.

The expected value is our forecast. Because we are discussing random quantities, we cannot count on an individual forecast being realized (although we hope that, on average, forecasts will be accurate). It is important, as a result, to measure the risk we face. Variance and standard deviation measure the dispersion of outcomes around the expected value or forecast.

- Definition of Variance. The **variance** of a random variable is the expected value (the probability-weighted average) of squared deviations from the random variable's expected value:

  $$\sigma^2(X) = E\left\{\left[X - E(X)\right]^2\right\} \tag{8}$$

  The two notations for variance are $\sigma^2(X)$ and $\mathrm{Var}(X)$.

Variance is a number greater than or equal to 0 because it is the sum of squared terms. If variance is 0, there is no dispersion or risk. The outcome is certain, and the quantity X is not random at all. Variance greater than 0 indicates dispersion of outcomes. Increasing variance indicates increasing dispersion, all else equal. Variance of X is a quantity in the squared units of X. For example, if the random variable is return in percent, variance of return is in units of percent squared. Standard deviation is easier to interpret than variance because it is in the same units as the random variable. If the random variable is return in percent, standard deviation of return is also in units of percent. In the following example, when the variance of returns is stated as a percent or amount of money, to conserve space, we may suppress showing the unit squared.

- Definition of Standard Deviation. **Standard deviation** is the positive square root of variance.

The best way to become familiar with these concepts is to work examples.

EXAMPLE 11

BankCorp's Earnings per Share (4)

In Example 10, you calculated the expected value of BankCorp's EPS as \$2.34, which is your forecast. Using the probability distribution of EPS from Exhibit 6, you want to measure the dispersion around your forecast. What are the variance and standard deviation of BankCorp's EPS for the current fiscal year?

The order of calculation is always expected value, then variance, then standard deviation. Expected value has already been calculated. Following the definition of variance above, calculate the deviation of each outcome from the mean or expected value, square each deviation, weight (multiply) each squared deviation by its probability of occurrence, and then sum these terms.

$$\begin{aligned}\sigma^2(\text{EPS}) &= P(\$2.60)\left[\$2.60 - E(\text{EPS})\right]^2 + P(\$2.45)\left[\$2.45 - E(\text{EPS})\right]^2\\ &\quad + P(\$2.20)\left[\$2.20 - E(\text{EPS})\right]^2 + P(\$2.00)\left[\$2.00 - E(\text{EPS})\right]^2\\ &= 0.15(2.60 - 2.34)^2 + 0.45(2.45 - 2.34)^2\\ &\quad + 0.24(2.20 - 2.34)^2 + 0.16(2.00 - 2.34)^2\\ &= 0.01014 + 0.005445 + 0.004704 + 0.018496 = 0.038785\end{aligned}$$

Standard deviation is the positive square root of 0.038785:

$$\sigma(\text{EPS}) = 0.038785^{1/2} = 0.196939, \text{ or approximately } 0.20.$$

An equation that summarizes your calculation of variance in Example 11 is

$$\sigma^2(X) = P(X_1)[X_1 - E(X)]^2 + P(X_2)[X_2 - E(X)]^2 + \ldots + P(X_n)[X_n - E(X)]^2 = \sum_{i=1}^{n} P(X_i)[X_i - E(X)]^2 \quad (9)$$

where X_i is one of n possible outcomes of the random variable X.

i. explain the use of conditional expectation in investment applications

In investments, we make use of any relevant information available in making our forecasts. When we refine our expectations or forecasts, we are typically making adjustments based on new information or events; in these cases, we are using **conditional expected values**. The expected value of a random variable X given an event or scenario S is denoted $E(X \mid S)$. Suppose the random variable X can take on any one of n distinct outcomes $X_1, X_2, \ldots, X_n$ (these outcomes form a set of mutually exclusive and exhaustive events). The expected value of X conditional on S is the first outcome, X_1, times the probability of the first outcome given S, $P(X_1 \mid S)$, plus the second outcome, X_2, times the probability of the second outcome given S, $P(X_2 \mid S)$, and so forth.

$$E(X \mid S) = P(X_1 \mid S)X_1 + P(X_2 \mid S)X_2 + \ldots + P(X_n \mid S)X_n \quad (10)$$

We will illustrate this equation shortly.

Parallel to the total probability rule for stating unconditional probabilities in terms of conditional probabilities, there is a principle for stating (unconditional) expected values in terms of conditional expected values. This principle is the **total probability rule for expected value**.

- **Total Probability Rule for Expected Value.**

$$E(X) = E(X \mid S)P(S) + E(X \mid S^C)P(S^C) \quad (11)$$

$$E(X) = E(X \mid S_1)P(S_1) + E(X \mid S_2)P(S_2) + \ldots + E(X \mid S_n)P(S_n) \quad (12)$$

where $S_1, S_2, \ldots, S_n$ are mutually exclusive and exhaustive scenarios or events.

The general case, Equation 12, states that the expected value of X equals the expected value of X given Scenario 1, $E(X \mid S_1)$, times the probability of Scenario 1, $P(S_1)$, plus the expected value of X given Scenario 2, $E(X \mid S_2)$, times the probability of Scenario 2, $P(S_2)$, and so forth.

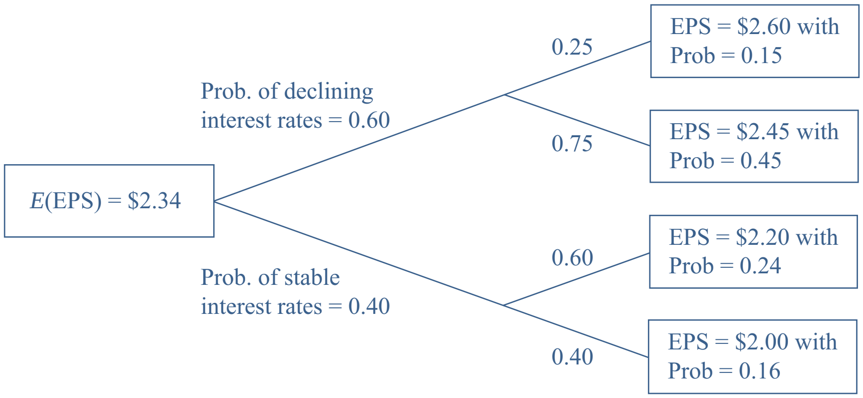
To use this principle, we formulate mutually exclusive and exhaustive scenarios that are useful for understanding the outcomes of the random variable. This approach was employed in developing the probability distribution of BankCorp's EPS in Examples 10 and 11, as we now discuss.

j interpret a probability tree and demonstrate its application to investment problems

The earnings of BankCorp are interest rate sensitive, benefiting from a declining interest rate environment. Suppose there is a 0.60 probability that BankCorp will operate in a *declining interest rate environment* in the current fiscal year and a 0.40 probability that it will operate in a *stable interest rate environment* (assessing the chance of an increasing interest rate environment as negligible). If a *declining interest rate environment* occurs, the probability that EPS will be \$2.60 is estimated at 0.25, and the probability that EPS will be \$2.45 is estimated at 0.75. Note that 0.60, the probability of *declining interest rate environment,* times 0.25, the probability of \$2.60 EPS given a *declining interest rate environment,* equals 0.15, the (unconditional) probability of

$2.60 given in the table in Exhibit 7. The probabilities are consistent. Also, 0.60(0.75) = 0.45, the probability of $2.45 EPS given in Exhibit 7. The **probability tree diagram** in Exhibit 8 shows the rest of the analysis.

Exhibit 8 BankCorp's Forecasted EPS


E(EPS) = $2.34
Prob. of declining interest rates = 0.60
Prob. of stable interest rates = 0.40
0.25
0.75
0.60
0.40
EPS = $2.60 with Prob = 0.15
EPS = $2.45 with Prob = 0.45
EPS = $2.20 with Prob = 0.24
EPS = $2.00 with Prob = 0.16

A declining interest rate environment points us to the **node** of the tree that branches off into outcomes of $2.60 and $2.45. We can find expected EPS given a declining interest rate environment as follows, using Equation 10:

$$\begin{aligned} E(\text{EPS} \mid \textit{declining interest rate environment}) &= 0.25(\$2.60) + 0.75(\$2.45) \\ &= \$2.4875 \end{aligned}$$

If interest rates are stable,

$$\begin{aligned} E(\text{EPS} \mid \textit{stable interest rate environment}) &= 0.60(\$2.20) + 0.40(\$2.00) \\ &= \$2.12 \end{aligned}$$

Once we have the new piece of information that interest rates are stable, for example, we revise our original expectation of EPS from $2.34 downward to $2.12. Now using the total probability rule for expected value,

$$\begin{aligned} &\text{E}(\text{EPS}) \\ &= \text{E}(\text{EPS} \mid \text{declining interest rate environment})\text{P}(\text{declining interest rateenvironment}) \\ &+ \text{E}(\text{EPS} \mid \text{stable interest rate environment})\text{P}(\text{stable interest rateenvironment}) \end{aligned}$$

So, E(EPS) = $2.4875(0.60) + $2.12(0.40) = $2.3405 or about $2.34.

This amount is identical to the estimate of the expected value of EPS calculated directly from the probability distribution in Example 10. Just as our probabilities must be consistent, so must our expected values, unconditional and conditional; otherwise our investment actions may create profit opportunities for other investors at our expense.

To review, we first developed the factors or scenarios that influence the outcome of the event of interest. After assigning probabilities to these scenarios, we formed expectations conditioned on the different scenarios. Then we worked backward to formulate an expected value as of today. In the problem just worked, EPS was the event of interest, and the interest rate environment was the factor influencing EPS.

We can also calculate the variance of EPS given each scenario:

$$\begin{aligned}&\sigma^2(\text{EPS} \mid \textit{declining interest rate environment})\\&\quad= P(\$2.60 \mid \textit{declining interest rate environment})\\&\quad\times\left[\$2.60 - E(\text{EPS} \mid \textit{declining interest rate environment})\right]^2\\&\quad+P(\$2.45 \mid \textit{declining interest rate environment})\\&\quad\times\left[\$2.45 - E(\text{EPS} \mid \textit{declining interest rate environment})\right]^2\\&\quad= 0.25(\$2.60 - \$2.4875)^2 + 0.75(\$2.45 - \$2.4875)^2\\&\quad= 0.004219\end{aligned}$$

Similarly, σ^2(EPS | *stable interest rate environment*) is found to be equal to

$$= 0.60(\$2.20 - \$2.12)^2 + 0.40(\$2.00 - \$2.12)^2 = 0.0096$$

These are **conditional variances**, the variance of EPS given a *declining interest rate environment* and the variance of EPS given a *stable interest rate environment*. The relationship between unconditional variance and conditional variance is a relatively advanced topic. The main points are 1) that variance, like expected value, has a conditional counterpart to the unconditional concept and 2) that we can use conditional variance to assess risk given a particular scenario.

EXAMPLE 12

BankCorp's Earnings per Share (5)

Continuing with BankCorp, you focus now on BankCorp's cost structure. One model, a simple linear regression model, you are researching for BankCorp's operating costs is

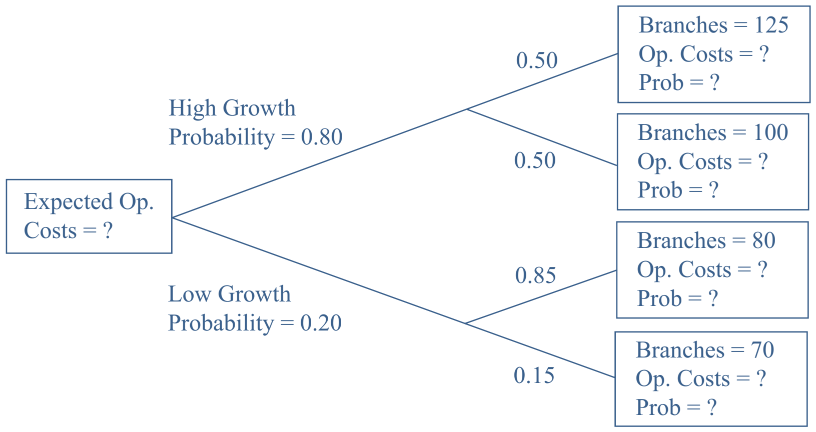
$$\hat{Y} = a + bX$$

where $\hat{Y}$ is a forecast of operating costs in millions of dollars and X is the number of branch offices. $\hat{Y}$ represents the expected value of Y given X, or $E(Y \mid X)$. You interpret the intercept a as fixed costs and b as variable costs. You estimate the equation as

$$\hat{Y} = 12.5 + 0.65X$$

BankCorp currently has 66 branch offices, and the equation estimates operating costs as 12.5 + 0.65(66) = \$55.4 million. You have two scenarios for growth, pictured in the tree diagram in Exhibit 9.

Exhibit 9 BankCorp's Forecasted Operating Costs


Expected Op. Costs = ?
High Growth Probability = 0.80
Low Growth Probability = 0.20
0.50
0.50
0.85
0.15
Branches = 125 Op. Costs = ? Prob = ?
Branches = 100 Op. Costs = ? Prob = ?
Branches = 80 Op. Costs = ? Prob = ?
Branches = 70 Op. Costs = ? Prob = ?

1. Compute the forecasted operating costs given the different levels of operating costs, using $\hat{Y} = 12.5 + 0.65X$. State the probability of each level of the number of branch offices. These are the answers to the questions in the terminal boxes of the tree diagram.
2. Compute the expected value of operating costs under the high growth scenario. Also calculate the expected value of operating costs under the low growth scenario.
3. Answer the question in the initial box of the tree: What are BankCorp's expected operating costs?

Solution to 1:

Using $\hat{Y} = 12.5 + 0.65X$, from top to bottom, we have

Operating Costs	Probability
$\hat{Y} = 12.5 + 0.65(125) = \$93.75$ million	$0.80(0.50) = 0.40$
$\hat{Y} = 12.5 + 0.65(100) = \$77.50$ million	$0.80(0.50) = 0.40$
$\hat{Y} = 12.5 + 0.65(80) = \$64.50$ million	$0.20(0.85) = 0.17$
$\hat{Y} = 12.5 + 0.65(70) = \$58.00$ million	$0.20(0.15) = 0.03$
	Sum = 1.00

Solution to 2:

Dollar amounts are in millions.

$$E(\text{operating costs} \mid \textit{high growth}) = 0.50(\$93.75) + 0.50(\$77.50)$$
$$= \$85.625$$

$$E(\text{operating costs} \mid \textit{low growth}) = 0.85(\$64.50) + 0.15(\$58.00)$$
$$= \$63.525$$

Solution to 3:

Dollar amounts are in millions.

$$\begin{aligned} E(\text{operating costs}) &= E(\text{operating costs} \mid \textit{high growth})P(\textit{high growth}) \\ &\quad + E(\text{operating costs} \mid \textit{low growth})P(\textit{low growth}) \\ &= \$85.625(0.80) + \$63.525(0.20) = \$81.205 \end{aligned}$$

BankCorp's expected operating costs are $81.205 million.

In this section, we have treated random variables such as EPS as standalone quantities. We have not explored how descriptors such as expected value and variance of EPS may be functions of other random variables. Portfolio return is one random variable that is clearly a function of other random variables, the random returns on the individual securities in the portfolio. To analyze a portfolio's expected return and variance of return, we must understand these quantities are a function of characteristics of the individual securities' returns. Looking at the variance of portfolio return, we see that the way individual security returns move together or covary is key. So, next we cover portfolio expected return, variance of return, and importantly, covariance and correlation.

EXPECTED VALUE, VARIANCE, STANDARD DEVIATION, COVARIANCES, AND CORRELATIONS OF PORTFOLIO RETURNS

4

k calculate and interpret the expected value, variance, standard deviation, covariances, and correlations of portfolio returns

Modern portfolio theory makes frequent use of the idea that investment opportunities can be evaluated using expected return as a measure of reward and variance of return as a measure of risk. In this section, we will develop an understanding of portfolio expected return and variance of return, which are functions of the returns on the individual portfolio holdings. To begin, the expected return on a portfolio is a weighted average of the expected returns on the securities in the portfolio, using exactly the same weights. When we have estimated the expected returns on the individual securities, we immediately have portfolio expected return.

- **Calculation of Portfolio Expected Return.** Given a portfolio with n securities, the **expected return on the portfolio** ($E(Rp)$) is a weighted average of the expected returns (R_1 to R_n) on the component securities using their respective weights (w_1 to w_n):

$$\begin{aligned} E(R_p) &= E(w_1R_1 + w_2R_2 + \ldots + w_nR_n) \\ &= w_1E(R_1) + w_2E(R_2) + \ldots + w_nE(R_n) \end{aligned} \quad (13)$$

Suppose we have estimated expected returns on assets in the three-asset portfolio shown in Exhibit 10.

Exhibit 10 Weights and Expected Returns

Asset Class	Weight	Expected Return (%)
S&P 500	0.50	13
US long-term corporate bonds	0.25	6
MSCI EAFE	0.25	15

We calculate the expected return on the portfolio as 11.75%:

$$\begin{aligned} E\left(R_p\right) &= w_1 E\left(R_1\right) + w_2 E\left(R_2\right) + w_3 E\left(R_3\right) \\ &= 0.50\left(13\%\right) + 0.25\left(6\%\right) + 0.25\left(15\%\right) = 11.75\% \end{aligned}$$

Here we are interested in portfolio variance of return as a measure of investment risk. Accordingly, portfolio variance is $\sigma^2(R_p) = E\{[R_p - E(R_p)]^2\}$, which is variance in a forward-looking sense. To implement this definition of portfolio variance, we use information about the individual assets in the portfolio, but we also need the concept of covariance. To avoid notational clutter, we write ER_p for $E(R_p)$.

- Definition of Covariance. Given two random variables R_i and R_j, the **covariance** between R_i and R_j is

$$\mathrm{Cov}\left(R_i, R_j\right) = E\left[\left(R_i - ER_i\right)\left(R_j - ER_j\right)\right] \tag{14}$$

Alternative notations are $\sigma(R_i,R_j)$ and σ_{ij}. Equation 14 states that the covariance between two random variables is the probability-weighted average of the cross-products of each random variable's deviation from its own expected value. The above measure is the population covariance and is forward-looking. The sample covariance between two random variables R_i and R_j, based on a sample of past data of size n is

$$\mathrm{Cov}\left(R_i, R_j\right) = \sum_{n=1}^{n}\left(R_{i,t} - \bar{R}_i\right)\left(R_{j,t} - \bar{R}_j\right)\Big/\left(n-1\right) \tag{15}$$

Start with the definition of variance for a three-asset portfolio and see how it decomposes into three variance terms and six covariance terms. Dispensing with the derivation, the result is Equation 16:

$$\begin{aligned} \sigma^2\left(R_p\right) &= E\left[\left(R_p - ER_p\right)^2\right] \\ &= E\left\{\left[w_1R_1 + w_2R_2 + w_3R_3 - E\left(w_1R_1 + w_2R_2 + w_3R_3\right)\right]^2\right\} \\ &= E\left\{\left[w_1R_1 + w_2R_2 + w_3R_3 - w_1ER_1 - w_2ER_2 - w_3ER_3\right]^2\right\} \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left(\text{using Equation 13}\right) \end{aligned}$$

$$\begin{aligned} &= w_1^2\sigma^2\left(R_1\right) + w_1w_2\mathrm{Cov}\left(R_1,R_2\right) + w_1w_3\mathrm{Cov}\left(R_1,R_3\right) \\ &\quad + w_1w_2Cov\left(R_1,R_2\right) + w_2^2\sigma^2\left(R_2\right) + w_2w_3\mathrm{Cov}\left(R_2,R_3\right) \\ &\quad + w_1w_3Cov\left(R_1,R_3\right) + w_2w_3Cov\left(R_2,R_3\right) + w_3^2\sigma^2\left(R_3\right) \end{aligned} \tag{16}$$

Noting that the order of variables in covariance does not matter, for example, $\text{Cov}(R_2,R_1) = \text{Cov}(R_1,R_2)$, and that diagonal variance terms $\sigma^2(R_1)$, $\sigma^2(R_2)$, and $\sigma^2(R_3)$ can be expressed as $\text{Cov}(R_1,R_1)$, $\text{Cov}(R_2,R_2)$, and $\text{Cov}(R_3,R_3)$, respectively, the most compact way to state Equation 16 is $\sigma^2\left(R_p\right) = \sum_{i=1}^{3}\sum_{j=1}^{3} w_i w_j \text{Cov}\left(R_i, R_j\right)$. Moreover, this expression generalizes for a portfolio of any size n to

$$\sigma^2\left(R_p\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \text{Cov}\left(R_i, R_j\right) \tag{17}$$

We see from Equation 16 that individual variances of return constitute part, but not all, of portfolio variance. The three variances are actually outnumbered by the six covariance terms off the diagonal. If there are 20 assets, there are 20 variance terms and 20(20) − 20 = 380 off-diagonal covariance terms. A first observation is that as the number of holdings increases, covariance becomes increasingly important, all else equal.

The covariance terms capture how the co-movements of returns affect portfolio variance. From the definition of covariance, we can establish two essential observations about covariance.

1. We can interpret the sign of covariance as follows:

   Covariance of returns is negative if, when the return on one asset is above its expected value, the return on the other asset tends to be below its expected value (an average inverse relationship between returns).

   Covariance of returns is 0 if returns on the assets are unrelated.

   Covariance of returns is positive when the returns on both assets tend to be on the same side (above or below) their expected values at the same time (an average positive relationship between returns).

2. The covariance of a random variable with itself (*own covariance*) is its own variance: $\text{Cov}(R,R) = E\{[R - E(R)][R - E(R)]\} = E\{[R - E(R)]^2\} = \sigma^2(R)$.

Exhibit 11 summarizes the inputs for portfolio expected return and variance of return. A complete list of the covariances constitutes all the statistical data needed to compute portfolio variance of return as shown in the **covariance matrix** in Panel B.

Exhibit 11 Inputs to Portfolio Expected Return and Variance

A. Inputs to Portfolio Expected Return

Asset	A	B	C
	$E(R_A)$	$E(R_B)$	$E(R_C)$

B. Covariance Matrix: The Inputs to Portfolio Variance of Return

Asset	A	B	C
A	**$\text{Cov}(R_A,R_A)$**	$\text{Cov}(R_A,R_B)$	$\text{Cov}(R_A,R_C)$
B	$\text{Cov}(R_B,R_A)$	**$\text{Cov}(R_B,R_B)$**	$\text{Cov}(R_B,R_C)$
C	$\text{Cov}(R_C,R_A)$	$\text{Cov}(R_C,R_B)$	**$\text{Cov}(R_C,R_C)$**

With three assets, the covariance matrix has $3^2 = 3 \times 3 = 9$ entries, but the diagonal terms, the variances (bolded in Exhibit 11), are treated separately from the off-diagonal terms. So there are 9 − 3 = 6 covariances, excluding variances. But $\text{Cov}(R_B,R_A) =$

$\text{Cov}(R_A,R_B)$, $\text{Cov}(R_C,R_A) = \text{Cov}(R_A,R_C)$, and $\text{Cov}(R_C,R_B) = \text{Cov}(R_B,R_C)$. The covariance matrix below the diagonal is the mirror image of the covariance matrix above the diagonal, so you only need to use one (i.e., either below or above the diagonal). As a result, there are only 6/2 = 3 distinct covariance terms to estimate. In general, for n securities, there are $n(n - 1)/2$ distinct covariances and n variances to estimate.

Suppose we have the covariance matrix shown in Exhibit 12. We will be working in returns stated as percents, and the table entries are in units of percent squared ($\%^2$). The terms $38\%^2$ and $400\%^2$ are 0.0038 and 0.0400, respectively, stated as decimals; correctly working in percents and decimals leads to identical answers.

Exhibit 12 Covariance Matrix

	S&P 500	US Long-Term Corporate Bonds	MSCI EAFE
S&P 500	400	45	189
US long-term corporate bonds	45	81	38
MSCI EAFE	189	38	441

Taking Equation 16 and grouping variance terms together produces the following:

$$\begin{aligned}\sigma^2\left(R_p\right) &= w_1^2\sigma^2\left(R_1\right) + w_2^2\sigma^2\left(R_2\right) + w_3^2\sigma^2\left(R_3\right) + 2w_1w_2\text{Cov}\left(R_1,R_2\right)\\ &\quad + 2w_1w_3\text{Cov}\left(R_1,R_3\right) + 2w_2w_3\text{Cov}\left(R_2,R_3\right)\\ &= \left(0.50\right)^2\left(400\right) + \left(0.25\right)^2\left(81\right) + \left(0.25\right)^2\left(441\right)\\ &\quad + 2\left(0.50\right)\left(0.25\right)\left(45\right) + 2\left(0.50\right)\left(0.25\right)\left(189\right)\\ &\quad + 2\left(0.25\right)\left(0.25\right)\left(38\right)\\ &= 100 + 5.0625 + 27.5625 + 11.25 + 47.25 + 4.75 = 195.875\end{aligned} \tag{18}$$

The variance is 195.875. Standard deviation of return is $195.875^{1/2}$ = 14%. To summarize, the portfolio has an expected annual return of 11.75% and a standard deviation of return of 14%.

Looking at the first three terms in the calculation above, their sum (100 + 5.0625 + 27.5625) is 132.625, the contribution of the individual variances to portfolio variance. If the returns on the three assets were independent, covariances would be 0 and the standard deviation of portfolio return would be $132.625^{1/2}$ = 11.52% as compared to 14% before, so a less risky portfolio. If the covariance terms were negative, then a negative number would be added to 132.625, so portfolio variance and risk would be even smaller, while expected return would not change. For the same expected portfolio return, the portfolio has less risk. This risk reduction is a diversification benefit, meaning a risk-reduction benefit from holding a portfolio of assets. The diversification benefit increases with decreasing covariance. This observation is a key insight of modern portfolio theory. This insight is even more intuitively stated when we can use the concept of correlation.

- Definition of Correlation. The **correlation** between two random variables, R_i and R_j, is defined as $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/[\sigma(R_i)\sigma(R_j)]$. Alternative notations are $\text{Corr}(R_i,R_j)$ and ρ_{ij}.

The above definition of correlation is forward-looking because it involves dividing the forward-looking covariance by the product of forward-looking standard deviations. Frequently, covariance is substituted out using the relationship $\text{Cov}(R_i,R_j) = \rho(R_i,R_j)\sigma(R_i)\sigma(R_j)$. Like covariance, the correlation coefficient is a measure of linear association.

However, the division in the definition makes correlation a pure number (without a unit of measurement) and places bounds on its largest and smallest possible values, which are +1 and −1, respectively.

If two variables have a strong positive linear relation, then their correlation will be close to +1. If two variables have a strong negative linear relation, then their correlation will be close to −1. If two variables have a weak linear relation, then their correlation will be close to 0. Using the above definition, we can state a correlation matrix from data in the covariance matrix alone. Exhibit 13 shows the correlation matrix.

Exhibit 13 Correlation Matrix of Returns

	S&P 500	US Long-Term Corporate Bonds	MSCI EAFE
S&P 500	1.00	0.25	0.45
US long-term corporate bonds	0.25	1.00	0.20
MSCI EAFE	0.45	0.20	1.00

For example, the covariance between long-term bonds and MSCI EAFE is 38, from Exhibit 12. The standard deviation of long-term bond returns is $81^{1/2}$ = 9%, that of MSCI EAFE returns is $441^{1/2}$ = 21%, from diagonal terms in Exhibit 12. The correlation $\rho(R_{\text{long-term bonds}}, R_{\text{EAFE}})$ is 38/[(9%)(21%)] = 0.201, rounded to 0.20. The correlation of the S&P 500 with itself equals 1: The calculation is its own covariance divided by its standard deviation squared.

EXAMPLE 13

Portfolio Expected Return and Variance of Return with Varying Portfolio Weights

Anna Cintara is constructing different portfolios from the following two stocks:

Exhibit 14 Description of Two-Stock Portfolio

	Stock 1	Stock 2
Expected return	4%	8%
Standard deviation	6%	15%
Current portfolio weights	0.40	0.60
Correlation between returns	0.30	

1 Calculate the covariance between the returns on the two stocks.

2 What is the portfolio expected return and standard deviation if Cintara puts 100% of her investment in Stock 1 (w_1 = 1.00 and w_2 = 0.00)? What is the portfolio expected return and standard deviation if Cintara puts 100% of her investment in Stock 2 (w_1 = 0.00 and w_2 = 1.00)?

3 What are the portfolio expected return and standard deviation using the current portfolio weights?

4 Calculate the expected return and standard deviation of the portfolios when w_1 goes from 0.00 to 1.00 in 0.10 increments (and $w_2 = 1 - w_1$). Place the results (stock weights, portfolio expected return, and portfolio standard deviation) in a table, and then sketch a graph of the results with the standard deviation on the horizontal axis and expected return on the vertical axis.

Solution to 1:

The correlation between two stock returns is $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/[\sigma(R_i)\ \sigma(R_j)]$, so the covariance is $\text{Cov}(R_i,R_j) = \rho(R_i,R_j)\ \sigma(R_i)\ \sigma(R_j)$. For these two stocks, the covariance is $\text{Cov}(R_1,R_2) = \rho(R_1,R_2)\ \sigma(R_1)\ \sigma(R_2) = 0.30\ (6)\ (15) = 27$.

Solution to 2:

If the portfolio is 100% invested in Stock 1, the portfolio has an expected return of 4% and a standard deviation of 6%. If the portfolio is 100% invested in Stock 2, the portfolio has an expected return of 8% and a standard deviation of 15%.

Solution to 3:

For the current 40/60 portfolio, the expected return is

$$E(R_p) = w_1 E(R_1) + (1 - w_1)E(R_2) = 0.40(4\%) + 0.60(8\%) = 6.4\%$$

The portfolio variance and standard deviation are

$$\begin{aligned}\sigma^2(R_p) &= w_1^2\sigma^2(R_1) + w_2^2\sigma^2(R_2) + 2w_1w_2\text{Cov}(R_1,R_2)\\ &= (0.40)^2(36) + (0.60)^2(225) + 2(0.40)(0.60)(27)\\ &= 5.76 + 81 + 12.94 = 99.72\end{aligned}$$

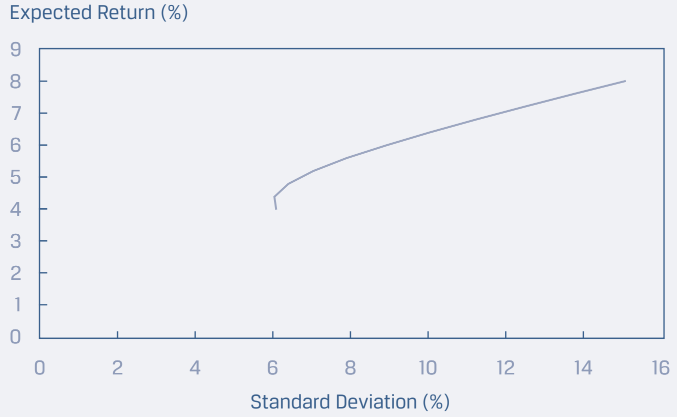
$$\sigma(R_p) = 99.72^{1/2} = 9.99\%$$

Solution to 4:

The portfolio expected returns, variances, and standard deviations for the different sets of portfolio weights are given in the following table. Three of the rows are already computed in the Solutions to 2 and 3, and the other rows are computed using the same expected return, variance, and standard deviation formulas as in the Solution to 3:

Stock 1 weight	Stock 2 weight	Expected return (%)	Variance (%^{2})	Standard deviation (%)
1.00	0.00	4.00	36.00	6.00
0.90	0.10	4.40	36.27	6.02
0.80	0.20	4.80	40.68	6.38
0.70	0.30	5.20	49.23	7.02
0.60	0.40	5.60	61.92	7.87
0.50	0.50	6.00	78.75	8.87
0.40	0.60	6.40	99.72	9.99
0.30	0.70	6.80	124.83	11.17
0.20	0.80	7.20	154.08	12.41
0.10	0.90	7.60	187.47	13.69
0.00	1.00	8.00	225.00	15.00

The graph of the expected return and standard deviation is



COVARIANCE GIVEN A JOINT PROBABILITY FUNCTION

5

l. calculate and interpret the covariances of portfolio returns using the joint probability function

How do we estimate return covariance and correlation? Frequently, we make forecasts on the basis of historical covariance or use other methods such as a market model regression based on historical return data. We can also calculate covariance using the **joint probability function** of the random variables, if that can be estimated. The joint probability function of two random variables X and Y, denoted $P(X,Y)$, gives the probability of joint occurrences of values of X and Y. For example, $P(X=3, Y=2)$, is the probability that X equals 3 and Y equals 2.

Suppose that the joint probability function of the returns on BankCorp stock (R_A) and the returns on NewBank stock (R_B) has the simple structure given in Exhibit 15.

Exhibit 15 Joint Probability Function of BankCorp and NewBank Returns (Entries Are Joint Probabilities)

	R_B = 20%	R_B = 16%	R_B = 10%
R_A = 25%	0.20	0	0
R_A = 12%	0	0.50	0
R_A = 10%	0	0	0.30

The expected return on BankCorp stock is 0.20(25%) + 0.50(12%) + 0.30(10%) = 14%. The expected return on NewBank stock is 0.20(20%) + 0.50(16%) + 0.30(10%) = 15%. The joint probability function above might reflect an analysis based on whether banking industry conditions are good, average, or poor. Exhibit 16 presents the calculation of covariance.

Exhibit 16 Covariance Calculations

Banking Industry Condition	Deviations BankCorp	Deviations NewBank	Product of Deviations	Probability of Condition	Probability-Weighted Product
Good	25–14	20–15	55	0.20	11
Average	12–14	16–15	−2	0.50	−1
Poor	10–14	10–15	20	0.30	6
					$\text{Cov}(R_A,R_B) = 16$

Note: Expected return for BankCorp is 14% and for NewBank, 15%.

The first and second columns of numbers show, respectively, the deviations of BankCorp and NewBank returns from their mean or expected value. The next column shows the product of the deviations. For example, for good industry conditions, (25 − 14)(20 − 15) = 11(5) = 55. Then, 55 is multiplied or weighted by 0.20, the probability that banking industry conditions are good: 55(0.20) = 11. The calculations for average and poor banking conditions follow the same pattern. Summing up these probability-weighted products, we find $\text{Cov}(R_A,R_B) = 16$.

A formula for computing the covariance between random variables R_A and R_B is

$$\text{Cov}\left(R_A,R_B\right) = \sum_i \sum_j P\left(R_{A,i},R_{B,j}\right)\left(R_{A,i} - ER_A\right)\left(R_{B,j} - ER_B\right) \tag{19}$$

The formula tells us to sum all possible deviation cross-products weighted by the appropriate joint probability.

Next, we take note of the fact that when two random variables are independent, their joint probability function simplifies.

- Definition of Independence for Random Variables. Two random variables X and Y are **independent** if and only if $P(X,Y) = P(X)P(Y)$.

For example, given independence, $P(3,2) = P(3)P(2)$. We multiply the individual probabilities to get the joint probabilities. *Independence* is a stronger property than *uncorrelatedness* because correlation addresses only linear relationships. The following condition holds for independent random variables and, therefore, also holds for uncorrelated random variables.

- **Multiplication Rule for Expected Value of the Product of Uncorrelated Random Variables**. The expected value of the product of uncorrelated random variables is the product of their expected values.

  $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated.

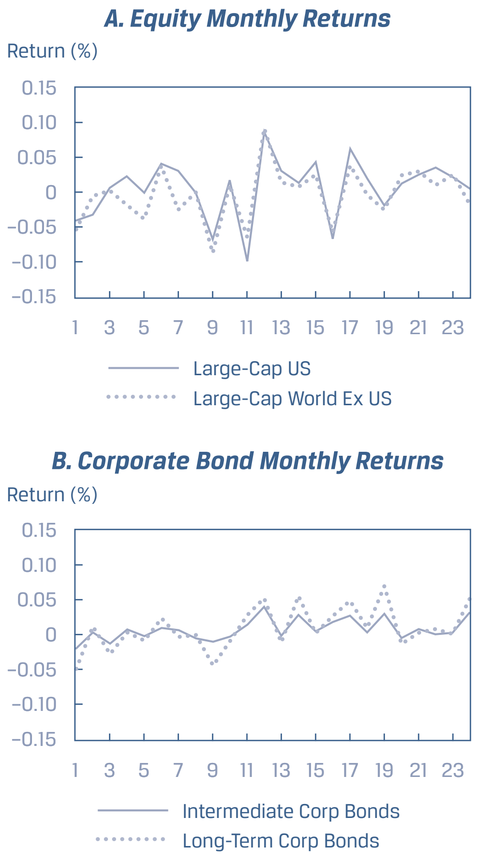
Many financial variables, such as revenue (price times quantity), are the product of random quantities. When applicable, the above rule simplifies calculating expected value of a product of random variables.

EXAMPLE 14

Covariances and Correlations of Security Returns

Isabel Vasquez is reviewing the correlations between four of the asset classes in her company portfolio. In Exhibit 17, she plots 24 recent monthly returns for large-cap US stocks versus for large-cap world ex-US stocks (Panel 1) and the 24 monthly returns for intermediate-term corporate bonds versus long-term corporate bonds (Panel 2). Vasquez presents the returns, variances, and covariances in decimal form instead of percentage form. Note the different ranges of their vertical axes (Return %).

Exhibit 17 Monthly Returns for Four Asset Classes



Selected data for the four asset classes are shown in Exhibit 18.

Exhibit 18 Selected Data for Four Asset Classes

Asset Classes	Large-Cap US Equities	World (ex US) Equities	Intermediate Corp Bonds	Long-Term Corp Bonds
Variance	0.001736	0.001488	0.000174	0.000699
Standard deviation	0.041668	0.038571	0.013180	0.026433
Covariance	0.001349		0.000318	
Correlation	0.87553		0.95133	

Vasquez noted, as shown in Exhibit 18, that although the two equity classes had much greater variances and covariance than the two bond classes, the correlation between the two equity classes was lower than the correlation between the two bond classes. She also noted that although long-term bonds were more volatile (higher variance) than intermediate-term bonds, long- and intermediate-term bond returns still had a high correlation.

6 BAYES' FORMULA

m calculate and interpret an updated probability using Bayes' formula

A topic that is often useful in solving investment problems is Bayes' formula: what probability theory has to say about learning from experience.

6.1 Bayes' Formula

When we make decisions involving investments, we often start with viewpoints based on our experience and knowledge. These viewpoints may be changed or confirmed by new knowledge and observations. **Bayes' formula** is a rational method for adjusting our viewpoints as we confront new information. Bayes' formula and related concepts have been applied in many business and investment decision-making contexts.

Bayes' formula makes use of Equation 6, the total probability rule. To review, that rule expressed the probability of an event as a weighted average of the probabilities of the event, given a set of scenarios. Bayes' formula works in reverse; more precisely, it reverses the "given that" information. Bayes' formula uses the occurrence of the event to infer the probability of the scenario generating it. For that reason, Bayes' formula is sometimes called an inverse probability. In many applications, including those illustrating its use in this section, an individual is updating his/her beliefs concerning the causes that may have produced a new observation.

- **Bayes' Formula.** Given a set of prior probabilities for an event of interest, if you receive new information, the rule for updating your probability of the event is

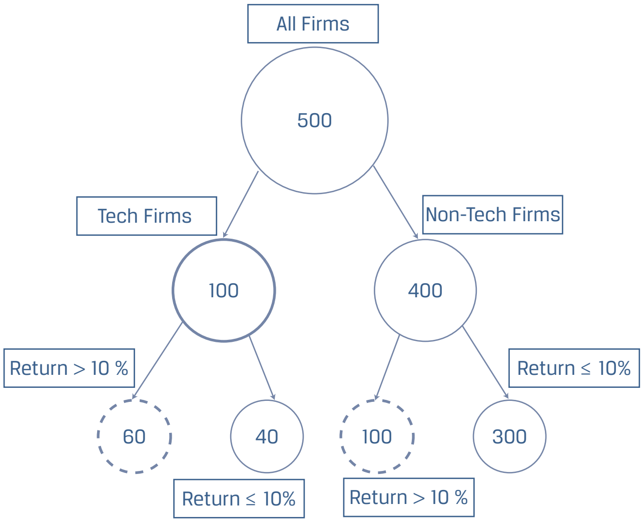
  $$\text{Updated probability of event given the new information} = \frac{\text{Probability of the new information given event}}{\text{Unconditional probability of the new information}} \times \text{Prior probability of event}$$

  In probability notation, this formula can be written concisely as

  $$P(\text{Event} \mid \text{Information}) = \frac{P(\text{Information} \mid \text{Event})}{P(\text{Information})} P(\text{Event}) \tag{20}$$

Consider the following example using frequencies—which may be more straightforward initially than probabilities—for illustrating and understanding Bayes' formula. Assume a hypothetical large-cap stock index has 500 member firms, of which 100 are technology firms, and 60 of these had returns of > 10%, and 40 had returns of ≤ 10%. Of the 400 non-technology firms in the index, 100 had returns of > 10%, and 300 had returns of ≤ 10%. The tree map in Exhibit 19 is useful for visualizing this example, which is summarized in the table in Exhibit 20.

Exhibit 19 Tree Map for Visualizing Bayes' Formula Using Frequencies



P (*Tech* | *R* > 10%) = 60/(60+100)

Exhibit 20 Summary of Returns for Tech and Non-Tech Firms in Hypothetical Large-Cap Equity Index

	Type of Firm in Stock Index		
Rate of Return (R)	**Non-Tech**	**Tech**	**Total**
R > 10%	100	60	**160**
R ≤ 10%	300	40	**340**
Total	**400**	**100**	**500**

What is the probability a firm is a tech firm given that it has a return of > 10% or *P*(*tech* | *R > 10%)?* Looking at the frequencies in the tree map and in the table, we can see many empirical probabilities, such as the following:

- *P(tech)* = 100 / 500 = 0.20,
- *P*(*non-tech*) = 400 / 500 = 0.80,
- *P(R > 10%* | *tech)* = 60 / 100 = 0.60,
- *P*(*R > 10%* | *non-tech*) = 100 / 400 = 0.25,
- *P(R > 10%)* = 160 / 500 = 0.32, and, finally,
- *P*(*tech* | *R > 10%*) = 60/ 160 = 0.375. This probability is the answer to our initial question.

Without looking at frequencies, let us use Bayes' formula to find the probability that a firm has a return of > 10% and then the probability that a firm with a return of > 10% is a tech firm, *P*(*tech* | *R > 10%*). First,

$$P(R > 10\%) = P(R > 10\% \mid tech) \times P(tech) + P(R > 10\% \mid non\text{-}tech) \times P(non\text{-}tech)$$

$$= 0.60 \times 0.20 + 0.25 \times 0.80 = 0.32.$$

Now we can implement the Bayes' formula answer to our question:

$$P(tech \mid R > 10\%) = \frac{P(R > 10\% \mid tech) \times P(tech)}{P(R > 10\%)} = \frac{0.60 \times 0.20}{0.32} = 0.375$$

The probability that a firm with a return of > 10% is a tech firm is 0.375, which is impressive because the probability that a firm is a tech firm (from the whole sample) is only 0.20. In sum, it can be readily seen from the tree map and the underlying frequency data (Exhibits 19 and 20, respectively) or from the probabilities in Bayes' formula that there are 160 firms with R > 10%, and 60 of them are tech firms, so $P(tech \mid R > 10\%) = 60/160 = .375$.

Users of Bayesian statistics do not consider probabilities (or likelihoods) to be known with certainty but that these should be subject to modification whenever new information becomes available. Our beliefs or probabilities are continually updated as new information arrives over time.

To further illustrate Bayes' formula, we work through an investment example that can be adapted to any actual problem. Suppose you are an investor in the stock of DriveMed, Inc. Positive earnings surprises relative to consensus EPS estimates often result in positive stock returns, and negative surprises often have the opposite effect. DriveMed is preparing to release last quarter's EPS result, and you are interested in which of these three events happened: *last quarter's EPS exceeded the consensus EPS estimate, last quarter's EPS exactly met the consensus EPS estimate,* or *last quarter's EPS fell short of the consensus EPS estimate.* This list of the alternatives is mutually exclusive and exhaustive.

On the basis of your own research, you write down the following **prior probabilities** (or priors, for short) concerning these three events:

- $P(EPS\ exceeded\ consensus) = 0.45$
- $P(EPS\ met\ consensus) = 0.30$
- $P(EPS\ fell\ short\ of\ consensus) = 0.25$

These probabilities are "prior" in the sense that they reflect only what you know now, before the arrival of any new information.

The next day, DriveMed announces that it is expanding factory capacity in Singapore and Ireland to meet increased sales demand. You assess this new information. The decision to expand capacity relates not only to current demand but probably also to the prior quarter's sales demand. You know that sales demand is positively related to EPS. So now it appears more likely that last quarter's EPS will exceed the consensus.

The question you have is, "In light of the new information, what is the updated probability that the prior quarter's EPS exceeded the consensus estimate?"

Bayes' formula provides a rational method for accomplishing this updating. We can abbreviate the new information as *DriveMed expands.* The first step in applying Bayes' formula is to calculate the probability of the new information (here: *DriveMed expands*), given a list of events or scenarios that may have generated it. The list of events should cover all possibilities, as it does here. Formulating these conditional probabilities is the key step in the updating process. Suppose your view, based on research of DriveMed and its industry, is

$$P(DriveMed\ expands \mid EPS\ exceeded\ consensus) = 0.75$$

$$P(DriveMed\ expands \mid EPS\ met\ consensus) = 0.20$$

$$P(DriveMed\ expands \mid EPS\ fell\ short\ of\ consensus) = 0.05$$

Conditional probabilities of an observation (here: *DriveMed expands*) are sometimes referred to as **likelihoods**. Again, likelihoods are required for updating the probability.

Next, you combine these conditional probabilities or likelihoods with your prior probabilities to get the unconditional probability for DriveMed expanding, $P(DriveMed\ expands)$, as follows:

$$
\begin{aligned}
&P(DriveMed\ expands)\\
&= P(DriveMed\ expands \mid EPS\ exceeded\ consensus)\\
&\quad \times P(EPS\ exceeded\ consensus)\\
&+ P(DriveMed\ expands \mid EPS\ met\ consensus)\\
&\quad \times P(EPS\ met\ consensus)\\
&+ P(DriveMed\ expands \mid EPS\ fell\ short\ of\ consensus)\\
&\quad \times P(EPS\ fell\ short\ of\ consensus)\\
&= 0.75(0.45) + 0.20(0.30) + 0.05(0.25) = 0.41, \text{ or } 41\%
\end{aligned}
$$

This is Equation 6, the total probability rule, in action. Now you can answer your question by applying Bayes' formula:

$$
\begin{aligned}
&P(EPS\ exceeded\ consensus \mid DriveMed\ expands)\\
&= \frac{P(DriveMed\ expands \mid EPS\ exceeded\ consensus)}{P(DriveMed\ expands)} P(EPS\ exceeded\ consensus)\\
&= (0.75/0.41)(0.45) = 1.829268(0.45) = 0.823171
\end{aligned}
$$

Prior to DriveMed's announcement, you thought the probability that DriveMed would beat consensus expectations was 45%. On the basis of your interpretation of the announcement, you update that probability to 82.3%. This updated probability is called your **posterior probability** because it reflects or comes after the new information.

The Bayes' calculation takes the prior probability, which was 45%, and multiplies it by a ratio—the first term on the right-hand side of the equal sign. The denominator of the ratio is the probability that DriveMed expands, as you view it without considering (conditioning on) anything else. Therefore, this probability is unconditional. The numerator is the probability that DriveMed expands, if last quarter's EPS actually exceeded the consensus estimate. This last probability is larger than unconditional probability in the denominator, so the ratio (1.83 roughly) is greater than 1. As a result, your updated or posterior probability is larger than your prior probability. Thus, the ratio reflects the impact of the new information on your prior beliefs.

EXAMPLE 15

Inferring Whether DriveMed's EPS Met Consensus EPS

You are still an investor in DriveMed stock. To review the givens, your prior probabilities are $P(EPS\ exceeded\ consensus) = 0.45$, $P(EPS\ met\ consensus) = 0.30$, and $P(EPS\ fell\ short\ of\ consensus) = 0.25$. You also have the following conditional probabilities:

$$P(DriveMed\ expands \mid EPS\ exceeded\ consensus) = 0.75$$

$$P(DriveMed\ expands \mid EPS\ met\ consensus) = 0.20$$

$$P(DriveMed\ expands \mid EPS\ fell\ short\ of\ consensus) = 0.05$$

Recall that you updated your probability that last quarter's EPS exceeded the consensus estimate from 45% to 82.3% after DriveMed announced it would expand. Now you want to update your other priors.

1 Update your prior probability that DriveMed's EPS met consensus.

2 Update your prior probability that DriveMed's EPS fell short of consensus.

3 Show that the three updated probabilities sum to 1. (Carry each probability to four decimal places.)

4 Suppose, because of lack of prior beliefs about whether DriveMed would meet consensus, you updated on the basis of prior probabilities that all three possibilities were equally likely: $P(EPS\ exceeded\ consensus) = P(EPS\ met\ consensus) = P(EPS\ fell\ short\ of\ consensus) = 1/3$.

What is your estimate of the probability $P(EPS\ exceeded\ consensus \mid DriveMed\ expands)$?

Solution to 1:

The probability is $P(EPS\ met\ consensus \mid DriveMed\ expands) =$

$$\frac{P(DriveMed\ expands \mid EPS\ met\ consensus)}{P(DriveMed\ expands)}P(EPS\ met\ consensus)$$

The probability $P(DriveMed\ expands)$ is found by taking each of the three conditional probabilities in the statement of the problem, such as $P(DriveMed\ expands \mid EPS\ exceeded\ consensus)$; multiplying each one by the prior probability of the conditioning event, such as $P(EPS\ exceeded\ consensus)$; then adding the three products. The calculation is unchanged from the problem in the text above: $P(DriveMed\ expands) = 0.75(0.45) + 0.20(0.30) + 0.05(0.25) = 0.41$, or 41%. The other probabilities needed, $P(DriveMed\ expands \mid EPS\ met\ consensus) = 0.20$ and $P(EPS\ met\ consensus) = 0.30$, are givens. So

$$P(EPS\ met\ consensus \mid DriveMed\ expands)$$

$$= [P(DriveMed\ expands \mid EPS\ met\ consensus)/P(DriveMed\ expands)]P(EPS\ met\ consensus)$$

$$= (0.20/0.41)(0.30) = 0.487805(0.30) = 0.146341$$

After taking account of the announcement on expansion, your updated probability that last quarter's EPS for DriveMed just met consensus is 14.6% compared with your prior probability of 30%.

Solution to 2:

$P(DriveMed\ expands)$ was already calculated as 41%. Recall that $P(DriveMed\ expands \mid EPS\ fell\ short\ of\ consensus) = 0.05$ and $P(EPS\ fell\ short\ of\ consensus) = 0.25$ are givens.

$$P(EPS\ fell\ short\ of\ consensus \mid DriveMed\ expands)$$

$$= \left[P(DriveMed\ expands \mid EPS\ fell\ short\ of\ consensus)\ /\ P(DriveMed\ expands)\right]P(EPS\ fell\ short\ of\ consensus)$$

$$= (0.05\ /\ 0.41)(0.25) = 0.121951(0.25) = 0.030488$$

As a result of the announcement, you have revised your probability that DriveMed's EPS fell short of consensus from 25% (your prior probability) to 3%.

Solution to 3:

The sum of the three updated probabilities is

$$P(EPS\ exceeded\ consensus \mid DriveMed\ expands) + P(EPS\ met\ consensus \mid DriveMed\ expands) + P(EPS\ fell\ short\ of\ consensus \mid DriveMed\ expands)$$

$$= 0.8232 + 0.1463 + 0.0305 = 1.0000$$

The three events (*EPS exceeded consensus*, *EPS met consensus*, *EPS fell short of consensus*) are mutually exclusive and exhaustive: One of these events or statements must be true, so the conditional probabilities must sum to 1. Whether we are talking about conditional or unconditional probabilities, whenever we have a complete set of distinct possible events or outcomes, the probabilities must sum to 1. This calculation serves to check your work.

Solution to 4:

Using the probabilities given in the question,

$$\begin{aligned} &P(DriveMed\ expands) \\ &= P(DriveMed\ expands \mid EPS\ exceeded\ consensus) \\ &\quad P(EPS\ exceeded\ consensus) + P(DriveMed\ expands \mid \\ &\quad EPS\ met\ consensus)P(EPS\ met\ consensus) + P(DriveMed\ expands \mid \\ &\quad EPS\ fell\ short\ of\ consensus)P(EPS\ fell\ short\ of\ consensus) \\ &= 0.75(1/3) + 0.20(1/3) + 0.05(1/3) = 1/3 \end{aligned}$$

Not surprisingly, the probability of DriveMed expanding is 1/3 because the decision maker has no prior beliefs or views regarding how well EPS performed relative to the consensus estimate.

Now we can use Bayes' formula to find $P(EPS\ exceeded\ consensus \mid DriveMed\ expands) = [P(DriveMed\ expands \mid EPS\ exceeded\ consensus)/P(DriveMed\ expands)]\ P(EPS\ exceeded\ consensus) = [(0.75/(1/3)](1/3) = 0.75$ or 75%. This probability is identical to your estimate of $P(DriveMed\ expands \mid EPS\ exceeded\ consensus)$.

When the prior probabilities are equal, the probability of information given an event equals the probability of the event given the information. When a decision maker has equal prior probabilities (called **diffuse priors**), the probability of an event is determined by the information.

Example 16 shows how Bayes' formula is used in credit granting where the probability of payment given credit information is higher than the probability of payment without the information.

EXAMPLE 16

Bayes' Formula and the Probability of Payment

Jake Bronson is predicting the probability that consumer finance applicants granted credit will repay in a timely manner (i.e., their accounts will not become "past due"). Using Bayes' formula, he has structured the problem as

$$P(\text{Event} \mid \text{Information}) = \frac{P(\text{Information} \mid \text{Event})}{P(\text{Information})} P(\text{Event})$$

where the event (A) is "timely repayment" and the information (B) is having a "good credit report."

Bronson estimates that the unconditional probability of receiving timely payment, $P(A)$, is 0.90 and that the unconditional probability of having a good credit report, $P(B)$, is 0.80. The probability of having a good credit report given that borrowers paid on time, $P(B \mid A)$, is 0.85.

What is the probability that applicants with good credit reports will repay in a timely manner?

A 0.720

B 0.944

C 0.956

Solution:

The correct answer is C. The probability of timely repayment given a good credit report, $P(A \mid B)$, is

$$P(A \mid B) = \frac{P(B \mid A)}{P(B)}P(A) = \frac{0.85}{0.80} \times 0.90 = 0.956$$

7 PRINCIPLES OF COUNTING

- **identify the most appropriate method to solve a particular counting problem and analyze counting problems using factorial, combination, and permutation concepts**

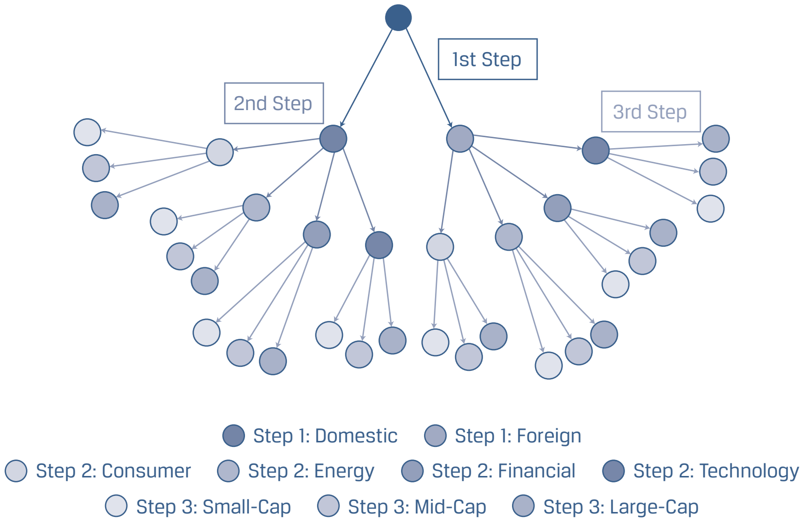
The first step in addressing a question often involves determining the different logical possibilities. We may also want to know the number of ways that each of these possibilities can happen. In the back of our mind is often a question about probability. How likely is it that I will observe this particular possibility? Records of success and failure are an example. For instance, the counting methods presented in this section have been used to evaluate a market timer's record. We can also use the methods in this section to calculate what we earlier called a priori probabilities. When we can assume that the possible outcomes of a random variable are equally likely, the probability of an event equals the number of possible outcomes favorable for the event divided by the total number of outcomes.

In counting, enumeration (counting the outcomes one by one) is of course the most basic resource. What we discuss in this section are shortcuts and principles. Without these shortcuts and principles, counting the total number of outcomes can be very difficult and prone to error. The first and basic principle of counting is the multiplication rule.

- **Multiplication Rule for Counting**. If one task can be done in n_1 ways, and a second task, given the first, can be done in n_2 ways, and a third task, given the first two tasks, can be done in n_3 ways, and so on for k tasks, then the number of ways the k tasks can be done is $(n_1)(n_2)(n_3) \ldots (n_k)$.

Exhibit 21 illustrates the multiplication rule where, for example, we have three steps in an investment decision process. In the first step, stocks are classified two ways, as domestic or foreign (represented by dark- and light-shaded circles, respectively). In the second step, stocks are assigned to one of four industries in our investment universe: consumer, energy, financial, or technology (represented by four circles with progressively darker shades, respectively). In the third step, stocks are classified three ways by size: small-cap, mid-cap, and large-cap (represented by light-, medium-, and dark-shaded circles, respectively). Because the first step can be done in two ways, the second in four ways, and the third in three ways, using the multiplication rule, we can carry out the three steps in (2)(4)(3) = 24 different ways.

Exhibit 21 **Investment Decision Process Using Multiplication Rule: $n_1 = 2$, $n_2 = 4$, $n_3 = 3$**



Another illustration is the assignment of members of a group to an equal number of positions. For example, suppose you want to assign three security analysts to cover three different industries. In how many ways can the assignments be made? The first analyst can be assigned in three different ways. Then two industries remain. The second analyst can be assigned in two different ways. Then one industry remains. The third and last analyst can be assigned in only one way. The total number of different assignments equals (3)(2)(1) = 6. The compact notation for the multiplication we have just performed is 3! (read: 3 factorial). If we had n analysts, the number of ways we could assign them to n tasks would be

$$n! = n(n - 1)(n - 2)(n - 3)\ldots 1$$

or ***n* factorial**. (By convention, 0! = 1.) To review, in this application, we repeatedly carry out an operation (here, job assignment) until we use up all members of a group (here, three analysts). With n members in the group, the multiplication formula reduces to n factorial.

The next type of counting problem can be called labeling problems.[1] We want to give each object in a group a label, to place it in a category. The following example illustrates this type of problem.

A mutual fund guide ranked 18 bond mutual funds by total returns for the last year. The guide also assigned each fund one of five risk labels: *high risk* (four funds), *above-average risk* (four funds), *average risk* (three funds), *below-average risk* (four funds), and *low risk* (three funds); as 4 + 4 + 3 + 4 + 3 = 18, all the funds are accounted for. How many different ways can we take 18 mutual funds and label 4 of them *high risk*, 4 *above-average risk*, 3 *average risk*, 4 *below-average risk*, and 3 *low risk*, so that each fund is labeled?

The answer is close to 13 billion. We can label any of 18 funds *high risk* (the first slot), then any of 17 remaining funds, then any of 16 remaining funds, then any of 15 remaining funds (now we have 4 funds in the *high risk* group); then we can label any of 14 remaining funds *above-average risk*, then any of 13 remaining funds, and so forth. There are 18! possible sequences. However, order of assignment within a category does not matter. For example, whether a fund occupies the first or third slot of the four funds labeled *high risk*, the fund has the same label (*high risk*). Thus, there are 4! ways to assign a given group of four funds to the four *high risk* slots. Making the same argument for the other categories, in total there are (4!)(4!)(3!)(4!)(3!) equivalent sequences. To eliminate such redundancies from the 18! total, we divide 18! by (4!)(4!)(3!)(4!)(3!). We have 18!/[(4!)(4!)(3!)(4!)(3!)] = 18!/[(24)(24)(6)(24)(6)] = 12,864,852,000. This procedure generalizes as follows.

---

[1] This discussion follows Kemeny, Schleifer, Snell, and Thompson (1972) in terminology and approach.

- **Multinomial Formula (General Formula for Labeling Problems)**. The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by

$$\frac{n!}{n_1!n_2!\ldots n_k!} \tag{21}$$

The multinomial formula with two different labels ($k = 2$) is especially important. This special case is called the combination formula. A **combination** is a listing in which the order of the listed items does not matter. We state the combination formula in a traditional way, but no new concepts are involved. Using the notation in the formula below, the number of objects with the first label is $r = n_1$ and the number with the second label is $n - r = n_2$ (there are just two categories, so $n_1 + n_2 = n$). Here is the formula:

- **Combination Formula (Binomial Formula)**. The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is

$$_nC_r = \binom{n}{r} = \frac{n!}{(n-r)!r!} \tag{22}$$

Here $_nC_r$ and $\binom{n}{r}$ are shorthand notations for $n!/(n - r)!r!$ (read: n choose r, or n combination r).

If we label the r objects as *belongs to the group* and the remaining objects as *does not belong to the group*, whatever the group of interest, the combination formula tells us how many ways we can select a group of size r. We can illustrate this formula with the binomial option pricing model. (The binomial pricing model is covered later in the CFA curriculum. The only intuition we are concerned with here is that a number of different pricing paths can end up with the same final stock price.) This model describes the movement of the underlying asset as a series of moves, price up (U) or price down (D). For example, two sequences of five moves containing three up moves, such as UUUDD and UDUUD, result in the same final stock price. At least for an option with a payoff dependent on final stock price, the number but not the order of up moves in a sequence matters. How many sequences of five moves *belong to the group with three up moves?* The answer is 10, calculated using the combination formula ("5 choose 3"):

$$_5C_3 = 5!/[(5 - 3)!3!]$$
$$= [(5)(4)(3)(2)(1)]/[(2)(1)(3)(2)(1)] = 120/12 = 10 \text{ ways}$$

A useful fact can be illustrated as follows: $_5C_3 = 5!/(2!3!)$ equals $_5C_2 = 5!/(3!2!)$, as 3 + 2 = 5; $_5C_4 = 5!/(1!4!)$ equals $_5C_1 = 5!/(4!1!)$, as 4 + 1 = 5. This symmetrical relationship can save work when we need to calculate many possible combinations.

Suppose jurors want to select three companies out of a group of five to receive the first-, second-, and third-place awards for the best annual report. In how many ways can the jurors make the three awards? Order does matter if we want to distinguish among the three awards (the rank within the group of three); clearly the question makes order important. On the other hand, if the question were "In how many ways can the jurors choose three winners, without regard to place of finish?" we would use the combination formula.

To address the first question above, we need to count ordered listings such as *first place*, *New Company*; *second place*, *Fir Company*; *third place*, *Well Company*. An ordered listing is known as a **permutation**, and the formula that counts the number of permutations is known as the permutation formula. A more formal definition states that a permutation is an ordered subset of n distinct objects.

- **Permutation Formula**. The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is

$$_{n}P_{r} = \frac{n!}{(n-r)!} \tag{23}$$

So the jurors have $_5P_3 = 5!/(5-3)! = [(5)(4)(3)(2)(1)]/[(2)(1)] = 120/2 = 60$ ways in which they can make their awards. To see why this formula works, note that [(5)(4)(3)(2)(1)]/[(2)(1)] reduces to (5)(4)(3), after cancellation of terms. This calculation counts the number of ways to fill three slots choosing from a group of five companies, according to the multiplication rule of counting. This number is naturally larger than it would be if order did not matter (compare 60 to the value of 10 for "5 choose 3" that we calculated above). For example, *first place*, *Well Company*; *second place*, *Fir Company*; *third place*, *New Company* contains the same three companies as *first place*, *New Company*; *second place*, *Fir Company*; *third place*, *Well Company*. If we were concerned only with award winners (without regard to place of finish), the two listings would count as one combination. But when we are concerned with the order of finish, the listings count as two permutations.

EXAMPLE 17

Permutations and Combinations for Two Out of Four Outcomes

There are four balls numbered 1, 2, 3, and 4 in a basket. You are running a contest in which two of the four balls are selected at random from the basket. To win, a player must have correctly chosen the numbers of the two randomly selected balls. Suppose the winning numbers are numbers 1 and 3. If the player must choose the balls in the same order in which they are drawn, she wins if she chose 1 first and 3 second. On the other hand, if order is not important, the player wins if the balls drawn are 1 and then 3 or if the balls drawn are 3 and then 1. The number of possible outcomes for permutations and combinations of choosing 2 out of 4 items is illustrated in Exhibit 22. If order is not important, for choosing 2 out of 4 items, the winner wins twice as often.

Exhibit 22 Permutations and Combinations for Choosing 2 out of 4 Items

Permutations: Order matters	Combinations: Order does not matter
List of all possible outcomes: (1 2) (2 1) (3 1) (4 1) (1 3) (2 3) (3 2) (4 2) (1 4) (2 4) (3 4) (4 3)	List of all possible outcomes: (1 2) (2 3) (3 4) (1 3) (2 4) (1 4)
Number of permutations: $_nP_r = \frac{n!}{(n-r)!}$ $_4P_2 = \frac{4!}{(4-2)!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1} = 12$	Number of combinations: $_nC_r = \frac{n!}{(n-r)!r!}$ $_4C_2 = \frac{4!}{(4-2)!2!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 2 \times 1} = 6$

If order is important, the number of permutations (possible outcomes) is much larger than the number of combinations when order is not important.

EXAMPLE 18

Reorganizing the Analyst Team Assignments

Gehr-Flint Investors classifies the stocks in its investment universe into 11 industries and is assigning each research analyst one or two industries. Five of the industries have been assigned, and you are asked to cover two industries from the remaining six.

How many possible pairs of industries remain?

A 12

B 15

C 36

Solution:

B is correct. The number of combinations of selecting two industries out of six is equal to

$$_nC_r = \begin{bmatrix} n \\ r \end{bmatrix} = \frac{n!}{(n-r)!r!} = \frac{6!}{4!2!} = 15$$

The number of possible combinations for picking two industries out of six is 15.

EXAMPLE 19

Australian Powerball Lottery

To win the Australian Powerball jackpot, you must match the numbers of seven balls pulled from a basket (the balls are numbered 1 through 35) plus the number of the Powerball (numbered 1 through 20). The order in which the seven balls are drawn is not important. The number of combinations of matching 7 out of 35 balls is

$$_nC_r = \begin{bmatrix} n \\ r \end{bmatrix} = \frac{n!}{(n-r)!r!} = \frac{35!}{28!7!} = 6{,}724{,}520$$

The number of combinations for picking the Powerball, 1 out of 20, is

$$_nC_r = \begin{bmatrix} n \\ r \end{bmatrix} = \frac{n!}{(n-r)!r!} = \frac{20!}{19!1!} = 20$$

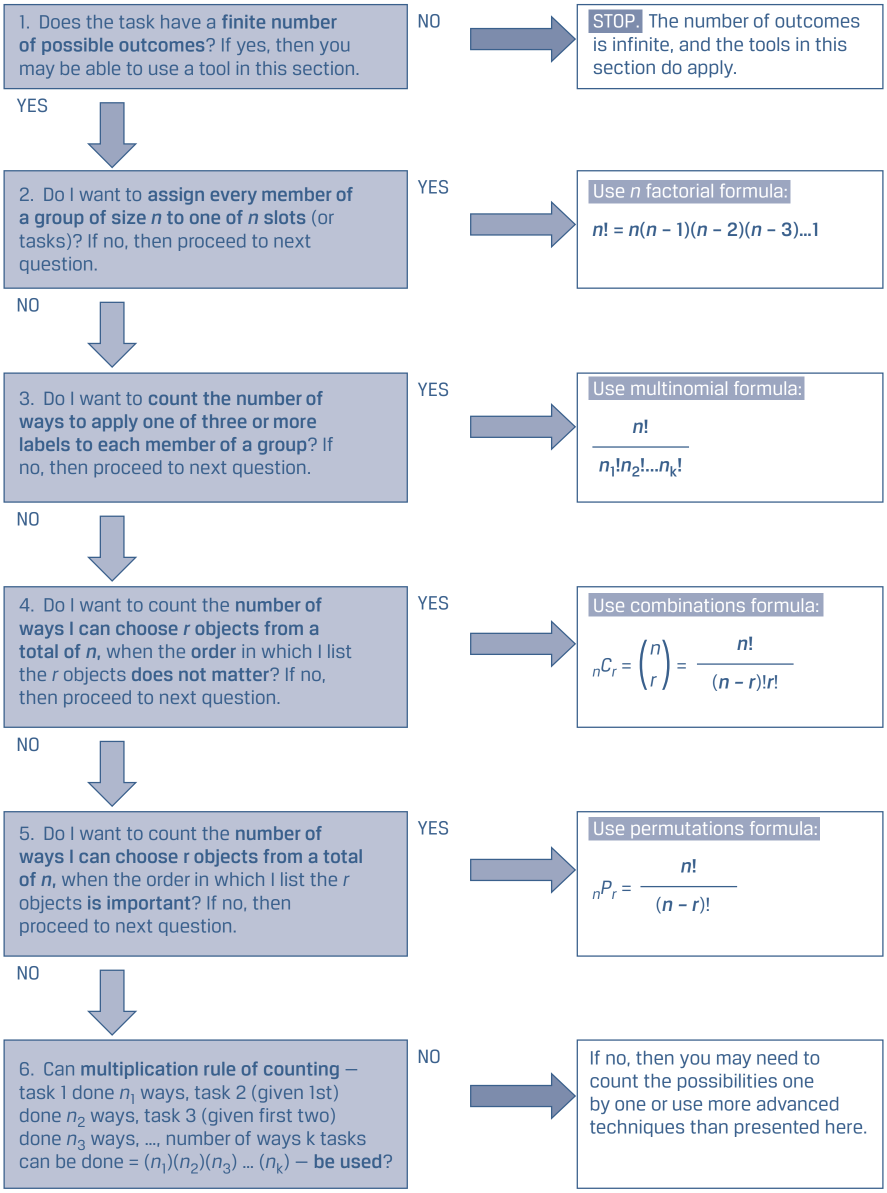
The number of ways to pick the seven balls plus the Powerball is

$$_{35}C_7 \times {}_{20}C_1 = 6{,}724{,}520 \times 20 = 134{,}490{,}400$$

Your probability of winning the Australian Powerball with one ticket is 1 in 134,490,400.

Exhibit 23 is a flow chart that may help you apply the counting methods we have presented in this section.

Exhibit 23 Summary of Counting Methods



1. Does the task have a **finite number of possible outcomes**? If yes, then you may be able to use a tool in this section.

NO → STOP. The number of outcomes is infinite, and the tools in this section do apply.

YES ↓

2. Do I want to **assign every member of a group of size *n* to one of *n* slots** (or tasks)? If no, then proceed to next question.

YES → Use *n* factorial formula: $n! = n(n-1)(n-2)(n-3)\ldots1$

NO ↓

3. Do I want to **count the number of ways to apply one of three or more labels to each member of a group**? If no, then proceed to next question.

YES → Use multinomial formula: $\dfrac{n!}{n_1!n_2!\ldots n_k!}$

NO ↓

4. Do I want to count the **number of ways I can choose *r* objects from a total of *n***, when the **order** in which I list the *r* objects **does not matter**? If no, then proceed to next question.

YES → Use combinations formula: ${}_nC_r = \binom{n}{r} = \dfrac{n!}{(n-r)!r!}$

NO ↓

5. Do I want to count the **number of ways I can choose r objects from a total of *n*,** when the order in which I list the *r* objects **is important**? If no, then proceed to next question.

YES → Use permutations formula: ${}_nP_r = \dfrac{n!}{(n-r)!}$

NO ↓

6. Can **multiplication rule of counting** — task 1 done n_1 ways, task 2 (given 1st) done n_2 ways, task 3 (given first two) done n_3 ways, ..., number of ways k tasks can be done = $(n_1)(n_2)(n_3)\ldots(n_k)$ — **be used**?

NO → If no, then you may need to count the possibilities one by one or use more advanced techniques than presented here.

SUMMARY

In this reading, we have discussed the essential concepts and tools of probability. We have applied probability, expected value, and variance to a range of investment problems.

- A random variable is a quantity whose outcome is uncertain.
- Probability is a number between 0 and 1 that describes the chance that a stated event will occur.
- An event is a specified set of outcomes of a random variable.

- Mutually exclusive events can occur only one at a time. Exhaustive events cover or contain all possible outcomes.
- The two defining properties of a probability are, first, that $0 \le P(E) \le 1$ (where $P(E)$ denotes the probability of an event E) and, second, that the sum of the probabilities of any set of mutually exclusive and exhaustive events equals 1.
- A probability estimated from data as a relative frequency of occurrence is an empirical probability. A probability drawing on personal or subjective judgment is a subjective probability. A probability obtained based on logical analysis is an a priori probability.
- A probability of an event E, $P(E)$, can be stated as odds for $E = P(E)/[1 - P(E)]$ or odds against $E = [1 - P(E)]/P(E)$.
- Probabilities that are inconsistent create profit opportunities, according to the Dutch Book Theorem.
- A probability of an event *not* conditioned on another event is an unconditional probability. The unconditional probability of an event A is denoted $P(A)$. Unconditional probabilities are also called marginal probabilities.
- A probability of an event given (conditioned on) another event is a conditional probability. The probability of an event A given an event B is denoted $P(A \mid B)$, and $P(A \mid B) = P(AB)/P(B)$, $P(B) \neq 0$.
- The probability of both A and B occurring is the joint probability of A and B, denoted $P(AB)$.
- The multiplication rule for probabilities is $P(AB) = P(A \mid B)P(B)$.
- The probability that A or B occurs, or both occur, is denoted by $P(A \text{ or } B)$.
- The addition rule for probabilities is $P(A \text{ or } B) = P(A) + P(B) - P(AB)$.
- When events are independent, the occurrence of one event does not affect the probability of occurrence of the other event. Otherwise, the events are dependent.
- The multiplication rule for independent events states that if A and B are independent events, $P(AB) = P(A)P(B)$. The rule generalizes in similar fashion to more than two events.
- According to the total probability rule, if $S_1, S_2, \ldots, S_n$ are mutually exclusive and exhaustive scenarios or events, then $P(A) = P(A \mid S_1)P(S_1) + P(A \mid S_2)P(S_2) + \ldots + P(A \mid S_n)P(S_n)$.
- The expected value of a random variable is a probability-weighted average of the possible outcomes of the random variable. For a random variable X, the expected value of X is denoted $E(X)$.
- The total probability rule for expected value states that $E(X) = E(X \mid S_1)P(S_1) + E(X \mid S_2)P(S_2) + \ldots + E(X \mid S_n)P(S_n)$, where $S_1, S_2, \ldots, S_n$ are mutually exclusive and exhaustive scenarios or events.
- The variance of a random variable is the expected value (the probability-weighted average) of squared deviations from the random variable's expected value $E(X)$: $\sigma^2(X) = E\{[X - E(X)]^2\}$, where $\sigma^2(X)$ stands for the variance of X.
- Variance is a measure of dispersion about the mean. Increasing variance indicates increasing dispersion. Variance is measured in squared units of the original variable.
- Standard deviation is the positive square root of variance. Standard deviation measures dispersion (as does variance), but it is measured in the same units as the variable.
- Covariance is a measure of the co-movement between random variables.

- The covariance between two random variables R_i and R_j in a forward-looking sense is the expected value of the cross-product of the deviations of the two random variables from their respective means: $\text{Cov}(R_i,R_j) = E\{[R_i - E(R_i)][R_j - E(R_j)]\}$. The covariance of a random variable with itself is its own variance.
- The historical or sample covariance between two random variables R_i and R_j based on a sample of past data of size n is the average value of the product of the deviations of observations on two random variables from their sample means:

$$\text{Cov}\left(R_i,R_j\right) = \sum_{n=1}^{n}\left(R_{i,t} - \bar{R}_i\right)\left(R_{j,t} - \bar{R}_j\right)\Big/(n-1)$$

- Correlation is a number between −1 and +1 that measures the co-movement (linear association) between two random variables: $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/[\sigma(R_i)\,\sigma(R_j)]$.
- If two variables have a very strong (inverse) linear relation, then the absolute value of their correlation will be close to 1 (-1). If two variables have a weak linear relation, then the absolute value of their correlation will be close to 0.
- If the correlation coefficient is positive, the two variables are positively related; if the correlation coefficient is negative, the two variables are inversely related.
- To calculate the variance of return on a portfolio of n assets, the inputs needed are the n expected returns on the individual assets, n variances of return on the individual assets, and $n(n - 1)/2$ distinct covariances.
- Portfolio variance of return is $\sigma^2\left(R_p\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \text{Cov}\left(R_i,R_j\right)$.
- The calculation of covariance in a forward-looking sense requires the specification of a joint probability function, which gives the probability of joint occurrences of values of the two random variables.
- When two random variables are independent, the joint probability function is the product of the individual probability functions of the random variables.
- Bayes' formula is a method for updating probabilities based on new information.
- Bayes' formula is expressed as follows: Updated probability of event given the new information = [(Probability of the new information given event)/(Unconditional probability of the new information)] × Prior probability of event.
- The multiplication rule of counting says, for example, that if the first step in a process can be done in 10 ways, the second step, given the first, can be done in 5 ways, and the third step, given the first two, can be done in 7 ways, then the steps can be carried out in (10)(5)(7) = 350 ways.
- The number of ways to assign every member of a group of size n to n slots is $n! = n\,(n-1)\,(n-2)(n-3)\ldots 1$. (By convention, $0! = 1$.)
- The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by $n!/(n_1!n_2!\ldots n_k!)$. This expression is the multinomial formula.
- A special case of the multinomial formula is the combination formula. The number of ways to choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is

$${}_nC_r = \binom{n}{r} = \frac{n!}{(n-r)!r!}$$

- The number of ways to choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is

  $${}_nP_r = \frac{n!}{(n-r)!}$$

  This expression is the permutation formula.

REFERENCES

Kemeny, John G., and Arthur Schleifer, Jr., J. Laurie Snell, and Gerald L. Thompson. 1972. Finite Mathematics with Business Applications, 2nd edition. Englewood Cliffs, NJ: Prentice-Hall.

PRACTICE PROBLEMS

1. Suppose that 5% of the stocks meeting your stock-selection criteria are in the telecommunications (telecom) industry. Also, dividend-paying telecom stocks are 1% of the total number of stocks meeting your selection criteria. What is the probability that a stock is dividend paying, given that it is a telecom stock that has met your stock selection criteria?

2. You are using the following three criteria to screen potential acquisition targets from a list of 500 companies:

Criterion	Fraction of the 500 Companies Meeting the Criterion
Product lines compatible	0.20
Company will increase combined sales growth rate	0.45
Balance sheet impact manageable	0.78

If the criteria are independent, how many companies will pass the screen?

3. Florence Hixon is screening a set of 100 stocks based on two criteria (Criterion 1 and Criterion 2). She set the passing level such that 50% of the stocks passed each screen. For these stocks, the values for Criterion 1 and Criterion 2 are not independent but are positively related. How many stocks should pass Hixon's two screens?

   A Less than 25

   B 25

   C More than 25

4. You apply both valuation criteria and financial strength criteria in choosing stocks. The probability that a randomly selected stock (from your investment universe) meets your valuation criteria is 0.25. Given that a stock meets your valuation criteria, the probability that the stock meets your financial strength criteria is 0.40. What is the probability that a stock meets both your valuation and financial strength criteria?

5. Suppose the prospects for recovering principal for a defaulted bond issue depend on which of two economic scenarios prevails. Scenario 1 has probability 0.75 and will result in recovery of $0.90 per $1 principal value with probability 0.45, or in recovery of $0.80 per $1 principal value with probability 0.55. Scenario 2 has probability 0.25 and will result in recovery of $0.50 per $1 principal value with probability 0.85, or in recovery of $0.40 per $1 principal value with probability 0.15.

   A Compute the probability of each of the four possible recovery amounts: $0.90, $0.80, $0.50, and $0.40.

   B Compute the expected recovery, given the first scenario.

   C Compute the expected recovery, given the second scenario.

   D Compute the expected recovery.

   E Graph the information in a probability tree diagram.

6. You have developed a set of criteria for evaluating distressed credits. Companies that do not receive a passing score are classed as likely to go bankrupt within 12 months. You gathered the following information when validating the criteria:

© 2021 CFA Institute. All rights reserved.

- Forty percent of the companies to which the test is administered will go bankrupt within 12 months: *P*(*non-survivor*) = 0.40.
- Fifty-five percent of the companies to which the test is administered pass it: *P*(*pass test*) = 0.55.
- The probability that a company will pass the test given that it will subsequently survive 12 months, is 0.85: *P*(*pass test* | *survivor*) = 0.85.

A What is P(*pass test* | *non-survivor*)?

B Using Bayes' formula, calculate the probability that a company is a survivor, given that it passes the test; that is, calculate *P*(*survivor* | *pass test*).

C What is the probability that a company is a *non-survivor*, given that it fails the test?

D Is the test effective?

7 In probability theory, exhaustive events are *best* described as the set of events that:

A have a probability of zero.

B are mutually exclusive.

C include all potential outcomes.

8 Which probability estimate *most likely* varies greatly between people?

A An *a priori* probability

B An empirical probability

C A subjective probability

9 If the probability that Zolaf Company sales exceed last year's sales is 0.167, the odds for exceeding sales are *closest* to:

A 1 to 5.

B 1 to 6.

C 5 to 1.

10 After six months, the growth portfolio that Rayan Khan manages has outperformed its benchmark. Khan states that his odds of beating the benchmark for the year are 3 to 1. If these odds are correct, what is the probability that Khan's portfolio will beat the benchmark for the year?

A 0.33

B 0.67

C 0.75

11 The probability of an event given that another event has occurred is a:

A joint probability.

B marginal probability.

C conditional probability.

12 After estimating the probability that an investment manager will exceed his benchmark return in each of the next two quarters, an analyst wants to forecast the probability that the investment manager will exceed his benchmark return over the two-quarter period in total. Assuming that each quarter's performance is independent of the other, which probability rule should the analyst select?

A Addition rule

B Multiplication rule

C Total probability rule

13 Which of the following is a property of two dependent events?

A The two events must occur simultaneously.

B The probability of one event influences the probability of the other event.

C The probability of the two events occurring is the product of each event's probability.

14 Which of the following *best* describes how an analyst would estimate the expected value of a firm under the scenarios of bankruptcy and survivorship? The analyst would use:

A the addition rule.

B conditional expected values.

C the total probability rule for expected value.

15 An analyst developed two scenarios with respect to the recovery of $100,000 principal from defaulted loans:

Scenario	Probability of Scenario (%)	Amount Recovered ($)	Probability of Amount (%)
1	40	50,000	60
		30,000	40
2	60	80,000	90
		60,000	10

The amount of the expected recovery is *closest* to:

A $36,400.

B $55,000.

C $63,600.

16 US and Spanish bonds have return standard deviations of 0.64 and 0.56, respectively. If the correlation between the two bonds is 0.24, the covariance of returns is *closest* to:

A 0.086.

B 0.335.

C 0.390.

17 The covariance of returns is positive when the returns on two assets tend to:

A have the same expected values.

B be above their expected value at different times.

C be on the same side of their expected value at the same time.

18 Which of the following correlation coefficients indicates the weakest linear relationship between two variables?

A –0.67

B –0.24

C 0.33

19 An analyst develops the following covariance matrix of returns:

	Hedge Fund	Market Index
Hedge fund	256	110
Market index	110	81

The correlation of returns between the hedge fund and the market index is *closest* to:

A 0.005.

B 0.073.

C 0.764.

20 All else being equal, as the correlation between two assets approaches +1.0, the diversification benefits:

A decrease.

B stay the same.

C increase.

21 Given a portfolio of five stocks, how many unique covariance terms, excluding variances, are required to calculate the portfolio return variance?

A 10

B 20

C 25

22 The probability distribution for a company's sales is:

Probability	Sales ($ millions)
0.05	70
0.70	40
0.25	25

The standard deviation of sales is *closest* to:

A $9.81 million.

B $12.20 million.

C $32.40 million.

23 Which of the following statements is *most* accurate? If the covariance of returns between two assets is 0.0023, then:

A the assets' risk is near zero.

B the asset returns are unrelated.

C the asset returns have a positive relationship.

24 An analyst produces the following joint probability function for a foreign index (FI) and a domestic index (DI).

	R_{DI} = 30%	R_{DI} = 25%	R_{DI} = 15%
R_{FI} = 25%	0.25		
R_{FI} = 15%		0.50	
R_{FI} = 10%			0.25

The covariance of returns on the foreign index and the returns on the domestic index is *closest* to:

A 26.39.

B 26.56.

C 28.12.

25 A two-stock portfolio includes stocks with the following characteristics:

	Stock 1	Stock 2
Expected return	7%	10%
Standard deviation	12%	25%
Portfolio weights	0.30	0.70
Correlation	0.20	

What is the standard deviation of portfolio returns?

A 14.91%

B 18.56%

C 21.10%

26 Lena Hunziger has designed the three-asset portfolio summarized below:

	Asset 1	Asset 2	Asset 3
Expected return	5%	6%	7%
Portfolio weight	**0.20**	0.30	0.50
Variance-Covariance Matrix			
	Asset 1	**Asset 2**	**Asset 3**
Asset 1	196	105	140
Asset 2	105	225	150
Asset 3	140	150	400

Hunziger estimated the portfolio return to be 6.3%. What is the portfolio standard deviation?

A 13.07%

B 13.88%

C 14.62%

27 An analyst estimates that 20% of high-risk bonds will fail (go bankrupt). If she applies a bankruptcy prediction model, she finds that 70% of the bonds will receive a "good" rating, implying that they are less likely to fail. Of the bonds that failed, only 50% had a "good" rating. Use Bayes' formula to predict the probability of failure given a "good" rating. (Hint, let $P(A)$ be the probability of failure, $P(B)$ be the probability of a "good" rating, $P(B \mid A)$ be the likelihood of a "good" rating given failure, and $P(A \mid B)$ be the likelihood of failure given a "good" rating.)

A 5.7%

B 14.3%

C 28.6%

28 In a typical year, 5% of all CEOs are fired for "performance" reasons. Assume that CEO performance is judged according to stock performance and that 50% of stocks have above-average returns or "good" performance. Empirically, 30% of all CEOs who were fired had "good" performance. Using Bayes' formula, what is the probability that a CEO will be fired given "good" performance? (Hint, let $P(A)$ be the probability of a CEO being fired, $P(B)$ be the probability of a "good" performance rating, $P(B \mid A)$ be the likelihood of a "good" performance rating given that the CEO was fired, and $P(A \mid B)$ be the likelihood of the CEO being fired given a "good" performance rating.)

A 1.5%

B 2.5%

C 3.0%

29 A manager will select 20 bonds out of his universe of 100 bonds to construct a portfolio. Which formula provides the number of possible portfolios?

A Permutation formula

B Multinomial formula

C Combination formula

30 A firm will select two of four vice presidents to be added to the investment committee. How many different groups of two are possible?

A 6

B 12

C 24

31 From an approved list of 25 funds, a portfolio manager wants to rank 4 mutual funds from most recommended to least recommended. Which formula is *most* appropriate to calculate the number of possible ways the funds could be ranked?

A Permutation formula

B Multinomial formula

C Combination formula

32 Himari Fukumoto has joined a new firm and is selecting mutual funds in the firm's pension plan. If 10 mutual funds are available, and she plans to select four, how many different sets of mutual funds can she choose?

A 210

B 720

C 5,040

Items 33, 34, and 35 are based on the following setup:

Gerd Sturm wants to sponsor a contest with a $1 million prize. The winner must pick the stocks that will be the top five performers next year among the 30 stocks in a well-known large-cap stock index. He asks you to estimate the chances that contestants can win the contest.

33 What are the chances of winning if the contestants must pick the five stocks in the correct order of their total return? If choosing five stocks randomly, a contestant's chance of winning is one out of:

A 142,506.

B 17,100,720.

C 24,300,000.

34 What are the chances of winning if the contestants must pick the top five stocks without regard to order? If choosing five stocks randomly, a contestant's chance of winning is one out of:

A 142,506.

B 17,100,720.

C 24,300,000.

35 Sturm asks, "Can we trust these probabilities of winning?"

SOLUTIONS

1 Use this equation to find this conditional probability: P(*stock is dividend paying* | *telecom stock that meets criteria*) = P(*stock is dividend paying and telecom stock that meets criteria*)/P(*telecom stock that meets criteria*) = 0.01/0.05 = 0.20.

2 According to the multiplication rule for independent events, the probability of a company meeting all three criteria is the product of the three probabilities. Labeling the event that a company passes the first, second, and third criteria, A, B, and C, respectively, $P(ABC) = P(A)P(B)P(C) = (0.20)(0.45)(0.78) = 0.0702$. As a consequence, (0.0702)(500) = 35.10, so 35 companies pass the screen.

3 C is correct. Let event A be a stock passing the first screen (Criterion 1) and event B be a stock passing the second screen (Criterion 2). The probability of passing each screen is $P(A) = 0.50$ and $P(B) = 0.50$. If the two criteria are independent, the joint probability of passing both screens is $P(AB) = P(A)P(B) = 0.50 \times 0.50 = 0.25$, so 25 out of 100 stocks would pass both screens. However, the two criteria are positively related, and $P(AB) \neq 0.25$. Using the multiplication rule for probabilities, the joint probability of A and B is $P(AB) = P(A \mid B)P(B)$. If the two criteria are not independent, and if $P(B) = 0.50$, then the contingent probability of $P(A \mid B)$ is greater than 0.50. So the joint probability of $P(AB) = P(A \mid B)P(B)$ is greater than 0.25. More than 25 stocks should pass the two screens.

4 Use the equation for the multiplication rule for probabilities $P(AB) = P(A \mid B)P(B)$, defining A as the event that *a stock meets the financial strength criteria* and defining B as the event that *a stock meets the valuation criteria*. Then $P(AB) = P(A \mid B)P(B) = 0.40 \times 0.25 = 0.10$. The probability that a stock meets both the financial and valuation criteria is 0.10.

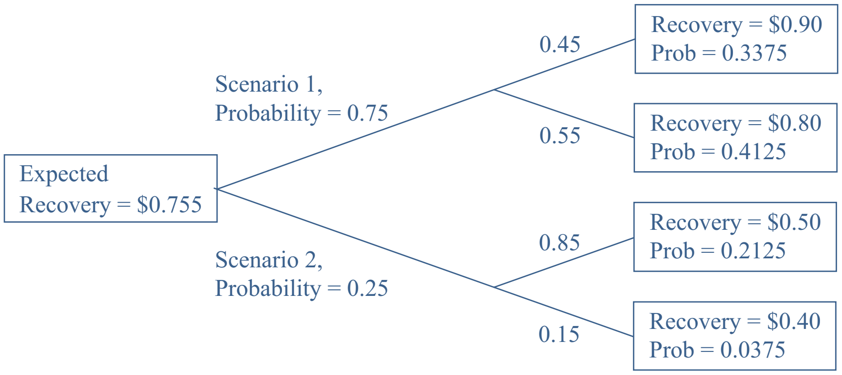
5 **A** *Outcomes associated with Scenario 1:* With a 0.45 probability of a \$0.90 recovery per \$1 principal value, given Scenario 1, and with the probability of Scenario 1 equal to 0.75, the probability of recovering \$0.90 is 0.45 (0.75) = 0.3375. By a similar calculation, the probability of recovering \$0.80 is 0.55(0.75) = 0.4125.

Outcomes associated with Scenario 2: With a 0.85 probability of a \$0.50 recovery per \$1 principal value, given Scenario 2, and with the probability of Scenario 2 equal to 0.25, the probability of recovering \$0.50 is 0.85(0.25) = 0.2125. By a similar calculation, the probability of recovering \$0.40 is 0.15(0.25) = 0.0375.

B $E(\textit{recovery} \mid \textit{Scenario 1}) = 0.45(\$0.90) + 0.55(\$0.80) = \$0.845$

C $E(\textit{recovery} \mid \textit{Scenario 2}) = 0.85(\$0.50) + 0.15(\$0.40) = \$0.485$

D $E(\textit{recovery}) = 0.75(\$0.845) + 0.25(\$0.485) = \$0.755$



6 **A** We can set up the equation using the total probability rule:

$$\begin{aligned} P(\mathit{pass\ test}) &= P(\mathit{pass\ test} \mid \mathit{survivor})P(\mathit{survivor}) \\ &\quad + P(\mathit{pass\ test} \mid \mathit{non\text{-}survivor})P(\mathit{non\text{-}survivor}) \end{aligned}$$

We know that $P(\mathit{survivor}) = 1 - P(\mathit{non\text{-}survivor}) = 1 - 0.40 = 0.60$. Therefore, $P(\mathit{pass\ test}) = 0.55 = 0.85(0.60) + P(\mathit{pass\ test} \mid \mathit{non\text{-}survivor})(0.40)$.

Thus, $P(\mathit{pass\ test} \mid \mathit{non\text{-}survivor}) = [0.55 - 0.85(0.60)]/0.40 = 0.10$.

B

$$\begin{aligned} P(\mathit{survivor} \mid \mathit{pass\ test}) &= \left[P(\mathit{pass\ test} \mid \mathit{survivor})/P(\mathit{pass\ test})\right]P(\mathit{survivor}) \\ &= (0.85/0.55)0.60 = 0.927273 \end{aligned}$$

The information that a company passes the test causes you to update your probability that it is a survivor from 0.60 to approximately 0.927.

C According to Bayes' formula, $P(\mathit{non\text{-}survivor} \mid \mathit{fail\ test}) = [P(\mathit{fail\ test} \mid \mathit{non\text{-}survivor})/ P(\mathit{fail\ test})]P(\mathit{non\text{-}survivor}) = [P(\mathit{fail\ test} \mid \mathit{non\text{-}survivor})/0.45]0.40$.

We can set up the following equation to obtain $P(\mathit{fail\ test} \mid \mathit{non\text{-}survivor})$:

$$\begin{aligned} P(\mathit{fail\ test}) &= P(\mathit{fail\ test} \mid \mathit{non\text{-}survivor})P(\mathit{non\text{-}survivor}) \\ &\quad + P(\mathit{fail\ test} \mid \mathit{survivor})P(\mathit{survivor}) \\ 0.45 &= P(\mathit{fail\ test} \mid \mathit{non\text{-}survivor})0.40 + 0.15(0.60) \end{aligned}$$

where $P(\mathit{fail\ test} \mid \mathit{survivor}) = 1 - P(\mathit{pass\ test} \mid \mathit{survivor}) = 1 - 0.85 = 0.15$. So $P(\mathit{fail\ test} \mid \mathit{non\text{-}survivor}) = [0.45 - 0.15(0.60)]/0.40 = 0.90$.

Using this result with the formula above, we find $P(\mathit{non\text{-}survivor} \mid \mathit{fail\ test}) = [0.90/0.45]0.40 = 0.80$. Seeing that a company fails the test causes us to update the probability that it is a non-survivor from 0.40 to 0.80.

D A company passing the test greatly increases our confidence that it is a survivor. A company failing the test doubles the probability that it is a non-survivor. Therefore, the test appears to be useful.

7 C is correct. The term "exhaustive" means that the events cover all possible outcomes.

8 C is correct. A subjective probability draws on personal or subjective judgment that may be without reference to any particular data.

9 A is correct. Given odds for E of a to b, the implied probability of $E = a/(a + b)$. Stated in terms of odds a to b with $a = 1$, $b = 5$, the probability of $E = 1/(1 + 5) = 1/6 = 0.167$. This result confirms that a probability of 0.167 for beating sales is odds of 1 to 5.

10 C is correct. The odds for beating the benchmark = $P(\mathit{beating\ benchmark}) / [1 - P(\mathit{beating\ benchmark})]$. Let $P(A) = P(\mathit{beating\ benchmark})$. Odds for beating the benchmark = $P(A) / [1 - P(A)]$.

$$3 = P(A) / [1 - P(A)]$$

Solving for $P(A)$, the probability of beating the benchmark is 0.75.

11 C is correct. A conditional probability is the probability of an event given that another event has occurred.

12 B is correct. Because the events are independent, the multiplication rule is most appropriate for forecasting their joint probability. The multiplication rule for independent events states that the joint probability of both A and B occurring is $P(AB) = P(A)P(B)$.

13 B is correct. The probability of the occurrence of one is related to the occurrence of the other. If we are trying to forecast one event, information about a dependent event may be useful.

14 C is correct. The total probability rule for expected value is used to estimate an expected value based on mutually exclusive and exhaustive scenarios.

15 C is correct. If Scenario 1 occurs, the expected recovery is 60% ($50,000) + 40% ($30,000) = $42,000, and if Scenario 2 occurs, the expected recovery is 90% ($80,000) + 10%($60,000) = $78,000. Weighting by the probability of each scenario, the expected recovery is 40%($42,000) + 60%($78,000) = $63,600. Alternatively, first calculating the probability of each amount occurring, the expected recovery is (40%)(60%)($50,000) + (40%)(40%)($30,000) + (60%)(90%)($80,000) + (60%)(10%)($60,000) = $63,600.

16 A is correct. The covariance is the product of the standard deviations and correlation using the formula Cov(US bond returns, Spanish bond returns) = σ(US bonds) × σ (Spanish bonds) × ρ(US bond returns, Spanish bond returns) = 0.64 × 0.56 × 0.24 = 0.086.

17 C is correct. The covariance of returns is positive when the returns on both assets tend to be on the same side (above or below) their expected values at the same time, indicating an average positive relationship between returns.

18 B is correct. Correlations near +1 exhibit strong positive linearity, whereas correlations near −1 exhibit strong negative linearity. A correlation of 0 indicates an absence of any linear relationship between the variables. The closer the correlation is to 0, the weaker the linear relationship.

19 C is correct. The correlation between two random variables R_i and R_j is defined as $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/[\sigma(R_i)\sigma(R_j)]$. Using the subscript i to represent hedge funds and the subscript j to represent the market index, the standard deviations are $\sigma(R_i) = 256^{1/2} = 16$ and $\sigma(R_j) = 81^{1/2} = 9$. Thus, $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/[\sigma(R_i)\sigma(R_j)] = 110/(16 \times 9) = 0.764$.

20 A is correct. As the correlation between two assets approaches +1, diversification benefits decrease. In other words, an increasingly positive correlation indicates an increasingly strong positive linear relationship and fewer diversification benefits.

21 A is correct. A covariance matrix for five stocks has 5 × 5 = 25 entries. Subtracting the 5 diagonal variance terms results in 20 off-diagonal entries. Because a covariance matrix is symmetrical, only 10 entries are unique (20/2 = 10).

22 A is correct. The analyst must first calculate expected sales as 0.05 × $70 + 0.70 × $40 + 0.25 × $25 = $3.50 million + $28.00 million + $6.25 million = $37.75 million.

After calculating expected sales, we can calculate the variance of sales:

$$\sigma^2 \text{ (Sales)} = P(\$70)[\$70 - E(\text{Sales})]^2 + P(\$40)[\$40 - E(\text{Sales})]^2 + P(\$25)[\$25 - E(\text{Sales})]^2$$

$$= 0.05(\$70 - 37.75)^2 + 0.70(\$40 - 37.75)^2 + 0.25(\$25 - 37.75)^2$$

$$= \$52.00 \text{ million} + \$3.54 \text{ million} + \$40.64 \text{ million} = \$96.18 \text{ million.}$$

The standard deviation of sales is thus $\sigma = (\$96.18)^{1/2} = \$9.81$ million.

23 C is correct. The covariance of returns is positive when the returns on both assets tend to be on the same side (above or below) their expected values at the same time.

24 B is correct. The covariance is 26.56, calculated as follows. First, expected returns are

$$E(R_{FI}) = (0.25 \times 25) + (0.50 \times 15) + (0.25 \times 10)$$

$$= 6.25 + 7.50 + 2.50 = 16.25 \text{ and}$$

$$E(R_{DI}) = (0.25 \times 30) + (0.50 \times 25) + (0.25 \times 15)$$

$$= 7.50 + 12.50 + 3.75 = 23.75.$$

Covariance is

$$\mathrm{Cov}(R_{FI},R_{DI}) = \sum_i \sum_j P\left(R_{FI,i}, R_{DI,j}\right)\left(R_{FI,i} - ER_{FI}\right)\left(R_{DI,j} - ER_{DI}\right)$$

$$= 0.25[(25 - 16.25)(30 - 23.75)] + 0.50[(15 - 16.25)(25 - 23.75)] + 0.25[(10 - 16.25)(15 - 23.75)]$$

$$= 13.67 + (-0.78) + 13.67 = 26.56.$$

25 B is correct. The covariance between the returns for the two stocks is $\mathrm{Cov}(R_1,R_2) = \rho(R_1,R_2)\,\sigma(R_1)\,\sigma(R_2) = 0.20\,(12)\,(25) = 60$. The portfolio variance is

$$\sigma^2(R_p) = w_1^2\sigma^2(R_1) + w_2^2\sigma^2(R_2) + 2w_1w_2\mathrm{Cov}(R_1,R_2)$$

$$= (0.30)^2(12)^2 + (0.70)^2(25)^2 + 2(0.30)(0.70)(60)$$

$$= 12.96 + 306.25 + 25.2 = 344.41$$

The portfolio standard deviation is

$$\sigma^2(R_p) = 344.41^{1/2} = 18.56\%$$

26 C is correct. For a three-asset portfolio, the portfolio variance is

$$\sigma^2(R_p) = w_1^2\sigma^2(R_1) + w_2^2\sigma^2(R_2) + w_3^2\sigma^2(R_3) + 2w_1w_2\mathrm{Cov}(R_1,R_2) + 2w_1w_3\mathrm{Cov}(R_1,R_3) + 2w_2w_3\mathrm{Cov}(R_2,R_3)$$

$$= (0.20)^2(196) + (0.30)^2(225) + (0.50)^2(400) + 2(0.20)(0.30)(105) + (2)(0.20)(0.50)(140) + (2)(0.30)(0.50)(150)$$

$$= 7.84 + 20.25 + 100 + 12.6 + 28 + 45 = 213.69$$

The portfolio standard deviation is

$$\sigma^2(R_p) = 213.69^{1/2} = 14.62\%$$

27 B is correct. With Bayes' formula, the probability of failure given a "good" rating is

$$P(A \mid B) = \frac{P(B \mid A)}{P(B)}P(A)$$

where

$P(A) = 0.20$ = probability of failure
$P(B) = 0.70$ = probability of a "good" rating
$P(B \mid A) = 0.50$ = probability of a "good" rating given failure

With these estimates, the probability of failure given a "good" rating is

$$P(A \mid B) = \frac{P(B \mid A)}{P(B)}P(A) = \frac{0.50}{0.70} \times 0.20 = 0.143$$

If the analyst uses the bankruptcy prediction model as a guide, the probability of failure declines from 20% to 14.3%.

28 C is correct. With Bayes' formula, the probability of the CEO being fired given a "good" rating is

$$P(A \mid B) = \frac{P(B \mid A)}{P(B)}P(A)$$

where

$P(A)$ = 0.05 = probability of the CEO being fired
$P(B)$ = 0.50 = probability of a "good" rating
$P(B \mid A)$ = 0.30 = probability of a "good" rating given that the CEO is fired

With these estimates, the probability of the CEO being fired given a "good" rating is

$$P(A \mid B) = \frac{P(B \mid A)}{P(B)}P(A) = \frac{0.30}{0.50} \times 0.05 = 0.03$$

Although 5% of all CEOs are fired, the probability of being fired given a "good" performance rating is 3%.

29 C is correct. The combination formula provides the number of ways that *r* objects can be chosen from a total of *n* objects, when the order in which the *r* objects are listed does not matter. The order of the bonds within the portfolio does not matter.

30 A is correct. The answer is found using the combination formula

$$_{n}C_{r} = \binom{n}{r} = \frac{n!}{(n-r)!r!}$$

Here, $n = 4$ and $r = 2$, so the answer is $4!/[(4-2)!2!] = 24/[(2) \times (2)] = 6$. This result can be verified by assuming there are four vice presidents, VP1–VP4. The six possible additions to the investment committee are VP1 and VP2, VP1 and VP3, VP1 and VP4, VP2 and VP3, VP2 and VP4, and VP3 and VP4.

31 A is correct. The permutation formula is used to choose *r* objects from a total of *n* objects when order matters. Because the portfolio manager is trying to rank the four funds from most recommended to least recommended, the order of the funds matters; therefore, the permutation formula is most appropriate.

32 A is correct. The number of combinations is the number of ways to pick four mutual funds out of 10 without regard to order, which is

$$_{n}C_{r} = \frac{n!}{(n-r)!r!}$$

$$_{10}C_{4} = \frac{10!}{(10-4)!4!} = \frac{10 \times 9 \times 8 \times 7}{4 \times 3 \times 2 \times 1} = 210$$

33 B is correct. The number of permutations is the number of ways to pick five stocks out of 30 in the correct order.

$$_{n}P_{r} = \frac{n!}{(n-r)!r!}$$

$$_{30}P_{5} = \frac{30!}{(30-5)!} = \frac{30!}{25!} = 30 \times 29 \times 28 \times 27 \times 26 = 17{,}100{,}720$$

The contestant's chance of winning is one out of 17,100,720.

34 A is correct. The number of combinations is the number of ways to pick five stocks out of 30 without regard to order.

$$_nC_r = \frac{n!}{(n-r)!r!}$$

$$_{30}C_5 = \frac{30!}{(30-5)!5!} = \frac{30 \times 29 \times 28 \times 27 \times 26}{5 \times 4 \times 3 \times 2 \times 1} = 142{,}506$$

The contestant's chance of winning is one out of 142,506.

35 This contest does not resemble a usual lottery. Each of the 30 stocks does not have an equal chance of having the highest returns. Furthermore, contestants may have some favored investments, and the 30 stocks will not be chosen with the same frequencies. To guard against more than one person selecting the winners correctly, Sturm may wish to stipulate that if there is more than one winner, the winners will share the $1 million prize.

QUANTITATIVE METHODS
STUDY SESSION
2

Quantitative Methods (2)

This study session introduces the common probability distributions used to describe the behavior of random variables, such as asset prices and returns. How to estimate measures of a population (mean, standard deviation) based on a population sample is shown. The study session provides a framework for hypothesis testing, used for validating dataset hypotheses, along with techniques to test a hypothesis. Finally, simple linear regression is presented as a method for understanding the relationship between two variables as a way of making predictions.

READING ASSIGNMENTS

Reading 4	Common Probability Distributions by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, DBA, CFA, Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA
Reading 5	Sampling and Estimation by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, DBA, CFA, Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA
Reading 6	Hypothesis Testing by Pamela Peterson Drake, PhD, CFA
Reading 7	Introduction to Linear Regression by Pamela Peterson Drake, PhD, CFA

© 2021 CFA Institute. All rights reserved.

READING

4

Common Probability Distributions

**by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, CFA,
Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA**

Richard A. DeFusco, PhD, CFA, is at the University of Nebraska–Lincoln (USA). Dennis W. McLeavey, DBA, CFA, is at the University of Rhode Island (USA). Jerald E. Pinto, PhD, CFA. David E. Runkle, PhD, CFA, is at Jacobs Levy Equity Management (USA).

CFA Institute would like to thank Adam Kobor, PhD, CFA, at New York University Investment Office (USA), for this major revision of "Common Probability Distributions," including new visuals, graphics, Microsoft Excel functions, code snippets, and related text content throughout the reading.

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** define a probability distribution and compare and contrast discrete and continuous random variables and their probability functions;
☐	**b.** calculate and interpret probabilities for a random variable given its cumulative distribution function;
☐	**c.** describe the properties of a discrete uniform random variable, and calculate and interpret probabilities given the discrete uniform distribution function;
☐	**d.** describe the properties of the continuous uniform distribution, and calculate and interpret probabilities given a continuous uniform distribution;
☐	**e.** describe the properties of a Bernoulli random variable and a binomial random variable, and calculate and interpret probabilities given the binomial distribution function;
☐	**f.** explain the key properties of the normal distribution;
☐	**g.** contrast a multivariate distribution and a univariate distribution, and explain the role of correlation in the multivariate normal distribution;
☐	**h.** calculate the probability that a normally distributed random variable lies inside a given interval;
☐	**i.** explain how to standardize a random variable;
☐	**j.** calculate and interpret probabilities using the standard normal distribution;

(continued)

© 2019 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**k.** define shortfall risk, calculate the safety-first ratio, and identify an optimal portfolio using Roy's safety-first criterion;
☐	**l.** explain the relationship between normal and lognormal distributions and why the lognormal distribution is used to model asset prices;
☐	**m.** calculate and interpret a continuously compounded rate of return, given a specific holding period return;
☐	**n.** describe the properties of the Student's *t*-distribution, and calculate and interpret its degrees of freedom;
☐	**o.** describe the properties of the chi-square distribution and the *F*-distribution, and calculate and interpret their degrees of freedom;
☐	**p.** describe Monte Carlo simulation.

1

INTRODUCTION AND DISCRETE RANDOM VARIABLES

a define a probability distribution and compare and contrast discrete and continuous random variables and their probability functions

Probabilities play a critical role in investment decisions. Although we cannot predict the future, informed investment decisions are based on some kind of probabilistic thinking. An analyst may put probability estimates behind the success of her high-conviction or low-conviction stock recommendations. Risk managers would typically think in probabilistic terms: What is the probability of not achieving the target return, or what kind of losses are we facing with high likelihood over the relevant time horizon? Probability distributions also underpin validating trade signal–generating models: For example, does earnings revision play a significant role in forecasting stock returns?

In nearly all investment decisions, we work with random variables. The return on a stock and its earnings per share are familiar examples of random variables. To make probability statements about a random variable, we need to understand its probability distribution. A **probability distribution** specifies the probabilities associated with the possible outcomes of a random variable.

In this reading, we present important facts about seven probability distributions and their investment uses. These seven distributions—the uniform, binomial, normal, lognormal, Student's *t*-, chi-square, and *F*-distributions—are used extensively in investment analysis. Normal and binomial distributions are used in such basic valuation models as the Black–Scholes–Merton option pricing model, the binomial option pricing model, and the capital asset pricing model. Student's *t*-, chi-square, and *F*-distributions are applied in validating statistical significance and in hypothesis testing. With the working knowledge of probability distributions provided in this reading, you will be better prepared to study and use other quantitative methods, such as regression analysis, time-series analysis, and hypothesis testing. After discussing probability distributions, we end with an introduction to Monte Carlo simulation, a computer-based tool for obtaining information on complex investment problems.

We start by defining basic concepts and terms, then illustrate the operation of these concepts through the simplest distribution, the uniform distribution, and then address probability distributions that have more applications in investment work but also greater complexity.

1.1 Discrete Random Variables

b calculate and interpret probabilities for a random variable given its cumulative distribution function

A **random variable** is a quantity whose future outcomes are uncertain. The two basic types of random variables are discrete random variables and continuous random variables. A **discrete random variable** can take on at most a countable (possibly infinite) number of possible values. For example, a discrete random variable X can take on a limited number of outcomes $x_1, x_2, \ldots, x_n$ (n possible outcomes), or a discrete random variable Y can take on an unlimited number of outcomes $y_1, y_2, \ldots$ (without end). The number of "yes" votes at a corporate board meeting, for example, is a discrete variable that is countable and finite (from 0 to the voting number of board members). The number of trades at a stock exchange is also countable but is infinite, since there is no limit to the number of trades by the market participants. Note that X refers to the random variable, and x refers to an outcome of X. We subscript outcomes, as in x_1 and x_2, when we need to distinguish among different outcomes in a list of outcomes of a random variable. Since we can count all the possible outcomes of X and Y (even if we go on forever in the case of Y), both X and Y satisfy the definition of a discrete random variable.

In contrast, we cannot count the outcomes of a **continuous random variable**. We cannot describe the possible outcomes of a continuous random variable Z with a list $z_1, z_2, \ldots$, because the outcome $(z_1 + z_2)/2$, not in the list, would always be possible. The volume of water in a glass is an example of a continuous random variable since we cannot "count" water on a discrete scale but can only measure its volume. In finance, unless a variable exhibits truly discrete behavior—for example, a positive or negative earnings surprise or the number of central bank board members voting for a rate hike—it is practical to work with a continuous distribution in many cases. The rate of return on an investment is an example of such a continuous random variable.

In working with a random variable, we need to understand its possible outcomes. For example, a majority of the stocks traded on the New Zealand Stock Exchange are quoted in increments of NZ$0.01. Quoted stock price is thus a discrete random variable with possible values NZ$0, NZ$0.01, NZ$0.02, . . ., but we can also model stock price as a continuous random variable (as a lognormal random variable, to look ahead). In many applications, we have a choice between using a discrete or a continuous distribution. We are usually guided by which distribution is most efficient for the task we face. This opportunity for choice is not surprising, because many discrete distributions can be approximated with a continuous distribution, and vice versa. In most practical cases, a probability distribution is only a mathematical idealization, or approximate model, of the relative frequencies of a random variable's possible outcomes.

Every random variable is associated with a probability distribution that describes the variable completely. We can view a probability distribution in two ways. The basic view is the **probability function**, which specifies the probability that the random variable takes on a specific value: $P(X = x)$ is the probability that a random variable X takes on the value x. For a discrete random variable, the shorthand notation for the probability function (sometimes referred to as the "probability mass function") is $p(x) = P(X = x)$. For continuous random variables, the probability function is denoted $f(x)$ and called the **probability density function** (pdf), or just the density.

A probability function has two key properties (which we state, without loss of generality, using the notation for a discrete random variable):

- $0 \le p(x) \le 1$, because probability is a number between 0 and 1.
- The sum of the probabilities $p(x)$ over all values of X equals 1. If we add up the probabilities of all the distinct possible outcomes of a random variable, that sum must equal 1.
- calculate and interpret probabilities for a random variable given its cumulative distribution function

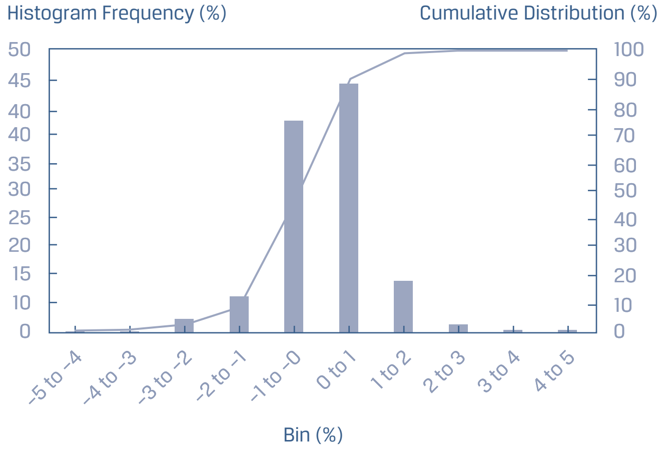
We are often interested in finding the probability of a range of outcomes rather than a specific outcome. In these cases, we take the second view of a probability distribution, the cumulative distribution function (cdf). The **cumulative distribution function**, or distribution function for short, gives the probability that a random variable X is less than or equal to a particular value x, $P(X \le x)$. For both discrete and continuous random variables, the shorthand notation is $F(x) = P(X \le x)$. How does the cumulative distribution function relate to the probability function? The word "cumulative" tells the story. To find $F(x)$, we sum up, or accumulate, values of the probability function for all outcomes less than or equal to x. The function of the cdf is parallel to that of cumulative relative frequency.

We illustrate the concepts of probability density functions and cumulative distribution functions with an empirical example using daily returns (i.e., percentage changes) of the fictitious Euro-Asia-Africa (EAA) Equity Index. This dataset spans five years and consists of 1,258 observations, with a minimum value of −4.1%, a maximum value of 5.0%, a range of 9.1%, and a mean daily return of 0.04%.

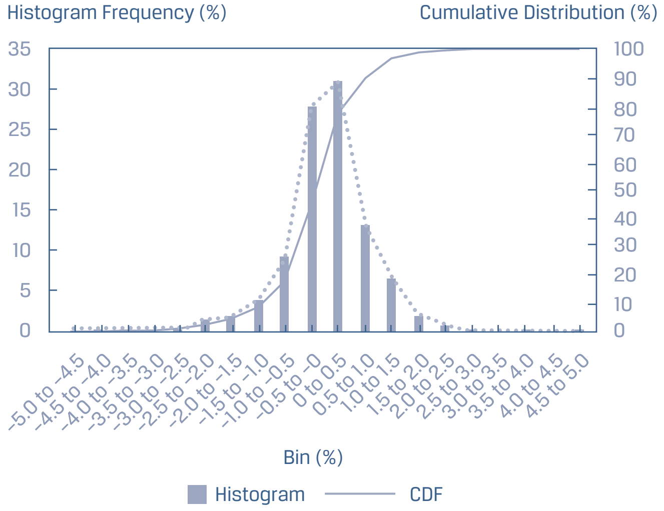
Exhibit 1 depicts the histograms, representing pdfs, and empirical cdfs (i.e., accumulated values of the bars in the histograms) based on daily returns of the EAA Equity Index. Panels A and B represent the same dataset; the only difference is the histogram bins used in Panel A are wider than those used in Panel B, so naturally Panel B has more bins. Note that in Panel A, we divided the range of observed daily returns (−5% to 5%) into 10 bins, so we chose the bin width to be 1.0%. In Panel B, we wanted a more granular histogram with a narrower range, so we divided the range into 20 bins, resulting in a bin width of 0.5%. Panel A gives a sense of the observed range of daily index returns, whereas Panel B is much more granular, so it more closely resembles continuous pdf and cdf graphs.

Exhibit 1 PDFs and CDFs of Daily Returns for EAA Equity Index

A. Wide Bin Widths



B. Narrow Bin Widths



EXAMPLE 1

Using PDFs and CDFs of Discrete and Continuous Random Variables to Calculate Probabilities

A. Discrete Random Variables: Rolling a Die

The example of rolling a six-sided die is an easy and intuitive way to illustrate a discrete random variable's pdf and cdf.

1 What is the probability that you would roll the number 3?

2 What is the probability that you would roll a number less than or equal to 3?

B. Continuous Random Variables: EAA Index Return

We use the EAA Index to illustrate the pdf and cdf for a continuously distributed random variable. Since daily returns can take on *any* numbers within a reasonable range rather than having discrete outcomes, we represent the pdf in the context of bins (as shown in the following table).

Bin	PDF	Bin	PDF
−5% to −4%	0.1%	0% to 1%	44.1%
−4% to −3%	0.6%	1% to 2%	8.8%
−3% to −2%	1.8%	2% to 3%	1.0%
−2% to −1%	6.1%	3% to 4%	0.1%
−1% to 0%	37.4%	4% to 5%	0.1%

In our sample, we did not find any daily returns below −5%, and we found only 0.1% of the total observations between −5% and −4%. In the next bin, −4% to −3%, we found 0.6% of the total observations, and so on.

3 If this empirical pdf is a guide for the future, what is the probability that we will see a daily return less than −2%?

Solution to 1:

Assuming the die is fair, rolling any number from 1 to 6 has a probability of 1/6 each, so the chance of rolling the number 3 would also equal 1/6. This outcome represents the pdf of this game; the pdf at number 3 takes a value of 1/6. In fact, it takes a value of 1/6 at every number from 1 to 6.

Solution to 2:

Answering this question involves the cdf of the die game. Three possible events would satisfy our criterion—namely, rolling 1, 2, or 3. The probability of rolling any of these numbers would be 1/6, so by accumulating them from 1 to 3, we get a probability of 3/6, or ½. The cdf of rolling a die takes a value of 1/6 at number 1, 3/6 (or 50%) at number 3, and 6/6 (or 100%) at number 6.

Solution to 3:

We must calculate the cdf up to −2%. The answer is the sum of the pdfs of the first three bins (see the shaded rectangle in the table provided); 0.1% + 0.6% + 1.8% = 2.5%. So, the probability that we will see a daily return less than −2% is 2.5%.

Next, we illustrate these concepts with examples and show how we use discrete and continuous distributions. We start with the simplest distribution, the discrete uniform distribution.

2 DISCRETE AND CONTINUOUS UNIFORM DISTRIBUTION

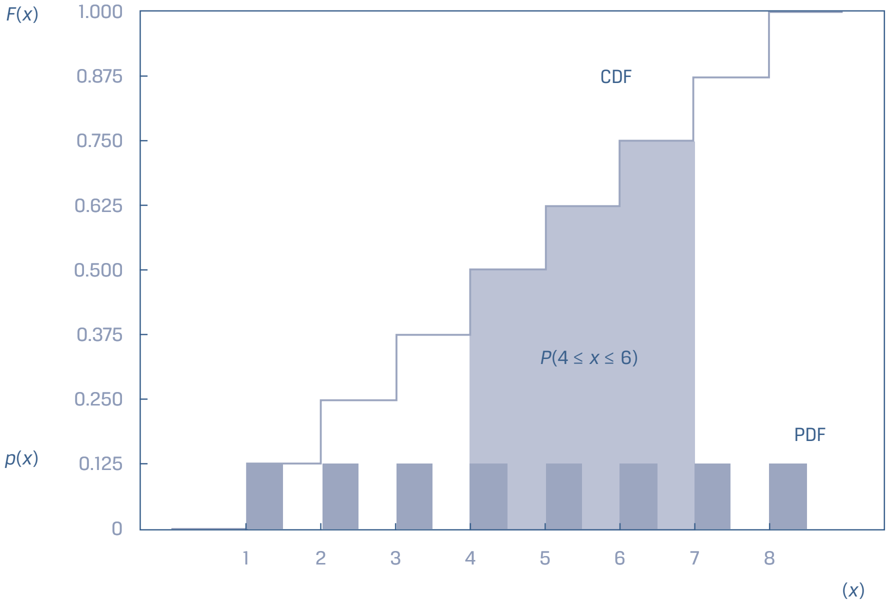
c describe the properties of a discrete uniform random variable, and calculate and interpret probabilities given the discrete uniform distribution function

The simplest of all probability distributions is the discrete uniform distribution. Suppose that the possible outcomes are the integers (whole numbers) 1–8, inclusive, and the probability that the random variable takes on any of these possible values is the same for all outcomes (that is, it is uniform). With eight outcomes, $p(x) = 1/8$, or 0.125, for all values of X ($X = 1, 2, 3, 4, 5, 6, 7, 8$); this statement is a complete description of this discrete uniform random variable. The distribution has a finite number of specified outcomes, and each outcome is equally likely. Exhibit 2 summarizes the two views of this random variable, the probability function and the cumulative distribution function, with Panel A in tabular form and Panel B in graphical form.

Exhibit 2 PDF and DCF for Discrete Uniform Random Variable

A. Probability Function and Cumulative Distribution Function for a Discrete Uniform Random Variable

$X = x$	Probability Function $p(x) = P(X = x)$	Cumulative Distribution Function $F(x) = P(X \le x)$
1	0.125	0.125
2	0.125	0.250
3	0.125	0.375
4	0.125	0.500
5	0.125	0.625
6	0.125	0.750
7	0.125	0.875
8	0.125	1.000



B. Graph of PDF and CDF for Discrete Uniform Random Variable

We can use the table in Panel A to find three probabilities: $P(X \le 7)$, $P(4 \le X \le 6)$, and $P(4 < X \le 6)$. The following examples illustrate how to use the cdf to find the probability that a random variable will fall in any interval (for any random variable, not only the uniform one). The results can also be gleaned visually from the graph in Panel B.

- The probability that X is less than or equal to 7, $P(X \le 7)$, is the next-to-last entry in the third column: 0.875, or 87.5%.
- To find $P(4 \le X \le 6)$, we need to find the sum of three probabilities: $p(4)$, $p(5)$, and $p(6)$. We can find this sum in two ways. We can add $p(4)$, $p(5)$, and $p(6)$ from the second column. Or we can calculate the probability as the difference between two values of the cumulative distribution function:

  $$F(6) = P(X \le 6) = p(6) + p(5) + p(4) + p(3) + p(2) + p(1),$$

  and

  $$F(3) = P(X \le 3) = p(3) + p(2) + p(1),$$

  so

  $$P(4 \le X \le 6) = F(6) - F(3) = p(6) + p(5) + p(4) = 3/8.$$

  So, we calculate the second probability as $F(6) - F(3) = 3/8$. This can be seen as the shaded area under the step function cdf graph in Panel B.
- The third probability, $P(4 < X \le 6)$, the probability that X is less than or equal to 6 but greater than 4, is $p(5) + p(6)$. We compute it as follows, using the cdf:

  $$P(4 < X \le 6) = P(X \le 6) - P(X \le 4) = F(6) - F(4) = p(6) + p(5) = 2/8.$$

  So we calculate the third probability as $F(6) - F(4) = 2/8$.

Suppose we want to check that the discrete uniform probability function satisfies the general properties of a probability function given earlier. The first property is $0 \le p(x) \le 1$. We see that $p(x) = 1/8$ for all x in the first column of in Panel A. [Note that $p(x)$ equals 0 for numbers x that are not in that column, such as −14 or 12.215.] The first property is satisfied. The second property is that the probabilities sum to 1. The entries in the second column of Panel A do sum to 1.

The cdf has two other characteristic properties:

- The cdf lies between 0 and 1 for any x: $0 \le F(x) \le 1$.
- As x increases, the cdf either increases or remains constant.

Check these statements by looking at the third column in the table in Panel A and at the graph in Panel B.

We now have some experience working with probability functions and cdfs for discrete random variables. Later, we will discuss Monte Carlo simulation, a methodology driven by random numbers. As we will see, the uniform distribution has an important technical use: It is the basis for generating random numbers, which, in turn, produce random observations for all other probability distributions.

2.1 Continuous Uniform Distribution

d describe the properties of the continuous uniform distribution, and calculate and interpret probabilities given a continuous uniform distribution

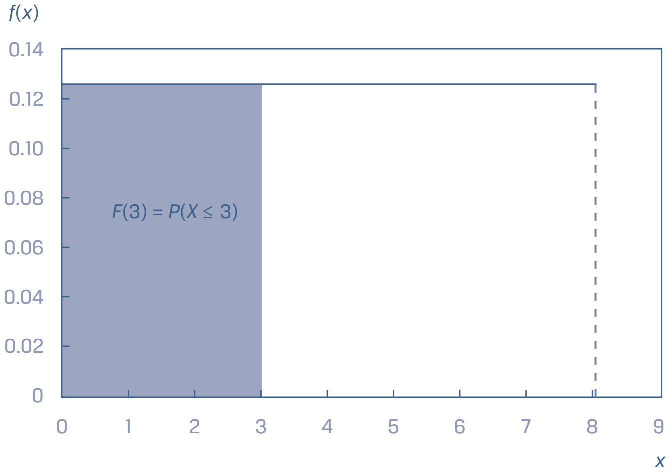
The continuous uniform distribution is the simplest continuous probability distribution. The uniform distribution has two main uses. As the basis of techniques for generating random numbers, the uniform distribution plays a role in Monte Carlo simulation. As the probability distribution that describes equally likely outcomes, the uniform distribution is an appropriate probability model to represent a particular kind of uncertainty in beliefs in which all outcomes appear equally likely.

The pdf for a uniform random variable is

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

For example, with $a = 0$ and $b = 8$, $f(x) = 1/8$, or 0.125. We graph this density in Exhibit 3.

Exhibit 3 Probability Density Function for a Continuous Uniform Distribution



The graph of the density function plots as a horizontal line with a value of 0.125.

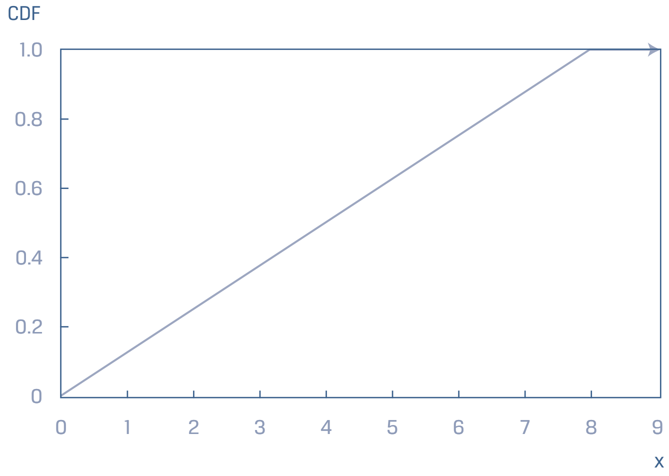
What is the probability that a uniform random variable with limits $a = 0$ and $b = 8$ is less than or equal to 3, or $F(3) = P(X \le 3)$? When we were working with the discrete uniform random variable with possible outcomes 1, 2, . . ., 8, we summed individual probabilities: $p(1) + p(2) + p(3) = 0.375$. In contrast, the probability that a continuous uniform random variable or any continuous random variable assumes any given fixed value is 0. To illustrate this point, consider the narrow interval 2.510–2.511. Because that interval holds an infinity of possible values, the sum of the probabilities of values in that interval alone would be infinite if each individual value in it had a positive probability. To find the probability $F(3)$, we find the area under the curve graphing the pdf, between 0 and 3 on the x-axis (shaded area in Exhibit 3). In calculus, this operation is called integrating the probability function $f(x)$ from 0 to 3. This area under the curve is a rectangle with base $3 - 0 = 3$ and height 1/8. The area of this rectangle equals base times height: $3(1/8) = 3/8$, or 0.375. So $F(3) = 3/8$, or 0.375.

The interval from 0 to 3 is three-eighths of the total length between the limits of 0 and 8, and $F(3)$ is three-eighths of the total probability of 1. The middle line of the expression for the cdf captures this relationship:

$$F(x) = \begin{cases} 0 \text{ for } x < a \\ \dfrac{x-a}{b-a} \text{ for } a \le x \le b. \\ 1 \text{ for } x < b \end{cases}$$

For our problem, $F(x) = 0$ for $x \le 0$, $F(x) = x/8$ for $0 < x < 8$, and $F(x) = 1$ for $x \ge 8$. Exhibit 4 shows a graph of this cdf.

Exhibit 4 Continuous Uniform Cumulative Distribution



The mathematical operation that corresponds to finding the area under the curve of a pdf $f(x)$ from a to b is the definite integral of $f(x)$ from a to b:

$$P(a \le X \le b) = \int_a^b f(x)dx, \tag{1}$$

where $\int dx$ is the symbol for summing $\int$ over small changes dx and the limits of integration (a and b) can be any real numbers or $-\infty$ and $+\infty$. All probabilities of continuous random variables can be computed using Equation 1. For the uniform distribution example considered previously, $F(7)$ is Equation 1 with lower limit $a = 0$ and upper limit $b = 7$. The integral corresponding to the cdf of a uniform distribution reduces to the three-line expression given previously. To evaluate Equation 1 for nearly all other continuous distributions, including the normal and lognormal, we rely on spreadsheet functions, computer programs, or tables of values to calculate probabilities. Those tools use various numerical methods to evaluate the integral in Equation 1.

Recall that the probability of a continuous random variable equaling any fixed point is 0. This fact has an important consequence for working with the cumulative distribution function of a continuous random variable: For any continuous random variable X, $P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$, because the probabilities at the endpoints a and b are 0. For discrete random variables, these relations of equality are not true, because for them probability accumulates at points.

EXAMPLE 2

Probability That a Lending Facility Covenant Is Breached

You are evaluating the bonds of a below-investment-grade borrower at a low point in its business cycle. You have many factors to consider, including the terms of the company's bank lending facilities. The contract creating a bank lending facility such as an unsecured line of credit typically has clauses known as covenants. These covenants place restrictions on what the borrower can do. The company will be in breach of a covenant in the lending facility if the interest coverage ratio, EBITDA/interest, calculated on EBITDA over the four trailing quarters, falls below 2.0. EBITDA is earnings before interest, taxes, depreciation, and amortization. Compliance with the covenants will be checked at the end of the current quarter. If the covenant is breached, the bank can demand immediate repayment of all borrowings on the facility. That action would probably trigger a liquidity crisis for the company. With a high degree of confidence, you forecast interest charges of $25 million. Your estimate of EBITDA runs from $40 million on the low end to $60 million on the high end.

Address two questions (treating projected interest charges as a constant):

1. If the outcomes for EBITDA are equally likely, what is the probability that EBITDA/interest will fall below 2.0, breaching the covenant?
2. Estimate the mean and standard deviation of EBITDA/interest. For a continuous uniform random variable, the mean is given by $\mu = (a + b)/2$ and the variance is given by $\sigma^2 = (b - a)^2/12$.

Solution to 1:

EBITDA/interest is a continuous uniform random variable because all outcomes are equally likely. The ratio can take on values between 1.6 = ($40 million)/($25 million) on the low end and 2.4 = ($60 million/$25 million) on the high end. The range of possible values is 2.4 − 1.6 = 0.8. The fraction of possible values falling below 2.0, the level that triggers default, is the distance between 2.0 and 1.6, or 0.40; the value 0.40 is one-half the total length of 0.8, or 0.4/0.8 = 0.50. So, the probability that the covenant will be breached is 50%.

Solution to 2:

In Solution 1, we found that the lower limit of EBITDA/interest is 1.6. This lower limit is a. We found that the upper limit is 2.4. This upper limit is b. Using the formula given previously,

$$\mu = (a + b)/2 = (1.6 + 2.4)/2 = 2.0.$$

The variance of the interest coverage ratio is

$$\sigma^2 = (b - a)^2/12 = (2.4 - 1.6)^2/12 = 0.053333.$$

The standard deviation is the positive square root of the variance, $0.230940 = (0.053333)^{1/2}$. However, the standard deviation is not particularly useful as a risk measure for a uniform distribution. The probability that lies within various standard deviation bands around the mean is sensitive to different specifications of the upper and lower. Here, a one standard deviation interval around the mean of 2.0 runs from 1.769 to 2.231 and captures 0.462/0.80 = 0.5775, or 57.8%, of the probability. A two standard deviation interval runs from 1.538 to 2.462, which extends past both the lower and upper limits of the random variable.

3 BINOMIAL DISTRIBUTION

e describe the properties of a Bernoulli random variable and a binomial random variable, and calculate and interpret probabilities given the binomial distribution function

In many investment contexts, we view a result as either a success or a failure or as binary (twofold) in some other way. When we make probability statements about a record of successes and failures or about anything with binary outcomes, we often use the binomial distribution. What is a good model for how a stock price moves over time? Different models are appropriate for different uses. Cox, Ross, and Rubinstein (1979) developed an option pricing model based on binary moves—price up or price down—for the asset underlying the option. Their binomial option pricing model was the first of a class of related option pricing models that have played an important role in the development of the derivatives industry. That fact alone would be sufficient reason for studying the binomial distribution, but the binomial distribution has uses in decision making as well.

The building block of the binomial distribution is the **Bernoulli random variable**, named after the Swiss probabilist Jakob Bernoulli (1654–1704). Suppose we have a trial (an event that may repeat) that produces one of two outcomes. Such a trial is a **Bernoulli trial**. If we let Y equal 1 when the outcome is success and Y equal 0 when the outcome is failure, then the probability function of the Bernoulli random variable Y is

$$p(1) = P(Y = 1) = p$$

and

$$p(0) = P(Y = 0) = 1 - p,$$

where p is the probability that the trial is a success.

In n Bernoulli trials, we can have 0 to n successes. If the outcome of an individual trial is random, the total number of successes in n trials is also random. A **binomial random variable** X is defined as the number of successes in n Bernoulli trials. A binomial random variable is the sum of Bernoulli random variables Y_i, where $i = 1, 2, \ldots, n$:

$$X = Y_1 + Y_2 + \ldots + Y_n,$$

where Y_i is the outcome on the ith trial (1 if a success, 0 if a failure). We know that a Bernoulli random variable is defined by the parameter p. The number of trials, n, is the second parameter of a binomial random variable. The binomial distribution makes these assumptions:

- The probability, p, of success is constant for all trials.
- The trials are independent.

The second assumption has great simplifying force. If individual trials were correlated, calculating the probability of a given number of successes in n trials would be much more complicated.

Under these two assumptions, a binomial random variable is completely described by two parameters, n and p. We write

$$X \sim B(n, p),$$

which we read as "X has a binomial distribution with parameters n and p." You can see that a Bernoulli random variable is a binomial random variable with $n = 1$: $Y \sim B(1, p)$.

Now we can find the general expression for the probability that a binomial random variable shows x successes in n trials (also known as the probability mass function). We can think in terms of a model of stock price dynamics that can be generalized to allow any possible stock price movements if the periods are made extremely small.

Each period is a Bernoulli trial: With probability p, the stock price moves up; with probability $1 - p$, the price moves down. A success is an up move, and x is the number of up moves or successes in n periods (trials). With each period's moves independent and p constant, the number of up moves in n periods is a binomial random variable. We now develop an expression for $P(X = x)$, the probability function for a binomial random variable.

Any sequence of n periods that shows exactly x up moves must show $n - x$ down moves. We have many different ways to order the up moves and down moves to get a total of x up moves, but given independent trials, any sequence with x up moves must occur with probability $p^x(1 - p)^{n-x}$. Now we need to multiply this probability by the number of different ways we can get a sequence with x up moves. Using a basic result in counting, there are

$$\frac{n!}{(n-x)!x!}$$

different sequences in n trials that result in x up moves (or successes) and $n - x$ down moves (or failures). Recall that for positive integers n, n factorial ($n!$) is defined as $n(n - 1)(n - 2) \ldots 1$ (and $0! = 1$ by convention). For example, $5! = (5)(4)(3)(2)(1) = 120$. The combination formula $n!/[(n - x)!x!]$ is denoted by

$$\binom{n}{x}$$

(read "n combination x" or "n choose x"). For example, over three periods, exactly three different sequences have two up moves: *uud*, *udu*, and *duu*. We confirm this by

$$\binom{3}{2} = \frac{3!}{(3-2)!2!} = \frac{(3)(2)(1)}{(1)(2)(1)} = 3.$$

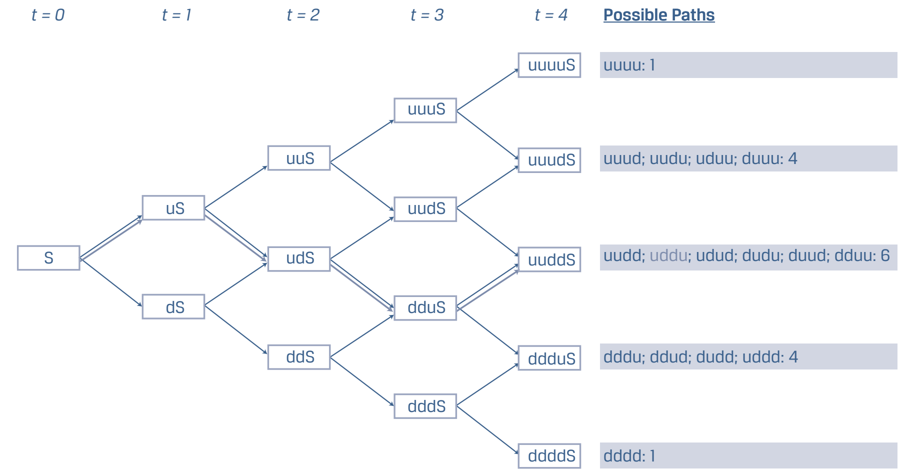
If, hypothetically, each sequence with two up moves had a probability of 0.15, then the total probability of two up moves in three periods would be 3 × 0.15 = 0.45. This example should persuade you that for X distributed $B(n, p)$, the probability of x successes in n trials is given by

$$p(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}. \tag{2}$$

Some distributions are always symmetric, such as the normal, and others are always asymmetric or skewed, such as the lognormal. The binomial distribution is symmetric when the probability of success on a trial is 0.50, but it is asymmetric or skewed otherwise.

We illustrate Equation 2 (the probability function) and the cdf through the symmetrical case by modeling the behavior of stock price movements on four consecutive trading days in a binomial tree framework. Each day is an independent trial. The stock moves up with constant probability p (the **up transition probability**); if it moves up, u is 1 plus the rate of return for an up move. The stock moves down with constant probability $1 - p$ (the **down transition probability**); if it moves down, d is 1 plus the rate of return for a down move. The **binomial tree** is shown in Exhibit 5, where we now associate each of the $n = 4$ stock price moves with time indexed by t; the shape of the graph suggests why it is a called a binomial tree. Each boxed value from which successive moves or outcomes branch out in the tree is called a **node**. The initial node, at $t = 0$, shows the beginning stock price, S. Each subsequent node represents a potential value for the stock price at the specified future time.

Exhibit 5 A Binomial Model of Stock Price Movement



We see from the tree that the stock price at $t = 4$ has five possible values: *uuuuS*, *uuudS*, *uuddS*, *ddduS*, and *ddddS*. The probability that the stock price equals *any* one of these five values is given by the binomial distribution. For example, four sequences of moves result in a final stock price of *uuudS*: These are *uuud*, *uudu*, *uduu*, and *duuu*. These sequences have three up moves out of four moves in total; the combination formula confirms that the number of ways to get three up moves (successes) in four periods (trials) is $4!/(4 - 3)!3! = 4$. Next, note that each of these sequences—*uuud*, *uudu*, *uduu*, and *duuu*—has probability $p^3(1 - p)^1$, which equals 0.0625 ($= 0.50^3 \times 0.50^1$). So, $P(S_4 = uuudS) = 4[p^3(1 - p)]$, or 0.25, where S_4 indicates the stock's price after four moves. This is shown numerically in Panel A of Exhibit 6, in the line indicating three up moves in x, as well as graphically in Panel B, as the height of the bar above $x = 3$. Note that in Exhibit 6, columns 5 and 6 in Panel A show the pdf and cdf, respectively, for this binomial distribution, and in Panel B, the pdf and cdf are represented by the bars and the line graph, respectively.

Exhibit 6 PDF and CDF of Binomial Probabilities for Stock Price Movements

A. Binomial Probabilities, $n = 4$ and $p = 0.50$

Col. 1 Number of Up Moves, x	Col. 2 Implied Number of Down Moves, $n - x$	Col. 3[A] Number of Possible Ways to Reach x Up Moves	Col. 4[B] Probability for Each Way, $p(x)$	Col. 5[C] Probability for x $p(x)$	Col. 6 $F(x) = P(X \le x)$
0	4	1	0.0625	0.0625	0.0625
1	3	4	0.0625	0.2500	0.3125
2	2	6	0.0625	0.3750	0.6875
3	1	4	0.0625	0.2500	0.9375
4	0	1	0.0625	0.0625	1.0000
				1.0000	

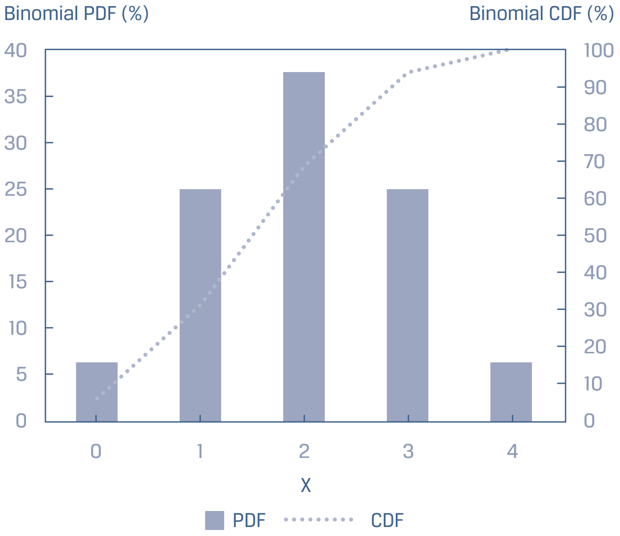
A: Column 3 = $n! / [(n - x)!\, x!]$

B: Column 4 = $p^x(1 - p)^{n-x}$

C: Column 5 = Column 3 × Column 4

Exhibit 6 (Continued)

B. Graphs of Binomial PDF and CDF



To be clear, the binomial random variable in this application is the number of up moves. Final stock price distribution is a function of the initial stock price, the *number* of up moves, and the *size* of the up moves and down moves. We cannot say that stock price itself is a binomial random variable; rather, it is a function of a binomial random variable, as well as of u and d, and initial price, S. This richness is actually one key to why this way of modeling stock price is useful: It allows us to choose values of these parameters to approximate various distributions for stock price (using a large number of time periods). One distribution that can be approximated is the lognormal, an important continuous distribution model for stock price that we will discuss later. The flexibility extends further. In the binomial tree shown in Exhibit 5, the transition probabilities are the same at each node: p for an up move and $1 - p$ for a down move. That standard formula describes a process in which stock return volatility is constant over time. Derivatives experts, however, sometimes model changing volatility over time using a binomial tree in which the probabilities for up and down moves differ at different nodes.

EXAMPLE 3

A Trading Desk Evaluates Block Brokers

Blocks are orders to sell or buy that are too large for the liquidity ordinarily available in dealer networks or stock exchanges. Your firm has known interests in certain kinds of stock. Block brokers call your trading desk when they want to sell blocks of stocks that they think your firm may be interested in buying. You know that these transactions have definite risks. For example, if the broker's client (the seller) has unfavorable information on the stock or if the total amount he or she is selling through all channels is not truthfully communicated, you may see an immediate loss on the trade. Your firm regularly audits the performance of block brokers by calculating the post-trade, market-risk-adjusted returns on stocks purchased from block brokers. On that basis, you classify each trade as unprofitable or profitable. You have summarized the performance of the brokers in a spreadsheet, excerpted in the following table for November of last year. The broker names are coded BB001 and BB002.

Block Trading Gains and Losses

	Profitable Trades	Losing Trades
BB001	3	9
BB002	5	3

You now want to evaluate the performance of the block brokers, and you begin with two questions:

1. If you are paying a fair price on average in your trades with a broker, what should be the probability of a profitable trade?
2. Did each broker meet or miss that expectation on probability?

You also realize that the brokers' performance has to be evaluated in light of the sample sizes, and for that you need to use the binomial probability function (Equation 2).

3. Under the assumption that the prices of trades were fair,
   - **A** calculate the probability of three or fewer profitable trades with broker BB001.
   - **B** calculate the probability of five or more profitable trades with broker BB002.

Solution to 1:

If the price you trade at is fair, then 50% of the trades you do with a broker should be profitable.

Solution to 2:

Your firm has logged 3 + 9 = 12 trades with block broker BB001. Since 3 of the 12 trades were profitable, the portion of profitable trades was 3/12, or 25%. With broker BB002, the portion of profitable trades was 5/8, or 62.5%. The rate of profitable trades with broker BB001 of 25% clearly missed your performance expectation of 50%. Broker BB002, at 62.5% profitable trades, exceeded your expectation.

Solution to 3:

A For broker BB001, the number of trades (the trials) was $n = 12$, and 3 were profitable. You are asked to calculate the probability of three or fewer profitable trades, $F(3) = p(3) + p(2) + p(1) + p(0)$.

Suppose the underlying probability of a profitable trade with BB001 is $p = 0.50$. With $n = 12$ and $p = 0.50$, according to Equation 2 the probability of three profitable trades is

$$p(3) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{12}{3}\left(0.50^3\right)\left(0.50^9\right)$$

$$= \frac{12!}{(12-3)!3!}0.50^{12} = 220(0.000244) = 0.053711.$$

The probability of exactly 3 profitable trades out of 12 is 5.4% if broker BB001 were giving you fair prices. Now you need to calculate the other probabilities:

$$p(2) = [12!/(12-2)!2!](0.50^2)(0.50^{10}) = 66(0.000244) = 0.016113.$$

$p(1) = [12!/(12 - 1)!1!](0.50^1)(0.50^{11}) = 12(0.000244) = 0.00293.$

$p(0) = [12!/(12 - 0)!0!](0.50^0)(0.50^{12}) = 1(0.000244) = 0.000244.$

Adding all the probabilities, $F(3) = 0.053711 + 0.016113 + 0.00293 + 0.000244 = 0.072998$, or 7.3%. The probability of making 3 or fewer profitable trades out of 12 would be 7.3% if your trading desk were getting fair prices from broker BB001.

B For broker BB002, you are assessing the probability that the underlying probability of a profitable trade with this broker was 50%, despite the good results. The question was framed as the probability of making five or more profitable trades if the underlying probability is 50%: $1 - F(4) = p(5) + p(6) + p(7) + p(8)$. You could calculate $F(4)$ and subtract it from 1, but you can also calculate $p(5) + p(6) + p(7) + p(8)$ directly.

You begin by calculating the probability that exactly five out of eight trades would be profitable if BB002 were giving you fair prices:

$$p(5) = \binom{8}{5}\left(0.50^5\right)\left(0.50^3\right)$$

$$= 56(0.003906) = 0.21875.$$

The probability is about 21.9%. The other probabilities are as follows:

$p(6) = 28(0.003906) = 0.109375.$

$p(7) = 8(0.003906) = 0.03125.$

$p(8) = 1(0.003906) = 0.003906.$

So, $p(5) + p(6) + p(7) + p(8) = 0.21875 + 0.109375 + 0.03125 + 0.003906 = 0.363281$, or 36.3%. A 36.3% probability is substantial; the underlying probability of executing a fair trade with BB002 might well have been 0.50 despite your success with BB002 in November of last year. If one of the trades with BB002 had been reclassified from profitable to unprofitable, exactly half the trades would have been profitable. In summary, your trading desk is getting at least fair prices from BB002; you will probably want to accumulate additional evidence before concluding that you are trading at better-than-fair prices.

The magnitude of the profits and losses in these trades is another important consideration. If all profitable trades had small profits but all unprofitable trades had large losses, for example, you might lose money on your trades even if the majority of them were profitable.

Two descriptors of a distribution that are often used in investments are the mean and the variance (or the standard deviation, the positive square root of variance). Exhibit 7 gives the expressions for the mean and variance of binomial random variables.

Exhibit 7 Mean and Variance of Binomial Random Variables

	Mean	Variance
Bernoulli, $B(1, p)$	P	$p(1 - p)$
Binomial, $B(n, p)$	Np	$np(1 - p)$

Because a single Bernoulli random variable, $Y \sim B(1, p)$, takes on the value 1 with probability p and the value 0 with probability $1 - p$, its mean or weighted-average outcome is p. Its variance is $p(1 - p)$. A general binomial random variable, $B(n, p)$, is the sum of n Bernoulli random variables, and so the mean of a $B(n, p)$ random variable is np. Given that a $B(1, p)$ variable has variance $p(1 - p)$, the variance of a $B(n, p)$ random variable is n times that value, or $np(1 - p)$, assuming that all the trials (Bernoulli random variables) are independent. We can illustrate the calculation for two binomial random variables with differing probabilities as follows:

Random Variable	Mean	Variance
$B(n = 5, p = 0.50)$	$2.50 = 5(0.50)$	$1.25 = 5(0.50)(0.50)$
$B(n = 5, p = 0.10)$	$0.50 = 5(0.10)$	$0.45 = 5(0.10)(0.90)$

For a $B(n = 5, p = 0.50)$ random variable, the expected number of successes is 2.5, with a standard deviation of $1.118 = (1.25)^{1/2}$; for a $B(n = 5, p = 0.10)$ random variable, the expected number of successes is 0.50, with a standard deviation of $0.67 = (0.45)^{1/2}$.

EXAMPLE 4

The Expected Number of Defaults in a Bond Portfolio

Suppose as a bond analyst you are asked to estimate the number of bond issues expected to default over the next year in an unmanaged high-yield bond portfolio with 25 US issues from distinct issuers. The credit ratings of the bonds in the portfolio are tightly clustered around Moody's B2/Standard & Poor's B, meaning that the bonds are speculative with respect to the capacity to pay interest and repay principal. The estimated annual default rate for B2/B rated bonds is 10.7%.

1. Over the next year, what is the expected number of defaults in the portfolio, assuming a binomial model for defaults?
2. Estimate the standard deviation of the number of defaults over the coming year.
3. Critique the use of the binomial probability model in this context.

Solution to 1:

For each bond, we can define a Bernoulli random variable equal to 1 if the bond defaults during the year and zero otherwise. With 25 bonds, the expected number of defaults over the year is $np = 25(0.107) = 2.675$, or approximately 3.

Solution to 2:

The variance is $np(1 - p) = 25(0.107)(0.893) = 2.388775$. The standard deviation is $(2.388775)^{1/2} = 1.55$. Thus, a two standard deviation confidence interval (± 3.10) about the expected number of defaults (≈ 3), for example, would run from approximately 0 to approximately 6.

Solution to 3:

An assumption of the binomial model is that the trials are independent. In this context, a trial relates to whether an individual bond issue will default over the next year. Because the issuing companies probably share exposure to common economic factors, the trials may not be independent. Nevertheless, for a quick estimate of the expected number of defaults, the binomial model may be adequate.

NORMAL DISTRIBUTION

4

f explain the key properties of the normal distribution

g contrast a multivariate distribution and a univariate distribution, and explain the role of correlation in the multivariate normal distribution

In this section, we focus on the two most important continuous distributions in investment work, the normal and lognormal.

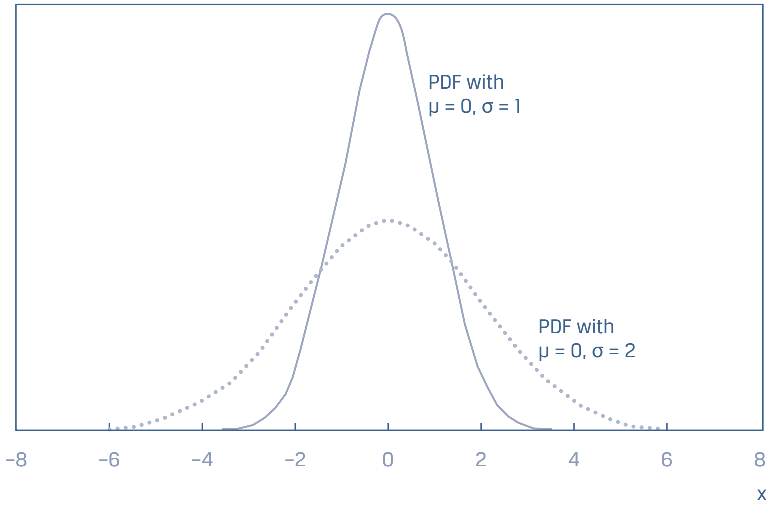
4.1 The Normal Distribution

The normal distribution may be the most extensively used probability distribution in quantitative work. It plays key roles in modern portfolio theory and in several risk management technologies. Because it has so many uses, the normal distribution must be thoroughly understood by investment professionals.

The role of the normal distribution in statistical inference and regression analysis is vastly extended by a crucial result known as the central limit theorem. The central limit theorem states that the sum (and mean) of a large number of independent random variables (with finite variance) is approximately normally distributed.

The French mathematician Abraham de Moivre (1667–1754) introduced the normal distribution in 1733 in developing a version of the central limit theorem. As Exhibit 8 shows, the normal distribution is symmetrical and bell-shaped. The range of possible outcomes of the normal distribution is the entire real line: all real numbers lying between $-\infty$ and $+\infty$. The tails of the bell curve extend without limit to the left and to the right.

Exhibit 8 PDFs of Two Different Normal Distributions



The defining characteristics of a normal distribution are as follows:

- The normal distribution is completely described by two parameters—its mean, μ, and variance, σ^2. We indicate this as $X \sim N(\mu, \sigma^2)$ (read "X follows a normal distribution with mean μ and variance σ^2"). We can also define a normal distribution in terms of the mean and the standard deviation, σ (this is often convenient because σ is measured in the same units as X and μ). As a consequence, we can answer any probability question about a normal random variable if we know its mean and variance (or standard deviation).
- The normal distribution has a skewness of 0 (it is symmetric). The normal distribution has a kurtosis of 3; its excess kurtosis (kurtosis − 3.0) equals 0. As a consequence of symmetry, the mean, the median, and the mode are all equal for a normal random variable.
- A linear combination of two or more normal random variables is also normally distributed.
- contrast a multivariate distribution and a univariate distribution, and explain the role of correlation in the multivariate normal distribution

The foregoing bullet points and descriptions concern a single variable or univariate normal distribution: the distribution of one normal random variable. A **univariate distribution** describes a single random variable. A **multivariate distribution** specifies the probabilities for a group of related random variables. You will encounter the **multivariate normal distribution** in investment work and readings and should know the following about it.

When we have a group of assets, we can model the distribution of returns on each asset individually or on the assets as a group. "As a group" implies that we take account of all the statistical interrelationships among the return series. One model that has often been used for security returns is the multivariate normal distribution. A multivariate normal distribution for the returns on n stocks is completely defined by three lists of parameters:

- the list of the mean returns on the individual securities (n means in total);
- the list of the securities' variances of return (n variances in total); and
- the list of all the distinct pairwise return correlations: $n(n - 1)/2$ distinct correlations in total.

The need to specify correlations is a distinguishing feature of the multivariate normal distribution in contrast to the univariate normal distribution.

The statement "assume returns are normally distributed" is sometimes used to mean a joint normal distribution. For a portfolio of 30 securities, for example, portfolio return is a weighted average of the returns on the 30 securities. A weighted average is a linear combination. Thus, portfolio return is normally distributed if the individual security returns are (joint) normally distributed. To review, in order to specify the normal distribution for portfolio return, we need the means, the variances, and the distinct pairwise correlations of the component securities.

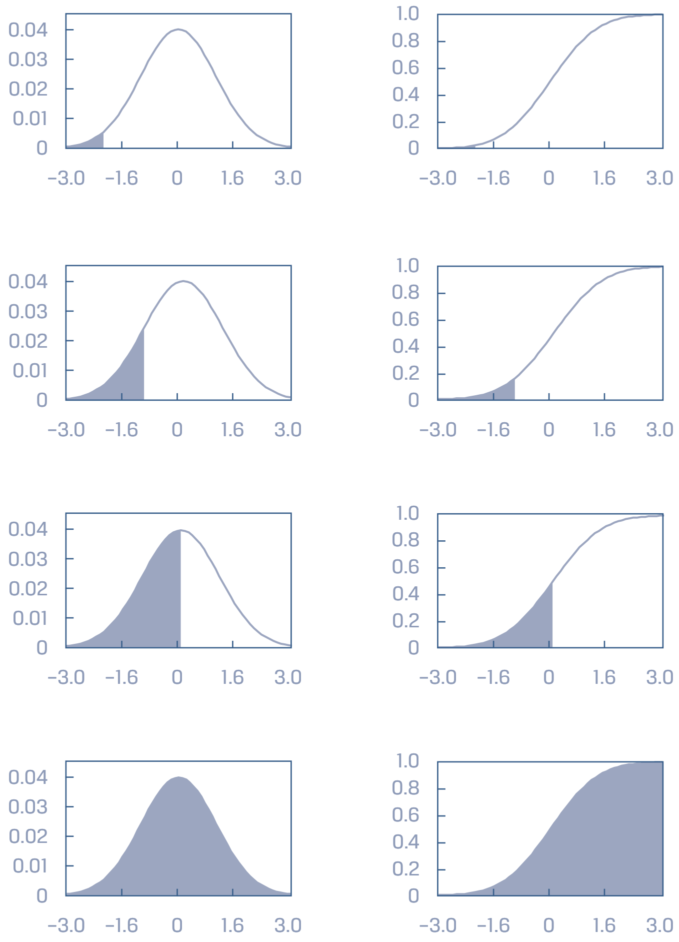
With these concepts in mind, we can return to the normal distribution for one random variable. The curves graphed in Exhibit 8 are the normal density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \text{ for } -\infty < x < +\infty. \tag{3}$$

The two densities graphed in Exhibit 8 correspond to a mean of $\mu = 0$ and standard deviations of $\sigma = 1$ and $\sigma = 2$. The normal density with $\mu = 0$ and $\sigma = 1$ is called the **standard normal distribution** (or **unit normal distribution**). Plotting two normal distributions with the same mean and different standard deviations helps us appreciate why standard deviation is a good measure of dispersion for the normal distribution: Observations are much more concentrated around the mean for the normal distribution with $\sigma = 1$ than for the normal distribution with $\sigma = 2$.

Exhibit 9 illustrates the relationship between the pdf (density function) and cdf (distribution function) of the standard normal distribution (mean = 0, standard deviation = 1). Most of the time, we associate a normal distribution with the "bell curve," which, in fact, is the probability density function of the normal distribution, depicted in Panel A. The cumulative distribution function, depicted in Panel B, in fact plots the size of the shaded areas of the pdfs. Let's take a look at the third row: In Panel A, we have shaded the bell curve up to $x = 0$, the mean of the standard normal distribution. This shaded area corresponds to 50% in the cdf graph, as seen in Panel B, meaning that 50% of the observations of a normally distributed random variable would be equal or less than the mean.

Exhibit 9 Density and Distribution Functions of the Standard Normal Distribution



A. PDFs

B. CDFs

Although not literally accurate, the normal distribution can be considered an approximate model for asset returns. Nearly all the probability of a normal random variable is contained within three standard deviations of the mean. For realistic values of mean return and return standard deviation for many assets, the normal probability of outcomes below −100% is very small.

Whether the approximation is useful in a given application is an empirical question. For example, Fama (1976) and Campbell, Lo, and MacKinlay (1997) showed that the normal distribution is a closer fit for quarterly and yearly holding period returns on a diversified equity portfolio than it is for daily or weekly returns. A persistent departure from normality in most equity return series is kurtosis greater than 3, the fat-tails problem. So when we approximate equity return distributions with the normal distribution, we should be aware that the normal distribution tends to underestimate the probability of extreme returns.

Fat tails can be modeled, among other things, by a mixture of normal random variables or by a Student’s *t*-distribution (which we shall cover shortly). In addition, since option returns are skewed, we should be cautious in using the symmetrical normal distribution to model the returns on portfolios containing significant positions in options.

The normal distribution is also less suitable as a model for asset prices than as a model for returns. An asset price can drop only to 0, at which point the asset becomes worthless. As a result, practitioners generally do not use the normal distribution to model the distribution of asset prices but work with the lognormal distribution, which we will discuss later.

4.2 Probabilities Using the Normal Distribution

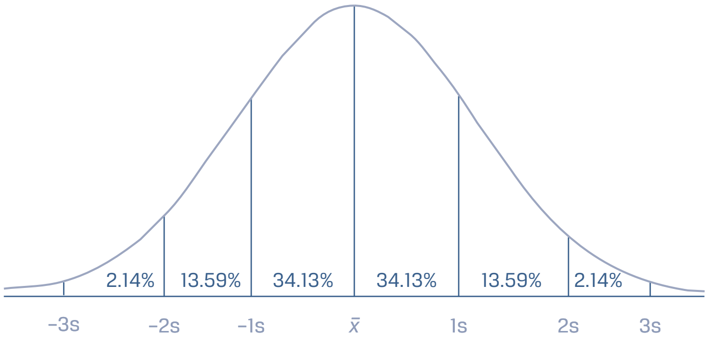
h calculate the probability that a normally distributed random variable lies inside a given interval

Having established that the normal distribution is the appropriate model for a variable of interest, we can use it to make the following probability statements:

- Approximately 50% of all observations fall in the interval $\mu \pm (2/3)\sigma$.
- Approximately 68% of all observations fall in the interval $\mu \pm \sigma$.
- Approximately 95% of all observations fall in the interval $\mu \pm 2\sigma$.
- Approximately 99% of all observations fall in the interval $\mu \pm 3\sigma$.

One, two, and three standard deviation intervals are illustrated in Exhibit 10. The intervals indicated are easy to remember but are only approximate for the stated probabilities. More precise intervals are $\mu \pm 1.96\sigma$ for 95% of the observations and $\mu \pm 2.58\sigma$ for 99% of the observations.

Exhibit 10 Units of Standard Deviation



In general, we do not observe the mean or the standard deviation of the distribution of the whole population, so we need to estimate them from an observable sample. We estimate the population mean, μ, using the sample mean, $\overline{X}$ (sometimes denoted as $\hat{\mu}$), and estimate the population standard deviation, σ, using the sample standard deviation, *s* (sometimes denoted as $\hat{\sigma}$).

EXAMPLE 5

Calculating Probabilities from the Normal Distribution

The chief investment officer of Fund XYZ would like to present some investment return scenarios to the Investment Committee, so she asks your assistance with some indicative numbers. Assuming daily asset returns are normally distributed, she would like to know the following:

1. What is the probability that returns would be less than or equal to 1 standard deviation below the mean?
2. What is the probability that returns would be between +1 and −1 standard deviation around the mean?
3. What is the probability that returns would be less than or equal to −2 standard deviations below the mean?
4. How far (in terms of standard deviation) must returns fall below the mean for the probability to equal 95%?

Note on Answering Questions 1–4:

Normal distribution–related functions are part of spreadsheets, R, Python, and all statistical packages. Here, we use Microsoft Excel functions to answer these questions. When we speak in terms of "number of standard deviations above or below the mean," we are referring to the standard normal distribution (i.e., mean of 0 and standard deviation of 1), so it is best to use Excel's "=NORM.S.DIST(Z, 0 or 1)" function. "Z" represents the distance in number of standard deviations away from the mean, and the second parameter of the function is either 0 (Excel returns pdf value) or 1 (Excel returns cdf value).

Solution to 1:

To answer Question 1, we need the normal cdf value (so, set the second parameter equal to 1) that is associated with a Z value of −1 (i.e., one standard deviation below the mean). Thus, "=NORM.S.DIST(-1,1)" returns 0.1587, or 15.9%.

Solution to 2:

Here, we need to calculate the area under the normal pdf within the range of the mean ±1 standard deviation. The area under the pdf is the cdf, so we must calculate the difference between the cdf one standard deviation above the mean and the cdf one standard deviation below the mean. Note that "=NORM.S.DIST(1,1)" returns 0.8413, or 84.1%, which means that 84.1% of all observations of a normally distributed random variable would fall below the mean plus one standard deviation. We already calculated 15.9% for the probability that observations for such a variable would fall less than one standard deviation below the mean in the Solution to 1, so the answer here is 84.1% − 15.9% = 68.3%.

Solution to 3:

Similar to Solution 1, use the Excel function "=NORM.S.DIST(-2,1)"—which returns a probability of 0.0228, or 2.3%.

Solution to 4:

This question is a typical way of phrasing "value at risk." In statistical terms, we want to know the lowest return value below which only 5% of the observations would fall. Thus, we need to find the Z value for which the normal cdf would be 5% probability. To do this, we use the inverse of the cdf function—that is,

"=NORM.S.INV(0.05)"—which results in −1.6449, or −1.64. In other words, only 5% of the observations should fall below the mean minus 1.64 standard deviations, or equivalently, 95% of the observations should exceed this threshold.

There are as many different normal distributions as there are choices for mean (μ) and variance (σ^2). We can answer all the previous questions in terms of any normal distribution. Spreadsheets, for example, have functions for the normal cdf for any specification of mean and variance. For the sake of efficiency, however, we would like to refer all probability statements to a single normal distribution. The standard normal distribution (the normal distribution with $\mu = 0$ and $\sigma = 1$) fills that role.

4.3 Standardizing a Random Variable

i. explain how to standardize a random variable

There are two steps in **standardizing** a normal random variable *X*: Subtract the mean of *X* from *X* and then divide that result by the standard deviation of *X* (this is also known as computing the *Z*-score). If we have a list of observations on a normal random variable, *X*, we subtract the mean from each observation to get a list of deviations from the mean and then divide each deviation by the standard deviation. The result is the standard normal random variable, *Z* (*Z* is the conventional symbol for a standard normal random variable). If we have $X \sim N(\mu, \sigma^2)$ (read "*X* follows the normal distribution with parameters μ and σ^2"), we standardize it using the formula

$$Z = (X - \mu)/\sigma. \tag{4}$$

Suppose we have a normal random variable, *X*, with $\mu = 5$ and $\sigma = 1.5$. We standardize *X* with $Z = (X - 5)/1.5$. For example, a value $X = 9.5$ corresponds to a standardized value of 3, calculated as $Z = (9.5 - 5)/1.5 = 3$. The probability that we will observe a value as small as or smaller than 9.5 for $X \sim N(5, 1.5)$ is exactly the same as the probability that we will observe a value as small as or smaller than 3 for $Z \sim N(0, 1)$.

4.4 Probabilities Using the Standard Normal Distribution

j calculate and interpret probabilities using the standard normal distribution

We can answer all probability questions about *X* using standardized values. We generally do not know the population mean and standard deviation, so we often use the sample mean $\overline{X}$ for μ and the sample standard deviation *s* for σ. Standard normal probabilities are computed with spreadsheets, statistical and econometric software, and programming languages. Tables of the cumulative distribution function for the standard normal random variable are also readily available.

To find the probability that a standard normal variable is less than or equal to 0.24, for example, calculate NORM.S.DIST(0.24,1)=0.5948; thus, $P(Z \leq 0.24) = 0.5948$, or 59.48%. If we want to find the probability of observing a value 1.65 standard deviations below the mean, calculate NORM.S.DIST(-1.65,1)=0.04947, or roughly 5%.

The following are some of the most frequently referenced values when using the normal distribution, and for these values, =NORM.S.INV(Probability) is a convenient Excel function:

- The 90th percentile point is 1.282, or NORM.S.INV(0.90)=1.28155. Thus, only 10% of values remain in the right tail beyond the mean plus 1.28 standard deviations;

- The 95th percentile point is 1.65, or NORM.S.INV(0.95)=1.64485, which means that $P(Z \leq 1.65) = N(1.65) = 0.95$, or 95%, and 5% of values remain in the right tail. The 5th percentile point, in contrast, is NORM.S.INV(0.05)=-1.64485—that is, the same number as for 95%, but with a negative sign.
- Note the difference between the use of a percentile point when dealing with one tail rather than two tails. We used 1.65 because we are concerned with the 5% of values that lie only on one side, the right tail. If we want to cut off both the left and right 5% tails, then 90% of values would stay within the mean ±1.65 standard deviations range.
- The 99th percentile point is 2.327: $P(Z \leq 2.327) = N(2.327) = 0.99$, or 99%, and 1% of values remain in the right tail.

EXAMPLE 6

Probabilities for a Common Stock Portfolio

Assume the portfolio mean return is 12% and the standard deviation of return estimate is 22% per year. Note also that if X is portfolio return, the standardized portfolio return is $Z = (X - \bar{X})/s = (X - 12\%)/22\%$. We use this expression throughout the solutions.
You want to calculate the following probabilities, assuming that a normal distribution describes returns.

1. What is the probability that portfolio return will exceed 20%?
2. What is the probability that portfolio return will be between 12% and 20%? In other words, what is $P(12\% \leq \text{portfolio return} \leq 20\%)$?
3. You can buy a one-year T-bill that yields 5.5%. This yield is effectively a one-year risk-free interest rate. What is the probability that your portfolio's return will be equal to or less than the risk-free rate?

Solution to 1:

For $X = 20\%$, $Z = (20\% - 12\%)/22\% = 0.363636$. You want to find $P(Z > 0.363636)$. First, note that $P(Z > x) = P(Z \geq x)$ because the normal distribution is a continuous distribution. Also, recall that $P(Z \geq x) = 1.0 - P(Z \leq x)$ or $1 - N(x)$. Next, NORM.S.DIST(0.363636,1)=0.64194, so, $1 - 0.6419 = 0.3581$. Therefore, the probability that portfolio return will exceed 20% is about 36% if your normality assumption is accurate.

Solution to 2:

$P(12\% \leq \text{Portfolio return} \leq 20\%) = N(Z \text{ corresponding to } 20\%) - N(Z \text{ corresponding to } 12\%)$. For the first term, $Z = (20\% - 12\%)/22\% = 0.363636$, and $N(0.363636) = 0.6419$ (as in Solution 1). To get the second term immediately, note that 12% is the mean, and for the normal distribution, 50% of the probability lies on either side of the mean. Therefore, $N(Z$ corresponding to 12%) must equal 50%. So $P(12\% \leq \text{Portfolio return} \leq 20\%) = 0.6419 - 0.50 = 0.1419$, or approximately 14%.

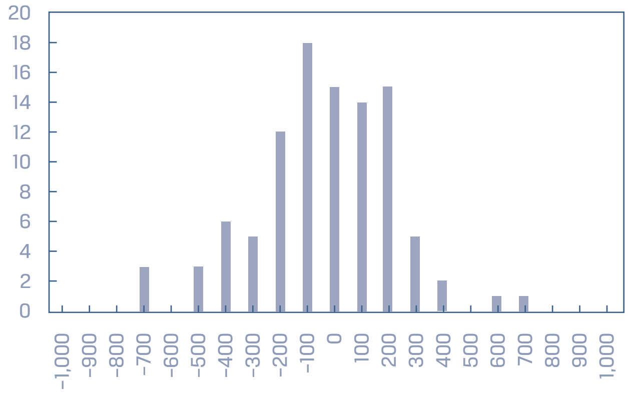
Solution to 3:

If X is portfolio return, then we want to find $P(\text{Portfolio return} \leq 5.5\%)$. For $X = 5.5\%$, $Z = (5.5\% - 12\%)/22\% = -0.2955$. Using NORM.S.DIST(-0.2955,1)=0.3838, we see an approximately 38% chance the portfolio's return will be equal to or less than the risk-free rate.

Next, we will briefly discuss and illustrate the concept of the **central limit theorem**, according to which the sum (as well as the mean) of a set of independent, identically distributed random variables with finite variances is normally distributed, whatever distribution the random variables follow.

To illustrate this concept, consider a sample of 30 observations of a random variable that can take a value of just −100, 0, or 100, with equal probability. Clearly, this sample is drawn from a simple discrete uniform distribution, where the possible values of −100, 0, and 100 each have 1/3 probability. We randomly pick 10 elements of this sample and calculate the sum of these elements, and then we repeat this process a total of 100 times. The histogram in Exhibit 11 shows the distribution of these sums: The underlying distribution is a very simple discrete uniform distribution, but the sums converge toward a normal distribution.

Exhibit 11 Central Limit Theorem: Sums of Elements from Discrete Uniform Distribution Converge to Normal Distribution



APPLICATIONS OF THE NORMAL DISTRIBUTION

5

k define shortfall risk, calculate the safety-first ratio, and identify an optimal portfolio using Roy's safety-first criterion

Modern portfolio theory (MPT) makes wide use of the idea that the value of investment opportunities can be meaningfully measured in terms of mean return and variance of return. In economic theory, **mean–variance analysis** holds exactly when investors are risk averse; when they choose investments so as to maximize expected utility, or satisfaction; and when either (1) returns are normally distributed or (2) investors have quadratic utility functions, a concept used in economics for a mathematical representation of attitudes toward risk and return. Mean–variance analysis, however, can still be useful—that is, it can hold approximately—when either Assumption 1 or 2 is violated. Because practitioners prefer to work with observables, such as returns, the proposition that returns are at least approximately normally distributed has played a key role in much of MPT.

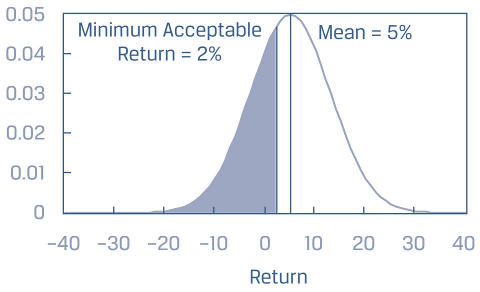
To illustrate this concept, assume an investor is saving for retirement, and although her goal is to earn the highest real return possible, she believes that the portfolio should at least achieve real capital preservation over the long term. Assuming a long-term expected inflation rate of 2%, the minimum acceptable return would be 2%.

Exhibit 12 compares three investment alternatives in terms of their expected returns and standard deviation of returns. The probability of falling below 2% is calculated on basis of the assumption of normally distributed returns. In the table, we see that Portfolio II, which combines the highest expected return and the lowest volatility, has the lowest probability of earning less than 2% (or equivalently, the highest probability of earning at least 2%). This can also be seen in Panel B, where Portfolio II has the smallest shaded area to the left of 2% (the probability of earning less than the minimum acceptable return).

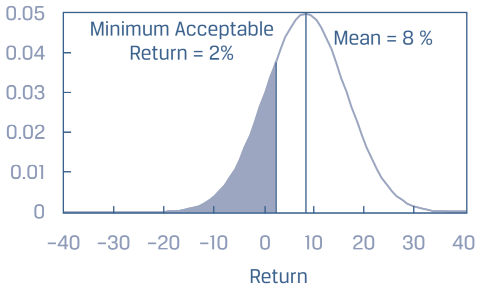
Exhibit 12 Probability of Earning a Minimum Acceptable Return

Portfolio	I	II	II
Expected return	5%	8%	5%
Standard deviation of return	8%	8%	12%
Probability of earning < 2% [$P(x < 2)$]	37.7%	24.6%	41.7%
Probability of earning ≥ 2% [$P(x \geq 2)$]	62.3%	75.4%	58.3%

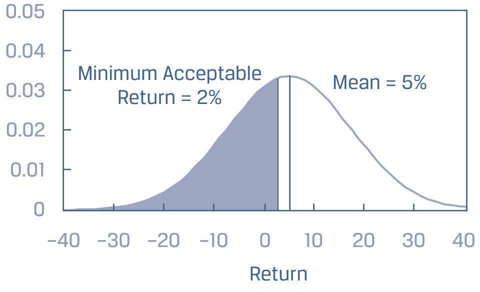
A. Portfolio I



B. Portfolio II



C. Portfolio III



Mean–variance analysis generally considers risk symmetrically in the sense that standard deviation captures variability both above and below the mean. An alternative approach evaluates only downside risk. We discuss one such approach, safety-first rules, because they provide an excellent illustration of the application of normal distribution theory to practical investment problems. **Safety-first rules** focus on **shortfall risk**, the risk that portfolio value (or portfolio return) will fall below some minimum acceptable level over some time horizon. The risk that the assets in a defined benefit plan will fall below plan liabilities is an example of a shortfall risk.

Suppose an investor views any return below a level of R_L as unacceptable. Roy's safety-first criterion (Roy 1952) states that the optimal portfolio minimizes the probability that portfolio return, R_P, will fall below the threshold level, R_L. In symbols, the investor's objective is to choose a portfolio that minimizes $P(R_P < R_L)$. When portfolio returns are normally distributed, we can calculate $P(R_P < R_L)$ using the number of standard deviations that R_L lies below the expected portfolio return, $E(R_P)$. The portfolio for which $E(R_P) - R_L$ is largest relative to standard deviation minimizes $P(R_P < R_L)$. Therefore, if returns are normally distributed, the safety-first optimal portfolio *maximizes* the safety-first ratio (SFRatio):

$$\text{SFRatio} = [E(R_P) - R_L]/\sigma_P \,. \tag{5}$$

The quantity $E(R_P) - R_L$ is the distance from the mean return to the shortfall level. Dividing this distance by σ_P gives the distance in units of standard deviation. There are two steps in choosing among portfolios using Roy's criterion (assuming normality):

1. Calculate each portfolio's SFRatio.
2. Choose the portfolio with the highest SFRatio.

For a portfolio with a given safety-first ratio, the probability that its return will be less than R_L is $N(-\text{SFRatio})$, and the safety-first optimal portfolio has the lowest such probability. For example, suppose an investor's threshold return, R_L, is 2%. He is presented with two portfolios. Portfolio 1 has an expected return of 12%, with a standard deviation of 15%. Portfolio 2 has an expected return of 14%, with a standard deviation of 16%. The SFRatios, using Equation 5, are 0.667 = (12 – 2)/15 and 0.75 = (14 – 2)/16 for Portfolios 1 and 2, respectively. For the superior Portfolio 2, the probability that portfolio return will be less than 2% is $N(-0.75) = 1 - N(0.75) = 1 - 0.7734 = 0.227$, or about 23%, assuming that portfolio returns are normally distributed.

You may have noticed the similarity of the SFRatio to the Sharpe ratio. If we substitute the risk-free rate, R_F, for the critical level R_L, the SFRatio becomes the Sharpe ratio. The safety-first approach provides a new perspective on the Sharpe ratio: When we evaluate portfolios using the Sharpe ratio, the portfolio with the highest Sharpe ratio is the one that minimizes the probability that portfolio return will be less than the risk-free rate (given a normality assumption).

EXAMPLE 7

The Safety-First Optimal Portfolio for a Client

You are researching asset allocations for a client in Canada with a C$800,000 portfolio. Although her investment objective is long-term growth, at the end of a year she may want to liquidate C$30,000 of the portfolio to fund educational expenses. If that need arises, she would like to be able to take out the C$30,000 without invading the initial capital of C$800,000. The table below shows three alternative allocations.

Mean and Standard Deviation for Three Allocations (in Percent)

Allocation	A	B	C
Expected annual return	25	11	14
Standard deviation of return	27	8	20

Address these questions (assume normality for Questions 2 and 3):

1 Given the client's desire not to invade the C$800,000 principal, what is the shortfall level, R_L? Use this shortfall level to answer Question 2.

2 According to the safety-first criterion, which of the three allocations is the best?

3 What is the probability that the return on the safety-first optimal portfolio will be less than the shortfall level?

Solution to 1:

Because C$30,000/C$800,000 is 3.75%, for any return less than 3.75% the client will need to invade principal if she takes out C$30,000. So, $R_L = 3.75\%$.

Solution to 2:

To decide which of the three allocations is safety-first optimal, select the alternative with the highest ratio $[E(R_P) - R_L]/\sigma_P$:

Allocation A: $0.787037 = (25 - 3.75)/27$.

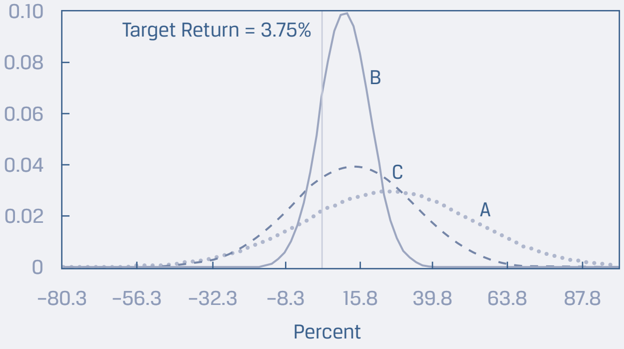
Allocation B: $0.90625 = (11 - 3.75)/8$.

Allocation C: $0.5125 = (14 - 3.75)/20$.

Allocation B, with the largest ratio (0.90625), is the best alternative according to the safety-first criterion.

Solution to 3:

To answer this question, note that $P(R_B < 3.75) = N(-0.90625)$. We can round 0.90625 to 0.91 for use with tables of the standard normal cdf. First, we calculate $N(-0.91) = 1 - N(0.91) = 1 - 0.8186 = 0.1814$, or about 18.1%. Using a spreadsheet function for the standard normal cdf on −0.90625 without rounding, we get 0.182402, or about 18.2%. The safety-first optimal portfolio has a roughly 18% chance of not meeting a 3.75% return threshold. This can be seen in the following graphic, where Allocation B has the smallest area under the distribution curve to the left of 3.75%.



Several points are worth noting. First, if the inputs were slightly different, we could get a different ranking. For example, if the mean return on B were 10% rather than 11%, Allocation A would be superior to B. Second, if meeting the 3.75% return threshold were a necessity rather than a wish, C$830,000 in one year could be modeled as a liability. Fixed-income strategies, such as cash flow matching, could be used to offset or immunize the C$830,000 quasi-liability.

In many investment contexts besides Roy's safety-first criterion, we use the normal distribution to estimate a probability. Another arena in which the normal distribution plays an important role is financial risk management. Financial institutions, such as investment banks, security dealers, and commercial banks, have formal systems to measure and control financial risk at various levels, from trading positions to the overall risk for the firm. Two mainstays in managing financial risk are value at risk (VaR) and stress testing/scenario analysis. **Stress testing** and **scenario analysis** refer to a set of techniques for estimating losses in extremely unfavorable combinations of events or scenarios. **Value at risk** (VaR) is a money measure of the minimum value of losses expected over a specified time period (for example, a day, a quarter, or a year) at a given level of probability (often 0.05 or 0.01). Suppose we specify a one-day time horizon and a level of probability of 0.05, which would be called a 95% one-day VaR. If this VaR equaled €5 million for a portfolio, there would be a 0.05 probability that the portfolio would lose €5 million or more in a single day (assuming our assumptions were correct). One of the basic approaches to estimating VaR, the variance–covariance or analytical method, assumes that returns follow a normal distribution.

LOGNORMAL DISTRIBUTION AND CONTINUOUS COMPOUNDING

6

6.1 The Lognormal Distribution

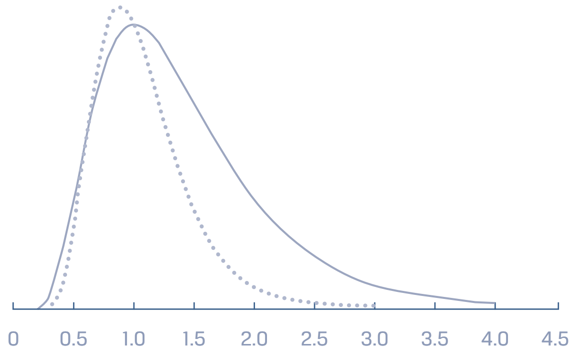
l. explain the relationship between normal and lognormal distributions and why the lognormal distribution is used to model asset prices

Closely related to the normal distribution, the lognormal distribution is widely used for modeling the probability distribution of share and other asset prices. For example, the lognormal distribution appears in the Black–Scholes–Merton option pricing model. The Black–Scholes–Merton model assumes that the price of the asset underlying the option is lognormally distributed.

A random variable Y follows a lognormal distribution if its natural logarithm, ln Y, is normally distributed. The reverse is also true: If the natural logarithm of random variable Y, ln Y, is normally distributed, then Y follows a lognormal distribution. If you think of the term lognormal as "the log is normal," you will have no trouble remembering this relationship.

The two most noteworthy observations about the lognormal distribution are that it is bounded below by 0 and it is skewed to the right (it has a long right tail). Note these two properties in the graphs of the pdfs of two lognormal distributions in Exhibit 13. Asset prices are bounded from below by 0. In practice, the lognormal distribution has been found to be a usefully accurate description of the distribution of prices for many financial assets. However, the normal distribution is often a good approximation for returns. For this reason, both distributions are very important for finance professionals.

Exhibit 13 Two Lognormal Distributions

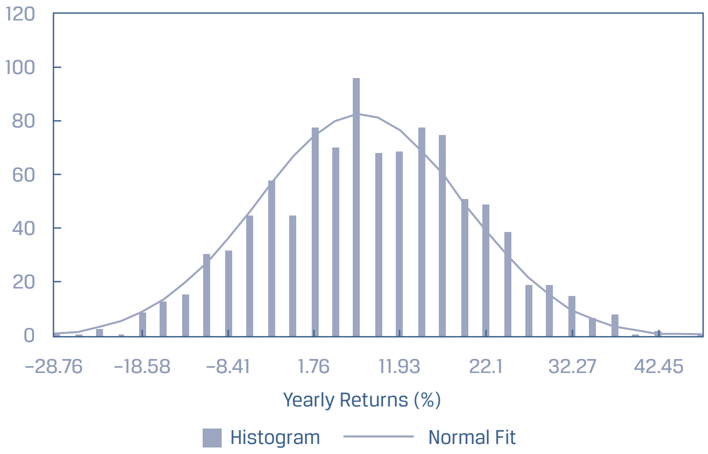


Like the normal distribution, the lognormal distribution is completely described by two parameters. Unlike the other distributions we have considered, a lognormal distribution is defined in terms of the parameters of a *different* distribution. The two parameters of a lognormal distribution are the mean and standard deviation (or variance) of its associated normal distribution: the mean and variance of ln Y, given that Y is lognormal. Remember, we must keep track of two sets of means and standard deviations (or variances): the mean and standard deviation (or variance) of the associated normal distribution (these are the parameters) and the mean and standard deviation (or variance) of the lognormal variable itself.

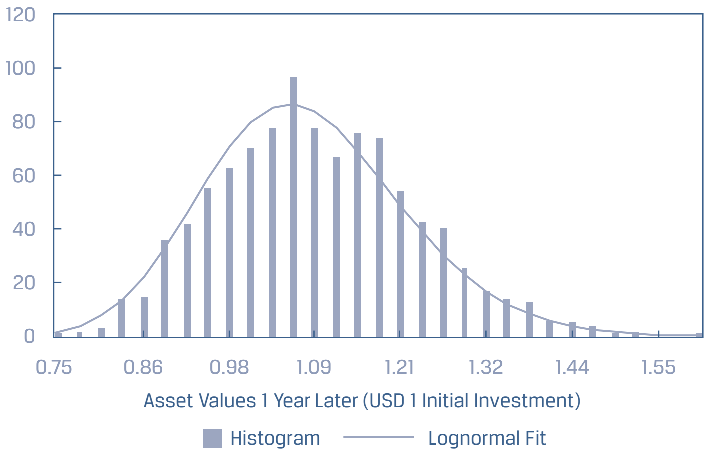
To illustrate this relationship, we simulated 1,000 scenarios of yearly asset returns, assuming that returns are normally distributed with 7% mean and 12% standard deviation. For each scenario i, we converted the simulated continuously compounded returns (r_i) to future asset prices with the formula Price(1 year later)$_i$ = \$1 x exp(r_i), where exp is the exponential function and assuming that the asset's price is \$1 today. In Exhibit 14, Panel A shows the distribution of the simulated returns together with the fitted normal pdf, whereas Panel B shows the distribution of the corresponding future asset prices together with the fitted lognormal pdf. Again, note that the lognormal distribution of future asset prices is bounded below by 0 and has a long right tail.

Exhibit 14 **Simulated Returns (Normal PDF) and Asset Prices (Lognormal PDF)**

A. Normal PDF



B. Lognormal PDF



The expressions for the mean and variance of the lognormal variable itself are challenging. Suppose a normal random variable X has expected value μ and variance σ^2. Define $Y = \exp(X)$. Remember that the operation indicated by $\exp(X)$ or e^X (where $e \approx 2.7183$) is the opposite operation from taking logs. Because $\ln Y = \ln[\exp(X)] = X$ is normal (we assume X is normal), Y is lognormal. What is the expected value of $Y = \exp(X)$? A guess might be that the expected value of Y is $\exp(\mu)$. The expected value is actually $\exp(\mu + 0.50\sigma^2)$, which is larger than $\exp(\mu)$ by a factor of $\exp(0.50\,\sigma^2) > 1$. To get some insight into this concept, think of what happens if we increase σ^2. The distribution spreads out; it can spread upward, but it cannot spread downward past 0. As a result, the center of its distribution is pushed to the right: The distribution's mean increases.

The expressions for the mean and variance of a lognormal variable are summarized below, where μ and σ^2 are the mean and variance of the associated normal distribution (refer to these expressions as needed, rather than memorizing them):

- Mean (μ_L) of a lognormal random variable = $\exp(\mu + 0.50\sigma^2)$.
- Variance ($\sigma_L{}^2$) of a lognormal random variable = $\exp(2\mu + \sigma^2) \times [\exp(\sigma^2) - 1]$.

6.2 Continuously Compounded Rates of Return

m calculate and interpret a continuously compounded rate of return, given a specific holding period return

We now explore the relationship between the distribution of stock return and stock price. In this section, we show that if a stock's continuously compounded return is normally distributed, then future stock price is necessarily lognormally distributed. Furthermore, we show that stock price may be well described by the lognormal distribution even when continuously compounded returns do not follow a normal distribution. These results provide the theoretical foundation for using the lognormal distribution to model asset prices.

To outline the presentation that follows, we first show that the stock price at some future time T, S_T, equals the current stock price, S_0, multiplied by e raised to power $r_{0,T}$, the continuously compounded return from 0 to T; this relationship is expressed as $S_T = S_0\exp(r_{0,T})$. We then show that we can write $r_{0,T}$ as the sum of shorter-term continuously compounded returns and that if these shorter-period returns are normally distributed, then $r_{0,T}$ is normally distributed (given certain assumptions) or approximately normally distributed (not making those assumptions). As S_T is proportional to the log of a normal random variable, S_T is lognormal.

To supply a framework for our discussion, suppose we have a series of equally spaced observations on stock price: S_0, S_1, S_2, . . ., S_T. Current stock price, S_0, is a known quantity and thus is nonrandom. The future prices (such as S_1), however, are random variables. The **price relative**, S_1/S_0, is an ending price, S_1, over a beginning price, S_0; it is equal to 1 plus the holding period return on the stock from $t = 0$ to $t = 1$:

$$S_1/S_0 = 1 + R_{0,1}.$$

For example, if S_0 = \$30 and S_1 = \$34.50, then S_1/S_0 = \$34.50/\$30 = 1.15. Therefore, $R_{0,1} = 0.15$, or 15%. In general, price relatives have the form

$$S_{t+1}/S_t = 1 + R_{t,t+1},$$

where $R_{t,t+1}$ is the rate of return from t to $t + 1$.

An important concept is the continuously compounded return associated with a holding period return, such as $R_{0,1}$. The **continuously compounded return** associated with a holding period return is the natural logarithm of 1 plus that holding period return, or equivalently, the natural logarithm of the ending price over the beginning price (the price relative). Note that here we are using lowercase r to refer specifically to continuously compounded returns. For example, if we observe a one-week holding period return of 0.04, the equivalent continuously compounded return, called the one-week continuously compounded return, is ln(1.04) = 0.039221; €1.00 invested for one week at 0.039221 continuously compounded gives €1.04, equivalent to a 4% one-week holding period return. The continuously compounded return from t to $t + 1$ is

$$r_{t,t+1} = \ln(S_{t+1}/S_t) = \ln(1 + R_{t,t+1}). \tag{6}$$

For our example, $r_{0,1} = \ln(S_1/S_0) = \ln(1 + R_{0,1}) = \ln(\$34.50/\$30) = \ln(1.15) = 0.139762$. Thus, 13.98% is the continuously compounded return from $t = 0$ to $t = 1$. The continuously compounded return is smaller than the associated holding period return. If our investment horizon extends from $t = 0$ to $t = T$, then the continuously compounded return to T is

$$r_{0,T} = \ln(S_T/S_0).$$

Applying the function exp to both sides of the equation, we have $\exp(r_{0,T}) = \exp[\ln(S_T/S_0)] = S_T/S_0$, so

$$S_T = S_0\exp(r_{0,T}).$$

We can also express S_T/S_0 as the product of price relatives:

$$S_T/S_0 = (S_T/S_{T-1})(S_{T-1}/S_{T-2}) \ldots (S_1/S_0).$$

Taking logs of both sides of this equation, we find that the continuously compounded return to time T is the sum of the one-period continuously compounded returns:

$$r_{0,T} = r_{T-1,T} + r_{T-2,T-1} + \ldots + r_{0,1}. \tag{7}$$

Using holding period returns to find the ending value of a \$1 investment involves the multiplication of quantities (1 + holding period return). Using continuously compounded returns involves addition (as shown in Equation 7).

A key assumption in many investment applications is that returns are **independently and identically distributed** (i.i.d.). Independence captures the proposition that investors cannot predict future returns using past returns. Identical distribution captures the assumption of stationarity, a property implying that the mean and variance of return do not change from period to period.

Assume that the one-period continuously compounded returns (such as $r_{0,1}$) are i.i.d. random variables with mean μ and variance σ^2 (but making no normality or other distributional assumption). Then,

$$E(r_{0,T}) = E(r_{T-1,T}) + E(r_{T-2,T-1}) + \ldots + E(r_{0,1}) = \mu T \tag{8}$$

(we add up μ for a total of T times), and

$$\sigma^2(r_{0,T}) = \sigma^2 T \tag{9}$$

(as a consequence of the independence assumption). The variance of the T holding period continuously compounded return is T multiplied by the variance of the one-period continuously compounded return; also, $\sigma(r_{0,T}) = \sigma\sqrt{T}$. If the one-period continuously compounded returns on the right-hand side of Equation 7 are normally distributed, then the T holding period continuously compounded return, $r_{0,T}$, is also normally distributed with mean μT and variance $\sigma^2 T$. This relationship is so because a linear combination of normal random variables is also normal. But even if the one-period continuously compounded returns are not normal, their sum, $r_{0,T}$, is approximately normal according to the central limit theorem. Now compare $S_T = S_0\exp(r_{0,T})$ to $Y = \exp(X)$, where X is normal and Y is lognormal (as we discussed previously). Clearly, we can model future stock price S_T as a lognormal random variable because $r_{0,T}$ should be at least approximately normal. This assumption of normally distributed returns is the basis in theory for the lognormal distribution as a model for the distribution of prices of shares and other assets.

Continuously compounded returns play a role in many asset pricing models, as well as in risk management. **Volatility** measures the standard deviation of the continuously compounded returns on the underlying asset; by convention, it is stated as an annualized measure. In practice, we very often estimate volatility using a historical series of continuously compounded daily returns. We gather a set of daily holding period returns and then use Equation 6 to convert them into continuously compounded daily returns. We then compute the standard deviation of the continuously compounded daily returns and annualize that number using Equation 9.

To compute the standard deviation of a set (or sample) of *n* returns, we sum the squared deviation of each return from the mean return and then divide that sum by $n - 1$. The result is the sample variance. Taking the square root of the sample variance gives the sample standard deviation. Annualizing is typically done on the basis of 250 days in a year, the approximate number of days markets are open for trading. Thus if daily volatility were 0.01, we would state volatility (on an annual basis) as $0.01\sqrt{250} = 0.1581$. Example 8 illustrates the estimation of volatility for the shares of Astra International.

EXAMPLE 8

Volatility of Share Price

Suppose you are researching Astra International (Indonesia Stock Exchange: ASII) and are interested in Astra's price action in a week in which international economic news had significantly affected the Indonesian stock market. You decide to use volatility as a measure of the variability of Astra shares during that week. The following shows closing prices during that week.

Astra International Daily Closing Prices

Day	Closing Price (IDR)
Monday	6,950
Tuesday	7,000
Wednesday	6,850
Thursday	6,600
Friday	6,350

Use the data provided to do the following:

1. Estimate the volatility of Astra shares. (Annualize volatility on the basis of 250 days in a year.)
2. Identify the probability distribution for Astra share prices if continuously compounded daily returns follow the normal distribution.

Solution to 1:

First, use Equation 6 to calculate the continuously compounded daily returns; then, find their standard deviation in the usual way. In calculating sample variance, to get sample standard deviation, the divisor is sample size minus 1.

$\ln(7{,}000/6{,}950) = 0.007168.$

$\ln(6{,}850/7{,}000) = -0.021661.$

$\ln(6{,}600/6{,}850) = -0.037179.$

$\ln(6{,}350/6{,}600) = -0.038615.$

Sum = −0.090287.

Mean = −0.022572.

Variance = 0.000452.

Standard deviation = 0.021261.

The standard deviation of continuously compounded daily returns is 0.021261. Equation 9 states that $\hat{\sigma}\left(r_{0,T}\right) = \hat{\sigma}\sqrt{T}$. In this example, $\hat{\sigma}$ is the sample standard deviation of one-period continuously compounded returns. Thus, $\hat{\sigma}$ refers to 0.021261. We want to annualize, so the horizon T corresponds to one year. Because $\hat{\sigma}$ is in days, we set T equal to the number of trading days in a year (250).

We find that annualized volatility for Astra stock that week was 33.6%, calculated as $0.021261\sqrt{250} = 0.336165$.

Note that the sample mean, −0.022572, is a possible estimate of the mean, μ, of the continuously compounded one-period or daily returns. The sample mean can be translated into an estimate of the expected continuously compounded annual return using Equation 8: $\hat{\mu}T = -0.022572(250)$ (using 250 to be consistent with the calculation of volatility). But four observations are far too few to estimate expected returns. The variability in the daily returns overwhelms any information about expected return in a series this short.

Solution to 2:

Astra share prices should follow the lognormal distribution if the continuously compounded daily returns on Astra shares follow the normal distribution.

We have shown that the distribution of stock price is lognormal, given certain assumptions. What are the mean and variance of S_T if S_T follows the lognormal distribution? Earlier we gave bullet-point expressions for the mean and variance of a lognormal random variable. In the bullet-point expressions, the $\hat{\mu}$ and $\hat{\sigma}^2$ would refer, in the context of this discussion, to the mean and variance of the T horizon (not the one-period) continuously compounded returns (assumed to follow a normal distribution), compatible with the horizon of S_T. Related to the use of mean and variance (or standard deviation), previously we used those quantities to construct intervals in which we expect to find a certain percentage of the observations of a normally distributed random variable. Those intervals were symmetric about the mean. Can we state similar symmetric intervals for a lognormal random variable? Unfortunately, we cannot; because the lognormal distribution is not symmetric, such intervals are more complicated than for the normal distribution, and we will not discuss this specialist topic here.

STUDENT'S *T*-, CHI-SQUARE, AND *F*-DISTRIBUTIONS

7

7.1 Student's *t*-Distribution

- describe the properties of the Student's *t*-distribution, and calculate and interpret its degrees of freedom

To complete the review of probability distributions commonly used in finance, we discuss Student's *t*-, chi-square, and *F*-distributions. Most of the time, these distributions are used to support statistical analyses, such as sampling, testing the statistical significance of estimated model parameters, or hypothesis testing. In addition, Student's *t*-distribution is also sometimes used to model asset returns in a manner similar to that of the normal distribution. However, since the *t*-distribution has "longer tails," it may provide a more reliable, more conservative downside risk estimate.

The standard *t*-distribution is a symmetrical probability distribution defined by a single parameter known as **degrees of freedom** (df), the number of independent variables used in defining sample statistics, such as variance, and the probability distributions they measure.

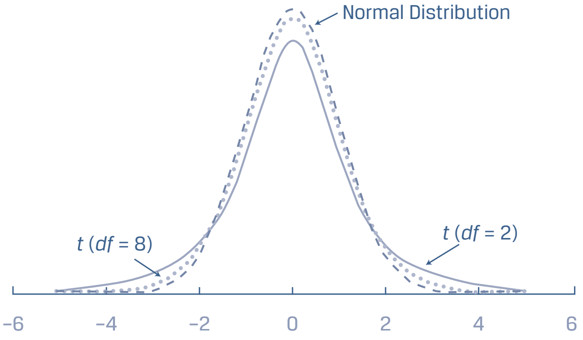
Each value for the number of degrees of freedom defines one distribution in this family of distributions. We will shortly compare *t*-distributions with the standard normal distribution, but first we need to understand the concept of degrees of freedom. We can do so by examining the calculation of the sample variance,

$$s^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}. \tag{10}$$

Equation 10 gives the unbiased estimator of the sample variance that we use. The term in the denominator, $n - 1$, which is the sample size minus 1, is the number of degrees of freedom in estimating the population variance when using Equation 10. We also use $n - 1$ as the number of degrees of freedom for determining reliability factors based on the *t*-distribution. The term "degrees of freedom" is used because in a random sample, we assume that observations are selected independently of each other. The numerator of the sample variance, however, uses the sample mean. How does the use of the sample mean affect the number of observations collected independently for the sample variance formula? With a sample size of 10 and a mean of 10%, for example, we can freely select only 9 observations. Regardless of the 9 observations selected, we can always find the value for the 10th observation that gives a mean equal to 10%. From the standpoint of the sample variance formula, then, there are nine degrees of freedom. Given that we must first compute the sample mean from the total of n independent observations, only $n - 1$ observations can be chosen independently for the calculation of the sample variance. The concept of degrees of freedom comes up frequently in statistics, and you will see it often later in the CFA Program curriculum.

Suppose we sample from a normal distribution. The ratio $z = \left(\bar{X} - \mu\right)\Big/\left(\sigma/\sqrt{n}\right)$ is distributed normally with a mean of 0 and standard deviation of 1; however, the ratio $t = \left(\bar{X} - \mu\right)\Big/\left(s/\sqrt{n}\right)$ follows the *t*-distribution with a mean of 0 and $n - 1$ degrees of freedom. The ratio represented by t is not normal because t is the ratio of two random variables, the sample mean and the sample standard deviation. The definition of the standard normal random variable involves only one random variable, the sample mean. As degrees of freedom increase (i.e., as sample size increases), however, the *t*-distribution approaches the standard normal distribution. Exhibit 15 shows the probability density functions for the standard normal distribution and two *t*-distributions, one with df = 2 and one with df = 8.

Exhibit 15 Student's *t*-Distributions vs. Standard Normal Distribution



Of the three distributions shown in Exhibit 15, the standard normal distribution has tails that approach zero faster than the tails of the two *t*-distributions. The *t*-distribution is also symmetrically distributed around its mean value of zero, just like the normal distribution. As the degrees of freedom increase, the *t*-distribution approaches the standard normal distribution. The *t*-distribution with df = 8 is closer to the standard normal distribution than the *t*-distribution with df = 2.

Beyond plus and minus four standard deviations from the mean, the area under the standard normal distribution appears to approach 0; both *t*-distributions, however, continue to show some area under each curve beyond four standard deviations. The *t*-distributions have fatter tails, but the tails of the *t*-distribution with df = 8 more closely resemble the normal distribution's tails. As the degrees of freedom increase, the tails of the *t*-distribution become less fat.

Probabilities for the *t*-distribution can be readily computed with spreadsheets, statistical software, and programming languages. As an example of the latter, see the final sidebar at the end of this section for sample code in the R programming language.

7.2 Chi-Square and *F*-Distribution

o describe the properties of the chi-square distribution and the *F*-distribution, and calculate and interpret their degrees of freedom

The chi-square distribution, unlike the normal and *t*-distributions, is asymmetrical. Like the *t*-distribution, the chi-square distribution is a family of distributions. The chi-square distribution with *k* degrees of freedom is the distribution of the sum of the squares of *k* independent standard normally distributed random variables; hence, this distribution does not take on negative values. A different distribution exists for each possible value of degrees of freedom, $n - 1$ (n is sample size).

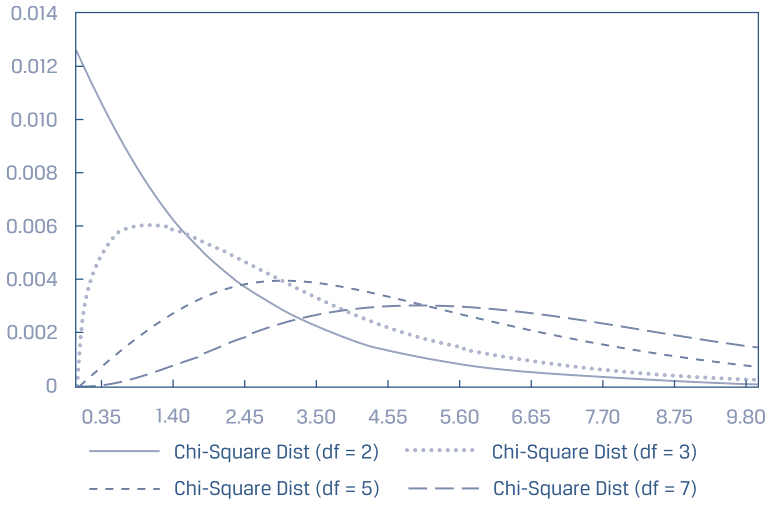
Like the chi-square distribution, the *F*-distribution is a family of asymmetrical distributions bounded from below by 0. Each *F*-distribution is defined by two values of degrees of freedom, called the numerator and denominator degrees of freedom.

The relationship between the chi-square and *F*-distributions is as follows: If χ_1^2 is one chi-square random variable with m degrees of freedom and χ_2^2 is another chi-square random variable with n degrees of freedom, then $F = \left(\chi_1^2/m\right)/\left(\chi_2^2/n\right)$ follows an *F*-distribution with m numerator and n denominator degrees of freedom.

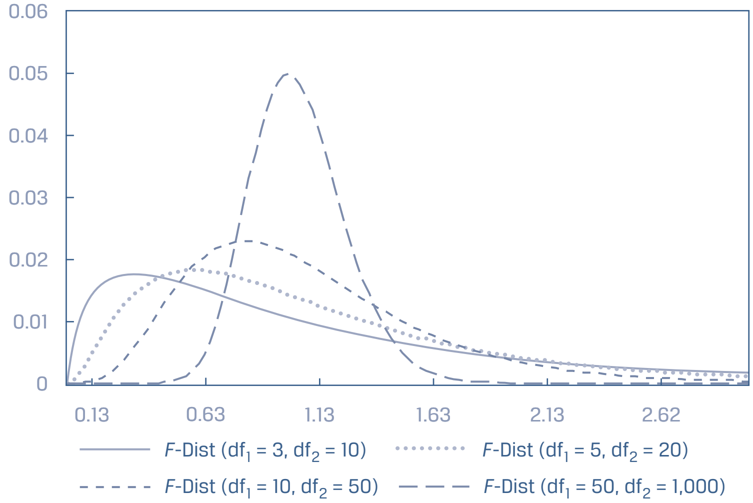
Chi-square and *F*-distributions are asymmetric, and as shown in Exhibit 16, the domain of their pdfs are positive numbers. Like Student's *t*-distribution, as the degrees of freedom of the chi-square distribution increase, the shape of its pdf becomes more similar to a bell curve (see Panel A). For the *F*-distribution, as both the numerator (df_1) and the denominator (df_2) degrees of freedom increase, the density function will also become more bell curve–like (see Panel B).

Exhibit 16 PDFs of Chi-Square and *F*-Distributions

A. Chi-Square Distributions



B. F-Distributions



As for typical investment applications, Student's *t*, chi-square, and *F*-distributions are the basis for test statistics used in performing various types of hypothesis tests on portfolio returns, such as those summarized in Exhibit 17.

Exhibit 17 Student's *t*, Chi-Square, and *F*-Distributions: Basis for Hypothesis Tests of Investment Returns

Distribution	Test Statistic	Hypothesis Tests of Returns
Student's *t*	*t*-Statistic	Tests of a single population mean, of differences between two population means, of mean difference between paired (dependent) populations, and of population correlation coefficient
Chi-square	Chi-square statistic	Test of variance of a normally distributed population
F	*F*-statistic	Test of equality of variances of two normally distributed populations from two independent random samples

EXAMPLE 9

Probabilities Using Student's-*t*, Chi-Square, and *F*-Distributions

1. Of the distributions we have covered in this reading, which can take values that are only positive numbers (i.e., no negative values)?
2. Interpret the degrees of freedom for a chi-square distribution, and describe how a larger value of df affects the shape of the chi-square pdf.
3. Generate cdf tables in Excel for values 1, 2, and 3 for the following distributions: standard normal, Student's *t*- (df = 5), chi-square (df = 5), and *F*-distribution ($df_1 = 5$, $df_2 = 1$). Then, calculate the distance from the mean for probability (p) = 90%, 95%, and 99% for each distribution.
4. You fit a Student's *t*-distribution to historically observed returns of stock market index ABC. Your best fit comes with five degrees of freedom. Compare this Student's *t*-distribution (df = 5) to a standard normal distribution on the basis of your answer to Question 3.

Solution to 1:

Of the probability distributions covered in this reading, the domains of the pdfs of the lognormal, the chi-square, and the *F*-distribution are only positive numbers.

Solution to 2:

A chi-square distribution with *k* degrees of freedom is the distribution of the sum of the squares of *k* independent standard normally distributed random variables. The greater the degrees of freedom, the more symmetrical and bell curve–like the pdf becomes.

Solution to 3:

In Excel, we can calculate cdfs using the NORM.S.DIST(value,1), T.DIST(value,DF,1), CHISQ.DIST(value,DF,1), and F.DIST(value,DF1,DF2,1) functions for the standard normal, Student's *t*-, chi-square, and *F*-distributions, respectively. At the end of this question set, we also show code snippets in the R language for generating cdfs for the requested values. For values 1, 2, and 3, the following are the results using the Excel functions:

CDF Values Using Different Probability Distributions

Value	Normal	Student's *t* (df = 5)	Chi-Square (df = 5)	*F* (df_1 = 5, df_2 = 1)
1	84.1%	81.8%	3.7%	36.3%
2	97.7%	94.9%	15.1%	51.1%
3	99.9%	98.5%	30.0%	58.9%

To calculate distances from the mean given probability *p*, we must use the inverse of the distribution functions: NORM.S.INV(p), T.INV(p,DF), CHISQ.INV(p,DF), and F.INV(p,DF1,DF2), respectively. At the end of this question set, we also show code snippets in the R language for calculating distances from the mean for the requested probabilities. The results using the inverse functions and the requested probabilities are as follows:

Distance from the Mean for a Given Probability (*p*)

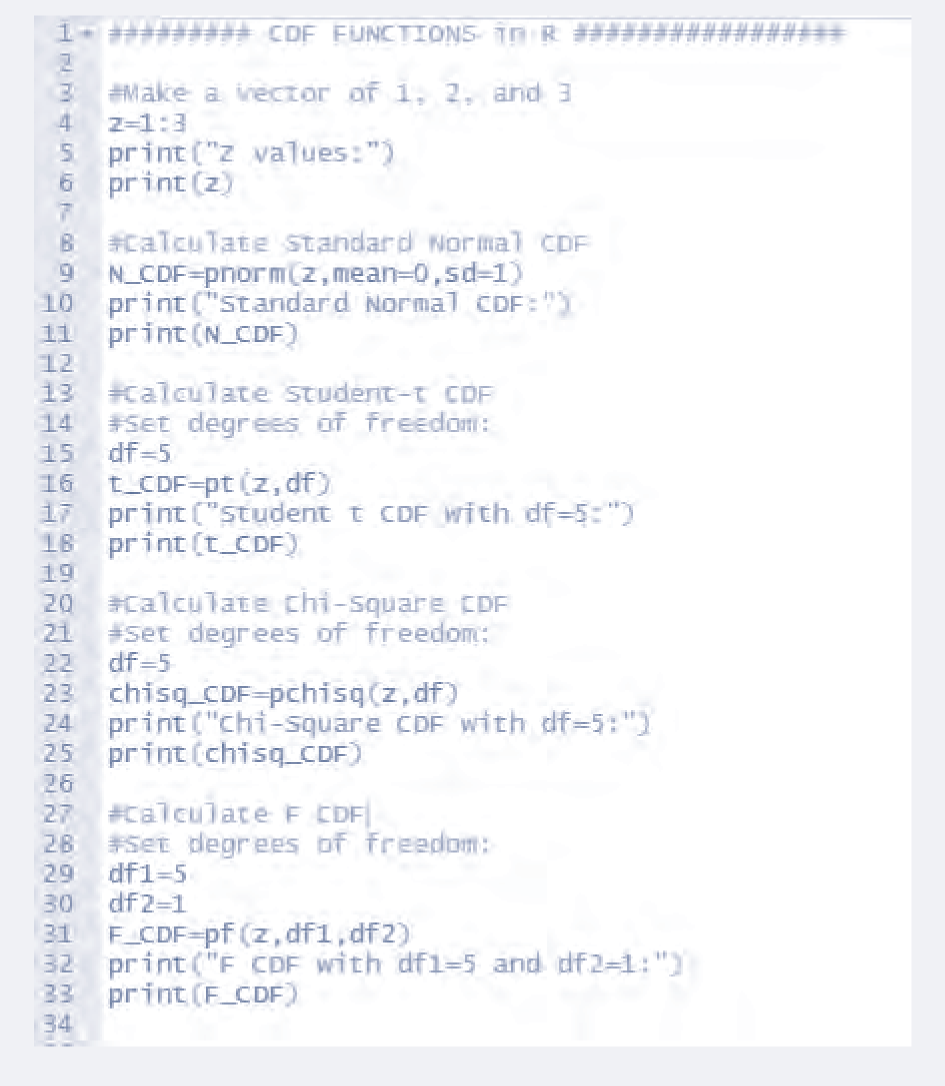
Probability	Normal	Student's *t* (df = 5)	Chi-Square (df = 5)	*F* (df_1 = 5, df_2 = 1)
90%	1.28	1.48	9.24	57.24
95%	1.64	2.02	11.07	230.16
99%	2.33	3.36	15.09	5,763.65

Solution to 4:

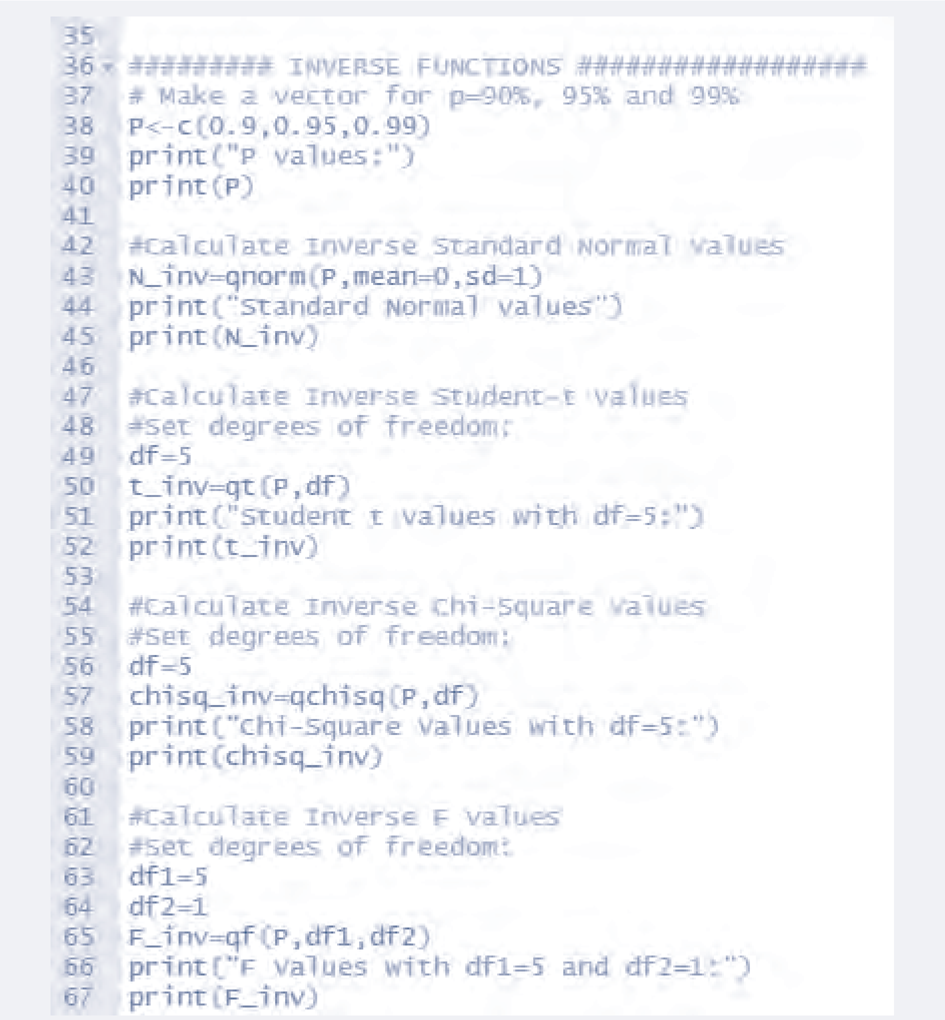
Student's *t*-distribution with df of 5 has longer tails than the standard normal distribution. For probabilities 90%, 95%, and 99%, such *t*-distributed random variables would fall farther away from their mean (1.48, 2.02, and 3.36 standard deviations, respectively) than a normally distributed random variable (1.28, 1.64, and 2.33 standard deviations, respectively).

R CODE FOR PROBABILITIES INVOLVING STUDENT'S *T*-, CHI-SQUARE, AND *F*-DISTRIBUTIONS

For those of you with a knowledge of (or interest in learning) readily accessible computer code to find probabilities involving Student's *t*-, chi-square, and *F*-distributions, you can try out the following program. Specifically, this program uses code in the R language to solve for the answers to Example 9, which you have just completed. Good luck and have fun!



```
######### CDF FUNCTIONS in R #################

#Make a vector of 1, 2, and 3
z=1:3
print("z values:")
print(z)

#Calculate Standard Normal CDF
N_CDF=pnorm(z,mean=0,sd=1)
print("Standard Normal CDF:")
print(N_CDF)

#Calculate Student-t CDF
#Set degrees of freedom:
df=5
t_CDF=pt(z,df)
print("Student t CDF with df=5:")
print(t_CDF)

#Calculate Chi-Square CDF
#Set degrees of freedom:
df=5
chisq_CDF=pchisq(z,df)
print("Chi-Square CDF with df=5:")
print(chisq_CDF)

#Calculate F CDF
#Set degrees of freedom:
df1=5
df2=1
F_CDF=pf(z,df1,df2)
print("F CDF with df1=5 and df2=1:")
print(F_CDF)

```



```
########## INVERSE FUNCTIONS ####################
# Make a vector for p=90%, 95% and 99%
P<-c(0.9,0.95,0.99)
print("P values:")
print(P)

#Calculate Inverse Standard Normal Values
N_inv=qnorm(P,mean=0,sd=1)
print("Standard Normal values")
print(N_inv)

#Calculate Inverse Student-t values
#Set degrees of freedom:
df=5
t_inv=qt(P,df)
print("Student t values with df=5:")
print(t_inv)

#Calculate Inverse Chi-Square values
#Set degrees of freedom:
df=5
chisq_inv=qchisq(P,df)
print("Chi-Square values with df=5:")
print(chisq_inv)

#Calculate Inverse F values
#Set degrees of freedom:
df1=5
df2=1
F_inv=qf(P,df1,df2)
print("F values with df1=5 and df2=1:")
print(F_inv)
```

8

MONTE CARLO SIMULATION

p describe Monte Carlo simulation

After gaining an understanding of probability distributions, we now learn about a technique in which probability distributions play an integral role. The technique is called **Monte Carlo simulation**, and in finance it involves the use of computer software to represent the operation of a complex financial system. A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Monte Carlo simulation is widely used to estimate risk and return in investment applications. In this setting, we simulate the portfolio's profit and loss performance for a specified time horizon. Repeated trials within the simulation (each trial involving a draw of random observations from a probability distribution) produce a simulated frequency distribution of portfolio returns from which performance and risk measures are derived.

Another important use of Monte Carlo simulation in investments is as a tool for valuing complex securities for which no analytic pricing formula is available. For other securities, such as mortgage-backed securities with complex embedded options, Monte Carlo simulation is also an important modeling resource. Since we control the assumptions when we carry out a simulation, we can run a model for valuing such securities through a Monte Carlo simulation to examine the model's sensitivity to a change in key assumptions.

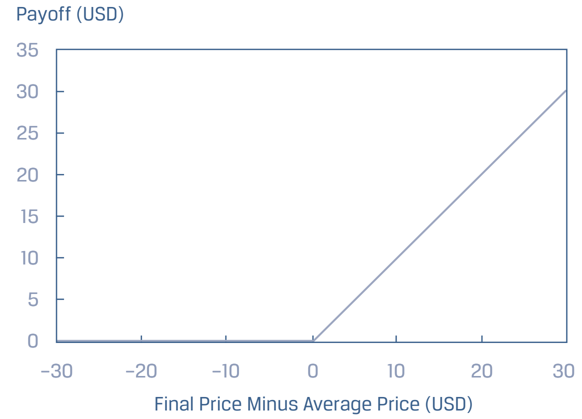
To understand the technique of Monte Carlo simulation, we present the process as a series of steps; these can be viewed as providing an overview rather than a detailed recipe for implementing a Monte Carlo simulation in its many varied applications. To illustrate the steps, we use Monte Carlo simulation to value a **contingent claim** security (a security whose value is based on some other underlying security) for which no analytic pricing formula is available. For our purposes, such a contingent claim security has a value at its maturity equal to the difference between the underlying stock price at that maturity and the average stock price during the life of the contingent claim or \$0, whichever is greater. For instance, if the final underlying stock price is \$34 and the average value over the life of the claim is \$31, the value of the contingent claim at its maturity is \$3 (the greater of \$34 – \$31 = \$3 and \$0).

Assume that the maturity of the claim is one year from today; we will simulate stock prices in monthly steps over the next 12 months and will generate 1,000 scenarios to evaluate this claim. The payoff diagram of this contingent claim security is depicted in Panel A of Exhibit 18, a histogram of simulated average and final stock prices is shown in Panel B, and a histogram of simulated payoffs of the contingent claim is presented in Panel C.

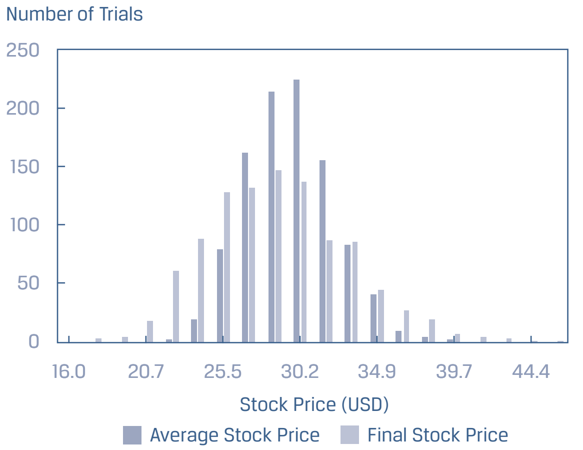
The payoff diagram (Panel A) is a snapshot of the contingent claim at maturity. If the stock's final price is less than or equal to its average over the life of the contingent claim, then the payoff would be zero. However, if the final price exceeds the average price, the payoff is equal to this difference. Panel B shows histograms of the simulated final and average stock prices. Note that the simulated final price distribution is wider than the simulated average price distribution. Also, note that the contingent claim's value depends on the difference between the final and average stock prices, which cannot be directly inferred from these histograms.

Exhibit 18 Payoff Diagram, Histogram of Simulated Average and Final Stock Prices, and Histogram of Simulated Payoffs for Contingent Claim

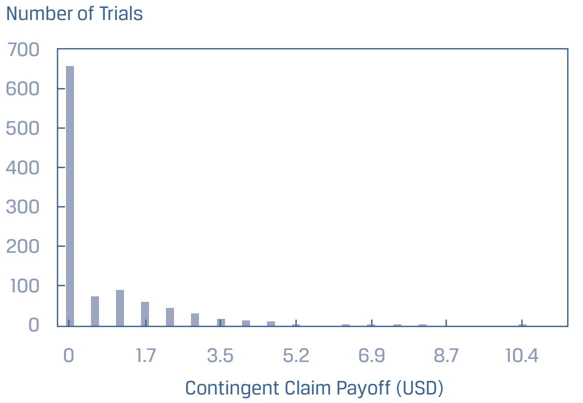
A. Contingent Claim Payoff Diagram



B. Histogram of Simulated Average and Final Stock Prices



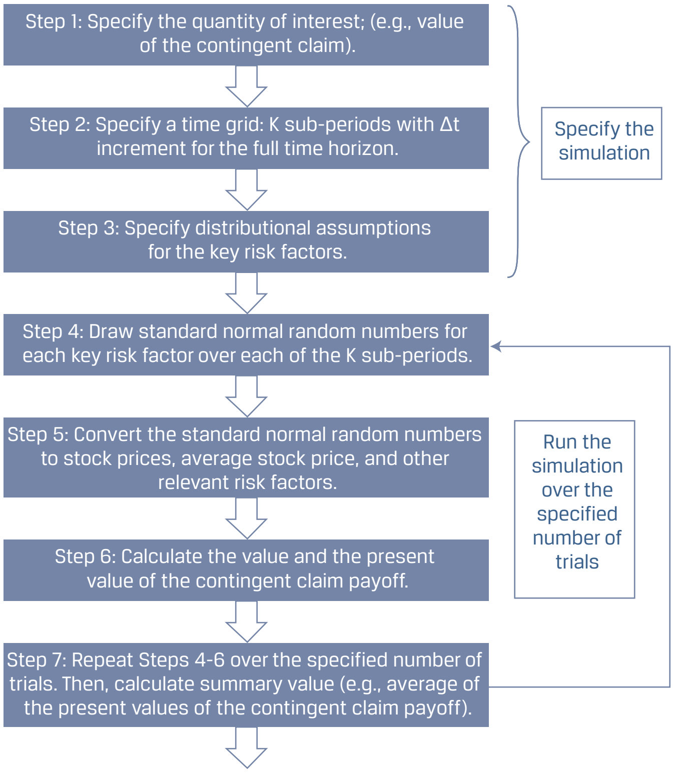
C. Histogram of Simulated Contingent Claim Payoffs



Finally, Panel C shows the histogram of the contingent claim's simulated payoffs. In 654 of 1,000 total trials, the final stock price was less than or equal to the average price, so in 65.4% of the trials the contingent claim paid off zero. In the remaining 34.6% of the trials, however, the claim paid the positive difference between the final and average prices, with the maximum payoff being $11.

The process flowchart in Exhibit 19 shows the steps for implementing the Monte Carlo simulation for valuing this contingent claim. Steps 1 through 3 of the process describe specifying the simulation; Steps 4 through 7 describe running the simulation.

Exhibit 19 Steps in Implementing the Monte Carlo Simulation


Step 1: Specify the quantity of interest; (e.g., value of the contingent claim).
Step 2: Specify a time grid: K sub-periods with Δt increment for the full time horizon.
Step 3: Specify distributional assumptions for the key risk factors.
Specify the simulation
Step 4: Draw standard normal random numbers for each key risk factor over each of the K sub-periods.
Step 5: Convert the standard normal random numbers to stock prices, average stock price, and other relevant risk factors.
Step 6: Calculate the value and the present value of the contingent claim payoff.
Step 7: Repeat Steps 4-6 over the specified number of trials. Then, calculate summary value (e.g., average of the present values of the contingent claim payoff).
Run the simulation over the specified number of trials

The mechanics of implementing the Monte Carlo simulation for valuing the contingent claim using the seven-step process are described as follows:

1 Specify the quantity of interest in terms of underlying variables. Here the quantity of interest is the contingent claim value, and the underlying variable is the stock price. Then, specify the starting value(s) of the underlying variable(s).

We use C_{iT} to represent the value of the claim at maturity, T. The subscript i in C_{iT} indicates that C_{iT} is a value resulting from the ith **simulation trial**, each simulation trial involving a drawing of random values (an iteration of Step 4).

2 Specify a time grid. Take the horizon in terms of calendar time and split it into a number of subperiods—say, K in total. Calendar time divided by the number of subperiods, K, is the time increment, Δt. In our example, calendar time is one year and K is 12, so Δt equals one month.

3 Specify distributional assumptions for the key risk factors that drive the underlying variables. For example, stock price is the underlying variable for the contingent claim, so we need a model for stock price movement. We choose the following model for changes in stock price, where Z_k stands for the standard normal random variable:

$$\Delta\text{Stock price} = (\mu \times \text{Prior stock price} \times \Delta t) + (\sigma \times \text{Prior stock price} \times Z_k).$$

The term Z_k is the key risk factor in the simulation. Through our choice of μ (mean) and σ (standard deviation), we control the distribution of the stock price variable. Although this example has one key risk factor, a given simulation may have multiple key risk factors.

4 Using a computer program or spreadsheet function, draw K random values of each risk factor. In our example, the spreadsheet function would produce a draw of K (= 12) values of the standard normal variable Z_k: $Z_1, Z_2, Z_3, \ldots, Z_K$. We will discuss generating standard normal random numbers (or, in fact, random numbers with any kind of distribution) after describing the sequence of simulation steps.

5 Convert the standard normal random numbers generated in Step 4 into stock price changes (ΔStock price) by using the model of stock price dynamics from Step 3. The result is K observations on possible changes in stock price over the K subperiods (remember, $K = 12$). An additional calculation is needed to convert those changes into a sequence of K stock prices, with the initial stock price as the starting value over the K subperiods. Another calculation produces the average stock price during the life of the contingent claim (the sum of K stock prices divided by K).

6 Compute the value of the contingent claim at maturity, C_{iT}, and then calculate its present value, C_{i0}, by discounting this terminal value using an appropriate interest rate as of today. (The subscript i in C_{i0} stands for the ith simulation trial, as it does in C_{iT}.) We have now completed one simulation trial.

7 Iteratively go back to Step 4 until the specified number of trials, I, is completed. Finally, produce summary values and statistics for the simulation. The quantity of interest in our example is the mean value of C_{i0} for the total number of simulation trials (I = 1,000). This mean value is the Monte Carlo estimate of the value of our contingent claim.

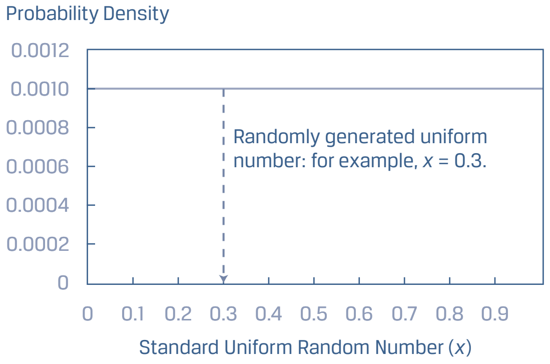
In Step 4 of our example, a computer function produced a set of random observations on a standard normal random variable. Recall that for a uniform distribution, all possible numbers are equally likely. The term **random number generator** refers to an algorithm that produces uniformly distributed random numbers between 0 and 1. In the context of computer simulations, the term **random number** refers to an observation drawn from a uniform distribution. For other distributions, the term "random observation" is used in this context.

It is a remarkable fact that random observations from any distribution can be produced using the uniform random variable with endpoints 0 and 1. The technique for producing random observations is known as the **inverse transformation method**. As a generalist, you do not need to address the technical details of converting random numbers into random observations, but you do need to know that random observations from any distribution can be generated using a uniform random variable.

Exhibit 20 provides a visual representation of the workings of the inverse transformation method. In essence, the randomly generated uniform number (0.30) lying on the continuous uniform probability density function, pdf, bounded by 0 and 1 (Panel A), is mapped onto the inverted cumulative density function, cdf, bounded by 0 and 1, of any distribution from which random observations are desired; here, we use the standard normal distribution (Panel B). The point on the given distribution's cdf is then mapped onto its pdf (Panel C), and the random observation is thereby identified (−0.5244). Note that in actuality, the random observation can be read off the y-axis in Panel B, but we include Panel C here to reinforce the intuition on how inverse transformation works. This method for generating random observations for the standard normal distribution is the same one used in the Monte Carlo simulation–based valuation of the contingent claim just described.

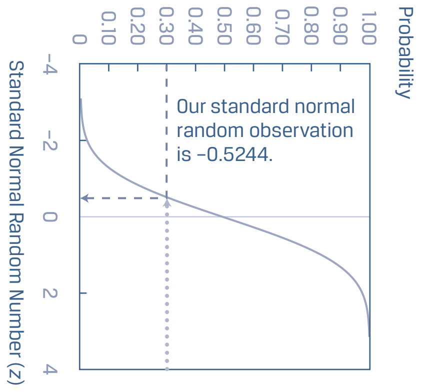
Exhibit 20 Inverse Transformation: Random Number from Uniform Distribution (PDF) Mapped to CDF and PDF of Standard Normal Distribution to Produce Random Observation

A. Uniform Random Number Is Generated



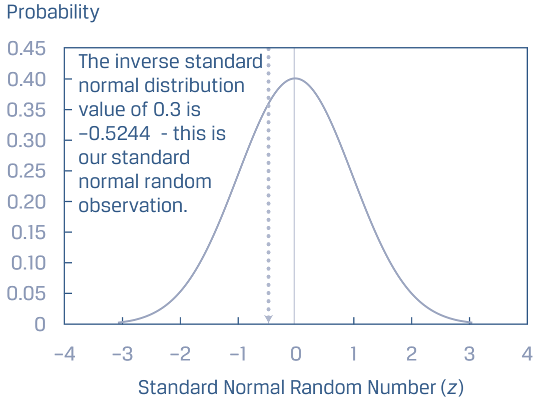
Randomly generated uniform number: for example, $x = 0.3$.

B. Random Number Mapped to Inverted Standard Normal CDF



Our standard normal random observation is −0.5244.

C. Point on Inverted CDF Mapped to Standard Normal PDF



The inverse standard normal distribution value of 0.3 is −0.5244 - this is our standard normal random observation.

In Example 10, we continue with the application of Monte Carlo simulation to value another type of contingent claim.

EXAMPLE 10

Valuing a Lookback Contingent Claim Using Monte Carlo Simulation

A standard lookback contingent claim on stock has a value at maturity equal to (Value of the stock at maturity – Minimum value of stock during the life of the claim prior to maturity) or \$0, whichever is greater. If the minimum value reached prior to maturity was \$20.11 and the value of the stock at maturity is \$23, for example, the contingent claim is worth \$23 – \$20.11 = \$2.89.

Briefly discuss how you might use Monte Carlo simulation in valuing a lookback contingent claim.

Solution:

We previously described how to use Monte Carlo simulation to value a certain type of contingent claim. Just as we can calculate the average value of the stock over a simulation trial to value that claim, for a lookback contingent claim, we can also calculate the minimum value of the stock over a simulation trial. Then, for a given simulation trial, we can calculate the terminal value of the claim, given the minimum value of the stock for the simulation trial. We can then discount this terminal value back to the present to get the value of the claim today (t = 0). The average of these t = 0 values over all simulation trials is the Monte Carlo simulated value of the lookback contingent claim.

Finally, it is important to note that Monte Carlo simulation is a complement to analytical methods. It provides only statistical estimates, not exact results. Analytical methods, where available, provide more insight into cause-and-effect relationships. However, as financial product innovations proceed, the applications for Monte Carlo simulation in investment management continue to grow.

SUMMARY

In this reading, we have presented the most frequently used probability distributions in investment analysis and Monte Carlo simulation.

- A probability distribution specifies the probabilities of the possible outcomes of a random variable.
- The two basic types of random variables are discrete random variables and continuous random variables. Discrete random variables take on at most a countable number of possible outcomes that we can list as $x_1, x_2, \ldots$. In contrast, we cannot describe the possible outcomes of a continuous random variable Z with a list $z_1, z_2, \ldots$, because the outcome $(z_1 + z_2)/2$, not in the list, would always be possible.
- The probability function specifies the probability that the random variable will take on a specific value. The probability function is denoted $p(x)$ for a discrete random variable and $f(x)$ for a continuous random variable. For any probability function $p(x)$, $0 \le p(x) \le 1$, and the sum of $p(x)$ over all values of X equals 1.
- The cumulative distribution function, denoted $F(x)$ for both continuous and discrete random variables, gives the probability that the random variable is less than or equal to x.

- The discrete uniform and the continuous uniform distributions are the distributions of equally likely outcomes.
- The binomial random variable is defined as the number of successes in n Bernoulli trials, where the probability of success, p, is constant for all trials and the trials are independent. A Bernoulli trial is an experiment with two outcomes, which can represent success or failure, an up move or a down move, or another binary (twofold) outcome.
- A binomial random variable has an expected value or mean equal to np and variance equal to $np(1 - p)$.
- A binomial tree is the graphical representation of a model of asset price dynamics in which, at each period, the asset moves up with probability p or down with probability $(1 - p)$. The binomial tree is a flexible method for modeling asset price movement and is widely used in pricing options.
- The normal distribution is a continuous symmetric probability distribution that is completely described by two parameters: its mean, μ, and its variance, σ^2.
- A univariate distribution specifies the probabilities for a single random variable. A multivariate distribution specifies the probabilities for a group of related random variables.
- To specify the normal distribution for a portfolio when its component securities are normally distributed, we need the means, the standard deviations, and all the distinct pairwise correlations of the securities. When we have those statistics, we have also specified a multivariate normal distribution for the securities.
- For a normal random variable, approximately 68% of all possible outcomes are within a one standard deviation interval about the mean, approximately 95% are within a two standard deviation interval about the mean, and approximately 99% are within a three standard deviation interval about the mean.
- A normal random variable, X, is standardized using the expression $Z = (X - \mu)/\sigma$, where μ and σ are the mean and standard deviation of X. Generally, we use the sample mean, $\bar{X}$, as an estimate of μ and the sample standard deviation, s, as an estimate of σ in this expression.
- The standard normal random variable, denoted Z, has a mean equal to 0 and variance equal to 1. All questions about any normal random variable can be answered by referring to the cumulative distribution function of a standard normal random variable, denoted $N(x)$ or $N(z)$.
- Shortfall risk is the risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.
- Roy's safety-first criterion, addressing shortfall risk, asserts that the optimal portfolio is the one that minimizes the probability that portfolio return falls below a threshold level. According to Roy's safety-first criterion, if returns are normally distributed, the safety-first optimal portfolio P is the one that maximizes the quantity $[E(R_P) - R_L]/\sigma_P$, where R_L is the minimum acceptable level of return.
- A random variable follows a lognormal distribution if the natural logarithm of the random variable is normally distributed. The lognormal distribution is defined in terms of the mean and variance of its associated normal distribution. The lognormal distribution is bounded below by 0 and skewed to the right (it has a long right tail).
- The lognormal distribution is frequently used to model the probability distribution of asset prices because it is bounded below by zero.
- Continuous compounding views time as essentially continuous or unbroken; discrete compounding views time as advancing in discrete finite intervals.

- The continuously compounded return associated with a holding period is the natural log of 1 plus the holding period return, or equivalently, the natural log of ending price over beginning price.
- If continuously compounded returns are normally distributed, asset prices are lognormally distributed. This relationship is used to move back and forth between the distributions for return and price. Because of the central limit theorem, continuously compounded returns need not be normally distributed for asset prices to be reasonably well described by a lognormal distribution.
- Student's *t*-, chi-square, and *F*-distributions are used to support statistical analyses, such as sampling, testing the statistical significance of estimated model parameters, or hypothesis testing.
- The standard *t*-distribution is a symmetrical probability distribution defined by degrees of freedom (df) and characterized by fat tails. As df increase, the *t*-distribution approaches the standard normal distribution.
- The chi-square distribution is asymmetrical, defined by degrees of freedom, and with k df is the distribution of the sum of the squares of k independent standard normally distributed random variables, so it does not take on negative values. A different distribution exists for each value of df, $n - 1$.
- The *F*-distribution is a family of asymmetrical distributions bounded from below by 0. Each *F*-distribution is defined by two values of degrees of freedom, the numerator df and the denominator df. If χ_1^2 is one chi-square random variable with m df and χ_2^2 is another chi-square random variable with n df, then $F = \left(\chi_1^2/m\right)/\left(\chi_2^2/n\right)$ follows an *F*-distribution with m numerator df and n denominator df.
- Monte Carlo simulation involves the use of a computer to represent the operation of a complex financial system. A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from specified probability distributions to represent the operation of risk in the system. Monte Carlo simulation is used in planning, in financial risk management, and in valuing complex securities. Monte Carlo simulation is a complement to analytical methods but provides only statistical estimates, not exact results.
- Random observations from any distribution can be produced using the uniform random variable with endpoints 0 and 1 via the inverse transformation method. The randomly generated uniform random number is mapped onto the inverted cdf of any distribution from which random observations are desired. The point on the given distribution's cdf is then mapped onto its pdf, and the random observation is identified.

REFERENCES

Campbell, John, Andrew Lo, and A. Craig MacKinlay. 1997. The Econometrics of Financial Markets. Princeton, NJ: Princeton University Press.

Cox, Jonathan, Stephen Ross, and Mark Rubinstein. 1979. “Options Pricing: A Simplified Approach.” Journal of Financial Economics 7: 229–63. doi:10.1016/0304-405X(79)90015-1

Fama, Eugene. 1976. Foundations of Finance. New York: Basic Books.

Kolb, Robert W., Gerald D. Gay, and William C. Hunter. 1985. “Liquidity Requirements for Financial Futures Investments.” Financial Analysts Journal 41 (May/June): 60–68.

Roy, A. D. 1952. “Safety-First and the Holding of Assets.” Econometrica 20: 431–39. doi:10.2307/1907413

PRACTICE PROBLEMS

1 A European put option on stock conveys the right to sell the stock at a prespecified price, called the exercise price, at the maturity date of the option. The value of this put at maturity is (exercise price – stock price) or \$0, whichever is greater. Suppose the exercise price is \$100 and the underlying stock trades in increments of \$0.01. At any time before maturity, the terminal value of the put is a random variable.

A Describe the distinct possible outcomes for terminal put value. (Think of the put's maximum and minimum values and its minimum price increments.)

B Is terminal put value, at a time before maturity, a discrete or continuous random variable?

C Letting Y stand for terminal put value, express in standard notation the probability that terminal put value is less than or equal to \$24. No calculations or formulas are necessary.

2 Define the term "binomial random variable." Describe the types of problems for which the binomial distribution is used.

3 The value of the cumulative distribution function $F(x)$, where x is a particular outcome, for a discrete uniform distribution:

A sums to 1.

B lies between 0 and 1.

C decreases as x increases.

4 For a binomial random variable with five trials and a probability of success on each trial of 0.50, the distribution will be:

A skewed.

B uniform.

C symmetric.

5 In a discrete uniform distribution with 20 potential outcomes of integers 1–20, the probability that X is greater than or equal to 3 but less than 6, $P(3 \leq X < 6)$, is:

A 0.10.

B 0.15.

C 0.20.

6 Over the last 10 years, a company's annual earnings increased year over year seven times and decreased year over year three times. You decide to model the number of earnings increases for the next decade as a binomial random variable.

A What is your estimate of the probability of success, defined as an increase in annual earnings?

For Parts B, C, and D of this problem, assume the estimated probability is the actual probability for the next decade.

B What is the probability that earnings will increase in exactly 5 of the next 10 years?

C Calculate the expected number of yearly earnings increases during the next 10 years.

© 2017 CFA Institute. All rights reserved.

D Calculate the variance and standard deviation of the number of yearly earnings increases during the next 10 years.

E The expression for the probability function of a binomial random variable depends on two major assumptions. In the context of this problem, what must you assume about annual earnings increases to apply the binomial distribution in Part B? What reservations might you have about the validity of these assumptions?

7 A portfolio manager annually outperforms her benchmark 60% of the time. Assuming independent annual trials, what is the probability that she will outperform her benchmark four or more times over the next five years?

A 0.26

B 0.34

C 0.48

8 You are examining the record of an investment newsletter writer who claims a 70% success rate in making investment recommendations that are profitable over a one-year time horizon. You have the one-year record of the newsletter's seven most recent recommendations. Four of those recommendations were profitable. If all the recommendations are independent and the newsletter writer's skill is as claimed, what is the probability of observing four or fewer profitable recommendations out of seven in total?

9 You are forecasting sales for a company in the fourth quarter of its fiscal year. Your low-end estimate of sales is €14 million, and your high-end estimate is €15 million. You decide to treat all outcomes for sales between these two values as equally likely, using a continuous uniform distribution.

A What is the expected value of sales for the fourth quarter?

B What is the probability that fourth-quarter sales will be less than or equal to €14,125,000?

10 State the approximate probability that a normal random variable will fall within the following intervals:

A Mean plus or minus one standard deviation.

B Mean plus or minus two standard deviations.

C Mean plus or minus three standard deviations.

11 If the probability that a portfolio outperforms its benchmark in any quarter is 0.75, the probability that the portfolio outperforms its benchmark in three or fewer quarters over the course of a year is *closest* to:

A 0.26

B 0.42

C 0.68

12 In futures markets, profits or losses on contracts are settled at the end of each trading day. This procedure is called marking to market or daily resettlement. By preventing a trader's losses from accumulating over many days, marking to market reduces the risk that traders will default on their obligations. A futures markets trader needs a liquidity pool to meet the daily mark to market. If liquidity is exhausted, the trader may be forced to unwind his position at an unfavorable time.

Suppose you are using financial futures contracts to hedge a risk in your portfolio. You have a liquidity pool (cash and cash equivalents) of λ dollars per contract and a time horizon of T trading days. For a given size liquidity pool, λ, Kolb, Gay, and Hunter developed an expression for the probability stating that you will exhaust your liquidity pool within a T-day horizon as a result of the daily marking to market. Kolb et al. assumed that the expected change in futures price is 0 and that futures price changes are normally distributed. With σ representing the standard deviation of daily futures price changes, the standard deviation of price changes over a time horizon to day T is $\sigma\sqrt{T}$, given continuous compounding. With that background, the Kolb et al. expression is

$$\text{Probability of exhausting liquidity pool} = 2[1 - N(x)],$$

where $x = \lambda/\left(\sigma\sqrt{T}\right)$. Here, x is a standardized value of λ. $N(x)$ is the standard normal cumulative distribution function. For some intuition about $1 - N(x)$ in the expression, note that the liquidity pool is exhausted if losses exceed the size of the liquidity pool at any time up to and including T; the probability of that event happening can be shown to be proportional to an area in the right tail of a standard normal distribution, $1 - N(x)$.

Using the Kolb et al. expression, answer the following questions:

A Your hedging horizon is five days, and your liquidity pool is $2,000 per contract. You estimate that the standard deviation of daily price changes for the contract is $450. What is the probability that you will exhaust your liquidity pool in the five-day period?

B Suppose your hedging horizon is 20 days but all the other facts given in Part A remain the same. What is the probability that you will exhaust your liquidity pool in the 20-day period?

13 Which of the following is characteristic of the normal distribution?

A Asymmetry

B Kurtosis of 3

C Definitive limits or boundaries

14 Which of the following assets *most likely* requires the use of a multivariate distribution for modeling returns?

A A call option on a bond

B A portfolio of technology stocks

C A stock in a market index

15 The total number of parameters that fully characterizes a multivariate normal distribution for the returns on two stocks is:

A 3.

B 4.

C 5.

16 A client has a portfolio of common stocks and fixed-income instruments with a current value of £1,350,000. She intends to liquidate £50,000 from the portfolio at the end of the year to purchase a partnership share in a business. Furthermore, the client would like to be able to withdraw the £50,000 without reducing the initial capital of £1,350,000. The following table shows four alternative asset allocations.

Mean and Standard Deviation for Four Allocations (in Percent)

	A	B	C	D
Expected annual return	16	12	10	9
Standard deviation of return	24	17	12	11

Address the following questions (assume normality for Parts B and C):

A Given the client's desire not to invade the £1,350,000 principal, what is the shortfall level, R_L? Use this shortfall level to answer Part B.

B According to the safety-first criterion, which of the allocations is the best?

C What is the probability that the return on the safety-first optimal portfolio will be less than the shortfall level, R_L?

17 A portfolio has an expected mean return of 8% and standard deviation of 14%. The probability that its return falls between 8% and 11% is *closest* to:

A 8.5%.

B 14.8%.

C 58.3%.

18 A portfolio has an expected return of 7%, with a standard deviation of 13%. For an investor with a minimum annual return target of 4%, the probability that the portfolio return will fail to meet the target is *closest* to:

A 33%.

B 41%.

C 59%.

19 **A** Define Monte Carlo simulation, and explain its use in investment management.

B Compared with analytical methods, what are the strengths and weaknesses of Monte Carlo simulation for use in valuing securities?

20 Which of the following is a continuous random variable?

A The value of a futures contract quoted in increments of $0.05

B The total number of heads recorded in 1 million tosses of a coin

C The rate of return on a diversified portfolio of stocks over a three-month period

21 X is a discrete random variable with possible outcomes $X = \{1, 2, 3, 4\}$. Three functions—$f(x)$, $g(x)$, and $h(x)$—are proposed to describe the probabilities of the outcomes in X.

	Probability Function		
$X = x$	**$f(x) = P(X = x)$**	**$g(x) = P(X = x)$**	**$h(x) = P(X = x)$**
1	–0.25	0.20	0.20
2	0.25	0.25	0.25
3	0.50	0.50	0.30
4	0.25	0.05	0.35

The conditions for a probability function are satisfied by:

A $f(x)$.

B $g(x)$.

C $h(x)$.

22 The cumulative distribution function for a discrete random variable is shown in the following table.

$X = x$	Cumulative Distribution Function $F(x) = P(X \leq x)$
1	0.15
2	0.25
3	0.50
4	0.60
5	0.95
6	1.00

The probability that X will take on a value of either 2 or 4 is *closest* to:

A 0.20.

B 0.35.

C 0.85.

23 Which of the following events can be represented as a Bernoulli trial?

A The flip of a coin

B The closing price of a stock

C The picking of a random integer between 1 and 10

24 The weekly closing prices of Mordice Corporation shares are as follows:

Date	Closing Price (€)
1 August	112
8 August	160
15 August	120

The continuously compounded return of Mordice Corporation shares for the period August 1 to August 15 is *closest* to:

A 6.90%.

B 7.14%.

C 8.95%.

25 A stock is priced at $100.00 and follows a one-period binomial process with an up move that equals 1.05 and a down move that equals 0.97. If 1 million Bernoulli trials are conducted and the average terminal stock price is $102.00, the probability of an up move (p) is *closest* to:

A 0.375.

B 0.500.

C 0.625.

26 A call option on a stock index is valued using a three-step binomial tree with an up move that equals 1.05 and a down move that equals 0.95. The current level of the index is $190, and the option exercise price is $200. If the option value is positive when the stock price exceeds the exercise price at expiration and $0 otherwise, the number of terminal nodes with a positive payoff is:

A one.

B two.

C three.

27 A random number between zero and one is generated according to a continuous uniform distribution. What is the probability that the first number generated will have a value of exactly 0.30?

A 0%

B 30%

C 70%

28 A Monte Carlo simulation can be used to:

A directly provide precise valuations of call options.

B simulate a process from historical records of returns.

C test the sensitivity of a model to changes in assumptions—for example, on distributions of key variables.

29 A limitation of Monte Carlo simulation is:

A its failure to do "what if" analysis.

B that it requires historical records of returns.

C its inability to independently specify cause-and-effect relationships.

30 Which parameter equals zero in a normal distribution?

A Kurtosis

B Skewness

C Standard deviation

31 An analyst develops the following capital market projections.

	Stocks	**Bonds**
Mean Return	10%	2%
Standard Deviation	15%	5%

Assuming the returns of the asset classes are described by normal distributions, which of the following statements is correct?

A Bonds have a higher probability of a negative return than stocks.

B On average, 99% of stock returns will fall within two standard deviations of the mean.

C The probability of a bond return less than or equal to 3% is determined using a Z-score of 0.25.

32 A client holding a £2,000,000 portfolio wants to withdraw £90,000 in one year without invading the principal. According to Roy's safety-first criterion, which of the following portfolio allocations is optimal?

	Allocation A	**Allocation B**	**Allocation C**
Expected annual return	6.5%	7.5%	8.5%
Standard deviation of returns	8.35%	10.21%	14.34%

A Allocation A

B Allocation B

C Allocation C

33 In contrast to normal distributions, lognormal distributions:

A are skewed to the left.

B have outcomes that cannot be negative.

C are more suitable for describing asset returns than asset prices.

34 The lognormal distribution is a more accurate model for the distribution of stock prices than the normal distribution because stock prices are:

A symmetrical.

B unbounded.

C non-negative.

35 The price of a stock at $t = 0$ is \$208.25 and at $t = 1$ is \$186.75. The continuously compounded rate of return for the stock from $t = 0$ to $t = 1$ is *closest* to:

A −10.90%.

B −10.32%.

C 11.51%.

36 Which one of the following statements about Student's *t*-distribution is *false*?

A It is symmetrically distributed around its mean value, like the normal distribution.

B It has shorter (i.e., thinner) tails than the normal distribution.

C As its degrees of freedom increase, Student's *t*-distribution approaches the normal distribution.

37 Which one of the following statements concerning chi-square and *F*-distributions is *false*?

A They are both asymmetric distributions.

B As their degrees of freedom increase, the shapes of their pdfs become more bell curve–like.

C The domains of their pdfs are positive and negative numbers.

SOLUTIONS

1 **A** The put's minimum value is \$0. The put's value is \$0 when the stock price is at or above \$100 at the maturity date of the option. The put's maximum value is \$100 = \$100 (the exercise price) – \$0 (the lowest possible stock price). The put's value is \$100 when the stock is worthless at the option's maturity date. The put's minimum price increments are \$0.01. The possible outcomes of terminal put value are thus \$0.00, \$0.01, \$0.02, . . . , \$100.

B The price of the underlying has minimum price fluctuations of \$0.01: These are the minimum price fluctuations for terminal put value. For example, if the stock finishes at \$98.20, the payoff on the put is \$100 – \$98.20 = \$1.80. We can specify that the nearest values to \$1.80 are \$1.79 and \$1.81. With a continuous random variable, we cannot specify the nearest values. So, we must characterize terminal put value as a discrete random variable.

C The probability that terminal put value is less than or equal to \$24 is $P(Y \leq 24)$, or $F(24)$ in standard notation, where F is the cumulative distribution function for terminal put value.

2 A binomial random variable is defined as the number of successes in n Bernoulli trials (a trial that produces one of two outcomes). The binomial distribution is used to make probability statements about a record of successes and failures or about anything with binary (twofold) outcomes.

3 B is correct. The value of the cumulative distribution function lies between 0 and 1 for any x: $0 \leq F(x) \leq 1$.

4 C is correct. The binomial distribution is symmetric when the probability of success on a trial is 0.50, but it is asymmetric or skewed otherwise. Here, it is given that $p = 0.50$.

5 B is correct. The probability of any outcome is 0.05, $P(1) = 1/20 = 0.05$. The probability that X is greater than or equal to 3 but less than 6 is expressed as $P(3 \leq X < 6) = P(3) + P(4) + P(5) = 0.05 + 0.05 + 0.05 = 0.15$.

6 **A** The probability of an earnings increase (success) in a year is estimated as 7/10 = 0.70, or 70%, based on the record of the past 10 years.

B The probability that earnings will increase in 5 of the next 10 years is about 10.3%. Define a binomial random variable X, counting the number of earnings increases over the next 10 years. From Part A, the probability of an earnings increase in a given year is $p = 0.70$ and the number of trials (years) is $n = 10$. Equation 2 gives the probability that a binomial random variable has x successes in n trials, with the probability of success on a trial equal to p:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}.$$

For this example,

$$\begin{aligned}\binom{10}{5} 0.7^5 0.3^{10-5} &= \frac{10!}{(10-5)!5!} 0.7^5 0.3^{10-5} \\ &= 252 \times 0.16807 \times 0.00243 = 0.102919.\end{aligned}$$

We conclude that the probability that earnings will increase in exactly 5 of the next 10 years is 0.1029, or approximately 10.3%.

C The expected number of yearly increases is $E(X) = np = 10 \times 0.70 = 7$.

D The variance of the number of yearly increases over the next 10 years is $\sigma^2 = np(1 - p) = 10 \times 0.70 \times 0.30 = 2.1$. The standard deviation is 1.449 (the positive square root of 2.1).

E You must assume that (1) the probability of an earnings increase (success) is constant from year to year and (2) earnings increases are independent trials. If current and past earnings help forecast next year's earnings, Assumption 2 is violated. If the company's business is subject to economic or industry cycles, neither assumption is likely to hold.

7 B is correct. To calculate the probability of four years of outperformance, use the formula

$$p(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} = \frac{n!}{(n - x)!x!} p^x (1 - p)^{n-x}.$$

Using this formula to calculate the probability in four of five years, $n = 5$, $x = 4$, and $p = 0.60$.

Therefore,

$$p(4) = \frac{5!}{(5 - 4)!4!} 0.6^4 (1 - 0.6)^{5-4} = [120/24](0.1296)(0.40) = 0.2592.$$

$$p(5) = \frac{5!}{(5 - 5)!5!} 0.6^5 (1 - 0.6)^{5-5} = [120/120](0.0778)(1) = 0.0778.$$

The probability of outperforming four or more times is $p(4) + p(5) = 0.2592 + 0.0778 = 0.3370$.

8 The observed success rate is 4/7 = 0.571, or 57.1%. The probability of four or fewer successes is $F(4) = p(4) + p(3) + p(2) + p(1) + p(0)$, where $p(4)$, $p(3)$, $p(2)$, $p(1)$, and $p(0)$ are, respectively, the probabilities of 4, 3, 2, 1, and 0 successes, according to the binomial distribution with $n = 7$ and $p = 0.70$. We have the following probabilities:

$p(4) = (7!/4!3!)(0.70^4)(0.30^3) = 35(0.006483) = 0.226895.$

$p(3) = (7!/3!4!)(0.70^3)(0.30^4) = 35(0.002778) = 0.097241.$

$p(2) = (7!/2!5!)(0.70^2)(0.30^5) = 21(0.001191) = 0.025005.$

$p(1) = (7!/1!6!)(0.70^1)(0.30^6) = 7(0.000510) = 0.003572.$

$p(0) = (7!/0!7!)(0.70^0)(0.30^7) = 1(0.000219) = 0.000219.$

Summing all these probabilities, you conclude that $F(4) = 0.226895 + 0.097241 + 0.025005 + 0.003572 + 0.000219 = 0.352931$, or 35.3%.

9 A The expected value of fourth-quarter sales is €14,500,000, calculated as (€14,000,000 + €15,000,000)/2. With a continuous uniform random variable, the mean or expected value is the midpoint between the smallest and largest values.

B The probability that fourth-quarter sales will be less than or equal to €14,125,000 is 0.125, or 12.5%, calculated as (€14,125,000 – €14,000,000)/(€15,000,000 – €14,000,000).

10 A Approximately 68% of all outcomes of a normal random variable fall within plus or minus one standard deviation of the mean.

B Approximately 95% of all outcomes of a normal random variable fall within plus or minus two standard deviations of the mean.

C Approximately 99% of all outcomes of a normal random variable fall within plus or minus three standard deviations of the mean.

11 C is correct. The probability that the performance is at or below the expectation is calculated by finding $F(3) = p(3) + p(2) + p(1) + p(0)$ using the formula:

$$p(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} = \frac{n!}{(n - x)!x!} p^x (1 - p)^{n-x}.$$

Using this formula,

$$p(3) = \frac{4!}{(4 - 3)!3!} 0.75^3 (1 - 0.75)^{4-3} = [24/6](0.42)(0.25) = 0.42.$$

$$p(2) = \frac{4!}{(4 - 2)!2!} 0.75^2 (1 - 0.75)^{4-2} = [24/4](0.56)(0.06) = 0.20.$$

$$p(1) = \frac{4!}{(4 - 1)!1!} 0.75^1 (1 - 0.75)^{4-1} = [24/6](0.75)(0.02) = 0.06.$$

$$p(0) = \frac{4!}{(4 - 0)!0!} 0.75^0 (1 - 0.75)^{4-0} = [24/24](1)(0.004) = 0.004.$$

Therefore,

$F(3) = p(3) + p(2) + p(1) + p(0) = 0.42 + 0.20 + 0.06 + 0.004 = 0.684$, or approximately 68%.

12 **A** The probability of exhausting the liquidity pool is 4.7%. First, calculate $x = \lambda/\left(\sigma\sqrt{T}\right) = \$2{,}000/\left(\$450\sqrt{5}\right) = 1.987616$. By using Excel's NORM.S.DIST() function, we get NORM.S.DIST(1.987616) = 0.9766. Thus, the probability of exhausting the liquidity pool is $2[1 - N(1.99)] = 2(1 - 0.9766) = 0.0469$, or about 4.7%.

B The probability of exhausting the liquidity pool is now 32.2%. The calculation follows the same steps as those in Part A. We calculate $x = \lambda/\left(\sigma\sqrt{T}\right) = \$2{,}000/\left(\$450\sqrt{20}\right) = 0.993808$. By using Excel's NORM.S.DIST() function, we get NORM.S.DIST(0.993808)= 0.8398. Thus, the probability of exhausting the liquidity pool is $2[1 - N(0.99)] = 2(1 - 0.8398) = 0.3203$, or about 32.0%. This is a substantial probability that you will run out of funds to meet marking to market.

In their paper, Kolb et al. called the probability of exhausting the liquidity pool the probability of ruin, a traditional name for this type of calculation.

13 B is correct. The normal distribution has a skewness of 0, a kurtosis of 3, and a mean, median, and mode that are all equal.

14 B is correct. Multivariate distributions specify the probabilities for a group of related random variables. A portfolio of technology stocks represents a group of related assets. Accordingly, statistical interrelationships must be considered, resulting in the need to use a multivariate normal distribution.

15 C is correct. A bivariate normal distribution (two stocks) will have two means, two variances, and one correlation. A multivariate normal distribution for the returns on n stocks will have n means, n variances, and $n(n - 1)/2$ distinct correlations.

16 **A** Because £50,000/£1,350,000 is 3.7%, for any return less than 3.7% the client will need to invade principal if she takes out £50,000. So $R_L = 3.7\%$.

B To decide which of the allocations is safety-first optimal, select the alternative with the highest ratio $[\mathrm{E}(R_P) - R_L]/\sigma_P$:

Allocation A: 0.5125 = (16 – 3.7)/24.

Allocation B: 0.488235 = (12 – 3.7)/17.

Allocation C: 0.525 = (10 – 3.7)/12.

Allocation D: 0.481818 = (9 – 3.7)/11.

Allocation C, with the largest ratio (0.525), is the best alternative according to the safety-first criterion.

C To answer this question, note that $P(R_C < 3.7) = N(0.037 - 0.10)/0.12$ $= N(-0.525)$. By using Excel's NORM.S.DIST() function, we get NORM.S.DIST((0.037 – 0.10)/0.12) = 29.98%, or about 30%. The safety-first optimal portfolio has a roughly 30% chance of not meeting a 3.7% return threshold.

17 A is correct. $P(8\% \leq \text{Portfolio return} \leq 11\%) = N(Z \text{ corresponding to } 11\%) - N(Z \text{ corresponding to } 8\%)$. For the first term, NORM.S.DIST((11% – 8%)/14%) = 58.48%. To get the second term immediately, note that 8% is the mean, and for the normal distribution, 50% of the probability lies on either side of the mean. Therefore, $N(Z$ corresponding to 8%) must equal 50%. So, $P(8\% \leq \text{Portfolio return} \leq 11\%) = 0.5848 - 0.50 = 0.0848$, or approximately 8.5%.

18 B is correct. By using Excel's NORM.S.DIST() function, we get NORM.S.DIST((4% – 7%)/13%) = 40.87%. The probability that the portfolio will underperform the target is about 41%.

19 **A** Monte Carlo simulation involves the use of computer software to represent the operation of a complex financial system. A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from a specified probability distribution (or distributions) to represent the role of risk in the system. Monte Carlo simulation is widely used to estimate risk and return in investment applications. In this setting, we simulate the portfolio's profit and loss performance for a specified time horizon. Repeated trials within the simulation produce a simulated frequency distribution of portfolio returns from which performance and risk measures are derived. Another important use of Monte Carlo simulation in investments is as a tool for valuing complex securities for which no analytic pricing formula is available. It is also an important modeling resource for securities with complex embedded options.

B *Strengths:* Monte Carlo simulation can be used to price complex securities for which no analytic expression is available, particularly European-style options.

Weaknesses: Monte Carlo simulation provides only statistical estimates, not exact results. Analytic methods, when available, provide more insight into cause-and-effect relationships than does Monte Carlo simulation.

20 C is correct. The rate of return is a random variable because the future outcomes are uncertain, and it is continuous because it can take on an unlimited number of outcomes.

21 B is correct. The function $g(x)$ satisfies the conditions of a probability function. All of the values of $g(x)$ are between 0 and 1, and the values of $g(x)$ all sum to 1.

22 A is correct. The probability that X will take on a value of 4 or less is $F(4) = P(X \leq 4) = p(1) + p(2) + p(3) + p(4) = 0.60$. The probability that X will take on a value of 3 or less is $F(3) = P(X \leq 3) = p(1) + p(2) + p(3) = 0.50$. So, the probability that X will take on a value of 4 is $F(4) - F(3) = p(4) = 0.10$. The probability of $X = 2$ can be found using the same logic: $F(2) - F(1) = p(2) = 0.25 - 0.15 = 0.10$. The probability of X taking on a value of 2 or 4 is $p(2) + p(4) = 0.10 + 0.10 = 0.20$.

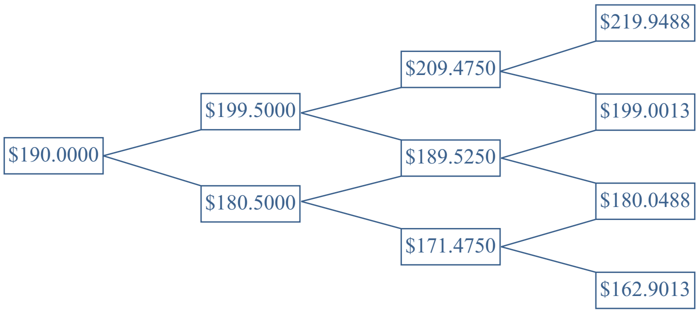
23 A is correct. A trial, such as a coin flip, will produce one of two outcomes. Such a trial is a Bernoulli trial.

24 A is correct. The continuously compounded return of an asset over a period is equal to the natural log of the asset's price change during the period. In this case,

$$\ln(120/112) = 6.90\%.$$

25 C is correct. The probability of an up move (p) can be found by solving the equation $(p)uS + (1 - p)dS = (p)105 + (1 - p)97 = 102$. Solving for p gives $8p = 5$, so $p = 0.625$.

26 A is correct. Only the top node value of \$219.9488 exceeds \$200.



27 A is correct. The probability of generating a random number equal to any fixed point under a continuous uniform distribution is zero.

28 C is correct. A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system. Therefore, it is very useful for investigating the sensitivity of a model to changes in assumptions—for example, on distributions of key variables.

29 C is correct. Monte Carlo simulation is a complement to analytical methods. Monte Carlo simulation provides statistical estimates and not exact results. Analytical methods, when available, provide more insight into cause-and-effect relationships.

30 B is correct. A normal distribution has a skewness of zero (it is symmetrical around the mean). A non-zero skewness implies asymmetry in a distribution.

31 A is correct. The chance of a negative return falls in the area to the left of 0% under a standard normal curve. By standardizing the returns and standard deviations of the two assets, the likelihood of either asset experiencing a negative return may be determined: Z-score (standardized value) = $(X - \mu)/\sigma$.

Z-score for a bond return of 0% = $(0 - 2)/5 = -0.40$.

Z-score for a stock return of 0% = $(0 - 10)/15 = -0.67$.

For bonds, a 0% return falls 0.40 standard deviations below the mean return of 2%. In contrast, for stocks, a 0% return falls 0.67 standard deviations below the mean return of 10%. A standard deviation of 0.40 is less than a standard deviation of 0.67. Negative returns thus occupy more of the left tail of the bond distribution than the stock distribution. Thus, bonds are more likely than stocks to experience a negative return.

32 B is correct. Allocation B has the highest safety-first ratio. The threshold return level, R_L, for the portfolio is £90,000/£2,000,000 = 4.5%;, thus, any return less than R_L = 4.5% will invade the portfolio principal. To compute the allocation that is safety-first optimal, select the alternative with the highest ratio:

$$\frac{\left[\mathrm{E}\left(R_P - R_L\right)\right]}{\sigma_P}.$$

$$\text{Allocation A} = \frac{6.5 - 4.5}{8.35} = 0.240.$$

$$\text{Allocation B} = \frac{7.5 - 4.5}{10.21} = 0.294.$$

$$\text{Allocation C} = \frac{8.5 - 4.5}{14.34} = 0.279.$$

33 B is correct. By definition, lognormal random variables cannot have negative values.

34 C is correct. A lognormal distributed variable has a lower bound of zero. The lognormal distribution is also right skewed, which is a useful property in describing asset prices.

35 A is correct. The continuously compounded return from $t = 0$ to $t = 1$ is $r_{0,1} = \ln(S_1/S_0) = \ln(186.75/208.25) = -0.10897 = -10.90\%$.

36 A is correct, since it is false. Student's t-distribution has longer (fatter) tails than the normal distribution and, therefore, it may provide a more reliable, more conservative downside risk estimate.

37 C is correct, since it is false. Both chi-square and F-distributions are bounded from below by zero, so the domains of their pdfs are restricted to positive numbers.

READING

5

Sampling and Estimation

by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, CFA, Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA

Richard A. DeFusco, PhD, CFA, is at the University of Nebraska-Lincoln (USA). Dennis W. McLeavey, DBA, CFA, is at the University of Rhode Island (USA). Jerald E. Pinto, PhD, CFA, is at CFA Institute (USA). David E. Runkle, PhD, CFA, is at Jacobs Levy Equity Management (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** compare and contrast probability samples with non-probability samples and discuss applications of each to an investment problem;
☐	**b.** explain sampling error;
☐	**c.** compare and contrast simple random, stratified random, cluster, convenience, and judgmental sampling;
☐	**d.** explain the central limit theorem and its importance;
☐	**e.** calculate and interpret the standard error of the sample mean;
☐	**f.** identify and describe desirable properties of an estimator;
☐	**g.** contrast a point estimate and a confidence interval estimate of a population parameter;
☐	**h.** calculate and interpret a confidence interval for a population mean, given a normal distribution with 1) a known population variance, 2) an unknown population variance, or 3) an unknown population variance and a large sample size;
☐	**i.** describe the use of resampling (bootstrap, jackknife) to estimate the sampling distribution of a statistic.
☐	**j.** describe the issues regarding selection of the appropriate sample size, data snooping bias, sample selection bias, survivorship bias, look-ahead bias, and time-period bias.

CFA Institute would like to thank Jian Wu, PhD, at State Street (USA), for this major revision of Sampling & Estimation, including new visuals, graphics, Excel functions and related text content throughout the reading.

© 2021 CFA Institute. All rights reserved.

1 INTRODUCTION

Each day, we observe the high, low, and close of stock market indexes from around the world. Indexes such as the S&P 500 Index and the Nikkei 225 Stock Average are samples of stocks. Although the S&P 500 and the Nikkei do not represent the populations of US or Japanese stocks, we view them as valid indicators of the whole population's behavior. As analysts, we are accustomed to using this sample information to assess how various markets from around the world are performing. Any statistics that we compute with sample information, however, are only estimates of the underlying population parameters. A sample, then, is a subset of the population—a subset studied to infer conclusions about the population itself.

We introduce and discuss **sampling**—the process of obtaining a sample. In investments, we continually make use of the mean as a measure of central tendency of random variables, such as return and earnings per share. Even when the probability distribution of the random variable is unknown, we can make probability statements about the population mean using the central limit theorem. We discuss and illustrate this key result. Following that discussion, we turn to statistical estimation. Estimation seeks precise answers to the question "What is this parameter's value?"

The central limit theorem and estimation, the core of the body of methods presented in the sections that follow, may be applied in investment applications. We often interpret the results for the purpose of deciding what works and what does not work in investments. We will also discuss the interpretation of statistical results based on financial data and the possible pitfalls in this process.

2 SAMPLING METHODS

a compare and contrast probability samples with non-probability samples and discuss applications of each to an investment problem

b explain sampling error

c compare and contrast simple random, stratified random, cluster, convenience, and judgmental sampling

In this section, we present the various methods for obtaining information on a population (all members of a specified group) through samples (part of the population). The information on a population that we try to obtain usually concerns the value of a **parameter**, a quantity computed from or used to describe a population of data. When we use a sample to estimate a parameter, we make use of sample statistics (statistics, for short). A statistic is a quantity computed from or used to describe a sample of data.

We take samples for one of two reasons. In some cases, we cannot possibly examine every member of the population. In other cases, examining every member of the population would not be economically efficient. Thus, savings of time and money are two primary factors that cause an analyst to use sampling to answer a question about a population.

There are two types of sampling methods: **probability sampling** and **non-probability sampling** Probability sampling gives every member of the population an equal chance of being selected. Hence it can create a sample that is representative of the population. In contrast, non-probability sampling depends on factors other than probability considerations, such as a sampler's judgment or the convenience to access data. Consequently, there is a significant risk that non-probability sampling might generate a non-representative sample. In general, all else being equal, probability sampling can yield more accuracy and reliability compared with non-probability sampling.

We first focus on probability sampling, particularly the widely used **simple random sampling** and **stratified random sampling**. We then turn our attention to non-probability sampling.

2.1 Simple Random Sampling

Suppose a telecommunications equipment analyst wants to know how much major customers will spend on average for equipment during the coming year. One strategy is to survey the population of telecom equipment customers and inquire what their purchasing plans are. In statistical terms, the characteristics of the population of customers' planned expenditures would then usually be expressed by descriptive measures such as the mean and variance. Surveying all companies, however, would be very costly in terms of time and money.

Alternatively, the analyst can collect a representative sample of companies and survey them about upcoming telecom equipment expenditures. In this case, the analyst will compute the sample mean expenditure, $\overline{X}$, a statistic. This strategy has a substantial advantage over polling the whole population because it can be accomplished more quickly and at lower cost.

Sampling, however, introduces error. The error arises because not all the companies in the population are surveyed. The analyst who decides to sample is trading time and money for sampling error.

When an analyst chooses to sample, he must formulate a sampling plan. A **sampling plan** is the set of rules used to select a sample. The basic type of sample from which we can draw statistically sound conclusions about a population is the **simple random sample** (random sample, for short).

- **Definition of Simple Random Sample.** A simple random sample is a subset of a larger population created in such a way that each element of the population has an equal probability of being selected to the subset.

The procedure of drawing a sample to satisfy the definition of a simple random sample is called **simple random sampling**. How is simple random sampling carried out? We need a method that ensures randomness—the lack of any pattern—in the selection of the sample. For a finite (limited) population, the most common method for obtaining a random sample involves the use of random numbers (numbers with assured properties of randomness). First, we number the members of the population in sequence. For example, if the population contains 500 members, we number them in sequence with three digits, starting with 001 and ending with 500. Suppose we want a simple random sample of size 50. In that case, using a computer random-number generator or a table of random numbers, we generate a series of three-digit random numbers. We then match these random numbers with the number codes of the population members until we have selected a sample of size 50. Simple random sampling is particularly useful when data in the population is homogeneous—that is, the characteristics of the data or observations (e.g., size or region) are broadly similar. We will see that if this condition is not satisfied other types of sampling may be more appropriate.

Sometimes we cannot code (or even identify) all the members of a population. We often use **systematic sampling** in such cases. With systematic sampling, we select every *k*th member until we have a sample of the desired size. The sample that results from this procedure should be approximately random. Real sampling situations may require that we take an approximately random sample.

Suppose the telecommunications equipment analyst polls a random sample of telecom equipment customers to determine the average equipment expenditure. The sample mean will provide the analyst with an estimate of the population mean expenditure. The mean obtained from the sample this way will differ from the population mean that we are trying to estimate. It is subject to error. An important part of this error is known as **sampling error**, which comes from sampling variation and occurs because we have data on only a subset of the population.

- **Definition of Sampling Error.** Sampling error is the difference between the observed value of a statistic and the quantity it is intended to estimate as a result of using subsets of the population.

A random sample reflects the properties of the population in an unbiased way, and sample statistics, such as the sample mean, computed on the basis of a random sample are valid estimates of the underlying population parameters.

A sample statistic is a random variable. In other words, not only do the original data from the population have a distribution but so does the sample statistic.

This distribution is the statistic's **sampling distribution**.

- **Definition of Sampling Distribution of a Statistic.** The sampling distribution of a statistic is the distribution of all the distinct possible values that the statistic can assume when computed from samples of the same size randomly drawn from the same population.

In the case of the sample mean, for example, we refer to the "sampling distribution of the sample mean" or the distribution of the sample mean. We will have more to say about sampling distributions later in this reading. Next, however, we look at another sampling method that is useful in investment analysis.

2.2 Stratified Random Sampling

The simple random sampling method just discussed may not be the best approach in all situations. One frequently used alternative is stratified random sampling.

- **Definition of Stratified Random Sampling.** In stratified random sampling, the population is divided into subpopulations (strata) based on one or more classification criteria. Simple random samples are then drawn from each stratum in sizes proportional to the relative size of each stratum in the population. These samples are then pooled to form a stratified random sample.

In contrast to simple random sampling, stratified random sampling guarantees that population subdivisions of interest are represented in the sample. Another advantage is that estimates of parameters produced from stratified sampling have greater precision—that is, smaller variance or dispersion—than estimates obtained from simple random sampling.

Bond indexing is one area in which stratified sampling is frequently applied. **Indexing** is an investment strategy in which an investor constructs a portfolio to mirror the performance of a specified index. In pure bond indexing, also called the full-replication approach, the investor attempts to fully replicate an index by owning all the bonds in the index in proportion to their market value weights. Many bond indexes consist of thousands of issues, however, so pure bond indexing is difficult to implement. In addition, transaction costs would be high because many bonds do not have liquid markets. Although a simple random sample could be a solution to the cost problem, the sample would probably not match the index's major risk factors—interest rate sensitivity, for example. Because the major risk factors of fixed-income portfolios are well known and quantifiable, stratified sampling offers a more effective approach. In this approach, we divide the population of index bonds into groups of similar duration (interest rate sensitivity), cash flow distribution, sector, credit quality, and call exposure. We refer to each group as a stratum or cell (a term frequently used in this context). Then, we choose a sample from each stratum proportional to the relative market weighting of the stratum in the index to be replicated.

EXAMPLE 1

Bond Indexes and Stratified Sampling

Suppose you are the manager of a portfolio of bonds indexed to the Bloomberg Barclays US Government/Credit Index, meaning that the portfolio returns should be similar to those of the index. You are exploring several approaches to indexing, including a stratified sampling approach. You first distinguish among agency bonds, US Treasury bonds, and investment-grade corporate bonds. For each of these three groups, you define 10 maturity intervals—1 to 2 years, 2 to 3 years, 3 to 4 years, 4 to 6 years, 6 to 8 years, 8 to 10 years, 10 to 12 years, 12 to 15 years, 15 to 20 years, and 20 to 30 years—and also separate the bonds with coupons (annual interest rates) of 6% or less from the bonds with coupons of more than 6%.

1 How many cells or strata does this sampling plan entail?

2 If you use this sampling plan, what is the minimum number of issues the indexed portfolio can have?

3 Suppose that in selecting among the securities that qualify for selection within each cell, you apply a criterion concerning the liquidity of the security's market. Is the sample obtained random? Explain your answer.

Solution to 1:

We have 3 issuer classifications, 10 maturity classifications, and 2 coupon classifications. So, in total, this plan entails 3(10)(2) = 60 different strata or cells.

Solution to 2:

One cannot have less than 1 issue for each cell, so the portfolio must include at least 60 issues.

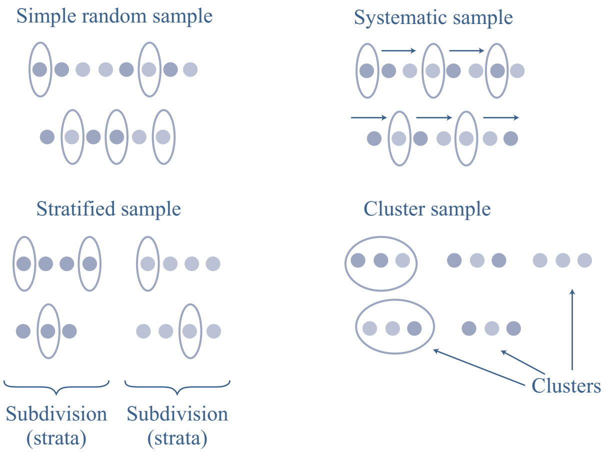
Solution to 3:

Applying any additional criteria to the selection of securities for the cells, not every security that might be included has an equal probability of being selected. As a result, the sampling is not random. In practice, indexing using stratified sampling usually does not strictly involve random sampling because the selection of bond issues within cells is subject to various additional criteria. Because the purpose of sampling in this application is not to make an inference about a population parameter but rather to index a portfolio, lack of randomness is not in itself a problem in this application of stratified sampling.

2.3 Cluster Sampling

Another sampling method, **cluster sampling**, also requires the division or classification of the population into subpopulation groups, called clusters. In this method, the population is divided into clusters, each of which is essentially a mini-representation of the entire populations. Then certain clusters are chosen as a whole using simple random sampling. If all the members in each sampled cluster are sampled, this sample plan is referred to as one-stage cluster sampling. If a subsample is randomly selected from each selected cluster, then the plan is referred as two-stage cluster sampling. Exhibit 1 (bottom right panel) shows how cluster sampling works and how it compares with the other probability sampling methods.

Exhibit 1 Probability Sampling



A major difference between cluster and stratified random samples is that in cluster sampling, a whole cluster is regarded as a sampling unit and only sampled clusters are included. In stratified random sampling, however, all the strata are included and only specific elements within each stratum are then selected as sampling units.

Cluster sampling is commonly used for market survey, and the most popular version identifies clusters based on geographic parameters. For example, a research institute is looking to survey if individual investors in the United States are bullish, bearish, or neutral on the stock market. It would be impossible to carry out the research by surveying all the individual investors in the country. The two-stage cluster sampling is a good solution in this case. At the first stage, a researcher can group the population by states and all the individual investors of each state represent a cluster. A handful of the clusters are then randomly selected. At the second stage, a simple random sample of individual investors is selected from each sampled cluster to conduct the survey.

Compared with other probability sampling methods, given equal sample size, cluster sampling usually yields lower accuracy because a sample from a cluster might be less representative of the entire population. Its major advantage, however, is offering the most time-efficient and cost-efficient probability sampling plan for analyzing a vast population.

2.4 Non-Probability Sampling

Non-probability sampling methods rely not on a fixed selection process but instead on a researcher's sample selection capabilities. We introduce two major types of non-probability sampling methods here.

- **Convenience Sampling:** In this method, an element is selected from the population based on whether or not it is accessible to a researcher or on how easy it is for a researcher to access the element. Because the samples are selected conveniently, they are not necessarily representative of the entire population, and hence the level of the sampling accuracy could be limited. But the advantage of **convenience sampling** is that data can be collected quickly at a low cost. In situations such as the preliminary stage of research or in circumstances subject to cost constraints, convenience sampling is often used as a time-efficient and cost-effective sampling plan for a small-scale pilot study before testing a large-scale and more representative sample.

- **Judgmental Sampling:** This sampling process involves selectively handpicking elements from the population based on a researcher's knowledge and professional judgment. Sample selection under **judgmental sampling** could be affected by the bias of the researcher and might lead to skewed results that do not represent the whole population. In circumstances where there is a time constraint, however, or when the specialty of researchers is critical to select a more representative sample than by using other probability or non-probability sampling methods, judgmental sampling allows researchers to go directly to the target population of interest. For example, when auditing financial statements, seasoned auditors can apply their sound judgment to select accounts or transactions that can provide sufficient audit coverage. Example 2 illustrates an application of these sampling methods.

EXAMPLE 2

Demonstrating the Power of Sampling

To demonstrate the power of sampling, we conduct two sampling experiments on a large dataset. The full dataset is the "population," representing daily returns of the fictitious Euro-Asia-Africa (EAA) Equity Index. This dataset spans a five-year period and consists of 1,258 observations of daily returns with a minimum value of −4.1% and a maximum value of 5.0%.

First, we calculate the mean daily return of the EAA Equity Index (using the population).

By taking the average of all the data points, the mean of the entire daily return series is computed as 0.035%.

First Experiment: Random Sampling

The sample size m is set to 5, 10, 20, 50, 100, 200, 500, and 1,000. At each sample size, we run random sampling multiple times (N = 100) to collect 100 samples to compute mean absolute error. The aim is to compute and plot the mean error versus the sample size.

For a given sample size m, we use the following procedure to compute mean absolute error in order to measure sampling error:

1. Randomly draw m observations from the entire daily return series to form a sample.
2. Compute the mean of this sample.
3. Compute the absolute error, the difference between the sample's mean and the population mean. Because we treat the whole five-year daily return series as a population, the population mean is 0.035% as we computed in Solution 1.
4. We repeat the previous three steps a hundred times (N =100) to collect 100 samples of the same size m and compute the absolute error of each sample.
5. Compute the mean of the 100 absolute errors as the mean absolute error for the sample size m.

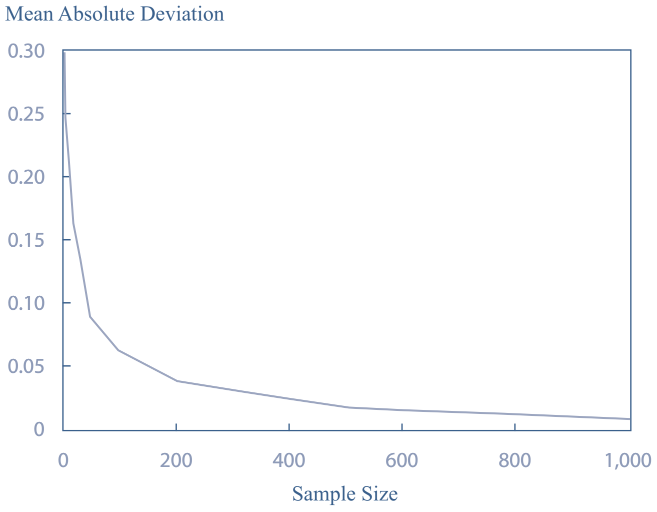
By applying this procedure, we compute mean absolute errors for eight different sample sizes: m = 5, 10, 20, 50, 100, 200, 500, and 1000. Exhibit 2 summarizes the results.

Exhibit 2 Mean Absolute Error of Random Sampling

Sample size	5	10	20	50	100	200	500	1,000
Mean absolute error	0.297%	0.218%	0.163%	0.091%	0.063%	0.039%	0.019%	0.009%

Mean absolute errors are plotted against sample size in Exhibit 3. The plot shows that the error quickly shrinks as the sample size increases. It also indicates that a minimum sample size is needed to limit sample error and achieve a certain level of accuracy. After a certain size, however (e.g., 200 to 400 in this case), there is little incremental benefit from adding more observations.

Exhibit 3 Mean Absolute Error of Random Sampling vs. Sample Size



Second Experiment: Stratified Random Sampling

We now conduct stratified random sampling by dividing daily returns into groups by year. The sample size m is again set to 5, 10, 20, 50, 100, 200, 500, and 1,000. At each sample size, run random sampling multiple times (N = 100) to collect 100 samples to compute mean absolute error.

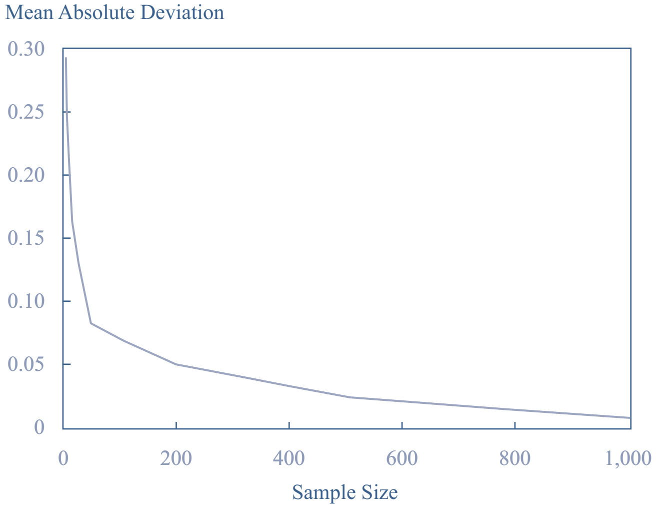
We follow the same steps as before, except for the first step. Rather than running a simple random sampling, we conduct stratified random sampling—that is, randomly selecting subsamples of equal number from daily return groups by year to generate a full sample. For example, for a sample of 50, 10 data points are randomly selected from daily returns of each year from 2014 to 2018, respectively. Exhibit 4 summarizes the results.

Exhibit 4 Mean Absolute Error of Stratified Random Sampling

Sample size	5	10	20	50	100	200	500	1,000
Mean absolute error	0.294%	0.205%	0.152%	0.083%	0.071%	0.051%	0.025%	0.008%

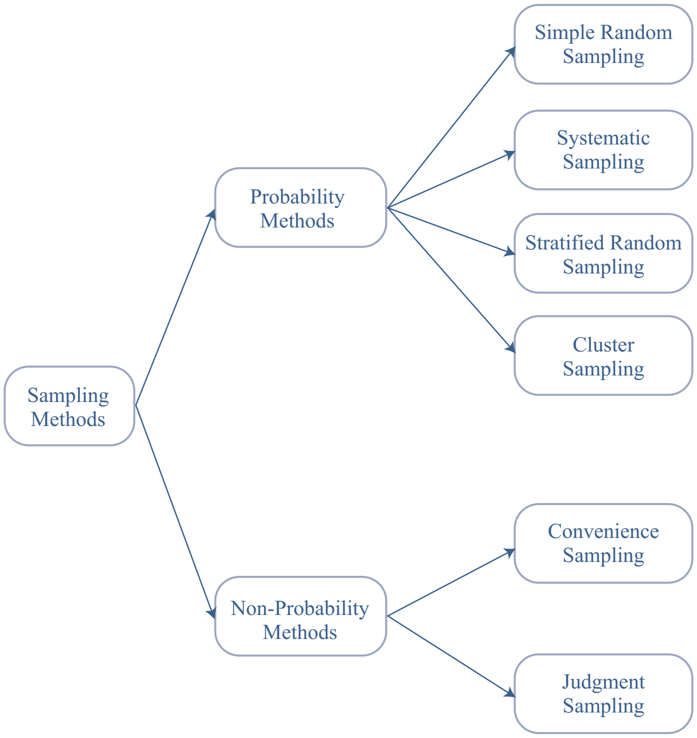
Mean absolute errors are plotted against sample size in Exhibit 5. Similar to random sampling, the plot shows rapid shrinking of errors with increasing sample size, but this incremental benefit diminishes after a certain sample size is reached.

Exhibit 5 Mean Absolute Error of Stratified Random Sampling vs. Sample Size



The sampling methods introduced here are summarized in the diagram in Exhibit 6.

Exhibit 6 Summary of Sampling Methods


Sampling Methods
Probability Methods
Non-Probability Methods
Simple Random Sampling
Systematic Sampling
Stratified Random Sampling
Cluster Sampling
Convenience Sampling
Judgment Sampling

EXAMPLE 3

An analyst is studying research and development (R&D) spending by pharmaceutical companies around the world. She considers three sampling methods for understanding a company's level of R&D. Method 1 is to simply use all the data available to her from an internal database that she and her colleagues built while researching several dozen representative stocks in the sector. Method 2 involves relying on a commercial database provided by a data vendor. She would select every fifth pharmaceutical company on the list to pull the data. Method 3 is to first divide pharmaceutical companies in the commercial database into three groups according to the region where a company is headquartered (e.g., Asia, Europe, or North America) and then randomly select a subsample of companies from each group, with the sample size proportional to the size of its associated group in order to form a complete sample.

1 Method 1 is an example of:

- **A** simple random sampling.
- **B** stratified random sampling.
- **C** convenience sampling.

2 Method 2 is an example of:

- **A** judgmental sampling.
- **B** systematic sampling.
- **C** cluster sampling.

3 Method 3 is an example of:

- **A** simple random sampling.

B stratified random sampling.

C cluster sampling.

Solution to 1:

C is correct. The analyst selects the data from the internal database because they are easy and convenient to access.

Solution to 2:

B is correct. The sample elements are selected with a fixed interval ($k = 5$) from the large population provided by data vendor.

Solution to 3:

B is correct. The population of pharmaceutical companies is divided into three strata by region to perform random sampling individually.

2.5 Sampling from Different Distributions

In practice, other than selecting appropriate sampling methods, we also need to be careful when sampling from a population that is not under one single distribution. Example 4 illustrates the problems that can arise when sampling from more than one distribution.

EXAMPLE 4

Calculating Sharpe Ratios: One or Two Years of Quarterly Data

Analysts often use the Sharpe ratio to evaluate the performance of a managed portfolio. The **Sharpe ratio** is the average return in excess of the risk-free rate divided by the standard deviation of returns. This ratio measures the return of a fund or a security over and above the risk-free rate (the excess return) earned per unit of standard deviation of return.

To compute the Sharpe ratio, suppose that an analyst collects eight quarterly excess returns (i.e., total return in excess of the risk-free rate). During the first year, the investment manager of the portfolio followed a low-risk strategy, and during the second year, the manager followed a high-risk strategy. For each of these years, the analyst also tracks the quarterly excess returns of some benchmark against which the manager will be evaluated. For each of the two years, the Sharpe ratio for the benchmark is 0.21. Exhibit 7 gives the calculation of the Sharpe ratio of the portfolio.

Exhibit 7 Calculation of Sharpe Ratios: Low-Risk and High-Risk Strategies

Quarter/Measure	Year 1 Excess Returns	Year 2 Excess Returns
Quarter 1	−3%	−12%
Quarter 2	5	20
Quarter 3	−3	−12
Quarter 4	5	20
Quarterly average	1%	4%

(continued)

Exhibit 7 (Continued)

Quarter/Measure	Year 1 Excess Returns	Year 2 Excess Returns
Quarterly standard deviation	4.62%	18.48%
Sharpe ratio = 0.22 = 1/4.62 = 4/18.48		

For the first year, during which the manager followed a low-risk strategy, the average quarterly return in excess of the risk-free rate was 1% with a standard deviation of 4.62%. The Sharpe ratio is thus 1/4.62 = 0.22. The second year's results mirror the first year except for the higher average return and volatility. The Sharpe ratio for the second year is 4/18.48 = 0.22. The Sharpe ratio for the benchmark is 0.21 during the first and second years. Because larger Sharpe ratios are better than smaller ones (providing more return per unit of risk), the manager appears to have outperformed the benchmark.

Now, suppose the analyst believes a larger sample to be superior to a small one. She thus decides to pool the two years together and calculate a Sharpe ratio based on eight quarterly observations. The average quarterly excess return for the two years is the average of each year's average excess return. For the two-year period, the average excess return is (1 + 4)/2 = 2.5% per quarter. The standard deviation for all eight quarters measured from the sample mean of 2.5% is 12.57%. The portfolio's Sharpe ratio for the two-year period is now 2.5/12.57 = 0.199; the Sharpe ratio for the benchmark remains 0.21. Thus, when returns for the two-year period are pooled, the manager appears to have provided less return per unit of risk than the benchmark and less when compared with the separate yearly results.

The problem with using eight quarters of return data is that the analyst has violated the assumption that the sampled returns come from the same population. As a result of the change in the manager's investment strategy, returns in Year 2 followed a different distribution than returns in Year 1. Clearly, during Year 1, returns were generated by an underlying population with lower mean and variance than the population of the second year. Combining the results for the first and second years yielded a sample that was representative of no population. Because the larger sample did not satisfy model assumptions, any conclusions the analyst reached based on the larger sample are incorrect. For this example, she was better off using a smaller sample than a larger sample because the smaller sample represented a more homogeneous distribution of returns.

3 DISTRIBUTION OF THE SAMPLE MEAN AND THE CENTRAL LIMIT THEOREM

d explain the central limit theorem and its importance

e calculate and interpret the standard error of the sample mean

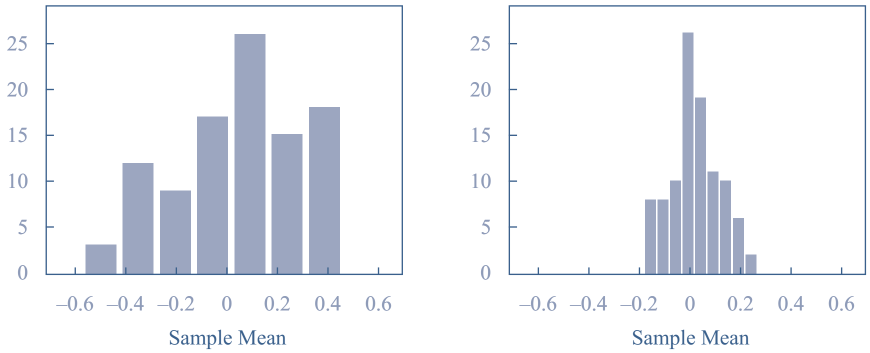
Earlier we presented a telecommunications equipment analyst who decided to sample in order to estimate mean planned capital expenditures by the customers of telecom equipment vendors. Supposing that the sample is representative of the underlying population, how can the analyst assess the sampling error in estimating the population mean? Viewed as a formula that takes a function of the random outcomes of a random variable, the sample mean is itself a random variable with a probability distribution. That probability distribution is called the statistic's sampling distribution. To estimate how closely the sample mean can be expected to match the underlying population mean, the analyst needs to understand the sampling distribution of the mean. Fortunately, we have a result, the central limit theorem, that helps us understand the sampling distribution of the mean for many of the estimation problems we face.

3.1 The Central Limit Theorem

To explain the central limit theorem, we will revisit the daily returns of the fictitious Euro-Asia-Africa Equity Index shown earlier. The dataset (the population) consists of daily returns of the index over a five-year period. The 1,258 return observations have a population mean of 0.035%.

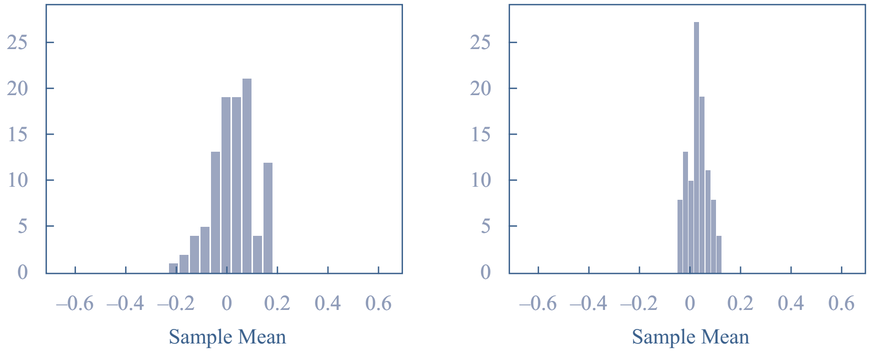
We conduct four different sets of random sampling from the population. We first draw a random sample of 10 daily returns and obtain a sample mean. We repeat this exercise 99 more times, drawing a total of 100 samples of 10 daily returns. We plot the sample mean results in a histogram, as shown in the top left panel of Exhibit 8. We then repeat the process with a larger sample size of 50 daily returns. We draw 100 samples of 50 daily returns and plot the results (the mean returns) in the histogram shown in the top right panel of Exhibit 8. We then repeat the process for sample sizes of 100 and 300 daily returns, respectively, again drawing 100 samples in each case. These results appear in the bottom left and bottom right panels of Exhibit 8. Looking at all four panels together, we observe that the larger the sample size, the more closely the histogram follows the shape of normal distribution.

Exhibit 8 Sampling Distribution with Increasing Sample Size



A. Sample Size $n = 10$

B. Sample Size $n = 50$



C. Sample Size $n = 100$

D. Sample Size $n = 300$

The results of this exercise show that as we increase the size of a random sample, the distribution of the sample means tends towards a normal distribution. This is a significant outcome and brings us to the central limit theorem concept, one of the most practically useful theorems in probability theory. It has important implications for how we construct confidence intervals and test hypotheses. Formally, it is stated as follows:

- **The Central Limit Theorem.** Given a population described by any probability distribution having mean μ and finite variance σ^2, the sampling distribution of the sample mean $\overline{X}$ computed from random samples of size n from this population will be approximately normal with mean μ (the population mean) and variance σ^2/n (the population variance divided by n) when the sample size n is large.

Consider what the expression σ^2/n signifies. Variance (σ^2) stays the same, but as n increases, the size of the fraction decreases. This dynamic suggests that it becomes progressively less common to obtain a sample mean that is far from the population mean with a larger sample size. For example, if we randomly pick returns of five stocks (trading on a market that features more than 1,000 stocks) on a particular day, their mean return is likely to be quite different from the market return. If we pick 100 stocks, the sample mean will be much closer to the market return. If we pick 1,000 stocks, the sample mean will likely be very close to the market return.

The central limit theorem allows us to make quite precise probability statements about the population mean by using the sample mean, *whatever the distribution of the population* (so long as it has finite variance), because the sample mean follows an approximate normal distribution for large-size samples. The obvious question is, "When is a sample's size large enough that we can assume the sample mean is normally distributed?" In general, when sample size n is greater than or equal to 30, we can assume that the sample mean is approximately normally distributed. When the underlying population is very non-normal, a sample size well in excess of 30 may be required for the normal distribution to be a good description of the sampling distribution of the mean.

EXAMPLE 5

A research analyst makes two statements about repeated random sampling:

1 When repeatedly drawing large samples from datasets, the sample means are approximately normally distributed.

2 The underlying population from which samples are drawn must be normally distributed in order for the sample mean to be normally distributed.

Which of the following best describes the analyst's statements?

A Statement 1 is false; Statement 2 is true.

B Both statements are true.

C Statement 1 is true; Statement 2 is false.

Solution:

C is correct. According to the central limit theorem, Statement 1 is true. Statement 2 is false because the underlying population does not need to be normally distributed in order for the sample mean to be normally distributed.

3.2 Standard Error of the Sample Mean

The central limit theorem states that the variance of the distribution of the sample mean is σ^2/n. The positive square root of variance is standard deviation. The standard deviation of a sample statistic is known as the standard error of the statistic. The standard error of the sample mean is an important quantity in applying the central limit theorem in practice.

- **Definition of the Standard Error of the Sample Mean.** For sample mean $\overline{X}$ calculated from a sample generated by a population with standard deviation σ, the standard error of the sample mean is given by one of two expressions:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \tag{1}$$

when we know σ, the population standard deviation, or by

$$s_{\overline{X}} = \frac{s}{\sqrt{n}} \tag{2}$$

when we do not know the population standard deviation and need to use the sample standard deviation, s, to estimate it.

In practice, we almost always need to use Equation 2. The estimate of s is given by the square root of the sample variance, s^2, calculated as follows:

$$s^2 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1} \tag{3}$$

It is worth noting that although the standard error is the standard deviation of the sampling distribution of the parameter, "standard deviation" in general and "standard error" are two distinct concepts, and the terms are not interchangeable. Simply put, standard deviation measures the dispersion of the data from the mean, whereas standard error measures how much inaccuracy of a population parameter estimate comes from sampling. The contrast between standard deviation and standard error reflects the distinction between data description and inference. If we want to draw conclusions about how spread out the data are, standard deviation is the term to quote. If we want to find out how precise the estimate of a population parameter from sampled data is relative to its true value, standard error is the metric to use.

We will soon see how we can use the sample mean and its standard error to make probability statements about the population mean by using the technique of confidence intervals. First, however, we provide an illustration of the central limit theorem's force.

EXAMPLE 6

The Central Limit Theorem

It is remarkable that the sample mean for large sample sizes will be distributed normally regardless of the distribution of the underlying population. To illustrate the central limit theorem in action, we specify in this example a distinctly non-normal distribution and use it to generate a large number of random samples of size 100. We then calculate the sample mean for each sample and observe the frequency distribution of the calculated sample means. Does that sampling distribution look like a normal distribution?

We return to the telecommunications analyst studying the capital expenditure plans of telecom businesses. Suppose that capital expenditures for communications equipment form a continuous uniform random variable with a **lower bound** equal to \$0 and an upper bound equal to \$100—for short, call this a uniform (0, 100) random variable. The probability function of this continuous uniform random variable has a rather simple shape that is anything but normal. It is a horizontal line with a vertical intercept equal to 1/100. Unlike a normal random variable, for which outcomes close to the mean are most likely, all possible outcomes are equally likely for a uniform random variable.

To illustrate the power of the central limit theorem, we conduct a Monte Carlo simulation to study the capital expenditure plans of telecom businesses. In this simulation, we collect 200 random samples of the capital expenditures of 100 companies (200 random draws, each consisting of the capital expenditures of 100 companies with $n = 100$). In each simulation trial, 100 values for capital expenditure are generated from the uniform (0, 100) distribution. For each random sample, we then compute the sample mean. We conduct 200 simulation trials in total. Because we have specified the continuous random distribution generating the samples, we know that the population mean capital expenditure is equal to (\$0 + \$100 million)/2 = \$50 million (i.e., $\mu = (a + b)/2$) ; the population variance of capital expenditures is equal to $(100 - 0)^2/12 = 833.33$ (i.e., $\sigma^2 = (b - a)^2/12$); thus, the standard deviation is \$28.87 million and the standard error is $28.87/\sqrt{100} = 2.887$ under the central limit theorem.

The results of this Monte Carlo experiment are tabulated in Exhibit 9 in the form of a frequency distribution. This distribution is the estimated sampling distribution of the sample mean.

Exhibit 9 Frequency Distribution: 200 Random Samples of a Uniform (0,100) Random Variable

Range of Sample Means ($ million)	Absolute Frequency
$42.5 \leq \overline{X} < 44$	1
$44 \leq \overline{X} < 45.5$	6
$45.5 \leq \overline{X} < 47$	22
$47 \leq \overline{X} < 48.5$	39
$48.5 \leq \overline{X} < 50$	41
$50 \leq \overline{X} < 51.5$	39
$51.5 \leq \overline{X} < 53$	23
$53 \leq \overline{X} < 54.5$	12
$54.5 \leq \overline{X} < 56$	12
$56 \leq \overline{X} < 57.5$	5

Note: $\overline{X}$ is the mean capital expenditure for each sample.

The frequency distribution can be described as bell-shaped and centered close to the population mean of 50. The most frequent, or modal, range, with 41 observations, is 48.5 to 50. The overall average of the sample means is $49.92, with a standard error equal to $2.80. The calculated standard error is close to the value of 2.887 given by the central limit theorem. The discrepancy between calculated and expected values of the mean and standard deviation under the central limit theorem is a result of random chance (sampling error).

In summary, although the distribution of the underlying population is very non-normal, the simulation has shown that a normal distribution well describes the estimated sampling distribution of the sample mean, with mean and standard error consistent with the values predicted by the central limit theorem.

To summarize, according to the central limit theorem, when we sample from any distribution, the distribution of the sample mean will have the following properties as long as our sample size is large:

- The distribution of the sample mean $\overline{X}$ will be approximately normal.

- The mean of the distribution of $\overline{X}$ will be equal to the mean of the population from which the samples are drawn.
- The variance of the distribution of $\overline{X}$ will be equal to the variance of the population divided by the sample size.

We next discuss the concepts and tools related to estimating the population parameters, with a special focus on the population mean. We focus on the population mean because analysts are more likely to meet interval estimates for the population mean than any other type of interval estimate.

4

POINT ESTIMATES OF THE POPULATION MEAN

f identify and describe desirable properties of an estimator

Statistical inference traditionally consists of two branches, hypothesis testing and estimation. Hypothesis testing addresses the question "Is the value of this parameter (say, a population mean) equal to some specific value (0, for example)?" In this process, we have a hypothesis concerning the value of a parameter, and we seek to determine whether the evidence from a sample supports or does not support that hypothesis. The topic of hypothesis testing will be discussed later.

The second branch of statistical inference, and what we focus on now, is estimation. Estimation seeks an answer to the question "What is this parameter's (for example, the population mean's) value?" In estimating, unlike in hypothesis testing, we do not start with a hypothesis about a parameter's value and seek to test it. Rather, we try to make the best use of the information in a sample to form one of several types of estimates of the parameter's value. With estimation, we are interested in arriving at a rule for best calculating a single number to estimate the unknown population parameter (a point estimate). In addition to calculating a point estimate, we may also be interested in calculating a range of values that brackets the unknown population parameter with some specified level of probability (a confidence interval). We first discuss point estimates of parameters and then turn our attention to the formulation of confidence intervals for the population mean.

4.1 Point Estimators

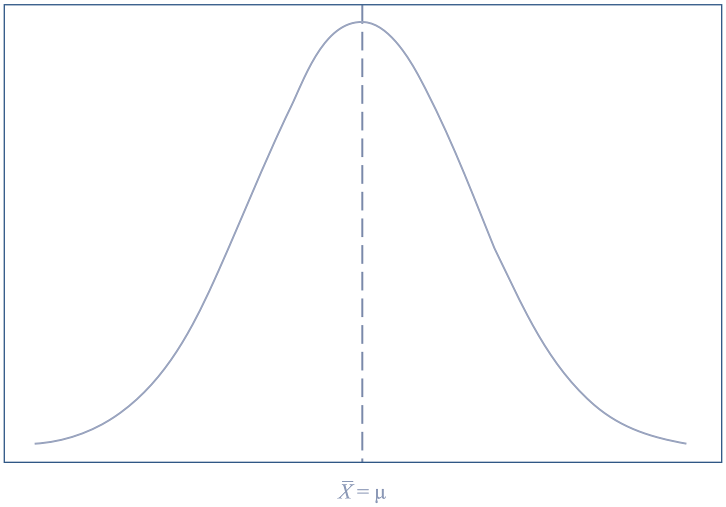
An important concept introduced here is that sample statistics viewed as formulas involving random outcomes are random variables. The formulas that we use to compute the sample mean and all the other sample statistics are examples of estimation formulas or **estimators**. The particular value that we calculate from sample observations using an estimator is called an **estimate**. An estimator has a sampling distribution; an estimate is a fixed number pertaining to a given sample and thus has no sampling distribution. To take the example of the mean, the calculated value of the sample mean in a given sample, used as an estimate of the population mean, is called a **point estimate** of the population mean. As we have seen earlier, the formula for the sample mean can and will yield different results in repeated samples as different samples are drawn from the population.

In many applications, we have a choice among a number of possible estimators for estimating a given parameter. How do we make our choice? We often select estimators because they have one or more desirable statistical properties. Following is a brief description of three desirable properties of estimators: unbiasedness (lack of bias), efficiency, and consistency.

- **Unbiasedness.** An unbiased estimator is one whose expected value (the mean of its sampling distribution) equals the parameter it is intended to estimate.

For example, as shown in Exhibit 10 of the sampling distribution of the sample mean, the expected value of the sample mean, $\overline{X}$, equals μ, the population mean, so we say that the sample mean is an unbiased estimator (of the population mean). The sample variance, s^2, calculated using a divisor of $n - 1$ (Equation 3), is an unbiased estimator of the population variance, σ^2. If we were to calculate the sample variance using a divisor of *n,* the estimator would be biased: Its expected value would be smaller than the population variance. We would say that sample variance calculated with a divisor of *n* is a biased estimator of the population variance.

Exhibit 10 Unbiasedness of an Estimator

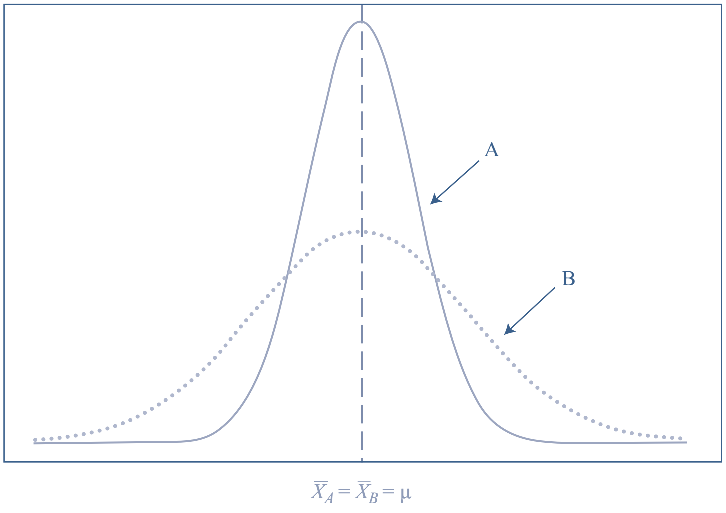


Whenever one unbiased estimator of a parameter can be found, we can usually find a large number of other unbiased estimators. How do we choose among alternative unbiased estimators? The criterion of efficiency provides a way to select from among unbiased estimators of a parameter.

- **Efficiency.** An unbiased estimator is efficient if no other unbiased estimator of the same parameter has a sampling distribution with smaller variance.

To explain the definition, in repeated samples we expect the estimates from an efficient estimator to be more tightly grouped around the mean than estimates from other unbiased estimators. For example, Exhibit 11 shows the sampling distributions of two different estimators of the population mean. Both estimators A and B are unbiased because their expected values are equal to the population mean ($\overline{X}_A = \overline{X}_B = \mu$), but estimator A is more efficient because it shows smaller variance. Efficiency is an important property of an estimator. Sample mean $\overline{X}$ is an efficient estimator of the population mean; sample variance s^2 is an efficient estimator of σ^2.

Exhibit 11 Efficiency of an Estimator

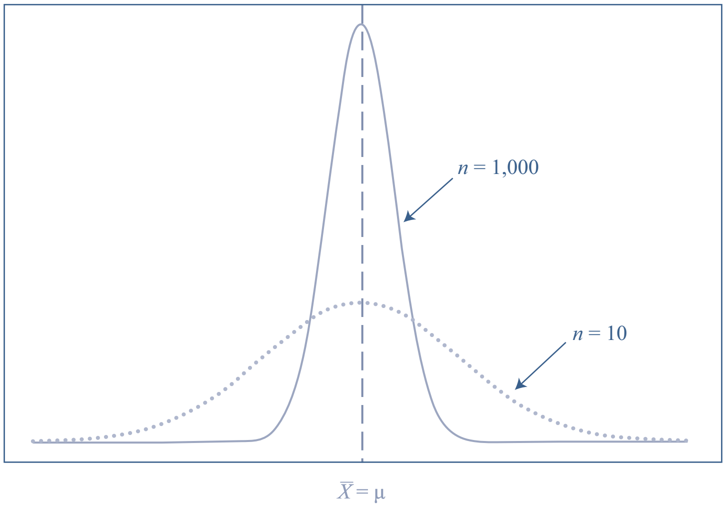


Recall that a statistic's sampling distribution is defined for a given sample size. Different sample sizes define different sampling distributions. For example, the variance of sampling distribution of the sample mean is smaller for larger sample sizes. Unbiasedness and efficiency are properties of an estimator's sampling distribution that hold for any size sample. An unbiased estimator is unbiased equally in a sample of size 100 and in a sample of size 1,000. In some problems, however, we cannot find estimators that have such desirable properties as unbiasedness in small samples. In this case, statisticians may justify the choice of an estimator based on the properties of the estimator's sampling distribution in extremely large samples, the estimator's so-called asymptotic properties. Among such properties, the most important is consistency.

- **Consistency.** A consistent estimator is one for which the probability of estimates close to the value of the population parameter increases as sample size increases.

Somewhat more technically, we can define a consistent estimator as an estimator whose sampling distribution becomes concentrated on the value of the parameter it is intended to estimate as the sample size approaches infinity. The sample mean, in addition to being an efficient estimator, is also a consistent estimator of the population mean: As sample size n goes to infinity, its standard error, $\sigma/\sqrt{n}$, goes to 0 and its sampling distribution becomes concentrated right over the value of population mean, μ. Exhibit 12 illustrates the consistency of the sample mean, in which the standard error of the estimator narrows as the sample size increases. To summarize, we can think of a consistent estimator as one that tends to produce more and more accurate estimates of the population parameter as we increase the sample's size. If an estimator is consistent, we may attempt to increase the accuracy of estimates of a population parameter by calculating estimates using a larger sample. For an inconsistent estimator, however, increasing sample size does not help to increase the probability of accurate estimates.

Exhibit 12 Consistency of an Estimator



It is worth noting that in a Big Data world, consistency is much more crucial than efficiency, because the accuracy of a population parameter's estimates can be increasingly improved with the availability of more sample data. In addition, given a big dataset, a biased but consistent estimator can offer considerably reduced error. For example, s^2/n is a biased estimator of variance. As n goes to infinity, the distinction between s^2/n and the unbiased estimator $s^2/(n - 1)$ diminish to zero.

EXAMPLE 7

Exhibit 13 plots several sampling distributions of an estimator for the population mean, and the vertical dash line represents the true value of population mean.

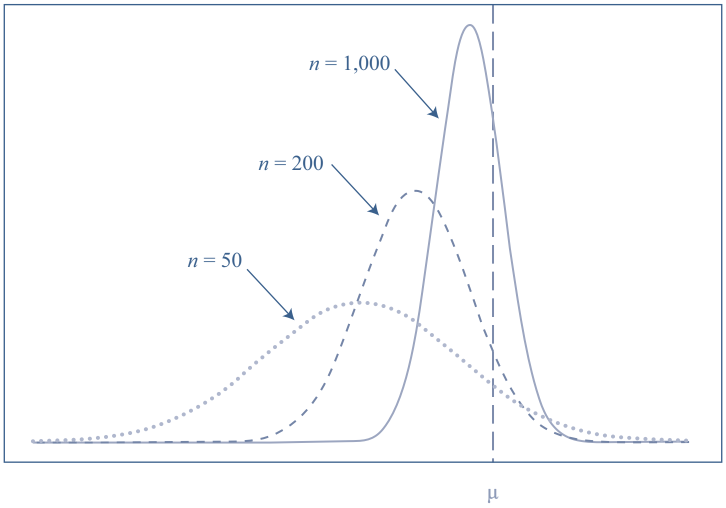
Which of the following statements *best* describes the estimator's properties?

A The estimator is unbiased.

B The estimator is biased and inconsistent.

C The estimator is biased but consistent.

Exhibit 13 Sampling Distributions of an Estimator



Solution:

C is correct. The chart shows three sampling distributions of the estimator at different sample sizes (n = 50, 200, and 1,000). We can observe that the means of each sampling distribution—that is, the expected value of the estimator—deviates from the population mean, so the estimator is biased. As the sample size increases, however, the mean of the sampling distribution draws closer to the population mean with smaller variance. So, it is a consistent estimator.

5

CONFIDENCE INTERVALS FOR THE POPULATION MEAN AND SELECTION OF SAMPLE SIZE

g contrast a point estimate and a confidence interval estimate of a population parameter

h calculate and interpret a confidence interval for a population mean, given a normal distribution with 1) a known population variance, 2) an unknown population variance, or 3) an unknown population variance and a large sample size

When we need a single number as an estimate of a population parameter, we make use of a point estimate. However, because of sampling error, the point estimate is not likely to equal the population parameter in any given sample. Often, a more useful approach than finding a point estimate is to find a range of values that we expect to bracket the parameter with a specified level of probability—an interval estimate of the parameter. A confidence interval fulfills this role.

- **Definition of Confidence Interval.** A confidence interval is a range for which one can assert with a given probability 1 − α, called the **degree of confidence**, that it will contain the parameter it is intended to estimate. This interval is often referred to as the 100(1 − α)% confidence interval for the parameter.

The endpoints of a confidence interval are referred to as the lower and upper confidence limits. In this reading, we are concerned only with two-sided confidence intervals—confidence intervals for which we calculate both lower and upper limits.

Confidence intervals are frequently given either a probabilistic interpretation or a practical interpretation. In the probabilistic interpretation, we interpret a 95% confidence interval for the population mean as follows. In repeated sampling, 95% of such confidence intervals will, in the long run, include or bracket the population mean. For example, suppose we sample from the population 1,000 times, and based on each sample, we construct a 95% confidence interval using the calculated sample mean. Because of random chance, these confidence intervals will vary from each other, but we expect 95%, or 950, of these intervals to include the unknown value of the population mean. In practice, we generally do not carry out such repeated sampling. Therefore, in the practical interpretation, we assert that we are 95% confident that a single 95% confidence interval contains the population mean. We are justified in making this statement because we know that 95% of all possible confidence intervals constructed in the same manner will contain the population mean. The confidence intervals that we discuss in this reading have structures similar to the following basic structure:

- **Construction of Confidence Intervals.** A $100(1 - \alpha)\%$ confidence interval for a parameter has the following structure:

  Point estimate ± Reliability factor × Standard error

  where

  Point estimate = a point estimate of the parameter (a value of a sample statistic)
  Reliability factor = a number based on the assumed distribution of the point estimate and the degree of confidence $(1 - \alpha)$ for the confidence interval
  Standard error = the standard error of the sample statistic providing the point estimate

  The quantity "Reliability factor × Standard error" is sometimes called the precision of the estimator; larger values of the product imply lower precision in estimating the population parameter.

The most basic confidence interval for the population mean arises when we are sampling from a normal distribution with known variance. The reliability factor in this case is based on the standard normal distribution, which has a mean of 0 and a variance of 1. A standard normal random variable is conventionally denoted by Z. The notation z_α denotes the point of the standard normal distribution such that α of the probability remains in the right tail. For example, 0.05 or 5% of the possible values of a standard normal random variable are larger than $z_{0.05} = 1.65$. Similarly, 0.025 or 2.5% of the possible values of a standard normal random variable are larger than $z_{0.025} = 1.96$.

Suppose we want to construct a 95% confidence interval for the population mean and, for this purpose, we have taken a sample of size 100 from a normally distributed population with known variance of $\sigma^2 = 400$ (so, $\sigma = 20$). We calculate a sample mean of $\overline{X} = 25$. Our point estimate of the population mean is, therefore, 25. If we move 1.96 standard deviations above the mean of a normal distribution, 0.025 or 2.5% of the probability remains in the right tail; by symmetry of the normal distribution, if we move 1.96 standard deviations below the mean, 0.025 or 2.5% of the probability remains in the left tail. In total, 0.05 or 5% of the probability is in the two tails and 0.95 or 95% lies in between. So, $z_{0.025} = 1.96$ is the reliability factor for this 95% confidence interval. Note the relationship $100(1 - \alpha)\%$ for the confidence interval and the $z_{\alpha/2}$ for the reliability factor. The standard error of the sample mean, given by Equation 1, is $\sigma_{\overline{X}} = 20/\sqrt{100} = 2$. The confidence interval, therefore, has a lower limit of $\overline{X} - 1.96\sigma_{\overline{X}}$

= 25 − 1.96(2) = 25 − 3.92 = 21.08. The upper limit of the confidence interval is $\overline{X} + 1.96\sigma_{\overline{X}}$ = 25 + 1.96(2) = 25 + 3.92 = 28.92. The 95% confidence interval for the population mean spans 21.08 to 28.92.

- **Confidence Intervals for the Population Mean (Normally Distributed Population with Known Variance).** A 100(1 − α)% confidence interval for population mean μ when we are sampling from a normal distribution with known variance σ^2 is given by

$$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{4}$$

The reliability factors for the most frequently used confidence intervals are as follows.

- **Reliability Factors for Confidence Intervals Based on the Standard Normal Distribution.** We use the following reliability factors when we construct confidence intervals based on the standard normal distribution:
  - 90% confidence intervals: Use $z_{0.05} = 1.65$
  - 95% confidence intervals: Use $z_{0.025} = 1.96$
  - 99% confidence intervals: Use $z_{0.005} = 2.58$

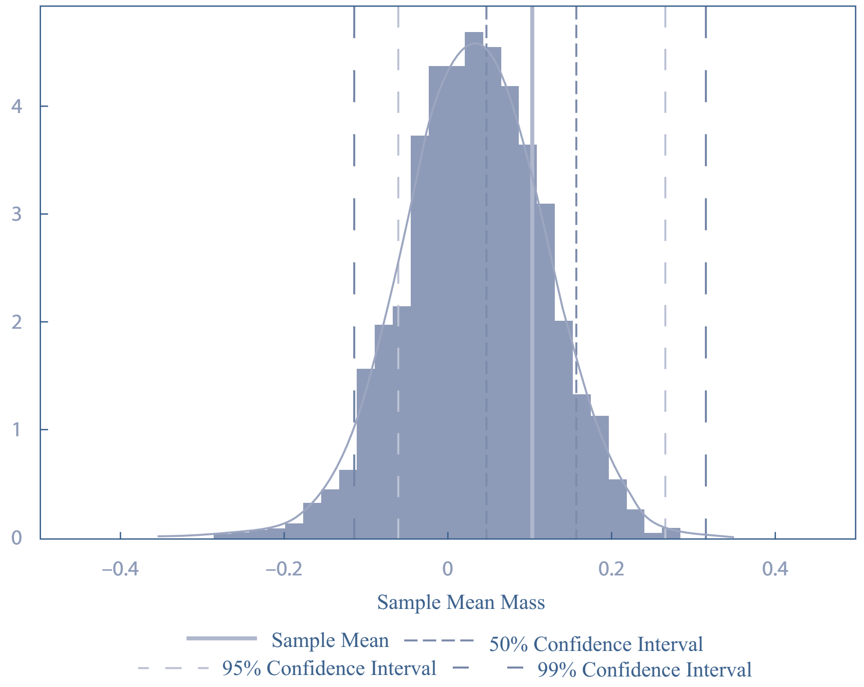
These reliability factors highlight an important fact about all confidence intervals. As we increase the degree of confidence, the confidence interval becomes wider and gives us less precise information about the quantity we want to estimate.

Exhibit 14 demonstrates how a confidence interval works. We again use the daily returns of the fictitious Euro-Asia-Africa Equity Index shown earlier. The dataset consists of 1,258 observations with a population mean of 0.035% and a population standard deviation of 0.834%. We conduct random sampling from the population 1,000 times, drawing a sample of a hundred daily returns (n = 100) each time.

We construct a histogram of the sample means, shown in Exhibit 14. The shape appears to be that of a normal distribution, in line with the central limit theorem. We next pick one random sample to construct confidence intervals around its sample mean. The mean of the selected sample is computed to be 0.103% (as plotted with a solid line). Next we construct 99%, 95%, and 50% confidence intervals around that sample mean. We use Equation 4, to compute the upper and lower bounds of each pair of confidence intervals and plot these bounds in dashed lines.

The resulting chart shows that confidence intervals narrow with decreasing confidence level, and vice versa. For example, the narrowest confidence interval in the chart corresponds to the lowest confidence level of 50%—that is, we are only 50% confident that the population mean falls within the 50% confidence interval around the sample mean. Importantly, as shown by Equation 4, given a fixed confidence level, the confidence interval narrows with smaller population deviation and greater sample size, indicating higher estimate accuracy.

Exhibit 14 Illustration of Confidence Intervals



In practice, the assumption that the sampling distribution of the sample mean is at least approximately normal is frequently reasonable, either because the underlying distribution is approximately normal or because we have a large sample and the central limit theorem applies. Rarely do we know the population variance in practice, however. When the population variance is unknown but the sample mean is at least approximately normally distributed, we have two acceptable ways to calculate the confidence interval for the population mean. We will soon discuss the more conservative approach, which is based on Student's *t*-distribution (the *t*-distribution, for short and covered earlier). In investment literature, it is the most frequently used approach in both estimation and hypothesis tests concerning the mean when the population variance is not known, whether sample size is small or large.

A second approach to confidence intervals for the population mean, based on the standard normal distribution, is the *z*-alternative. It can be used only when sample size is large. (In general, a sample size of 30 or larger may be considered large.) In contrast to the confidence interval given in Equation 4, this confidence interval uses the sample standard deviation, *s*, in computing the standard error of the sample mean (Equation 2).

- **Confidence Intervals for the Population Mean—The *z*-Alternative (Large Sample, Population Variance Unknown).** A 100(1 − α)% confidence interval for population mean μ when sampling from any distribution with unknown variance and when sample size is large is given by

$$\overline{X} \pm z_{\alpha/2}\frac{s}{\sqrt{n}} \tag{5}$$

Because this type of confidence interval appears quite often, we illustrate its calculation in Example 8.

EXAMPLE 8

Confidence Interval for the Population Mean of Sharpe Ratios—*z*-Statistic

Suppose an investment analyst takes a random sample of US equity mutual funds and calculates the average Sharpe ratio. The sample size is 100, and the average Sharpe ratio is 0.45. The sample has a standard deviation of 0.30. Calculate and interpret the 90% confidence interval for the population mean of all US equity mutual funds by using a reliability factor based on the standard normal distribution.

The reliability factor for a 90% confidence interval, as given earlier, is $z_{0.05} = 1.65$. The confidence interval will be

$$\overline{X} \pm z_{0.05}\frac{s}{\sqrt{n}} = 0.45 \pm 1.65\frac{0.30}{\sqrt{100}} = 0.45 \pm 1.65(0.03) = 0.45 \pm 0.0495$$

The confidence interval spans 0.4005 to 0.4995, or 0.40 to 0.50, carrying two decimal places. The analyst can say with 90% confidence that the interval includes the population mean.

In this example, the analyst makes no specific assumption about the probability distribution describing the population. Rather, the analyst relies on the central limit theorem to produce an approximate normal distribution for the sample mean.

As Example 8 shows, even if we are unsure of the underlying population distribution, we can still construct confidence intervals for the population mean as long as the sample size is large because we can apply the central limit theorem.

We now turn to the conservative alternative, using the *t*-distribution, for constructing confidence intervals for the population mean when the population variance is not known. For confidence intervals based on samples from normally distributed populations with unknown variance, the theoretically correct reliability factor is based on the *t*-distribution. Using a reliability factor based on the *t*-distribution is essential for a small sample size. Using a *t* reliability factor is appropriate when the population variance is unknown, even when we have a large sample and could use the central limit theorem to justify using a *z* reliability factor. In this large sample case, the *t*-distribution provides more-conservative (wider) confidence intervals.

Suppose we sample from a normal distribution. The ratio $z = \left(\overline{X} - \mu\right)\big/\left(\sigma/\sqrt{n}\right)$ is distributed normally with a mean of 0 and standard deviation of 1; however, the ratio $t = \left(\overline{X} - \mu\right)\big/\left(s/\sqrt{n}\right)$ follows the *t*-distribution with a mean of 0 and $n - 1$ degrees of freedom. The ratio represented by *t* is not normal because *t* is the ratio of two random variables, the sample mean and the sample standard deviation. The definition of the standard normal random variable involves only one random variable, the sample mean.

Values for the *t*-distribution are available from Excel, using the function T.INV(p,DF). For each degree of freedom, five values are given: $t_{0.10}$, $t_{0.05}$, $t_{0.025}$, $t_{0.01}$, and $t_{0.005}$. The values for $t_{0.10}$, $t_{0.05}$, $t_{0.025}$, $t_{0.01}$, and $t_{0.005}$ are such that, respectively, 0.10, 0.05, 0.025, 0.01, and 0.005 of the probability remains in the right tail, for the specified number of degrees of freedom. For example, for df = 30, $t_{0.10} = 1.310$, $t_{0.05} = 1.697$, $t_{0.025} = 2.042$, $t_{0.01} = 2.457$, and $t_{0.005} = 2.750$.

We now give the form of confidence intervals for the population mean using the t-distribution.

- **Confidence Intervals for the Population Mean (Population Variance Unknown)—*t*-Distribution.** If we are sampling from a population with unknown variance and either of the conditions below holds:
  - the sample is large, or
  - the sample is small, but the population is normally distributed, or approximately normally distributed,

  then a 100(1 − α)% confidence interval for the population mean μ is given by

  $$\overline{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} \tag{6}$$

  where the number of degrees of freedom for $t_{\alpha/2}$ is $n - 1$ and n is the sample size.

Example 9 reprises the data of Example 8 but uses the t-statistic rather than the z-statistic to calculate a confidence interval for the population mean of Sharpe ratios.

EXAMPLE 9

Confidence Interval for the Population Mean of Sharpe Ratios—*t*-Statistic

As in Example 8, an investment analyst seeks to calculate a 90% confidence interval for the population mean Sharpe ratio of US equity mutual funds based on a random sample of 100 US equity mutual funds. The sample mean Sharpe ratio is 0.45, and the sample standard deviation of the Sharpe ratios is 0.30. Now recognizing that the population variance of the distribution of Sharpe ratios is unknown, the analyst decides to calculate the confidence interval using the theoretically correct t-statistic.

Because the sample size is 100, df = 99. Using the Excel function T.INV(0.05,99), $t_{0.05} = 1.66$. This reliability factor is slightly larger than the reliability factor $z_{0.05} = 1.65$ that was used in Example 8. The confidence interval will be

$$\overline{X} \pm t_{0.05}\frac{s}{\sqrt{n}} = 0.45 \pm 1.66\frac{0.30}{\sqrt{100}} = 0.45 \pm 1.66(0.03) = 0.45 \pm 0.0498$$

The confidence interval spans 0.4002 to 0.4998, or 0.40 to 0.50, carrying two decimal places. To two decimal places, the confidence interval is unchanged from the one computed in Example 8.

Exhibit 15 summarizes the various reliability factors that we have used.

Exhibit 15 Basis of Computing Reliability Factors

Sampling from	Statistic for Small Sample Size	Statistic for Large Sample Size
Normal distribution with known variance	z	z
Normal distribution with unknown variance	t	t^*

(continued)

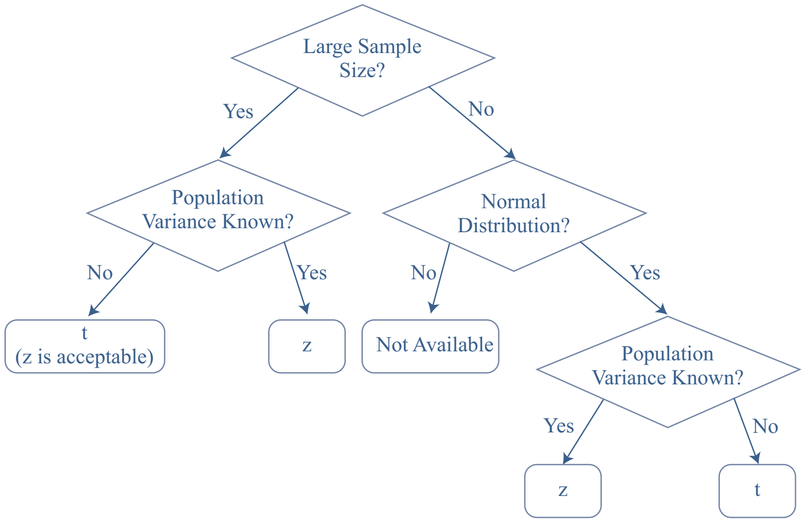
Exhibit 15 (Continued)

Sampling from	Statistic for Small Sample Size	Statistic for Large Sample Size
Non-normal distribution with known variance	not available	*z*
Non-normal distribution with unknown variance	not available	*t**

* Use of *z* also acceptable.

Exhibit 16 shows a flowchart that helps determine what statistics should be used to produce confidence intervals under different conditions.

Exhibit 16 Determining Statistics for Confidence Intervals


Large Sample Size?
Yes
No
Population Variance Known?
Normal Distribution?
No
Yes
No
Yes
t (z is acceptable)
z
Not Available
Population Variance Known?
Yes
No
z
t

5.1 Selection of Sample Size

What choices affect the width of a confidence interval? To this point we have discussed two factors that affect width: the choice of statistic (*t* or *z*) and the choice of degree of confidence (affecting which specific value of *t* or *z* we use). These two choices determine the reliability factor. (Recall that a confidence interval has the structure Point estimate ± Reliability factor × Standard error.)

The choice of sample size also affects the width of a confidence interval. All else equal, a larger sample size decreases the width of a confidence interval. Recall the expression for the standard error of the sample mean:

$$\text{Standard error of the sample mean} = \frac{\text{Sample standard deviation}}{\sqrt{\text{Sample size}}}$$

We see that the standard error varies inversely with the square root of sample size. As we increase sample size, the standard error decreases and consequently the width of the confidence interval also decreases. The larger the sample size, the greater precision with which we can estimate the population parameter.

At a given degree of confidence (1 − α), we can determine the sample size needed to obtain a desired width for a confidence interval. Define E = Reliability factor × Standard error; then $2E$ is the confidence interval's width. The smaller E is, the smaller the width of the confidence interval. Accordingly, the sample size to obtain a desired value of E at a given degree of confidence (1 − α) can be derived as $n = [(t \times s)/E]^2$. It is worth noting that appropriate sample size is also needed for performing a valid power analysis and determining the minimum detectable effect in hypothesis testing, concepts that will be covered at a later stage.

All else equal, larger samples are good, in that sense. In practice, however, two considerations may operate against increasing sample size. First, as we saw earlier concerning the Sharpe ratio, increasing the size of a sample may result in sampling from more than one population. Second, increasing sample size may involve additional expenses that outweigh the value of additional precision. Thus three issues that the analyst should weigh in selecting sample size are the need for precision, the risk of sampling from more than one population, and the expenses of different sample sizes.

EXAMPLE 10

A Money Manager Estimates Net Client Inflows

A money manager wants to obtain a 95% confidence interval for fund inflows and outflows over the next six months for his existing clients. He begins by calling a random sample of 10 clients and inquiring about their planned additions to and withdrawals from the fund. The manager then computes the change in cash flow for each client sampled as a percentage change in total funds placed with the manager. A positive percentage change indicates a net cash inflow to the client's account, and a negative percentage change indicates a net cash outflow from the client's account. The manager weights each response by the relative size of the account within the sample and then computes a weighted average.

As a result of this process, the money manager computes a weighted average of 5.5%. Thus, a point estimate is that the total amount of funds under management will increase by 5.5% in the next six months. The standard deviation of the observations in the sample is 10%. A histogram of past data looks fairly close to normal, so the manager assumes the population is normal.

1 Calculate a 95% confidence interval for the population mean and interpret your findings.

2 The manager decides to see what the confidence interval would look like if he had used a sample size of 20 or 30 and found the same mean (5.5%) and standard deviation (10%).

Compute the confidence interval for sample sizes of 20 and 30. For the sample size of 30, use Equation 6.

3 Interpret your results from Parts 1 and 2.

Solution to 1:

Because the population variance is unknown and the sample size is small, the manager must use the t-statistic in Equation 6 to calculate the confidence interval. Based on the sample size of 10, df = n − 1 = 10 − 1 = 9. For a 95% confidence interval, he needs to use the value of $t_{0.025}$ for df = 9. This value is 2.262, using Excel function T.INV(0.025,9). Therefore, a 95% confidence interval for the population mean is

$$\begin{aligned}\overline{X} \pm t_{0.025}\frac{s}{\sqrt{n}} &= 5.5\% \pm 2.262\frac{10\%}{\sqrt{10}}\\ &= 5.5\% \pm 2.262(3.162)\\ &= 5.5\% \pm 7.15\%\end{aligned}$$

The confidence interval for the population mean spans −1.65% to +12.65%. The manager can be confident at the 95% level that this range includes the population mean.

Solution to 2:

Exhibit 17 gives the calculations for the three sample sizes.

Exhibit 17 The 95% Confidence Interval for Three Sample Sizes

Distribution	95% Confidence Interval	Lower Bound	Upper Bound	Relative Size
$t(n = 10)$	5.5% ± 2.262(3.162)	−1.65%	12.65%	100.0%
$t(n = 20)$	5.5% ± 2.093(2.236)	0.82	10.18	65.5
$t(n = 30)$	5.5% ± 2.045(1.826)	1.77	9.23	52.2

Solution to 3:

The width of the confidence interval decreases as we increase the sample size. This decrease is a function of the standard error becoming smaller as n increases. The reliability factor also becomes smaller as the number of degrees of freedom increases. The last column of Exhibit 17 shows the relative size of the width of confidence intervals based on n = 10 to be 100%. Using a sample size of 20 reduces the confidence interval's width to 65.5% of the interval width for a sample size of 10. Using a sample size of 30 cuts the width of the interval almost in half. Comparing these choices, the money manager would obtain the most precise results using a sample of 30.

6 RESAMPLING

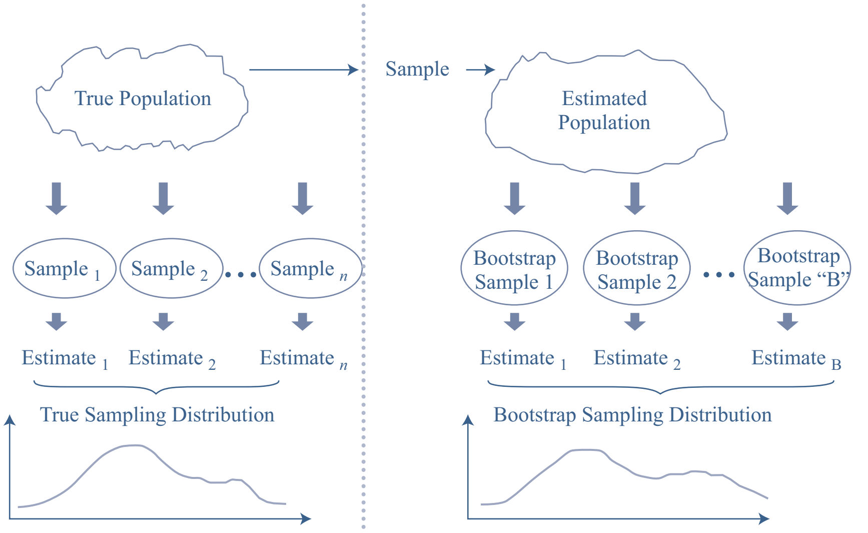
i. Describe the use of resampling (bootstrap, jackknife) to estimate the sampling distribution of a statistic

Earlier, we demonstrated how to find the standard error of the sample mean, which can be computed using Equation 4 based on the central limit theorem. We now introduce a computational tool called **resampling**, which repeatedly draws samples from the original observed data sample for the statistical inference of population parameters. **Bootstrap**, one of the most popular resampling methods, uses computer simulation for statistical inference without using an analytical formula such as a z-statistic or t-statistic.

The idea behind bootstrap is to mimic the process of performing random sampling from a population, similar to what we have shown earlier, to construct the sampling distribution of the sample mean. The difference lies in the fact that we have no knowledge of what the population looks like, except for a sample with size *n* drawn from the population. Because a random sample offers a good representation of the population, we can simulate sampling from the population by sampling from the observed sample. In other words, the bootstrap mimics the process by treating the randomly drawn sample as if it were the population.

The right-hand side of Exhibit 18 illustrates the process. In bootstrap, we repeatedly draw samples from the original sample, and each resample is of the same size as the original sample. Note that each item drawn is replaced for the next draw (i.e., the identical element is put back into the group so that it can be drawn more than once). Assuming we are looking to find the standard error of sample mean, we take many resamples and then compute the mean of each resample. Note that although some items may appear several times in the resamples, other items may not appear at all.

Exhibit 18 Bootstrap Resampling


True Population
Sample
Estimated Population
Sample 1
Sample 2
Sample n
Bootstrap Sample 1
Bootstrap Sample 2
Bootstrap Sample "B"
Estimate 1
Estimate 2
Estimate n
Estimate 1
Estimate 2
Estimate B
True Sampling Distribution
Bootstrap Sampling Distribution

Subsequently, we construct a sampling distribution with these resamples. The bootstrap sampling distribution (right-hand side of Exhibit 18) will approximate the true sampling distribution. We estimate the standard error of the sample mean using Equation 7. Note that to distinguish the foregoing resampling process from other types of resampling, it is often called model-free resampling or non-parametric resampling.

$$s_{\overline{X}} = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}_b - \overline{\theta})^2} \tag{7}$$

where

$S_{\overline{X}}$ is the estimate of the standard error of the sample mean

B denotes the number of resamples drawn from the original sample.

$\hat{\theta}_b$ denotes the mean of a resample

$\overline{\theta}$ denotes the mean across all the resample means

Bootstrap is one of the most powerful and widely used tools for statistical inference. As we have explained, it can be used to estimate the standard error of sample mean. Similarly, bootstrap can be used to find the standard error or construct confidence intervals for the statistic of other population parameters, such as the median, which does not apply to the previously discussed methodologies. Compared with conventional statistical methods, bootstrap does not rely on an analytical formula to estimate the distribution of the estimators. It is a simple but powerful method for any complicated estimators and particularly useful when no analytical formula is available. In addition, bootstrap has potential advantages in accuracy. Given these advantages, bootstrap can be applied widely in finance, such as for historical simulations in asset allocation or in gauging an investment strategy's performance against a benchmark.

EXAMPLE 11

Bootstrap Resampling Illustration

The following table displays a set of 12 monthly returns of a rarely traded stock, shown in Column A. Our aim is to calculate the standard error of the sample mean. Using the bootstrap resampling method, a series of bootstrap samples, labelled as "resamples" (with replacement) are drawn from the sample of 12 returns. Notice how some of the returns from data sample in Column A feature more than once in some of the resamples (for example, 0.055 features twice in Resample 1).

Column A	Resample 1	Resample 2	Resample 3		Resample 1,000
−0.096	0.055	−0.096	−0.033	.....	−0.072
−0.132	−0.033	0.055	−0.132	.....	0.255
−0.191	0.255	0.055	−0.157	.....	0.055
−0.096	−0.033	−0.157	0.255	.....	0.296
0.055	0.255	−0.096	−0.132	.....	0.055
−0.053	−0.157	−0.053	−0.191	.....	−0.096
−0.033	−0.053	−0.096	0.055	.....	0.296
0.296	−0.191	−0.132	0.255	.....	−0.132
0.055	−0.132	−0.132	0.296	.....	0.055
−0.072	−0.096	0.055	−0.096	.....	−0.096
0.255	0.055	−0.072	0.055	.....	−0.191
−0.157	−0.157	−0.053	−0.157	.....	0.055
Sample mean	−0.019	−0.060	0.001	.....	0.040

Drawing 1,000 such samples, we obtain 1,000 sample means. The mean across all resample means is −0.01367. The sum of squares of the differences between each sample mean and the mean across all resample means ($\sum_{b=1}^{B}(\hat{\theta}_b - \bar{\theta})^2$) is 1.94143. Using Equation 7, we calculate an estimate of the standard error of the sample mean:

$$s_{\bar{X}} = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}_b - \bar{\theta})^2} = \sqrt{\frac{1}{999} \times 1.94143} = 0.04408$$

Jackknife is another resampling technique for statistical inference of population parameters. Unlike bootstrap, which repeatedly draws samples with replacement, jackknife samples are selected by taking the original observed data sample and leaving out one observation at a time from the set (and not replacing it). Jackknife method is often used to reduce the bias of an estimator, and other applications include finding the standard error and confidence interval of an estimator. According to its computation procedure, we can conclude that jackknife produces similar results for every run, whereas bootstrap usually gives different results because bootstrap resamples are randomly drawn. For a sample of size *n*, jackknife usually requires *n* repetitions, whereas with bootstrap, we are left to determine how many repetitions are appropriate.

EXAMPLE 12

An analyst in a real estate investment company is researching the housing market of the Greater Boston area. From a sample of collected house sale price data in the past year, she estimates the median house price of the area. To find the standard error of the estimated median, she is considering two options:

Option 1 The standard error of the sample median can be given by $\frac{s}{\sqrt{n}}$, where s denotes the sample standard deviation and n denotes the sample size.

Option 2 Apply the bootstrap method to construct the sampling distribution of the sample median, and then compute the standard error by using Equation 7.

Which of the following statements is accurate?

A Option 1 is suitable to find the standard error of the sample median.

B Option 2 is suitable to find the standard error of the sample median.

C Both options are suitable to find the standard error of the sample median.

Solution:

B is correct. Option 1 is valid for estimating the standard error of the sample mean but not for that of the sample median, which is not based on the given formula. Thus, both A and C are incorrect. The bootstrap method is a simple way to find the standard error of an estimator even if no analytical formula is available or it is too complicated.

Having covered many of the fundamental concepts of sampling and estimation, we are in a good position to focus on sampling issues of special concern to analysts. The quality of inferences depends on the quality of the data as well as on the quality of the sampling plan used. Financial data pose special problems, and sampling plans frequently reflect one or more biases. The next section discusses these issues.

7 DATA SNOOPING BIAS, SAMPLE SELECTION BIAS, LOOK-AHEAD BIAS, AND TIME-PERIOD BIAS

j describe the issues regarding selection of the appropriate sample size, data snooping bias, sample selection bias, survivorship bias, look-ahead bias, and time-period bias

We have already seen that the selection of sample period length may raise the issue of sampling from more than one population. There are, in fact, a range of challenges to valid sampling that arise in working with financial data. In this section we discuss several such sampling-related issues: data snooping bias, sample selection group of biases (including survivorship bias), look-ahead bias, and time-period bias. All of these issues are important for point and interval estimation and hypothesis testing. As we will see, if the sample is biased in any way, then point and interval estimates and any other conclusions that we draw from the sample will be in error.

7.1 Data Snooping Bias

Data snooping relates to overuse of the same or related data in ways that we shall describe shortly. Data snooping bias refers to the errors that arise from such misuse of data. Investment strategies that reflect data snooping biases are often not successful if applied in the future. Nevertheless, both investment practitioners and researchers in general have frequently engaged in data snooping. Analysts thus need to understand and guard against this problem.

Data snooping is the practice of determining a model by extensive searching through a dataset for statistically significant patterns (that is, repeatedly “drilling” in the same data until finding something that appears to work). In exercises involving statistical significance, we set a significance level, which is the probability of rejecting the hypothesis we are testing when the hypothesis is in fact correct. Because rejecting a true hypothesis is undesirable, the investigator often sets the significance level at a relatively small number, such as 0.05 or 5%.

Suppose we test the hypothesis that a variable does not predict stock returns, and we test in turn 100 different variables. Let us also suppose that in truth, none of the 100 variables has the ability to predict stock returns. Using a 5% significance level in our tests, we would still expect that 5 out of 100 variables would appear to be significant predictors of stock returns because of random chance alone. We have mined the data to find some apparently significant variables. In essence, we have explored the same data again and again until we found some after-the-fact pattern or patterns in the dataset. This is the sense in which data snooping involves overuse of data. If we were to report only the significant variables without also reporting the total number of variables tested that were unsuccessful as predictors, we would be presenting a very misleading picture of our findings. Our results would appear to be far more significant than they actually were, because a series of tests such as the one just described invalidates the conventional interpretation of a given significance level (such as 5%), according to the theory of inference. Datasets in the Big Data space are often blindly used to make statistical inferences without a proper hypothesis testing framework, which may lead to inferring higher-than-justified significance.

How can we investigate the presence of data snooping bias? Typically we can split the data into three separate datasets: the training dataset, the validation dataset, and the test dataset. The training dataset is used to build a model and fit the model parameters. The validation dataset is used to evaluate the model fit while tuning the model parameters. The test dataset is to provide an **out-of-sample test** to evaluate the final model fit. If a variable or investment strategy is the result of data snooping, it should generally not be significant in out-of-sample tests.

A variable or investment strategy that is statistically and economically significant in out-of-sample tests, and that has a plausible economic basis, may be the basis for a valid investment strategy. Caution is still warranted, however. The most crucial out-of-sample test is future investment success. It should be noted that if the strategy becomes known to other investors, prices may adjust so that the strategy, however well tested, does not work in the future. To summarize, the analyst should be aware that many apparently profitable investment strategies may reflect data snooping bias and thus be cautious about the future applicability of published investment research results.

UNTANGLING THE EXTENT OF DATA SNOOPING

To assess the significance of an investment strategy, we need to know how many unsuccessful strategies were tried not only by the current investigator but also by *previous* investigators using the same or related datasets. Much research, in practice, closely builds on what other investigators have done, and so reflects intergenerational data mining (McQueen and Thorley, 1999) that involves using information developed by previous researchers using a dataset to guide current research using the same or a related dataset. Analysts have accumulated many observations about the peculiarities of many financial datasets, and other analysts may develop models or investment strategies that will tend to be supported within a dataset based on their familiarity with the prior experience of other analysts. As a consequence, the importance of those new results may be overstated. Research has suggested that the magnitude of this type of data-mining bias may be considerable.

McQueen and Thorley (1999) explored data mining in the context of the popular Motley Fool “Foolish Four” investment strategy, a version of the Dow Dividend Strategy tuned by its developers to exhibit an even higher arithmetic mean return than the original Dow Dividend Strategy. The Foolish Four strategy claimed to show significant investment returns over 20 years starting in 1973, and its proponents claimed that the strategy should have similar returns in the future. McQueen and Thorley highlighted the data-mining issues in that research and presented two signs that can warn analysts about the potential existence of data mining:

- *Too much digging/too little confidence*. The testing of many variables by the researcher is the “too much digging” warning sign of a data-mining problem. Although the number of variables examined may not be reported, we should look closely for verbal hints that the researcher searched over many variables. The use of terms such as “we noticed (or noted) that” or “someone noticed (or noted) that,” with respect to a pattern in a dataset, should raise suspicions that the researchers were trying out variables based on their own or others’ observations of the data.
- *No story/no future*. The absence of an explicit economic rationale for a variable or trading strategy is the “no story” warning sign of a data-mining problem. Without a plausible economic rationale or story for why a variable should work, the variable is unlikely to have predictive power. What if we do have a plausible economic explanation for a significant variable? McQueen and Thorley caution that a plausible economic rationale is a necessary but not a sufficient condition for a trading strategy to have value. As we mentioned earlier, if the strategy is publicized, market prices may adjust to reflect the new information as traders seek to exploit it; as a result, the strategy may no longer work.

7.2 Sample Selection Bias

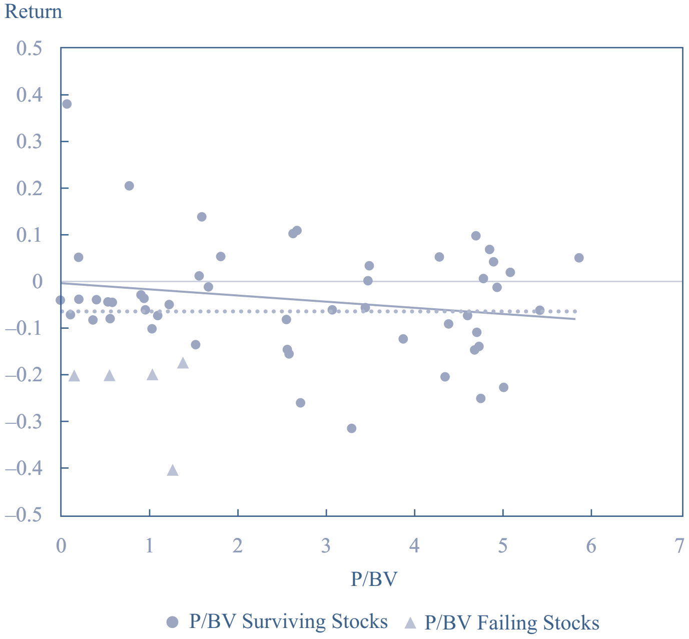
When researchers look into questions of interest to analysts or portfolio managers, they may exclude certain stocks, bonds, portfolios, or periods from the analysis for various reasons—perhaps because of data availability. When data availability leads to certain assets being excluded from the analysis, we call the resulting problem **sample selection bias**. For example, you might sample from a database that tracks only companies currently in existence. Many mutual fund databases, for instance, provide historical information about only those funds that currently exist. Databases that report historical balance sheet and income statement information suffer from the same sort of bias as the mutual fund databases: Funds or companies that are no longer in business do not appear there. So, a study that uses these types of databases suffers from a type of sample selection bias known as **survivorship bias**.

The issue of survivorship bias has also been raised in relation to international indexes, particularly those representing less established markets. Some of these markets have suffered complete loss of value as a result of hyperinflation, nationalization or confiscation of industries, or market failure. Measuring the performance of markets or particular investments that survive over time will overstate returns from investing. There is, of course, no way of determining in advance which markets will fail or survive.

Survivorship bias sometimes appears when we use both stock price and accounting data. For example, many studies in finance have used the ratio of a company's market price to book equity per share (i.e., the price-to-book ratio, P/B) and found that P/B is inversely related to a company's returns. P/B is also used to create many popular value and growth indexes. The "value" indexes, for example, would include companies trading on relatively low P/B. If the database that we use to collect accounting data excludes failing companies, however, a survivorship bias might result. It can be argued that failing stocks would be expected to have low returns and low P/Bs. If we exclude failing stocks, then those stocks with low P/Bs that are included in the index will have returns that are higher on average than if all stocks with low P/Bs were included. As shown in Exhibit 19, without failing stocks (shown in the bottom left part), we can fit a line with a negative slope indicating that P/B is inversely related to a company's stock return. With all the companies included, however, the fitted line (horizontal, dotted) shows an insignificant slope coefficient.

This bias would then be responsible for some of the traditional findings of an inverse relationship between average return and P/B. Researchers should be aware of any biases potentially inherent in a sample.

Exhibit 19 Survivorship Bias



DELISTINGS AND BIAS

A sample can also be biased because of the removal (or delisting) of a company's stock from an exchange. For example, the Center for Research in Security Prices at the University of Chicago is a major provider of return data used in academic research. When a delisting occurs, CRSP attempts to collect returns for the delisted company. Many times, however, it cannot do so because of the difficulty involved; CRSP must simply list delisted company returns as missing. A study in the *Journal of Finance* by Shumway and Warther (1999) documented the bias caused by delisting for CRSP NASDAQ return data. The authors showed that delistings associated with poor company performance (e.g., bankruptcy) are missed more often than delistings associated with good or neutral company performance (e.g., merger or moving to another exchange). In addition, delistings occur more frequently for small companies.

Sample selection bias occurs even in markets where the quality and consistency of the data are quite high. Newer asset classes such as hedge funds may present even greater problems of sample selection bias. Hedge funds are a heterogeneous group of investment vehicles typically organized so as to be free from regulatory oversight. In general, hedge funds are not required to publicly disclose performance (in contrast to, say, mutual funds). Hedge funds themselves decide whether they want to be included in one of the various databases of hedge fund performance. Hedge funds with poor track records clearly may not wish to make their records public, creating a problem of self-selection bias in hedge fund databases. Further, as pointed out by Fung and Hsieh (2002), because only hedge funds with good records will volunteer to enter a database, in general, overall past hedge fund industry performance will tend to appear better than it really is. Furthermore, many hedge fund databases drop funds that go out of business, creating survivorship bias in the database. Even if the database does not drop defunct hedge funds, in the attempt to eliminate survivorship bias, the problem remains of hedge funds that stop reporting performance because of poor results or because successful funds no longer want new cash inflows. In some circumstances, **implicit selection bias** may exist because of a threshold enabling self-selection. For example, compared with smaller exchanges, the NYSE has higher stock listing requirements. Choosing NYSE-listed stocks may introduce an implicit quality bias into the analysis. Although the bias is less obvious, it is important for generalizing findings.

A variation of selection bias is **backfill bias**. For example, when a new hedge fund is added to a given index, the fund's past performance may be backfilled into the index's database, even though the fund was not included in the database in the previous year. Usually a new fund starts contributing data after a period of good performance, so adding the fund's instant history into the index database may inflate the index performance.

7.3 Look-Ahead Bias

A test design is subject to **look-ahead bias** if it uses information that was not available on the test date. For example, tests of trading rules that use stock market returns and accounting balance sheet data must account for look-ahead bias. In such tests, a company's book value per share is commonly used to construct the P/B variable. Although the market price of a stock is available for all market participants at the same point in time, fiscal year-end book equity per share might not become publicly available until sometime in the following quarter. One solution to mitigate the look-ahead bias is to use point-in-time (PIT) data when possible. PIT data is stamped with the date when it was recorded or released. In the previous example, the PIT data of P/B would be accompanied with the date of company filing date or press release date, rather than the end date of the fiscal quarter the P/B data represents. It is worth noting that the look-ahead bias could also be implicitly introduced. For example, when normalizing input data by deducting the mean and dividing it by standard deviation, we must ensure that the standard deviation of the training data is used as the proxy for standard deviation in validation and test data sets. Using standard deviation of validation or test data to normalize them will implicitly introduce a look-ahead bias as the variance of future data is inappropriately used.

7.4 Time-Period Bias

A test design is subject to **time-period bias** if it is based on a period that may make the results period specific. A short time series is likely to give period-specific results that may not reflect a longer period. A long time series may give a more accurate picture of true investment performance; its disadvantage lies in the potential for a structural change occurring during the time frame that would result in two different return distributions. In this situation, the distribution that would reflect conditions before the change differs from the distribution that would describe conditions after the change. Regime changes, such as low versus high volatility regimes or low versus high interest rate regimes, are highly influential to asset classes. Inferences based on data influenced by one regime, and thus not appropriately distributed, should account for how the regime may bias the inferences.

EXAMPLE 13

Biases in Investment Research

An analyst is reviewing the empirical evidence on historical equity returns in the Eurozone (European countries that use the euro). She finds that value stocks (i.e., those with low P/Bs) outperformed growth stocks (i.e., those with high P/Bs) in recent periods. After reviewing the Eurozone market, the analyst wonders whether value stocks might be attractive in the United Kingdom. She investigates the performance of value and growth stocks in the UK market for a 10-year period. To conduct this research, the analyst does the following:

- obtains the current composition of the Financial Times Stock Exchange (FTSE) All Share Index, a market-capitalization-weighted index;
- eliminates the companies that do not have December fiscal year-ends;
- uses year-end book values and market prices to rank the remaining universe of companies by P/Bs at the end of the year;
- based on these rankings, divides the universe into 10 portfolios, each of which contains an equal number of stocks;
- calculates the equal-weighted return of each portfolio and the return for the FTSE All Share Index for the 12 months following the date each ranking was made; and
- subtracts the FTSE returns from each portfolio's returns to derive excess returns for each portfolio.

She discusses the research process with her supervisor, who makes two comments:

- The proposed process may introduce survivorship bias into her analysis.
- The proposed research should cover a longer period.

1 Which of the following best describes the supervisor's first comment?

A The comment is false. The proposed method is designed to avoid survivorship bias.

B The comment is true, because she is planning to use the current list of FTSE stocks rather than the actual list of stocks that existed at the start of each year.

C The comment is true, because the test design uses information unavailable on the test date.

2 What bias is the supervisor concerned about when making the second comment?

A Time period bias, because the results may be period specific

B Look-ahead bias, because the bias could be reduced or eliminated if one uses a longer period

C Survivorship bias, because the bias would become less relevant over longer periods

Solution to 1:

B is correct because the research design is subject to survivorship bias if it fails to account for companies that have gone bankrupt, merged, or otherwise departed the database. Using the current list of FTSE stocks rather than the actual list of stocks that existed at the start of each year means that the computation of returns excluded companies removed from the index. The performance of the portfolios with the lowest P/B is subject to survivorship bias and may be overstated. At some time during the testing period, those companies not currently in existence were eliminated from testing. They would probably have had low prices (and low P/Bs) and poor returns.

A is incorrect because the method is not designed to avoid survivorship bias. C is incorrect because the fact that the test design uses information unavailable on the test date relates to look-ahead bias. A test design is subject to look-ahead bias if it uses information unavailable on the test date. This bias would make a strategy based on the information appear successful, but it assumes perfect forecasting ability.

Solution to 2:

A is correct. A test design is subject to time-period bias if it is based on a period that may make the results period specific. Although the research covered a period of 10 years, that period may be too short for testing an anomaly. Ideally, an analyst should test market anomalies over several market cycles to ensure that results are not period specific. This bias can favor a proposed strategy if the period chosen was favorable to the strategy.

SUMMARY

In this reading, we have presented basic concepts and results in sampling and estimation. We have also emphasized the challenges faced by analysts in appropriately using and interpreting financial data. As analysts, we should always use a critical eye when evaluating the results from any study. The quality of the sample is of the utmost importance: If the sample is biased, the conclusions drawn from the sample will be in error.

- To draw valid inferences from a sample, the sample should be random.
- In simple random sampling, each observation has an equal chance of being selected. In stratified random sampling, the population is divided into subpopulations, called strata or cells, based on one or more classification criteria; simple random samples are then drawn from each stratum.
- Stratified random sampling ensures that population subdivisions of interest are represented in the sample. Stratified random sampling also produces more-precise parameter estimates than simple random sampling.
- Convenience sampling selects an element from the population on the basis of whether or not it is accessible to a researcher or how easy it is to access. Because convenience sampling presents the advantage of collecting data quickly at a low cost, it is a suitable sampling plan for small-scale pilot studies.
- Judgmental sampling may yield skewed results because of the bias of researchers, but its advantages lie in the fact that in some circumstances, the specialty of researchers and their judgmental can lead them directly to the target population of interest within time constraints.
- The central limit theorem states that for large sample sizes, for any underlying distribution for a random variable, the sampling distribution of the sample mean for that variable will be approximately normal, with mean equal to the population mean for that random variable and variance equal to the population variance of the variable divided by sample size.
- Based on the central limit theorem, when the sample size is large, we can compute confidence intervals for the population mean based on the normal distribution regardless of the distribution of the underlying population. In general, a sample size of 30 or larger can be considered large.

- An estimator is a formula for estimating a parameter. An estimate is a particular value that we calculate from a sample by using an estimator.
- Because an estimator or statistic is a random variable, it is described by some probability distribution. We refer to the distribution of an estimator as its sampling distribution. The standard deviation of the sampling distribution of the sample mean is called the standard error of the sample mean.
- The desirable properties of an estimator are *unbiasedness* (the expected value of the estimator equals the population parameter), *efficiency* (the estimator has the smallest variance), and *consistency* (the probability of accurate estimates increases as sample size increases).
- The two types of estimates of a parameter are point estimates and interval estimates. A point estimate is a single number that we use to estimate a parameter. An interval estimate is a range of values that brackets the population parameter with some probability.
- A confidence interval is an interval for which we can assert with a given probability $1 - \alpha$, called the degree of confidence, that it will contain the parameter it is intended to estimate. This measure is often referred to as the $100(1 - \alpha)\%$ confidence interval for the parameter.
- A $100(1 - \alpha)\%$ confidence interval for a parameter has the following structure: Point estimate ± Reliability factor × Standard error, where the reliability factor is a number based on the assumed distribution of the point estimate and the degree of confidence $(1 - \alpha)$ for the confidence interval and where standard error is the standard error of the sample statistic providing the point estimate.
- A $100(1 - \alpha)\%$ confidence interval for population mean μ when sampling from a normal distribution with known variance σ^2 is given by $\overline{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$, where $z_{\alpha/2}$ is the point of the standard normal distribution such that $\alpha/2$ remains in the right tail.
- A random sample of size n is said to have $n - 1$ degrees of freedom for estimating the population variance, in the sense that there are only $n - 1$ independent deviations from the mean on which to base the estimate.
- A $100(1 - \alpha)\%$ confidence interval for the population mean μ when sampling from a normal distribution with unknown variance (a *t*-distribution confidence interval) is given by $\overline{X} \pm t_{\alpha/2}\left(s/\sqrt{n}\right)$, where $t_{\alpha/2}$ is the point of the t-distribution such that $\alpha/2$ remains in the right tail and s is the sample standard deviation. This confidence interval can also be used, because of the central limit theorem, when dealing with a large sample from a population with unknown variance that may not be normal.
- We may use the confidence interval $\overline{X} \pm z_{\alpha/2}\left(s/\sqrt{n}\right)$ as an alternative to the t-distribution confidence interval for the population mean when using a large sample from a population with unknown variance. The confidence interval based on the z-statistic is less conservative (narrower) than the corresponding confidence interval based on a t-distribution.
- Bootstrap and jackknife are simple but powerful methods for statistical inference, and they are particularly useful when no analytical formula is available. Bootstrap constructs the sampling distribution of an estimator by repeatedly drawing samples from the original sample to find standard error and confidence interval. Jackknife draws repeated samples while leaving out one observation at a time from the set, without replacing it.

- Three issues in the selection of sample size are the need for precision, the risk of sampling from more than one population, and the expenses of different sample sizes.
- *Data snooping bias* comes from finding models by repeatedly searching through databases for patterns.
- *Sample selection bias* occurs when data availability leads to certain assets being excluded from the analysis, we call the resulting problem
- *Survivorship bias* is a subset of sample selection bias and occurs if companies are excluded from the analysis because they have gone out of business or because of reasons related to poor performance.
- *Self-selection* bias reflects the ability of entities to decide whether or not they wish to report their attributes or results and be included in databases or samples. *Implicit selection bias* is one type of selection bias introduced through the presence of a threshold that filters out some unqualified members. A subset of selection bias is *backfill bias*, in which past data, not reported or used before, is backfilled into an existing database.
- *Look-ahead bias* exists if the model uses data not available to market participants at the time the market participants act in the model.
- *Time-period bias* is present if the period used makes the results period specific or if the period used includes a point of structural change.

REFERENCES

Fung, William, and David Hsieh. 2002. "Hedge-Fund Benchmarks: Information Content and Biases." Financial Analysts Journal, vol. 58, no. 1:22–34. doi:10.2469/faj.v58.n1.2507

McQueen, Grant, and Steven Thorley. 1999. "Mining Fools Gold." Financial Analysts Journal, vol. 55, no. 2:61–72. doi:10.2469/faj.v55.n2.2261

Shumway, Tyler, and Vincent A. Warther. 1999. "The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect." Journal of Finance, vol. 54, no. 6:2361–2379. doi:10.1111/0022-1082.00192

PRACTICE PROBLEMS

1 Perkiomen Kinzua, a seasoned auditor, is auditing last year's transactions for Conemaugh Corporation. Unfortunately, Conemaugh had a very large number of transactions last year, and Kinzua is under a time constraint to finish the audit. He decides to audit only the small subset of the transaction population that is of interest and to use sampling to create that subset.

The most appropriate sampling method for Kinzua to use is:

A judgmental sampling.

B systematic sampling.

C convenience sampling.

2 Which one of the following statements is true about non-probability sampling?

A There is significant risk that the sample is not representative of the population.

B Every member of the population has an equal chance of being selected for the sample.

C Using judgment guarantees that population subdivisions of interest are represented in the sample.

3 The *best* approach for creating a stratified random sample of a population involves:

A drawing an equal number of simple random samples from each subpopulation.

B selecting every *k*th member of the population until the desired sample size is reached.

C drawing simple random samples from each subpopulation in sizes proportional to the relative size of each subpopulation.

4 Although he knows security returns are not independent, a colleague makes the claim that because of the central limit theorem, if we diversify across a large number of investments, the portfolio standard deviation will eventually approach zero as n becomes large. Is he correct?

5 Why is the central limit theorem important?

6 What is wrong with the following statement of the central limit theorem?

> *Central Limit Theorem.* "If the random variables $X_1, X_2, X_3, \ldots, X_n$ are a random sample of size n from any distribution with finite mean μ and variance σ^2, then the distribution of $\overline{X}$ will be approximately normal, with a standard deviation of $\sigma/\sqrt{n}$."

7 Peter Biggs wants to know how growth managers performed last year. Biggs assumes that the population cross-sectional standard deviation of growth manager returns is 6% and that the returns are independent across managers.

A How large a random sample does Biggs need if he wants the standard deviation of the sample means to be 1%?

B How large a random sample does Biggs need if he wants the standard deviation of the sample means to be 0.25%?

© 2021 CFA Institute. All rights reserved.

8 Petra Munzi wants to know how value managers performed last year. Munzi estimates that the population cross-sectional standard deviation of value manager returns is 4% and assumes that the returns are independent across managers.

A Munzi wants to build a 95% confidence interval for the population mean return. How large a random sample does Munzi need if she wants the 95% confidence interval to have a total width of 1%?

B Munzi expects a cost of about $10 to collect each observation. If she has a $1,000 budget, will she be able to construct the confidence interval she wants?

9 Find the reliability factors based on the *t*-distribution for the following confidence intervals for the population mean (df = degrees of freedom, n = sample size):

A A 99% confidence interval, df = 20

B A 90% confidence interval, df = 20

C A 95% confidence interval, $n = 25$

D A 95% confidence interval, $n = 16$

10 Assume that monthly returns are normally distributed with a mean of 1% and a sample standard deviation of 4%. The population standard deviation is unknown. Construct a 95% confidence interval for the sample mean of monthly returns if the sample size is 24.

11 Explain the differences between constructing a confidence interval when sampling from a normal population with a known population variance and sampling from a normal population with an unknown variance.

12 Suppose we take a random sample of 30 companies in an industry with 200 companies. We calculate the sample mean of the ratio of cash flow to total debt for the prior year. We find that this ratio is 23%. Subsequently, we learn that the population cash flow to total debt ratio (taking account of all 200 companies) is 26%. What is the explanation for the discrepancy between the sample mean of 23% and the population mean of 26%?

A Sampling error.

B Bias.

C A lack of consistency.

13 Alcorn Mutual Funds is placing large advertisements in several financial publications. The advertisements prominently display the returns of 5 of Alcorn's 30 funds for the past 1-, 3-, 5-, and 10-year periods. The results are indeed impressive, with all of the funds beating the major market indexes and a few beating them by a large margin. Is the Alcorn family of funds superior to its competitors?

14 Julius Spence has tested several predictive models in order to identify undervalued stocks. Spence used about 30 company-specific variables and 10 market-related variables to predict returns for about 5,000 North American and European stocks. He found that a final model using eight variables applied to telecommunications and computer stocks yields spectacular results. Spence wants you to use the model to select investments. Should you? What steps would you take to evaluate the model?

15 A population has a non-normal distribution with mean μ and variance σ^2. The sampling distribution of the sample mean computed from samples of large size from that population will have:

A the same distribution as the population distribution.

B its mean approximately equal to the population mean.

C its variance approximately equal to the population variance.

16 A sample mean is computed from a population with a variance of 2.45. The sample size is 40. The standard error of the sample mean is *closest* to:

A 0.039.

B 0.247.

C 0.387.

17 An estimator with an expected value equal to the parameter that it is intended to estimate is described as:

A efficient.

B unbiased.

C consistent.

18 If an estimator is consistent, an increase in sample size will increase the:

A accuracy of estimates.

B efficiency of the estimator.

C unbiasedness of the estimator.

19 For a two-sided confidence interval, an increase in the degree of confidence will result in:

A a wider confidence interval.

B a narrower confidence interval.

C no change in the width of the confidence interval.

20 For a sample size of 17, with a mean of 116.23 and a variance of 245.55, the width of a 90% confidence interval using the appropriate t-distribution is *closest to*:

A 13.23.

B 13.27.

C 13.68.

21 For a sample size of 65 with a mean of 31 taken from a normally distributed population with a variance of 529, a 99% confidence interval for the population mean will have a lower limit *closest* to:

A 23.64.

B 25.41.

C 30.09.

22 An increase in sample size is *most likely* to result in a:

A wider confidence interval.

B decrease in the standard error of the sample mean.

C lower likelihood of sampling from more than one population.

23 Otema Chi has a spreadsheet with 108 monthly returns for shares in Marunou Corporation. He writes a software program that uses bootstrap resampling to create 200 resamples of this Marunou data by sampling with replacement. Each resample has 108 data points. Chi's program calculates the mean of each of the 200 resamples, and then it calculates that the mean of these 200 resample means is 0.0261. The program subtracts 0.0261 from each of the 200 resample means, squares each of these 200 differences, and adds the squared differences together. The result is 0.835. The program then calculates an estimate of the standard error of the sample mean.

The estimated standard error of the sample mean is closest to:

A 0.0115

B 0.0648

C 0.0883

24 Compared with bootstrap resampling, jackknife resampling:

A is done with replacement.

B usually requires that the number of repetitions is equal to the sample size.

C produces dissimilar results for every run because resamples are randomly drawn.

25 A report on long-term stock returns focused exclusively on all currently publicly traded firms in an industry is *most likely* susceptible to:

A look-ahead bias.

B survivorship bias.

C intergenerational data mining.

26 Which sampling bias is *most likely* investigated with an out-of-sample test?

A Look-ahead bias

B Data-mining bias

C Sample selection bias

27 Which of the following characteristics of an investment study *most likely* indicates time-period bias?

A The study is based on a short time-series.

B Information not available on the test date is used.

C A structural change occurred prior to the start of the study's time series.

SOLUTIONS

1. A is correct. With judgmental sampling, Kinzua will use his knowledge and professional judgment as a seasoned auditor to select transactions of interest from the population. This approach will allow Kinzua to create a sample that is representative of the population and that will provide sufficient audit coverage. Judgmental sampling is useful in cases that have a time constraint or in which the specialty of researchers is critical to select a more representative sample than by using other probability or non-probability sampling methods. Judgement sampling, however, entails the risk that Kinzua is biased in his selections, leading to skewed results that are not representative of the whole population.
2. A is correct. Because non-probability sampling is dependent on factors other than probability considerations, such as a sampler's judgment or the convenience to access data, there is a significant risk that non-probability sampling might generate a non-representative sample
3. C is correct. Stratified random sampling involves dividing a population into subpopulations based on one or more classification criteria. Then, simple random samples are drawn from each subpopulation in sizes proportional to the relative size of each subpopulation. These samples are then pooled to form a stratified random sample.
4. No. First the conclusion on the limit of zero is wrong; second, the support cited for drawing the conclusion (i.e., the central limit theorem) is not relevant in this context.
5. In many instances, the distribution that describes the underlying population is not normal or the distribution is not known. The central limit theorem states that if the sample size is large, regardless of the shape of the underlying population, the distribution of the sample mean is approximately normal. Therefore, even in these instances, we can still construct confidence intervals (and conduct tests of inference) as long as the sample size is large (generally $n \geq 30$).
6. The statement makes the following mistakes:
   - Given the conditions in the statement, the distribution of $\overline{X}$ will be approximately normal only for large sample sizes.
   - The statement omits the important element of the central limit theorem that the distribution of $\overline{X}$ will have mean μ.
7. **A** The standard deviation or standard error of the sample mean is $\sigma_{\overline{X}} = \sigma/\sqrt{n}$. Substituting in the values for $\sigma_{\overline{X}}$ and σ, we have $1\% = 6\%/\sqrt{n}$, or $\sqrt{n} = 6$. Squaring this value, we get a random sample of $n = 36$.

   B As in Part A, the standard deviation of sample mean is $\sigma_{\overline{X}} = \sigma/\sqrt{n}$. Substituting in the values for $\sigma_{\overline{X}}$ and σ, we have $0.25\% = 6\%/\sqrt{n}$, or $\sqrt{n} = 24$. Squaring this value, we get a random sample of $n = 576$, which is substantially larger than for Part A of this question.

8 **A** Assume the sample size will be large and thus the 95% confidence interval for the population mean of manager returns is $\overline{X} \pm 1.96 s_{\overline{X}}$, where $s_{\overline{X}} = s/\sqrt{n}$. Munzi wants the distance between the upper limit and lower limit in the confidence interval to be 1%, which is

$$\left(\overline{X} + 1.96 s_{\overline{X}}\right) - \left(\overline{X} - 1.96 s_{\overline{X}}\right) = 1\%$$

Simplifying this equation, we get $2\left(1.96 s_{\overline{X}}\right) = 1\%$. Finally, we have $3.92 s_{\overline{X}} =$ 1%, which gives us the standard deviation of the sample mean, $s_{\overline{X}} = 0.255\%$. The distribution of sample means is $s_{\overline{X}} = s/\sqrt{n}$. Substituting in the values for $s_{\overline{X}}$ and s, we have $0.255\% = 4\%/\sqrt{n}$, or $\sqrt{n} = 15.69$. Squaring this value, we get a random sample of $n = 246$.

B With her budget, Munzi can pay for a sample of up to 100 observations, which is far short of the 246 observations needed. Munzi can either proceed with her current budget and settle for a wider confidence interval or she can raise her budget (to around $2,460) to get the sample size for a 1% width in her confidence interval.

9 **A** For a 99% confidence interval, the reliability factor we use is $t_{0.005}$; for df = 20, this factor is 2.845.

B For a 90% confidence interval, the reliability factor we use is $t_{0.05}$; for df = 20, this factor is 1.725.

C Degrees of freedom equals $n - 1$, or in this case $25 - 1 = 24$. For a 95% confidence interval, the reliability factor we use is $t_{0.025}$; for df = 24, this factor is 2.064.

D Degrees of freedom equals $16 - 1 = 15$. For a 95% confidence interval, the reliability factor we use is $t_{0.025}$; for df = 15, this factor is 2.131.

10 Because this is a small sample from a normal population and we have only the sample standard deviation, we use the following model to solve for the confidence interval of the population mean:

$$\overline{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

where we find $t_{0.025}$ (for a 95% confidence interval) for df = $n - 1 = 24 - 1 = 23$; this value is 2.069. Our solution is $1\% \pm 2.069(4\%)/\sqrt{24} = 1\% \pm 2.069(0.8165) =$ $1\% \pm 1.69$. The 95% confidence interval spans the range from −0.69% to +2.69%.

11 If the population variance is known, the confidence interval is

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

The confidence interval for the population mean is centered at the sample mean, $\overline{X}$. The population standard deviation is σ, and the sample size is n. The population standard deviation divided by the square root of n is the standard error of the estimate of the mean. The value of z depends on the desired degree of confidence. For a 95% confidence interval, $z_{0.025} = 1.96$ and the confidence interval estimate is

$$\overline{X} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

If the population variance is not known, we make two changes to the technique used when the population variance is known. First, we must use the sample standard deviation instead of the population standard deviation. Second, we use the *t*-distribution instead of the normal distribution. The critical *t*-value will depend on degrees of freedom $n - 1$. If the sample size is large, we have the alternative of using the *z*-distribution with the sample standard deviation.

12 A is correct. The discrepancy arises from sampling error. Sampling error exists whenever one fails to observe every element of the population, because a sample statistic can vary from sample to sample. As stated in the reading, the sample mean is an unbiased estimator, a consistent estimator, and an efficient estimator of the population mean. Although the sample mean is an unbiased estimator of the population mean—the expected value of the sample mean equals the population mean—because of sampling error, we do not expect the sample mean to exactly equal the population mean in any one sample we may take.

13 No, we cannot say that Alcorn Mutual Funds as a group is superior to competitors. Alcorn Mutual Funds' advertisement may easily mislead readers because the advertisement does not show the performance of all its funds. In particular, Alcorn Mutual Funds is engaging in sample selection bias by presenting the investment results from its best-performing funds only.

14 Spence may be guilty of data mining. He has used so many possible combinations of variables on so many stocks, it is not surprising that he found some instances in which a model worked. In fact, it would have been more surprising if he had not found any. To decide whether to use his model, you should do two things: First, ask that the model be tested on out-of-sample data—that is, data that were not used in building the model. The model may not be successful with out-of-sample data. Second, examine his model to make sure that the relationships in the model make economic sense, have a story, and have a future.

15 B is correct. Given a population described by any probability distribution (normal or non-normal) with finite variance, the central limit theorem states that the sampling distribution of the sample mean will be approximately normal, with the mean approximately equal to the population mean, when the sample size is large.

16 B is correct. Taking the square root of the known population variance to determine the population standard deviation (σ) results in

$$\sigma = \sqrt{2.45} = 1.565$$

The formula for the standard error of the sample mean (σ_X), based on a known sample size (n), is

$$\sigma_X = \frac{\sigma}{\sqrt{n}}$$

Therefore,

$$\sigma_X = \frac{1.565}{\sqrt{40}} = 0.247$$

17 B is correct. An unbiased estimator is one for which the expected value equals the parameter it is intended to estimate.

18 A is correct. A consistent estimator is one for which the probability of estimates close to the value of the population parameter increases as sample size increases. More specifically, a consistent estimator's sampling distribution becomes concentrated on the value of the parameter it is intended to estimate as the sample size approaches infinity.

19 A is correct. As the degree of confidence increases (e.g., from 95% to 99%), a given confidence interval will become wider. A confidence interval is a range for which one can assert with a given probability 1 – α, called the degree of confidence, that it will contain the parameter it is intended to estimate.

20 B is correct. The confidence interval is calculated using the following equation:

$$\overline{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Sample standard deviation (s) = $\sqrt{245.55}$ = 15.670.

For a sample size of 17, degrees of freedom equal 16, so $t_{0.05}$ = 1.746.

The confidence interval is calculated as

$$116.23 \pm 1.746 \frac{15.67}{\sqrt{17}} = 116.23 \pm 6.6357$$

Therefore, the interval spans 109.5943 to 122.8656, meaning its width is equal to approximately 13.271. (This interval can be alternatively calculated as 6.6357 × 2.)

21 A is correct. To solve, use the structure of Confidence interval = Point estimate ± Reliability factor × Standard error, which, for a normally distributed population with known variance, is represented by the following formula:

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

For a 99% confidence interval, use $z_{0.005}$ = 2.58.

Also, σ = $\sqrt{529}$ = 23.

Therefore, the lower limit = $31 - 2.58 \frac{23}{\sqrt{65}} = 23.6398$.

22 B is correct. All else being equal, as the sample size increases, the standard error of the sample mean decreases and the width of the confidence interval also decreases.

23 B is correct.

The estimate of the standard error of the sample mean with bootstrap resampling is calculated as follows:

$$s_{\overline{X}} = \sqrt{\frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{,}_{b} - \bar{,}\right)^2} = \sqrt{\frac{1}{200-1} \sum_{b=1}^{200} \left(\hat{,}_{b} - 0.0261\right)^2} = \sqrt{\frac{1}{199} \times 0.835}$$

$$s_{\overline{X}} = 0.0648$$

24 B is correct. For a sample of size n, jackknife resampling usually requires n repetitions. In contrast, with bootstrap resampling, we are left to determine how many repetitions are appropriate.

25 B is correct. A report that uses a current list of stocks does not account for firms that failed, merged, or otherwise disappeared from the public equity market in previous years. As a consequence, the report is biased. This type of bias is known as survivorship bias.

26 B is correct. An out-of-sample test is used to investigate the presence of data-mining bias. Such a test uses a sample that does not overlap the time period of the sample on which a variable, strategy, or model was developed.

27 A is correct. A short time series is likely to give period-specific results that may not reflect a longer time period.

READING

6

Hypothesis Testing

by Pamela Peterson Drake, PhD, CFA

Pamela Peterson Drake, PhD, CFA, is at James Madison University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** define a hypothesis, describe the steps of hypothesis testing, and describe and interpret the choice of the null and alternative hypotheses;
☐	**b.** compare and contrast one-tailed and two-tailed tests of hypotheses;
☐	**c.** explain a test statistic, Type I and Type II errors, a significance level, how significance levels are used in hypothesis testing, and the power of a test;
☐	**d.** explain a decision rule and the relation between confidence intervals and hypothesis tests, and determine whether a statistically significant result is also economically meaningful.
☐	**e.** explain and interpret the *p*-value as it relates to hypothesis testing;
☐	**f.** describe how to interpret the significance of a test in the context of multiple tests;
☐	**g.** identify the appropriate test statistic and interpret the results for a hypothesis test concerning the population mean of both large and small samples when the population is normally or approximately normally distributed and the variance is (1) known or (2) unknown;
☐	**h.** identify the appropriate test statistic and interpret the results for a hypothesis test concerning the equality of the population means of two at least approximately normally distributed populations based on independent random samples with equal assumed variances;
☐	**i.** identify the appropriate test statistic and interpret the results for a hypothesis test concerning the mean difference of two normally distributed populations;

(continued)

© 2021 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**j.** identify the appropriate test statistic and interpret the results for a hypothesis test concerning (1) the variance of a normally distributed population and (2) the equality of the variances of two normally distributed populations based on two independent random samples;
☐	**k.** compare and contrast parametric and nonparametric tests, and describe situations where each is the more appropriate type of test;
☐	**l.** explain parametric and nonparametric tests of the hypothesis that the population correlation coefficient equals zero, and determine whether the hypothesis is rejected at a given level of significance;
☐	**m.** explain tests of independence based on contingency table data.

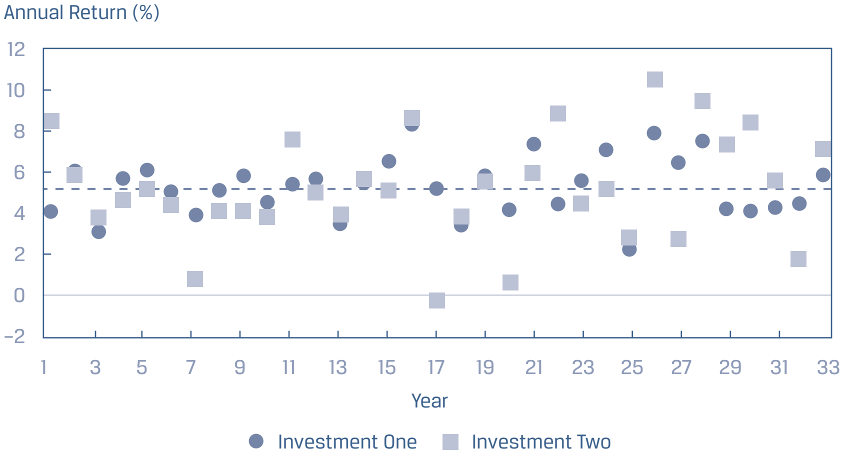
1 INTRODUCTION

a define a hypothesis, describe the steps of hypothesis testing, and describe and interpret the choice of the null and alternative hypotheses

1.1 Why Hypothesis Testing?

Faced with an overwhelming amount of data, analysts must deal with the task of wrangling those data into something that provides a clearer picture of what is going on. Consider an analyst evaluating the returns on two investments over 33 years, as we show in Exhibit 1.

Exhibit 1 Returns for Investments One and Two over 33 Years



Although "a picture is worth a thousand words," what can we actually glean from this plot? Can we tell if each investment's returns are different from an average of 5%? Can we tell whether the returns are different for Investment One and Investment Two? Can we tell whether the standard deviations of the two investments are each different from 2%? Can we tell whether the variability is different for the two investments? For these questions, we need to have more precise tools than simply a plot over time. What we need is a set of tools that aid us in making decisions based on the data.

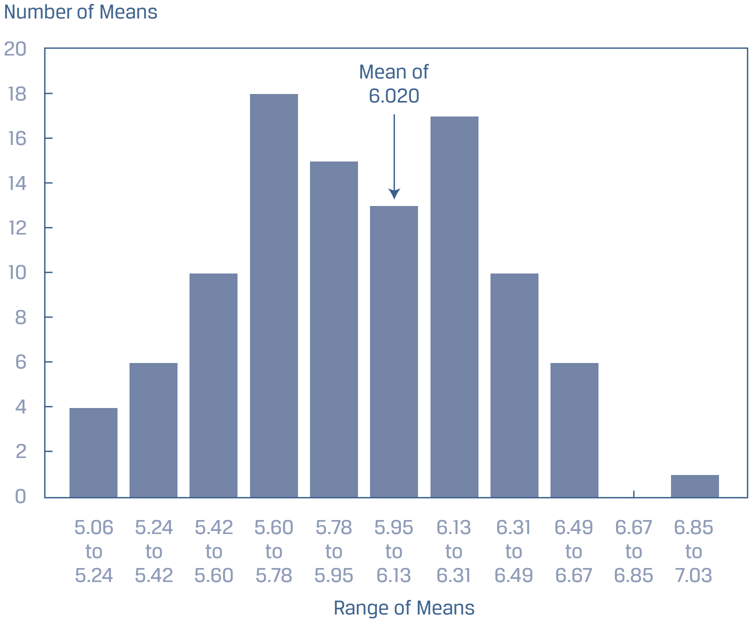
We use the concepts and tools of **hypothesis testing** to address these questions. Hypothesis testing is part of statistical inference, the process of making judgments about a larger group (a population) based on a smaller group of observations (that is, a sample).

1.2 Implications from a Sampling Distribution

Consider a set of 1,000 asset returns with a mean of 6% and a standard deviation of 2%. If we draw a sample of returns from this population, what is the chance that the mean of this sample will be 6%? What we know about sampling distributions is that how close any given sample mean will be to the population mean depends on the sample size, the variability within the population, and the quality of our sampling methodology.

For example, suppose we draw a sample of 30 observations and the sample mean is 6.13%. Is this close enough to 6% to alleviate doubt that the sample is drawn from a population with a mean of 6%? Suppose we draw another sample of 30 and find a sample mean of 4.8%. Does this bring into doubt whether the population mean is 6%? If we keep drawing samples of 30 observations from this population, we will get a range of possible sample means, as we show in Exhibit 2 for 100 different samples of size 30 from this population, with a range of values from 5.06 to 7.03%. All these sample means are a result of sampling from the 1,000 asset returns.

Exhibit 2 Distribution of Sample Means of 100 Samples Drawn from a Population of 1,000 Returns



As you can see in Exhibit 2, a sample mean that is quite different from the population mean can occur; this situation is not as likely as drawing a sample with a mean closer to the population mean, but it can still happen. *In hypothesis testing, we test to see whether a sample statistic is likely to come from a population with the hypothesized value of the population parameter.*

The concepts and tools of hypothesis testing provide an objective means to gauge whether the available evidence supports the hypothesis. After applying a statistical test of a hypothesis, we should have a clearer idea of the probability that a hypothesis is true or not, although our conclusion always stops short of certainty.

The main focus of this reading is on the framework of hypothesis testing and tests concerning mean, variance, and correlation, three quantities frequently used in investments.

2

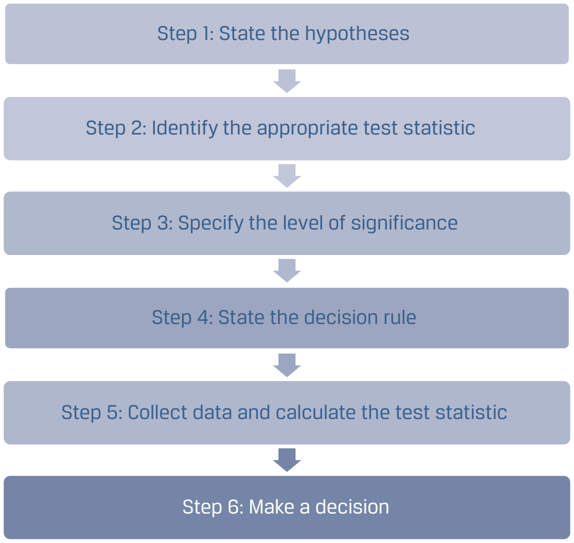
THE PROCESS OF HYPOTHESIS TESTING

Hypothesis testing is part of the branch of statistics known as statistical inference. In statistical inference, there is estimation and hypothesis testing. Estimation involves point estimates and interval estimates. Consider a sample mean, which is a point estimate, that we can use to form a confidence interval. In hypothesis testing, the focus is examining how a sample statistic informs us about a population parameter. A **hypothesis** is a statement about one or more populations that we test using sample statistics.

The process of hypothesis testing begins with the formulation of a theory to organize and explain observations. We judge the correctness of the theory by its ability to make accurate predictions—for example, to predict the results of new observations. If the predictions are correct, we continue to maintain the theory as a possibly correct explanation of our observations. Risk plays a role in the outcomes of observations in finance, so we can only try to make unbiased, probability-based judgments about whether the new data support the predictions. Statistical hypothesis testing fills that key role of testing hypotheses when there is uncertainty. When an analyst correctly formulates the question into a testable hypothesis and carries out a test of hypotheses, the use of well-established scientific methods supports the conclusions and decisions made on the basis of this test.

We organize this introduction to hypothesis testing around the six steps in Exhibit 3, which illustrate the standard approach to hypothesis testing.

Exhibit 3 The Process of Hypothesis Testing


Step 1: State the hypotheses
Step 2: Identify the appropriate test statistic
Step 3: Specify the level of significance
Step 4: State the decision rule
Step 5: Collect data and calculate the test statistic
Step 6: Make a decision

2.1 Stating the Hypotheses

For each hypothesis test, we always state two hypotheses: the **null hypothesis** (or null), designated H_0, and the **alternative hypothesis**, designated H_a. For example, our null hypothesis may concern the value of a population mean, μ, in relation to one possible value of the mean, μ_0. As another example, our null hypothesis may concern the population variance, σ^2, compared with a possible value of this variance, $\sigma_0{}^2$. The null hypothesis is a statement concerning a population parameter or parameters considered to be true unless the sample we use to conduct the hypothesis test gives convincing evidence that the null hypothesis is false. In fact, the null hypothesis is what we want to reject. If there is sufficient evidence to indicate that the null hypothesis is not true, we reject it in favor of the alternative hypothesis.

Importantly, the null and alternative hypotheses are stated in terms of population parameters, and we use sample statistics to test these hypotheses.

2.2 Two-Sided vs. One-Sided Hypotheses

b compare and contrast one-tailed and two-tailed tests of hypotheses

Suppose we want to test whether the population mean return is equal to 6%. We would state the hypotheses as

H_0: $\mu = 6$

and the alternative as

H_a: $\mu \neq 6$.

What we just created was a **two-sided hypothesis test**. We are testing whether the mean is equal to 6%; it could be greater than or less than that because we are simply asking whether the mean is different from 6%. If we find that the sample mean is far enough away from the hypothesized value, considering the risk of drawing a sample that is not representative of the population, then we would reject the null in favor of the alternative hypothesis.

What if we wanted to test whether the mean is greater than 6%. This presents us with a **one-sided hypothesis test**, and we specify the hypotheses as follows:

H_0: $\mu \leq 6$.

H_a: $\mu > 6$.

If we find that the sample mean is greater than the hypothesized value of 6 by a sufficient margin, then we would reject the null hypothesis. Why is the null hypothesis stated with a "≤" sign? First, if the sample mean is less than or equal to 6%, this would not support the alternative. Second, the null and alternative hypotheses must be mutually exclusive and collectively exhaustive; in other words, all possible values are contained in either the null or the alternative hypothesis.

Despite the different ways to formulate hypotheses, we always conduct a test of the null hypothesis at the point of equality; for example, $\mu = \mu_0$. Whether the null is H_0: $\mu = \mu_0$, H_0: $\mu \leq \mu_0$, or H_0: $\mu \geq \mu_0$, we actually test $\mu = \mu_0$. The reasoning is straightforward: Suppose the hypothesized value of the mean is 6. Consider H_0: $\mu \leq 6$, with a "greater than" alternative hypothesis, H_a: $\mu > 6$. If we have enough evidence to reject H_0: $\mu = 6$ in favor of H_a: $\mu > 6$, we definitely also have enough evidence to reject the hypothesis that the parameter μ is some smaller value, such as 4.5 or 5.

Using hypotheses regarding the population mean as an example, the three possible formulations of hypotheses are as follows:

Two-sided alternative: H_0: $\mu = \mu_0$ versus H_a: $\mu \neq \mu_0$

One-sided alternative (right side): $H_{0:}$ $\mu \leq \mu_0$ versus $H_{a:}$ $\mu > \mu_0$

One-sided alternative (left side): $H_{0:}$ $\mu \geq \mu_0$ versus $H_{a:}$ $\mu < \mu_0$

The reference to the side (right or left) refers to where we reject the null in the probability distribution. For example, if the alternative is H_a: $\mu > 6$, this means that we will reject the null hypothesis if the sample mean is sufficiently higher than (or on the right side of the distribution of) the hypothesized value.

Importantly, the calculation to test the null hypothesis is the same for all three formulations. What is different for the three formulations is how the calculation is evaluated to decide whether to reject the null.

2.3 Selecting the Appropriate Hypotheses

How do we choose the null and alternative hypotheses? The null is what we are hoping to reject. The most common alternative is the "not equal to" hypothesis. However, economic or financial theory may suggest a one-sided alternative hypothesis. For example, if the population parameter is the mean risk premium, financial theory may argue that this risk premium is positive. Following the principle of stating the alternative as the "hoped for" condition and using μ_{rp} for the population mean risk premium, we formulate the following hypotheses:

H_0: $\mu_{rp} \leq 0$ versus H_a: $\mu_{rp} > 0$

Note that the sign in the alternative hypotheses reflects the belief of the researcher more strongly than a two-sided alternative hypothesis. However, the researcher may sometimes select a two-sided alternative hypothesis to emphasize an attitude of neutrality when a one-sided alternative hypothesis is also reasonable. Typically, the easiest way to formulate the hypotheses is to specify the alternative hypothesis first and then specify the null.

EXAMPLE 1

Specifying the Hypotheses

An analyst suspects that in the most recent year excess returns on stocks have fallen below 5%. She wants to study whether the excess returns are less than 5%. Designating the population mean as μ, which hypotheses are most appropriate for her analysis?

A H_0: $\mu = 5$ versus H_a: $\mu \neq 5$

B H_0: $\mu \geq 5$ versus H_a: $\mu < 5$

C H_0: $\mu \leq 5$ versus H_a: $\mu > 5$

Solution

B is correct. The null hypothesis is what she wants to reject in favor of the alternative, which is that population mean excess return is less than 5%. This is a one-sided (left-side) alternative hypothesis.

IDENTIFY THE APPROPRIATE TEST STATISTIC

3

c explain a test statistic, Type I and Type II errors, a significance level, how significance levels are used in hypothesis testing, and the power of a test

A test statistic is a value calculated on the basis of a sample that, when used in conjunction with a decision rule, is the basis for deciding whether to reject the null hypothesis.

3.1 Test Statistics

The focal point of our statistical decision is the value of the test statistic. The test statistic that we use depends on what we are testing. As an example, let us examine the test of a population mean risk premium. Consider the sample mean, $\overline{X}$, calculated from a sample of returns drawn from the population. If the population standard deviation is known, the standard error of the distribution of sample means, $\sigma_{\overline{X}}$, is the ratio of the population standard deviation to the square root of the sample size:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}. \tag{1}$$

The test statistic for the test of the mean when the population variance is known is a z-distributed (that is, normally distributed) test statistic:

$$z = \frac{\overline{X}_{rp} - \mu_0}{\sigma/\sqrt{n}}. \tag{2}$$

If the hypothesized value of the mean population risk premium is 6 (that is, $\mu_0 = 6$), we calculate this as $z = \frac{\overline{X}_{rp} - 6}{\sigma/\sqrt{n}}$. If, however, the hypothesized value of the mean risk premium is zero (that is, $\mu_0 = 0$), we can simplify this test statistic as

$$z = \frac{\overline{X}_{rp}}{\sigma/\sqrt{n}}.$$

Notably, the key to hypothesis testing is identifying the appropriate test statistic for the hypotheses and the underlying distribution of the population.

3.2 Identifying the Distribution of the Test Statistic

Following the identification of the appropriate test statistic, we must be concerned with the distribution of the test statistic. We show examples of the test statistics and their corresponding distributions in Exhibit 4.

Exhibit 4 Test Statistics and Their Distributions

What We Want to Test	Test Statistic	Probability Distribution of the Statistic	Degrees of Freedom
Test of a single mean	$t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$	*t*-Distributed	$n - 1$
Test of the difference in means	$t = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}}$	*t*-Distributed	$n_1 + n_2 - 2$
Test of the mean of differences	$t = \dfrac{\bar{d} - \mu_{d0}}{s_{\bar{d}}}$	*t*-Distributed	$n - 1$
Test of a single variance	$\chi^2 = \dfrac{s^2(n-1)}{\sigma_0^2}$	Chi-square distributed	$n - 1$
Test of the difference in variances	$F = \dfrac{s_1^2}{s_2^2}$	*F*-distributed	$n_1 - 1, n_2 - 1$
Test of a correlation	$t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$	*t*-Distributed	$n - 2$
Test of independence (categorical data)	$\chi^2 = \sum_{i=1}^{m} \dfrac{(O_{ij} - E_{ij})^2}{E_{ij}}$	Chi-square distributed	$(r - 1)(c - 1)$

Note: μ_0, μ_{d0}, and σ_0^2 denote hypothesized values of the mean, mean difference, and variance, respectively. The $\bar{x}$, $\bar{d}$, s^2, s, and r denote for a sample the mean, mean of the differences, variance, standard deviation, and correlation, respectively, with subscripts indicating the sample, if appropriate. The sample size is indicated as n, and the subscript indicates the sample, if appropriate. O_{ij} and E_{ij} are observed and expected frequencies, respectively, with r indicating the number of rows and c indicating the number of columns in the contingency table.

4

SPECIFY THE LEVEL OF SIGNIFICANCE

The level of significance reflects how much sample evidence we require to reject the null hypothesis. The required standard of proof can change according to the nature of the hypotheses and the seriousness of the consequences of making a mistake. There are four possible outcomes when we test a null hypothesis, as shown in Exhibit 5. A **Type I error** is a false positive (reject when the null is true), whereas a **Type II error** is a false negative (fail to reject when the null is false).

Exhibit 5 Correct and Incorrect Decisions in Hypothesis Testing

	True Situation	
Decision	**H_0 True**	**H_0 False**
Fail to reject H_0	Correct decision: Do not reject a true null hypothesis.	Type II error: Fail to reject a false null hypothesis. False negative
Reject H_0	Type I error: Reject a true null hypothesis. False positive	Correct decision: Reject a false null hypothesis.

When we make a decision in a hypothesis test, we run the risk of making either a Type I or a Type II error. As you can see in Exhibit 5, these errors are mutually exclusive: If we mistakenly reject the true null, we can only be making a Type I error; if we mistakenly fail to reject the false null, we can only be making a Type II error.

Consider a test of a hypothesis of whether the mean return of a population is equal to 6%. How far away from 6% could a sample mean be before we believe it to be different from 6%, the hypothesized population mean? We are going to tolerate sample means that are close to 6%, but we begin doubting that the population mean is equal to 6% when we calculate a sample mean that is much different from 0.06. How do we determine "much different"? We do this by setting a risk tolerance for a Type I error and determining the critical value or values at which we believe that the sample mean is much different from the population mean. These critical values depend on (1) the alternative hypothesis, whether one sided or two sided, and (2) the probability distribution of the test statistic, which, in turn, depends on the sample size and the level of risk tolerance in making a Type I error.

The probability of a Type I error in testing a hypothesis is denoted by the lowercase Greek letter alpha, α. This probability is also known as the **level of significance** of the test, and its complement, $(1 - \alpha)$, is the **confidence level**. For example, a level of significance of 5% for a test means that there is a 5% probability of rejecting a true null hypothesis and corresponds to the 95% confidence level.

Controlling the probabilities of the two types of errors involves a trade-off. All else equal, if we decrease the probability of a Type I error by specifying a smaller significance level (say, 1% rather than 5%), we increase the probability of making a Type II error because we will reject the null less frequently, including when it is false. Both Type I and Type II errors are risks of being wrong. Whether to accept more of one type versus the other depends on the consequences of the errors, such as costs. The only way to reduce the probabilities of both types of errors simultaneously is to increase the sample size, n.

Quantifying the trade-off between the two types of errors in practice is challenging because the probability of a Type II error is itself difficult to quantify because there may be many different possible false hypotheses. Because of this, we specify only α, the probability of a Type I error, when we conduct a hypothesis test.

Whereas the significance level of a test is the probability of incorrectly rejecting the true null, the **power of a test** is the probability of *correctly* rejecting the null—that is, the probability of rejecting the null when it is false. The power of a test is, in fact, the complement of the Type II error. The probability of a Type II error is often denoted by the lowercase Greek letter beta, β. We can classify the different probabilities in Exhibit 6 to reflect the notation that is often used.

Exhibit 6 Probabilities Associated with Hypothesis Testing Decisions

	True Situation	
Decision	**H_0 True**	**H_0 False**
Fail to reject H_0	Confidence level (1 − α)	β
Reject H_0	Level of significance α	Power of the test (1 − β)

The standard approach to hypothesis testing involves choosing the test statistic with the most power and then specifying a level of significance. It is more appropriate to specify this significance level prior to calculating the test statistic because if we specify it after calculating the test statistic, we may be influenced by the result of the calculation. The researcher is free to specify the probability of a Type I error, but the most common are 10%, 5%, and 1%.

EXAMPLE 2

Significance Level

If a researcher selects a 5% level of significance for a hypothesis test, the confidence level is:

A 2.5%.

B 5%.

C 95%.

Solution

C is correct. The 5% level of significance (i.e., probability of a Type I error) corresponds to 1 − 0.05 = 0.95, or a 95% confidence level (i.e., probability of not rejecting a true null hypothesis). The level of significance is the complement to the confidence level; in other words, they sum to 1.00, or 100%.

5 STATE THE DECISION RULE

d explain a decision rule and the relation between confidence intervals and hypothesis tests, and determine whether a statistically significant result is also economically meaningful

The fourth step in hypothesis testing is stating the decision rule. Before any sample is drawn and before a test statistic is calculated, we need to set up a decision rule: When do we reject the null hypothesis, and when do we not? The action we take is based on comparing the calculated test statistic with a specified value or values, which we refer to as **critical values**. The critical value or values we choose are based on the level of significance and the probability distribution associated with the test statistic. If we find that the calculated value of the test statistic is more extreme than the critical value or values, then we reject the null hypothesis; we say the result is **statistically significant**. Otherwise, we fail to reject the null hypothesis; there is not sufficient evidence to reject the null hypothesis.

5.1 Determining Critical Values

For a two-tailed test, we indicate two critical values, splitting the level of significance, α, equally between the left and right tails of the distribution. Using a z-distributed (standard normal) test statistic, for example, we would designate these critical values as $\pm z_{\alpha/2}$. For a one-tailed test, we indicate a single rejection point using the symbol for the test statistic with a subscript indicating the specified probability of a Type I error— for example, z_{α}.

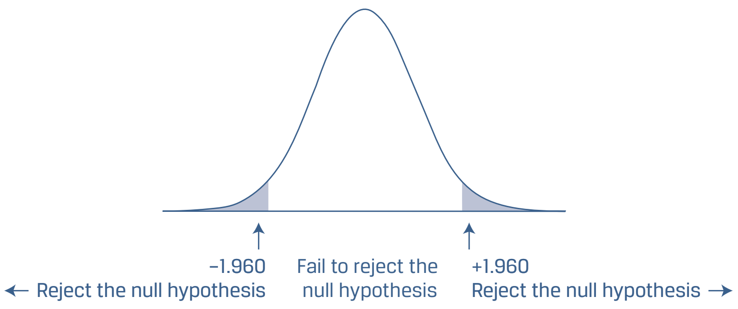
As we noted in our discussion of Exhibit 2, it is possible to draw a sample that has a mean different from the true population mean. In fact, it is likely that a given sample mean is different from the population mean because of sampling error. The issue becomes whether a given sample mean is far enough away from what is hypothesized to be the population mean that there is doubt about whether the hypothesized population mean is true. Therefore, we need to decide how far is too far for a sample mean in comparison to a population mean. That is where the critical values come into the picture.

Suppose we are using a z-test and have chosen a 5% level of significance. In Exhibit 7, we illustrate two tests at the 5% significance level using a z-statistic: A two-sided alternative hypothesis test in Panel A and a one-sided alternative hypothesis test in Panel B, with the white area under the curve indicating the confidence level and the shaded areas indicating the significance level. In Panel A, if the null hypothesis that $\mu = \mu_0$ is true, the test statistic has a 2.5% chance of falling in the left rejection region and a 2.5% chance of falling in the right rejection region. Any calculated value of the test statistic that falls in either of these two regions causes us to reject the null hypothesis at the 5% significance level.

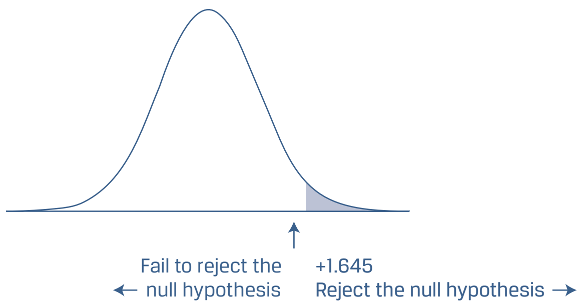
We determine the cut-off values for the reject and fail-to-reject regions on the basis of the distribution of the test statistic. For a test statistic that is normally distributed, we determine these cut-off points on the basis of the area under the normal curve; with a Type I error (i.e., level of significance, α) of 5% and a two-sided test, there is 2.5% of the area on either side of the distribution. This results in rejection points of −1.960 and +1.960, dividing the distribution between the rejection and fail-to-reject regions. Similarly, if we have a one-sided test involving a normal distribution, we would have 5% area under the curve with a demarcation of 1.645 for a right-side test, as we show in Exhibit 7 (or −1.645 for a left-side test).

Exhibit 7 Decision Criteria Using a 5% Level of Significance

A. H_0: $\mu = \mu_0$ versus H_a: $\mu \neq \mu_0$



B. H_0: $\mu \leq \mu_0$ versus H_a: $\mu > \mu_0$



Determining the cut-off points using programming

The programs in Microsoft Excel, Python, and R differ slightly, depending on whether the user specifies the area to the right or the left of the cut-off in the code:

Cut-off for . . .	Excel	Python	R
Right tail, 2.5%	NORM.S.INV(0.975)	norm.ppf(.975)	qnorm(.025,lower.tail=FALSE)
Left tail, 2.5%	NORM.S.INV(0.025)	norm.ppf(.025)	qnorm(.025,lower.tail=TRUE)
Right tail, 5%	NORM.S.INV(0.95)	norm.ppf(.95)	qnorm(.05,lower.tail=FALSE)
Left tail, 5%	NORM.S.INV(0.05)	norm.ppf(.05)	qnorm(.05,lower.tail=TRUE)

For Python, install scipy.stats and import: from scipy.stats import norm.

5.2 Decision Rules and Confidence Intervals

Exhibit 7 provides an opportunity to highlight the relationship between confidence intervals and hypothesis tests. A 95% confidence interval for the population mean, μ, based on sample mean, $\overline{X}$, is given by:

$$\left\{\overline{X} - 1.96\frac{\sigma}{\sqrt{n}}, \overline{X} + 1.96\frac{\sigma}{\sqrt{n}}\right\}, \tag{3}$$

or, more compactly, $\overline{X} \pm 1.96\frac{\sigma}{\sqrt{n}}$.

Now consider the conditions for rejecting the null hypothesis:

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -1.96 \text{ or } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > 1.96, \text{ where } z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

As you can see by comparing these conditions with the confidence interval, we can address the question of whether $\bar{X}$ is far enough away from μ_0 by either comparing the calculated test statistic with the critical values or comparing the hypothesized population parameter ($\mu = \mu_0$) with the bounds of the confidence interval, as we show in Exhibit 8. Thus, a significance level in a two-sided hypothesis test can be interpreted in the same way as a $(1 - \alpha)$ confidence interval.

Exhibit 8 Making a Decision Based on Critical Values and Confidence Intervals for a Two-Sided Alternative Hypothesis

Method	Procedure	Decision
1	Compare the calculated test statistic with the critical values.	If the calculated test statistic is less than the lower critical value or greater than the upper critical value, reject the null hypothesis.
2	Compare the calculated test statistic with the bounds of the confidence interval.	If the hypothesized value of the population parameter under the null is outside the corresponding confidence interval, the null hypothesis is rejected.

5.3 Collect the Data and Calculate the Test Statistic

The fifth step in hypothesis testing is collecting the data and calculating the test statistic. The quality of our conclusions depends on not only the appropriateness of the statistical model but also the quality of the data we use in conducting the test. First, we need to ensure that the sampling procedure does not include biases, such as sample selection or time bias. Second, we need to cleanse the data, checking inaccuracies and other measurement errors in the data. Once assured that the sample is unbiased and accurate, the sample information is used to calculate the appropriate test statistic.

EXAMPLE 3

Using a Confidence Interval in Hypothesis Testing

Consider the hypotheses H_0: $\mu = 3$ versus H_a: $\mu \neq 3$. If the confidence interval based on sample information has a lower bound of 2.75 and an upper bound of 4.25, the *most* appropriate decision is:

A reject the null hypothesis.

B accept the null hypothesis.

C fail to reject the null hypothesis.

Solution

C is correct. Since the hypothesized population mean ($\mu = 3$) is within the bounds of the confidence interval (2.75, 4.25), the correct decision is to fail to reject the null hypothesis. It is only when the hypothesized value is outside these bounds that the null hypothesis is rejected. Note that the null hypothesis is never accepted; either the null is rejected on the basis of the evidence or there is a failure to reject the null hypothesis.

6 MAKE A DECISION

d explain a decision rule and the relation between confidence intervals and hypothesis tests, and determine whether a statistically significant result is also economically meaningful

6.1 Make a Statistical Decision

The sixth step in hypothesis testing is making the decision. Consider a test of the mean risk premium, comparing the population mean with zero. If the calculated z-statistic is 2.5 and with a two-sided alternative hypothesis and a 5% level of significance, we reject the null hypothesis because 2.5 is outside the bounds of ±1.96. This is a statistical decision: The evidence indicates that the mean risk premium is not equal to zero.

6.2 Make an Economic Decision

Another part of the decision making is making the economic or investment decision. The economic or investment decision takes into consideration not only the statistical decision but also all pertinent economic issues. If, for example, we reject the null that the risk premium is zero in favor of the alternative hypothesis that the risk premium is greater than zero, we have found evidence that the US risk premium is different from zero. The question then becomes whether this risk premium is economically meaningful. On the basis of these considerations, an investor might decide to commit funds to US equities. A range of non-statistical considerations, such as the investor's tolerance for risk and financial position, might also enter the decision-making process.

6.3 Statistically Significant but Not Economically Significant?

We frequently find that slight differences between a variable and its hypothesized value are statistically significant but not economically meaningful. For example, we may be testing an investment strategy and reject a null hypothesis that the mean return to the strategy is zero based on a large sample. In the case of a test of the mean, the smaller the standard error of the mean, the larger the value of the test statistic and the greater the chance the null will be rejected, all else equal. The standard error decreases as the sample size, n, increases, so that for very large samples, we can reject the null for small departures from it. We may find that although a strategy provides a statistically significant positive mean return, the results may not be economically significant when we account for transaction costs, taxes, and risk. Even if we conclude that a strategy's results are economically meaningful, we should explore the logic of why the strategy might work in the future before implementing it. Such considerations cannot be incorporated into a hypothesis test.

EXAMPLE 4

Decisions and Significance

An analyst is testing whether there are positive risk-adjusted returns to a trading strategy. He collects a sample and tests the hypotheses of H_0: $\mu \leq 0\%$ versus H_a: $\mu > 0\%$, where μ is the population mean risk-adjusted return. The mean risk-adjusted return for the sample is 0.7%. The calculated t-statistic is 2.428, and the critical t-value is 2.345. He estimates that the transaction costs are 0.3%. The results are *most likely*:

A statistically and economically significant.

B statistically significant but not economically significant.

C economically significant but not statistically significant.

Solution

A is correct. The results indicate that the mean risk-adjusted return is greater than 0% because the calculated test statistic of 2.428 is greater than the critical value of 2.245. The results are also economically significant because the risk-adjusted return exceeds the transaction cost associated with this strategy by 0.4% (= 0.7 − 0.3).

THE ROLE OF *P*-VALUES

7

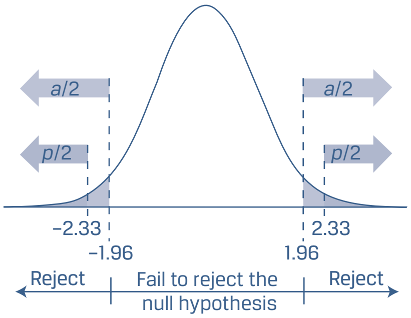
e explain and interpret the *p*-value as it relates to hypothesis testing

Analysts, researchers, and statistical software often report the *p*-value associated with hypothesis tests. The ***p*-value** is the area in the probability distribution outside the calculated test statistic; for a two-sided test, this is the area outside ± the calculated test statistic, but for a one-sided test, this is the area outside the calculated test statistic on the appropriate side of the probability distribution. We illustrated in Exhibit 7 the rejection region, which corresponds to the probability of a Type I error. However, the *p*-value is the area under the curve (so, the probability) associated with the calculated test statistic. Stated another way, the *p*-value is the smallest level of significance at which the null hypothesis can be rejected.

Consider the calculated z-statistic of 2.33 in a two-sided test: The *p*-value is the area in the z-distribution that lies outside ±2.33. Calculation of this area requires a bit of calculus, but fortunately statistical programs and other software calculate the *p*-value for us. For the value of the test statistic of 2.33, the *p*-value is approximately 0.02, or 2%. Using Excel, we can get the precise value of 0.019806 [(1-NORM.S.DIST(2.33,TRUE))*2]. We can reject the null hypothesis because we were willing to tolerate up to 5% outside the calculated value. The smaller the *p*-value, the stronger the evidence against the null hypothesis and in favor of the alternative hypothesis; if the *p*-value is less than the level of significance, we reject the null hypothesis.

We illustrate the comparison of the level of significance and the *p*-value in Exhibit 9. The fail-to-reject region is determined by the critical values of ±1.96, as we saw in Exhibit 7. There is 5% of the area under the distribution in the rejection regions—2.5% on the left side, 2.5% on the right. But now we introduce the area outside the calculated test statistic. For the calculated z-statistic of 2.33, there is 0.01, or 1%, of the area under the normal distribution above 2.33 and 1% of the area below −2.33 (or, in other words, 98% of the area between ±2.33). Since we are willing to tolerate a 5% Type I error, we reject the null hypothesis in the case of a calculated test statistic of 2.33 because there is a *p*-value of 2%; there is 2% of the distribution *outside* the calculated test statistic.

Exhibit 9 Comparison of the Level of Significance and the *p*-Value



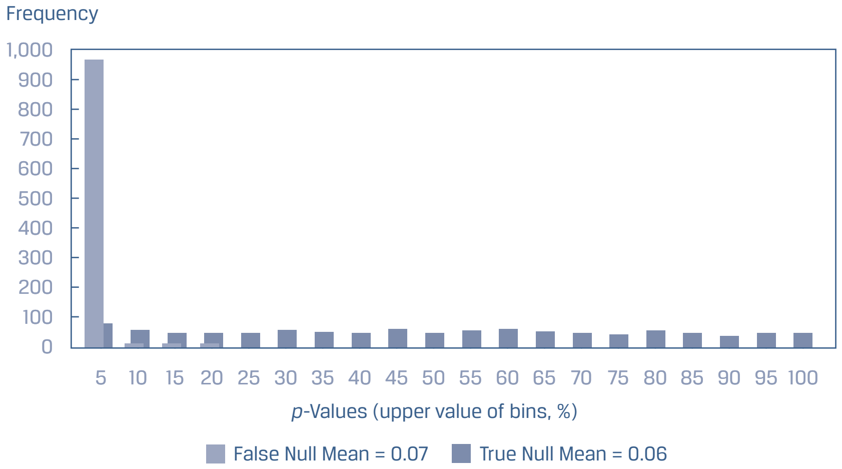
What if we are testing a one-sided alternative hypothesis? We focus solely on the area outside the calculated value on the side indicated by the alternative hypothesis. For example, if we are testing an alternative hypothesis that the population mean risk premium is greater than zero, calculate a *z*-statistic of 2.5, and have a level of significance of 5%, the *p*-value is the area in the probability distribution that is greater than 2.5. This area is 0.00621, or 0.621%. Since this is less than what we tolerate if the α is 5%, we reject the null hypothesis.

Consider a population of 1,000 assets that has a mean return of 6% and a standard deviation of 2%. Suppose we draw a sample of 50 returns and test whether the mean is equal to 6%, calculating the *p*-value for the calculated *z*-statistic. Then, suppose we repeat this process, draw 1,000 different samples, and, therefore, get 1,000 different calculated *z*-statistics and 1,000 different *p*-values. If we use a 5% level of significance, we should expect to reject the true null 5% of the time; if we use a 10% level of significance, we should expect to reject the true null 10% of the time.

Now suppose that with this same population, whose mean return is 6%, we test the hypothesis that the population mean is 7%, with the same standard deviation. As before, we draw 1,000 different samples and calculate 1,000 different *p*-values. If we use a 5% level of significance and a two-sided alternative hypothesis, we should expect to reject this false null hypothesis on the basis of the power of the test.

Putting this together, consider the histograms of *p*-values for the two different tests in Exhibit 10. The first bin is the *p*-values of 5% or less. What we see is that with the true null hypothesis of 6%, we reject the null approximately 5% of the time. For the false null hypothesis, that the mean is equal to 7%, we reject the null approximately 0.973, or 97.3%, of the time.

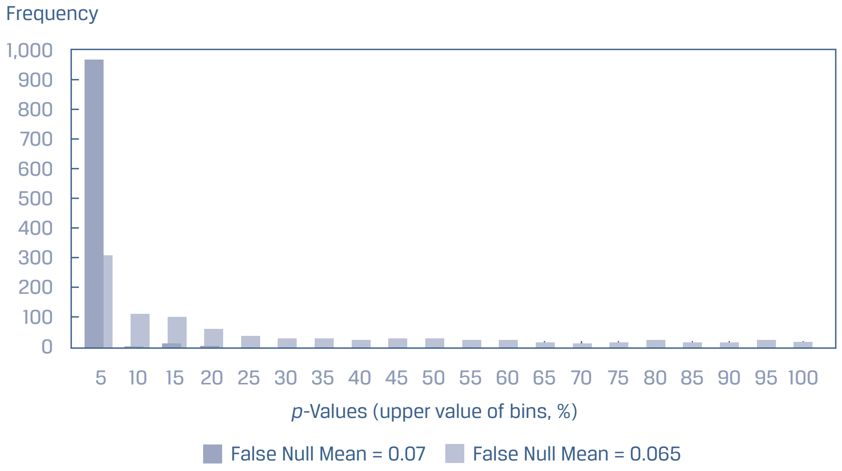
Exhibit 10 Distribution of *p*-values for 1,000 Different Samples of Size 50 Drawn from a Population with a Mean of 6%



The *p*-values for the true null hypothesis are generally uniformly distributed between 0% and 100% because under the null hypothesis, there is a 5% chance of the *p*-values being less than 5%, a 10% chance being less than 10%, and so on. Why is it not completely uniform? Because we took 1,000 samples of 50; taking more samples or larger samples would result in a more uniform distribution of *p*-values. When looking at the *p*-values for the false null hypothesis in Exhibit 10, we see that this is not a uniform distribution; rather, there is a peak around 0% and very little elsewhere. You can see the difference in the *p*-values for two false hypothesized means of 6.5% and 7% in Exhibit 11. It shows that the further the false hypothesis is away from the truth (i.e., mean of 6%), the greater the power of the test and the better the ability to detect the false hypothesis.

Software, such as Excel, Python, and R, is available for calculating *p*-values for most distributions.

Exhibit 11 Comparison of the Distribution of *p*-Values for the False Null Hypotheses H_0: μ = 7% and H_0: μ = 6.5%



EXAMPLE 5

Making a Decision Using *p*-Values

An analyst is testing the hypotheses H_0: $\sigma^2 = 0.01$ versus H_a: $\sigma^2 \neq 0.01$. Using software, she determines that the *p*-value for the test statistic is 0.03, or 3%. Which of the following statements are correct?

A Reject the null hypothesis at both the 1% and 5% levels of significance.

B Reject the null hypothesis at the 5% level but not at the 1% level of significance.

C Fail to reject the null hypothesis at both the 1% and 5% levels of significance.

Solution

B is correct. Rejection of the null hypothesis requires that the *p*-value be less than the level of significance. On the basis of this requirement, the null is rejected at the 5% level of significance but not at the 1% level of significance.

8 MULTIPLE TESTS AND INTERPRETING SIGNIFICANCE

f describe how to interpret the significance of a test in the context of multiple tests

A Type I error is the risk of rejection of a true null hypothesis. Another way of phrasing this is that it is a false positive result; that is, the null is rejected (the positive), yet the null is true (hence, a false positive). The expected portion of false positives is the **false discovery rate** (FDR). In the previous example of drawing 1,000 samples of 50 observations each, it is the case that there are samples in which we reject the true null hypothesis of a population mean of 6%. If we draw enough samples with a level of significance of 0.05, approximately 0.05 of the time you will reject the null hypothesis, even if the null is true. In other words, if you run 100 tests and use a 5% level of significance, you get five false positives, on average. This is referred to as the **multiple testing problem**.

The **false discovery approach** to testing requires adjusting the *p*-value when you have a series of tests. The idea of adjusting for the likelihood of significant results being false positives was first introduced by Benjamini and Hochberg (BH) in 1995. What they proposed is that the researcher rank the *p*-values from the various tests, from lowest to highest, and then make the following comparison, starting with the lowest *p*-value (with $k = 1$), $p(1)$:

$$p(1) \leq \alpha \frac{\text{Rank of } i}{\text{Number of tests}}.$$

This comparison is repeated, such that k is determined by the highest ranked $p(k)$ for which this is a true statement. If, say, k is 4, then the first four tests (ranked on the basis of the lowest *p*-values) are said to be significant.

Suppose we test the hypothesis that the population mean is equal to 6% and repeat the sampling process by drawing 20 samples and calculating 20 test statistics; the six test statistics with the lowest *p*-values are shown in . Using a significance level of 5%, if we simply relied on each test and its *p*-value, then there are five tests in which we would reject the null. However, using the BH criteria, only one test is considered significant, as shown in Exhibit 12.

Exhibit 12 Applying the Benjamini and Hochberg Criteria

(1) $\overline{X}$	(2) Calculated *z*-Statistic	(3) *p*-Value	(4) Rank of *p*-Value (lowest to highest)	(5) $\alpha \frac{\text{Rank of } i}{\text{Number of tests}}$	(6) Is value in (3) less than or equal to value in (5)?
0.0664	3.1966	0.0014	1	0.0025	Yes
0.0645	2.2463	0.0247	2	0.0050	No
0.0642	2.0993	0.0358	3	0.0075	No
0.0642	2.0756	0.0379	4	0.0100	No
0.0641	2.0723	0.0382	5	0.0125	No
0.0637	1.8627	0.0625	6	0.0150	No

Note: Level of significance = 5%.

So, what is the conclusion from looking at *p*-values and the multiple testing problem?

- First, if we sample, test, and find a result that is not statistically significant, this result is not wrong; in fact, the null hypothesis may well be true.
- Second, if the power of the test is low or the sample size is small, we should be cautious because there is a good chance of a false positive.
- Third, when we perform a hypothesis test and determine the critical values, these values are based on the assumption that the test is run once. Running multiple tests on data risks data snooping, which may result in spurious results. Determine the dataset and perform the test, but do not keep performing tests repeatedly to search out statistically significant results, because you may, by chance, find them (i.e., false positives).
- Fourth, in very large samples, we will find that nearly every test is significant. The approach to use to address this issue is to draw different samples; if the results are similar, the results are more robust.

EXAMPLE 6

False Discovery and Multiple Tests

A researcher is examining the mean return on assets of publicly traded companies that constitute an index of 2,000 mid-cap stocks and is testing hypotheses concerning whether the mean is equal to 15%: H_0: $\mu_{ROA} = 15$ versus Ha: $\mu_{ROA} \neq 15$. She uses a 10% level of significance and collects a sample of 50 firms. She wants to examine the robustness of her analysis, so she repeats the collection and test of the return on assets 30 times. The results for the samples with the five lowest *p*-values are given in Exhibit 13.

Exhibit 13 Five Lowest *p*-Values of the 30 Samples Tested

Sample	Calculated *z*-Statistic	*p*-Value	Ranked *p*-Value
1	3.203	0.00136	1
5	3.115	0.00184	2
14	2.987	0.00282	3
25	2.143	0.03211	4
29	1.903	0.05704	5

1 Of the 30 samples tested, how many should the researcher expect, on average, to have *p*-values less than the level of significance?

2 What are the corrected *p*-values based on her selected level of significance, and what is the effect on her decision?

Solution to 1

Of the 30 samples tested, she should expect $30 \times 0.10 = 3$ to have significant results just by chance. Consider why she ended up with more than three. Three is based on large sample sizes and large numbers of samples. Using a limited sample size (i.e., 50) and number of samples (i.e., 30), there is a risk of a false discovery with repeated samples and tests.

Solution to 2

Applying the BH criteria, the researcher determines the adjusted p-values shown in Exhibit 14.

Exhibit 14 Adjusted *p*-Values for Five Lowest *p*-Values from 30 Samples Tested

Calculated *z*-Statistic	*p*-Value	Rank of *p*-Value	$\alpha \dfrac{\text{Rank of } i}{\text{Number of tests}}$	Is *p*-value less than or equal to adjusted *p*-value?
3.203	0.00136	1	0.00333	Yes
3.115	0.00184	2	0.00667	Yes
2.987	0.00282	3	0.01000	Yes
2.143	0.03211	4	0.01333	No
1.903	0.05704	5	0.01667	No

On the basis of the results in Exhibit 14, there are three samples with *p*-values less than their adjusted *p*-values. So, the number of significant sample results is the same as would be expected from chance, given the 10% level of significance. The researcher concludes that the results for the samples with Ranks 4 and 5 are false discoveries, and she has not uncovered any evidence from her testing that supports rejecting the null hypothesis.

TESTS CONCERNING A SINGLE MEAN

g identify the appropriate test statistic and interpret the results for a hypothesis test concerning the population mean of both large and small samples when the population is normally or approximately normally distributed and the variance is (1) known or (2) unknown

Hypothesis tests concerning the mean are among the most common in practice. The sampling distribution of the mean when the population standard deviation is unknown is *t*-distributed, and when the population standard deviation is known, it is normally distributed, or *z*-distributed. Since the population standard deviation is unknown in almost all cases, we will focus on the use of a *t*-distributed test statistic.

The *t*-distribution is a probability distribution defined by a single parameter known as degrees of freedom (df). Like the standard normal distribution, a *t*-distribution is symmetrical with a mean of zero, but it has a standard deviation greater than 1 and generally fatter tails. As the number of degrees of freedom increases with the sample size, the *t*-distribution approaches the standard normal distribution.

For hypothesis tests concerning the population mean of a normally distributed population with unknown variance, the theoretically correct test statistic is the *t*-statistic. What if a normal distribution does not describe the population? The *t*-statistic is robust to moderate departures from normality, except for outliers and strong skewness. When we have large samples, departures of the underlying distribution from the normal case are of increasingly less concern. The sample mean is approximately normally distributed in large samples according to the central limit theorem, whatever the distribution describing the population. A traditional rule of thumb is that the normal distribution is used in cases when the sample size is larger than 30, but the more precise testing uses the *t*-distribution. Moreover, with software that aids us in such testing, we do not need to resort to rules of thumb.

If the population sampled has unknown variance, then the test statistic for hypothesis tests concerning a single population mean, μ, is

$$t_{n-1} = \frac{\overline{X} - \mu_0}{s/\sqrt{n}}, \tag{4}$$

where

$\overline{X}$ = sample mean

μ_0 = hypothesized value of the population mean

s = sample standard deviation

n = sample size

$s_{\overline{X}} = s/\sqrt{n}$ = estimate of the sample mean standard error

This test statistic is *t*-distributed with $n - 1$ degrees of freedom, which we can write as t_{n-1}. For simplicity, we often drop the subscript $n - 1$ because each particular test statistic has specified degrees of freedom, as we presented in Exhibit 4.

Consider testing whether Investment One's returns (from Exhibit 1) are different from 6%; that is, we are testing H_0: $\mu = 6$ versus H_a: $\mu \neq 6$. If the calculated *t*-distributed test statistic is outside the bounds of the critical values based on the level of significance, we will reject the null hypothesis in favor of the alternative. If we have a sample size of 33, there are $n - 1 = 32$ degrees of freedom. At a 5% significance level (two tailed) and 32 degrees of freedom, the critical *t*-values are ±2.037. We can determine the critical values from software:

- Excel [T.INV(0.025,32) and T.INV(0.975,32)]

- R [qt(c(.025,.975),32)]
- Python [from scipy.stats import t and t.ppf(.025,32) and t.ppf(.975,32)]

Suppose that the sample mean return is 5.2990% and the sample standard deviation is 1.4284%. The calculated test statistic is

$$t = \frac{5.2990 - 6}{1.4284/\sqrt{33}} = -2.8192$$

with 32 degrees of freedom. The calculated value is less than −2.037, so we reject the null that the population mean is 6%, concluding that it is different from 6%.

EXAMPLE 7

Risk and Return Characteristics of an Equity Mutual Fund

Suppose you are analyzing Sendar Equity Fund, a midcap growth fund that has been in existence for 24 months. During this period, it has achieved a mean monthly return of 1.50%, with a sample standard deviation of monthly returns of 3.60%. Given its level of market risk and according to a pricing model, this mutual fund was expected to have earned a 1.10% mean monthly return during that time period. Assuming returns are normally distributed, are the actual results consistent with an underlying or population mean monthly return of 1.10%?

1 Test the hypothesis using a 5% level of significance.

2 Test the hypothesis using the 95% confidence interval.

Solution to 1

Step 1	State the hypotheses.	H_0: $\mu = 1.1\%$ versus H_a: $\mu \neq 1.1\%$
Step 2	Identify the appropriate test statistic.	$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ with 24 – 1 =23 degrees of freedom.
Step 3	Specify the level of significance.	$\alpha = 5\%$ (two tailed).
Step 4	State the decision rule.	Critical *t*-values = ±2.069. Reject the null if the calculated *t*-statistic is less than −2.069, and reject the null if the calculated *t*-statistic is greater than +2.069. **Excel** *Lower:* T.INV(0.025,23) *Upper:* T.INV(0.975,23) **R** qt(c(.025,.975),23) **Python** from scipy.stats import t *Lower:* t.ppf(.025,23) *Upper:* t.ppf(.975,23)
Step 5	Calculate the test statistic.	$t = \frac{1.5 - 1.1}{3.6/\sqrt{24}} = 0.54433$
Step 6	Make a decision.	Fail to reject the null hypothesis because the calculated *t*-statistic falls between the two critical values. There is not sufficient evidence to indicate that the population mean monthly return is different from 1.10%.

Solution to 2

The 95% confidence interval is $\overline{X}$ ± Critical value$\left(\frac{s}{\sqrt{n}}\right)$, so

$$\left\{1.5 - 2.069\left(3.6\big/\sqrt{24}\right),\ 1.5 + 2.069\left(3.6\big/\sqrt{24}\right)\right\}$$

$$\{1.5 - 1.5204,\ 1.5 + 1.5204\}$$

$$\{-0.0204,\ 3.0204\}$$

The hypothesized value of 1.1% is within the bounds of the 95% confidence interval, so we fail to reject the null hypothesis.

We stated previously that when population variance is not known, we use a *t*-test for tests concerning a single population mean. Given at least approximate normality, the *t*-distributed test statistic is always called for when we deal with small samples and do not know the population variance. For large samples, the central limit theorem states that the sample mean is approximately normally distributed, whatever the distribution of the population. The *t*-statistic is still appropriate, but an alternative test may be more useful when sample size is large.

For large samples, practitioners sometimes use a *z*-test in place of a *t*-test for tests concerning a mean. The justification for using the *z*-test in this context is twofold. First, in large samples, the sample mean should follow the normal distribution at least approximately, as we have already stated, fulfilling the normality assumption of the *z*-test. Second, the difference between the rejection points for the *t*-test and *z*-test becomes quite small when the sample size is large. Since the *t*-test is readily available as statistical program output and theoretically correct for unknown population variance, we present it as the test of choice.

In a very limited number of cases, we may know the population variance; in such cases, the *z*-statistic is theoretically correct. In this case, the appropriate test statistic is what we used earlier (Equation 2):

$$z = \frac{\overline{X} - \mu_0}{\sigma\big/\sqrt{n}}.$$

In cases of large samples, a researcher may use the *z*-statistic, substituting the sample standard deviation (s) for the population standard deviation (σ) in the formula. When we use a *z*-test, we usually refer to a rejection point in Exhibit 15.

Exhibit 15 Critical Values for Common Significance Levels for the Standard Normal Distribution

Level of Significance	Alternative	Reject the Null if . . . below the Critical Value	Reject the Null if . . . above the Critical Value
0.01	Two sided: H_0: $\mu = \mu_0$, H_a: $\mu \neq \mu_0$	−2.576	2.576
	One sided: H_0: $\mu \leq \mu_0$, H_a: $\mu > \mu_0$		2.326
	One sided: H_0: $\mu \geq \mu_0$, H_a: $\mu < \mu_0$	−2.326	

(continued)

Exhibit 15 (Continued)

Level of Significance	Alternative	Reject the Null if . . . below the Critical Value	Reject the Null if . . . above the Critical Value
0.05	Two sided: H_0: $\mu = \mu_0$, H_a: $\mu \neq \mu_0$	−1.960	1.960
	One sided: H_0: $\mu \leq \mu_0$, H_a: $\mu > \mu_0$		1.645
	One sided: H_0: $\mu \geq \mu_0$, H_a: $\mu < \mu_0$	−1.645	

EXAMPLE 8

Testing the Returns on the ACE High Yield Index

Suppose we want to test whether the daily return in the ACE High Yield Total Return Index is different from zero. Collecting a sample of 1,304 daily returns, we find a mean daily return of 0.0157%, with a standard deviation of 0.3157%.

1. Test whether the mean daily return is different from zero at the 5% level of significance.

2. Using the *z*-distributed test statistic as an approximation, test whether the mean daily return is different from zero at the 5% level of significance.

Solution to 1

Step 1 State the hypotheses. H_0: μ = 0% versus H_a: $\mu \neq$ 0%

Step 2 Identify the appropriate test statistic.

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

with 1,304 − 1 = 1,303 degrees of freedom.

Step 3 Specify the level of significance. α = 5%.

Step 4 State the decision rule. Critical *t*-values = ±1.962.

Reject the null if the calculated *t*-statistic is less than −1.962, and reject the null if the calculated *t*-statistic is greater than +1.962.

Excel

Lower: T.INV(0.025,1303)

Upper: T.INV(0.975,1303)

R qt(c(.025,.975),1303)

Python from scipy.stats import t

Lower: t.ppf(.025,1303)

Upper: t.ppf(.975,1303)

Step 5 Calculate the test statistic.

$$t = \frac{0.0157 - 0}{0.3157/\sqrt{1{,}304}} = 1.79582$$

Step 6 Make a decision. Fail to reject the null because the calculated *t*-statistic falls between the two critical values. There is not sufficient evidence to indicate that the mean daily return is different from 0%.

Solution to 2

Step 1 State the hypotheses. H_0: μ = 0% versus H_a: μ ≠ 0%

Step 2 Identify the appropriate test statistic.

$$z = \frac{\overline{X} - \mu_0}{s / \sqrt{n}}$$

with 1,304 − 1 = 1,303 degrees of freedom.

Step 3 Specify the level of significance. α = 5%.

Step 4 State the decision rule. Critical *t*-values = ±1.960.

Reject the null if the calculated *z*-statistic is less than −1.960, and reject the null if the calculated *z*-statistic is greater than +1.960.

Excel	*Lower:* NORM.S.INV(0.025)
	Upper: NORM.S.INV(0.975)
R	qnorm(.025,lower.tail=TRUE)
	qnorm(.975,lower.tail=FALSE)
Python	from scipy.stats import norm
	Lower: norm.ppf(.025,23)
	Upper: norm.ppf(.975,23)

$$z = \frac{0.0157 - 0}{0.3157 / \sqrt{1{,}304}} = 1.79582$$

Fail to reject the null because the calculated *z*-statistic falls between the two critical values. There is not sufficient evidence to indicate that the mean daily return is different from 0%.

Step 5 Calculate the test statistic.

Step 6 Make a decision.

TEST CONCERNING DIFFERENCES BETWEEN MEANS WITH INDEPENDENT SAMPLES

10

h identify the appropriate test statistic and interpret the results for a hypothesis test concerning the equality of the population means of two at least approximately normally distributed populations based on independent random samples with equal assumed variances

We often want to know whether a mean value—for example, a mean return—differs for two groups. Is an observed difference due to chance or to different underlying values for the mean? We test this by drawing a sample from each group. When it is reasonable to believe that the samples are from populations at least approximately normally distributed and that the samples are also independent of each other, we use the test of the differences in the means. We may assume that population variances are equal or unequal. However, our focus in discussing the test of the differences of means is using the assumption that the population variances are equal. In the calculation of the test statistic, we combine the observations from both samples to obtain a pooled estimate of the common population variance.

Let μ_1 and μ_2 represent, respectively, the population means of the first and second populations. Most often we want to test whether the population means are equal or whether one is larger than the other. Thus, we formulate the following sets of hypotheses:

Two sided:

$$H_0: \mu_1 - \mu_2 = 0 \text{ versus } H_a: \mu_1 - \mu_2 \neq 0,$$

or, equivalently,

$$H_0: \mu_1 = \mu_2 \text{ versus } H_a: \mu_1 \neq \mu_2$$

One sided (right side):

$$H_0: \mu_1 - \mu_2 \leq 0 \text{ versus } H_a: \mu_1 - \mu_2 > 0,$$

or, equivalently,

$$H_0: \mu_1 \leq \mu_2 \text{ versus } H_a: \mu_1 > \mu_2$$

One sided (left side):

$$H_0: \mu_1 - \mu_2 \geq 0 \text{ versus } H_a: \mu_1 - \mu_2 < 0,$$

or, equivalently,

$$H_0: \mu_1 \geq \mu_2 \text{ versus } H_a: \mu_1 < \mu_2$$

We can, however, formulate other hypotheses, where the difference is something other than zero, such as $H_0: \mu_1 - \mu_2 = 2$ versus $H_a: \mu_1 - \mu_2 \neq 2$. The procedure is the same.

When we can assume that the two populations are normally distributed and that the unknown population variances are equal, we use a t-distributed test statistic based on independent random samples:

$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}, \tag{5}$$

Where $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ is a pooled estimator of the common variance. As you can see, the pooled estimate is a weighted average of the two samples' variances, with the degrees of freedom for each sample as the weight. The number of degrees of freedom for this t-distributed test statistic is $n_1 + n_2 - 2$.

EXAMPLE 9

Returns on the ACE High Yield Index Compared for Two Periods

Continuing the example of the returns in the ACE High Yield Total Return Index, suppose we want to test whether these returns, shown in Exhibit 16, are different for two different time periods, Period 1 and Period 2.

Exhibit 16 Descriptive Statistics for ACE High Yield Total Return Index for Periods 1 and 2

	Period 1	Period 2
Mean	0.01775%	0.01134%
Standard deviation	0.31580%	0.38760%
Sample size	445 days	859 days

Note that these periods are of different lengths and the samples are independent; that is, there is no pairing of the days for the two periods.

Test whether there is a difference between the mean daily returns in Period 1 and in Period 2 using a 5% level of significance.

Step 1 State the hypotheses.

H_0: $\mu_{Period1} = \mu_{Period2}$ versus H_a: $\mu_{Period1} \neq \mu_{Period2}$

Step 2 Identify the appropriate test statistic.

$$t = \frac{(\bar{X}_{Period1} - \bar{X}_{Period2}) - (\mu_{Period1} - \mu_{Period2})}{\sqrt{\frac{s_p^2}{n_{period1}} + \frac{s_p^2}{n_{period2}}}},$$

where $s_p^2 = \dfrac{(n_{period1} - 1)s_{Period1}^2 + (n_{period2} - 1)s_{Period2}^2}{n_{period1} + n_{period2} - 2}$

with 445 + 859 – 2 =1,302 degrees of freedom.

Step 3 Specify the level of significance.

$\alpha = 5\%$.

Step 4 State the decision rule.

Critical *t*-values = ±1.962.

Reject the null if the calculated *t*-statistic is less than −1.962, and reject the null if the calculated *t*-statistic is greater than +1.962.

Excel

Lower: T.INV(0.025,1302)

Upper: T.INV(0.975,1302)

R qt(c(.025,.975),1302)

Python from scipy.stats import t

Lower: t.ppf(.025,1302)

Upper: t.ppf(.975,1302)

Step 5 Calculate the test statistic.

$$s_p^2 = \frac{(445 - 1)0.09973 + (859 - 1)0.15023}{445 + 859 - 2} = 0.1330$$

$$t = \frac{(0.01775 - 0.01134) - 0}{\sqrt{\frac{0.1330}{445} + \frac{0.1330}{859}}} = \frac{0.0064}{0.0213} = 0.3009.$$

Step 6 Make a decision.

Fail to reject the null because the calculated *t*-statistic falls within the bounds of the two critical values. We conclude that there is insufficient evidence to indicate that the returns are different for the two time periods.

11 TEST CONCERNING DIFFERENCES BETWEEN MEANS WITH DEPENDENT SAMPLES

i. identify the appropriate test statistic and interpret the results for a hypothesis test concerning the mean difference of two normally distributed populations

When we compare two independent samples, we use a t-distributed test statistic that uses the difference in the means and a pooled variance. An assumption for the validity of those tests is that the samples are independent—that is, unrelated to each other. When we want to conduct tests on two means based on samples that we believe are dependent, we use the **test of the mean of the differences**.

The t-test in this section is based on data arranged in paired observations, and the test itself is sometimes referred to as the **paired comparisons test**. Paired observations are observations that are dependent because they have something in common. For example, we may be concerned with the dividend policy of companies before and after a change in the tax law affecting the taxation of dividends. We then have pairs of observations for the same companies; these are dependent samples because we have pairs of the sample companies before and after the tax law change. We may test a hypothesis about the mean of the differences that we observe across companies. For example, we may be testing whether the mean returns earned by two investment strategies were equal over a study period. The observations here are dependent in the sense that there is one observation for each strategy in each month, and both observations depend on underlying market risk factors. What is being tested are the differences, and the paired comparisons test assumes that the differences are normally distributed. By calculating a standard error based on differences, we can use a t-distributed test statistic to account for correlation between the observations.

How is this test of paired differences different from the test of the differences in means in independent samples? The test of paired comparisons is more powerful than the test of the differences in the means because by using the common element (such as the same periods or companies), we eliminate the variation between the samples that could be caused by something other than what we are testing.

Suppose we have observations for the random variables X_A and X_B and that the samples are dependent. We arrange the observations in pairs. Let d_i denote the difference between two paired observations. We can use the notation $d_i = x_{Ai} - x_{Bi}$, where x_{Ai} and x_{Bi} are the ith pair of observations, $i = 1, 2, \ldots, n$, on the two variables. Let μ_d stand for the population mean difference. We can formulate the following hypotheses, where μ_{d0} is a hypothesized value for the population mean difference:

Two sided: H_0: $\mu_d = \mu_{d0}$ versus H_a: $\mu_d \neq \mu_{d0}$

One sided (right side): H_0: $\mu_d \leq \mu_{d0}$ versus H_a: $\mu_d > \mu_{d0}$

One sided (left side) H_0: $\mu_d \geq \mu_{d0}$ versus H_a: $\mu_d < \mu_{d0}$

In practice, the most commonly used value for μ_{d0} is zero.

We are concerned with the case of normally distributed populations with unknown population variances, and we use a t-distributed test statistic. We begin by calculating $\bar{d}$, the sample mean difference, or the mean of the differences, d_i:

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i, \tag{6}$$

where n is the number of pairs of observations. The sample standard deviation, s_d, is the standard deviation of the differences, and the standard error of the mean differences is $s_{\bar{d}} = \dfrac{s_d}{\sqrt{n}}$.

When we have data consisting of paired observations from samples generated by normally distributed populations with unknown variances, the *t*-distributed test statistic is

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}} \tag{7}$$

with *n* – 1 degrees of freedom, where *n* is the number of paired observations.

For example, suppose we want to see if there is a difference between the returns for Investments One and Two (from Exhibit 1), for which we have returns in each of 33 years. Using a 1% level of significance, the critical values for a two-sided hypothesis test are ±2.7385. Lining up these returns by the years and calculating the differences, we find a sample mean difference ($\bar{d}$) of 0.10353% and a standard deviation of these differences (s_d) of 2.35979%. Therefore, the calculated *t*-statistic for testing whether the mean of the differences is equal to zero is

$$t = \frac{0.10353 - 0}{2.35979/\sqrt{33}} = 0.25203$$

with 32 degrees of freedom. In this case, we fail to reject the null because the *t*-statistic falls within the bounds of the two critical values. We conclude that there is not sufficient evidence to indicate that the returns for Investment One and Investment Two are different.

Importantly, if we think of the differences between the two samples as a single sample, then the test of the mean of differences is identical to the test of a single sample mean.

EXAMPLE 10

Testing for the Mean of the Differences

In Exhibit 17, we report the quarterly returns for a three-year period for two actively managed portfolios specializing in precious metals. The two portfolios are similar in risk and had nearly identical expense ratios. A major investment services company rated Portfolio B more highly than Portfolio A. In investigating the portfolios' relative performance, suppose we want to test the hypothesis that the mean quarterly return on Portfolio A is equal to the mean quarterly return on Portfolio B during the three-year period. Since the two portfolios share essentially the same set of risk factors, their returns are not independent, so a paired comparisons test is appropriate. Use a 10% level of significance.

Exhibit 17 Quarterly Returns for Two Actively Managed Precious Metals Portfolios

Year	Quarter	Portfolio A (%)	Portfolio B (%)	Difference (Portfolio A – Portfolio B)
1	1	4.50	0.50	4.00
1	2	−4.10	−3.10	−1.00
1	3	−14.50	−16.80	2.30
1	4	−5.50	−6.78	1.28
2	1	12.00	−2.00	14.00
2	2	−7.97	−8.96	0.99
2	3	−14.01	−10.01	−4.00

(continued)

Exhibit 17 (Continued)

Year	Quarter	Portfolio A (%)	Portfolio B (%)	Difference (Portfolio A – Portfolio B)
2	4	4.11	−6.31	10.42
3	1	2.34	−5.00	7.34
3	2	26.36	12.77	13.59
3	3	10.72	9.23	1.49
3	4	3.60	1.20	2.40
Average		1.46	-2.94	4.40083
Standard deviation		11.18	7.82	5.47434

Using this sample information, we can summarize the test as follows:

Step 1 State the hypotheses.

H_0: $\mu_{d0} = 0$ versus H_a: $\mu_{d0} \neq 0$

Step 2 Identify the appropriate test statistic.

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}}$$

Step 3 Specify the level of significance.

10%

Step 3 State the decision rule.

With 12 − 1 = 11 degrees of freedom, the critical values are ±1.796.

We reject the null hypothesis if the calculated test statistic is below −1.796 or above +1.796.

Excel

Lower: T.INV(0.05,11)

Upper: T.INV(0.95,11)

R qt(c(.05,.95),11)

Python from scipy.stats import t

Lower: t.ppf(.05,11)

Upper: t.ppf(.95,11)

Step 4 Calculate the test statistic.

$$\bar{d} = 4.40083$$

$$s_{\bar{d}} = \frac{5.47434}{\sqrt{12}} = 1.58031$$

$$t = \frac{4.40083 - 0}{1.58031} = 2.78480$$

Step 5 Make a decision.

Reject the null hypothesis because the calculated *t*-statistic falls outside the bounds of the two critical values. There is sufficient evidence to indicate that the mean of the differences of returns is not zero.

The following example illustrates the application of this test to evaluate two competing investment strategies.

EXAMPLE 11

A Comparison of the Returns of Two Indexes

Suppose we want to compare the returns of the ACE High Yield Index with those of the ACE BBB Index. We collect data over 1,304 days for both indexes and calculate the means and standard deviations as shown in Exhibit 18.

Exhibit 18 Mean and Standard Deviations for the ACE High Yield Index and the ACE BBB Index

	ACE High Yield Index (%)	ACE BBB Index (%)	Difference (%)
Mean return	0.0157	0.0135	−0.0021
Standard deviation	0.3157	0.3645	0.3622

Using a 5% level of significance, determine whether the mean of the differences is different from zero.

Solution

Step 1	State the hypotheses.	H_0: $\mu_{d0} = 0$ versus H_a: $\mu_{d0} \neq 0$
Step 2	Identify the appropriate test statistic.	$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}}$
Step 3	Specify the level of significance.	5%
Step 4	State the decision rule.	With 1,304 − 1 = 1,303 degrees of freedom, the critical values are ±1.962. We reject the null hypothesis if the calculated *t*-statistic is less than −1.962 or greater than +1.962. **Excel** *Lower:* T.INV(0.025,1303) *Upper:* T.INV(0.975,1303) **R** qt(c(.025,.975),1303) **Python** from scipy.stats import t *Lower:* t.ppf(.025,1303) *Upper:* t.ppf(.975,1303)
Step 5	Calculate the test statistic.	$\bar{d} = -0.0021\%$ $s_{\bar{d}} = \frac{s_d}{\sqrt{n}} = \frac{0.3622}{\sqrt{1,304}} = 0.01003\%$ $t = \frac{-0.00210 - 0}{0.01003} = -0.20937$
Step 6	Make a decision.	Fail to reject the null hypothesis because the calculated *t*-statistic falls within the bounds of the two critical values. There is insufficient evidence to indicate that the mean of the differences of returns is different from zero.

12 TESTING CONCERNING TESTS OF VARIANCES (CHI-SQUARE TEST)

j identify the appropriate test statistic and interpret the results for a hypothesis test concerning (1) the variance of a normally distributed population and (2) the equality of the variances of two normally distributed populations based on two independent random samples

Often, we are interested in the volatility of returns or prices, and one approach to examining volatility is to evaluate variances. We examine two types of tests involving variance: tests concerning the value of a single population variance and tests concerning the difference between two population variances.

12.1 Tests of a Single Variance

Suppose there is a goal to keep the variance of a fund's returns below a specified target. In this case, we would want to compare the observed sample variance of the fund with the target. Performing a test of a population variance requires specifying the hypothesized value of the variance, σ_0^2 . We can formulate hypotheses concerning whether the variance is equal to a specific value or whether it is greater than or less than a hypothesized value:

Two-sided alternative: $H_0 : \sigma^2 = \sigma_0^2$ versus $H_a : \sigma^2 \neq \sigma_0^2$

One-sided alternative (right tail): $H_0 : \sigma^2 \leq \sigma_0^2$ versus $H_a : \sigma^2 > \sigma_0^2$

One-sided alternative (left tail): $H_0 : \sigma^2 \geq \sigma_0^2$ versus $H_a : \sigma^2 < \sigma_0^2$

In tests concerning the variance of a single normally distributed population, we make use of a chi-square test statistic, denoted χ^2. The chi-square distribution, unlike the normal distribution and *t*-distribution, is asymmetrical. Like the *t*-distribution, the chi-square distribution is a family of distributions, with a different distribution for each possible value of degrees of freedom, $n - 1$ (n is sample size). Unlike the *t*-distribution, the chi-square distribution is bounded below by zero; χ^2 does not take on negative values.

If we have n independent observations from a normally distributed population, the appropriate test statistic is

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \tag{8}$$

with $n - 1$ degrees of freedom. The sample variance (s^2) is in the numerator, and the hypothesized variance (σ_0^2) is in the denominator.

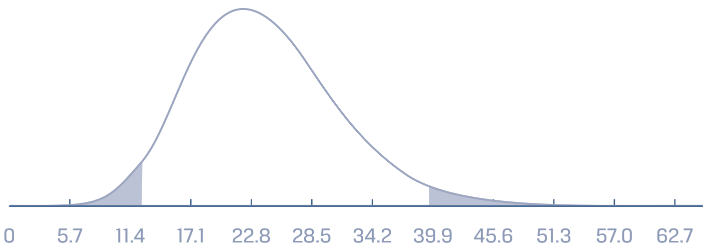
In contrast to the *t*-test, for example, the chi-square test is sensitive to violations of its assumptions. If the sample is not random or if it does not come from a normally distributed population, inferences based on a chi-square test are likely to be faulty.

Since the chi-square distribution is asymmetric and bounded below by zero, we no longer have the convenient ± for critical values as we have with the *z*- and the *t*-distributions, so we must either use a table of chi-square values or use software to generate the critical values. Consider a sample of 25 observations, so we have 24 degrees of freedom. We illustrate the rejection regions for the two- and one-sided tests at the 5% significance level in Exhibit 19.

Exhibit 19 Rejection Regions (Shaded) for the Chi-Square Distribution (df = 24) at 5% Significance

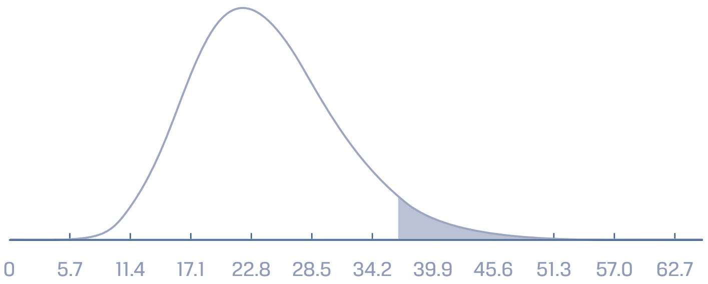
A. H_0: $\sigma^2 = \sigma_0^2$ versus H_a: $\sigma^2 \neq \sigma_0^2$

Critical values are 12.40115 and 39.36408



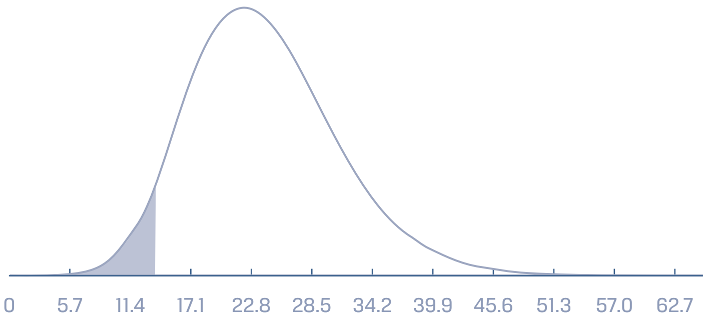
B. H_0: $\sigma^2 \leq \sigma_0^2$ versus H_a: $\sigma^2 > \sigma_0^2$

Critical value is 36.41503



C. H_0: $\sigma^2 \geq \sigma_0^2$ versus H_a: $\sigma^2 < \sigma_0^2$

Critical value is 13.84843



EXAMPLE 12

Risk and Return Characteristics of an Equity Mutual Fund

You continue with your analysis of Sendar Equity Fund, a midcap growth fund that has been in existence for only 24 months. During this period, Sendar Equity achieved a mean monthly return of 1.50% and a standard deviation of monthly returns of 3.60%.

1 Using a 5% level of significance, test whether the standard deviation of returns is different from 4%.

2 Using a 5% level of significance, test whether the standard deviation of returns is less than 4%.

Solution to 1

Step 1 State the hypotheses. H_0: $\sigma^2 = 16$ versus H_a: $\sigma^2 \neq 16$

Step 2 Identify the appropriate test statistic.

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Step 3 Specify the level of significance. 5%

Step 4 State the decision rule. With 24 – 1 = 23 degrees of freedom, the critical values are 11.68855 and 38.07563.

We reject the null hypothesis if the calculated χ^2 statistic is less than 11.68855 or greater than 38.07563.

Excel

Lower: CHISQ.INV(0.025,23)

Upper: CHISQ.INV(0.975,23)

R qchisq(c(.025,.975),23)

Python from scipy.stats import chi2

Lower: chi2.ppf(.025,23)

Upper: chi2.ppf(.975,23)

Step 5 Calculate the test statistic.

$$\chi^2 = \frac{(24-1)12.96}{16} = 18.63000$$

Step 6 Make a decision Fail to reject the null hypothesis because the calculated χ^2 statistic falls within the bounds of the two critical values. There is insufficient evidence to indicate that the variance is different from 16% (or, equivalently, that the standard deviation is different from 4%).

Solution to 2:

Step 1 State the hypotheses. H_0: $\sigma^2 \geq 16$ versus H_a: $\sigma^2 < 16$

Step 2 Identify the appropriate test statistic.

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Step 3 Specify the level of significance. 5%

Step 4 State the decision rule. With 24 – 1 = 23 degrees of freedom, the critical value is 13.09051.

We reject the null hypothesis if the calculated χ^2 statistic is less than 13.09051.

Excel CHISQ.INV(0.05,23)

R qchisq(.05,23)

Python from scipy.stats import chi2

chi2.ppf(.05,23)

Step 5 Calculate the test statistic.

$$\chi^2 = \frac{(24-1)12.96}{16} = 18.63000$$

Step 6 Make a decision.

Fail to reject the null hypothesis because the calculated χ^2 statistic is greater than the critical value. There is insufficient evidence to indicate that the variance is less than 16% (or, equivalently, that the standard deviation is less than 4%).

12.2 Test Concerning the Equality of Two Variances (*F*-Test)

j identify the appropriate test statistic and interpret the results for a hypothesis test concerning (1) the variance of a normally distributed population and (2) the equality of the variances of two normally distributed populations based on two independent random samples

There are many instances in which we want to compare the volatility of two samples, in which case we can test for the equality of two variances. Examples include comparisons of baskets of securities against indexes or benchmarks, as well as comparisons of volatility in different periods. Suppose we have a hypothesis about the relative values of the variances of two normally distributed populations with variances of σ_1^2 and σ_2^2, distinguishing the two populations as 1 or 2. We can formulate the hypotheses as two sided or one sided:

Two-sided alternative:

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ versus } H_a : \sigma_1^2 \neq \sigma_2^2$$

or, equivalently,

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \text{ versus } H_a : \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

One-sided alternative (right side):

$$H_0 : \sigma_1^2 \leq \sigma_2^2 \text{ versus } H_a : \sigma_1^2 > \sigma_2^2$$

or, equivalently,

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} \leq 1 \text{ versus } H_a : \frac{\sigma_1^2}{\sigma_2^2} > 1$$

One-sided alternative (left side):

$$H_0 : \sigma_1^2 \geq \sigma_2^2 \text{ versus } H_a : \sigma_1^2 < \sigma_2^2$$

or, equivalently,

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} \geq 1 \text{ versus } H_a : \frac{\sigma_1^2}{\sigma_2^2} < 1$$

Given independent random samples from these populations, tests related to these hypotheses are based on an *F*-test, which is the ratio of sample variances. Tests concerning the difference between the variances of two populations make use of the *F*-distribution. Like the chi-square distribution, the *F*-distribution is a family of asymmetrical distributions bounded from below by zero. Each *F*-distribution is defined by two values of degrees of freedom, which we refer to as the numerator and denominator degrees of freedom. The *F*-test, like the chi-square test, is not robust to violations of its assumptions.

Suppose we have two samples, the first with n_1 observations and a sample variance s_1^2 and the second with n_2 observations and a sample variance s_2^2. The samples are random, independent of each other, and generated by normally distributed populations. A test concerning differences between the variances of the two populations is based on the ratio of sample variances, as follows:

$$F = \frac{s_1^2}{s_2^2} \tag{9}$$

with $\text{df}_1 = (n_1 - 1)$ numerator degrees of freedom and $\text{df}_2 = (n_2 - 1)$ denominator degrees of freedom. Note that df_1 and df_2 are the divisors used in calculating s_1^2 and s_2^2, respectively.

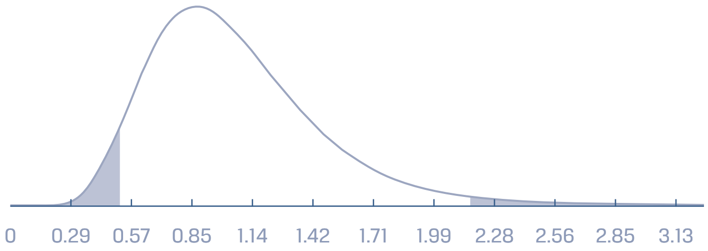
When we rely on tables to arrive at critical values, a convention is to use the larger of the two sample variances in the numerator in Equation 9; doing so reduces the number of *F*-tables needed. The key is to be consistent with how the alternative hypothesis is specified and the order of the sample sizes for the degrees of freedom.

Consider two samples, the first with 25 observations and the second with 40 observations. We show the rejection region and critical values in Exhibit 20 for two- and one-sided alternative hypotheses at the 5% significance level.

Exhibit 20 **Rejection Regions (Shaded) for the *F*-Distribution Based on Sample Sizes of 25 and 40 at 5% Significance**

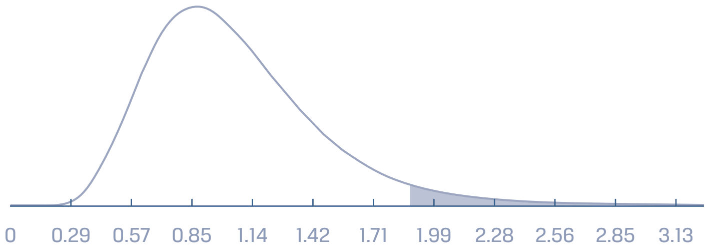
A. H_0: $\sigma_1^2 = \sigma_2^2$ versus H_a: $\sigma_1^2 \neq \sigma_2^2$

Critical values are 0.49587 and 2.15095



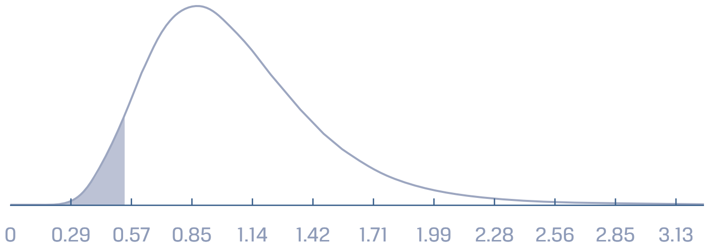
B. H_0: $\sigma_1^2 \leq \sigma_2^2$ versus H_a: $\sigma_1^2 > \sigma_2^2$

Critical value is 1.89566



C. H_0: $\sigma_1^2 \geq \sigma_2^2$ versus H_a: $\sigma_1^2 < \sigma_2^2$

Critical value is 0.55551



Consider Investments One and Two (from Exhibit 1), with standard deviations of returns of 1.4284 and 2.5914, respectively, calculated over the 33-year period. If we want to know whether the variance of Investment One is different from that of Investment Two, we use the *F*-distributed test statistic. With 32 and 32 degrees of freedom, the critical values are 0.49389 and 2.02475 at the 5% significance level. The calculated *F*-statistic is

$$F = \frac{2.5914^2}{1.4284^2} = 3.29131.$$

Therefore, we reject the null hypothesis that the variances of these two investments are the same because the calculated *F*-statistic is outside of the critical values. We can conclude that one investment is riskier than the other.

EXAMPLE 13

Volatility and Regulation

You are investigating whether the population variance of returns on a stock market index changed after a change in market regulation. The first 418 weeks occurred before the regulation change, and the second 418 weeks occurred after the regulation change. You gather the data in Exhibit 21 for 418 weeks of returns both before and after the change in regulation. You have specified a 5% level of significance.

Exhibit 21 Index Returns and Variances before and after the Market Regulation Change

	n	Mean Weekly Return (%)	Variance of Returns
Before regulation change	418	0.250	4.644
After regulation change	418	0.110	3.919

1. Test whether the variance of returns is different before the regulation change versus after the regulation change, using a 5% level of significance.
2. Test whether the variance of returns is greater before the regulation change versus after the regulation change, using a 5% level of significance.

Solution to 1

Step 1	State the hypotheses.	$H_0: \sigma^2_{Before} = \sigma^2_{After}$ versus $H_a: \sigma^2_{Before} \neq \sigma^2_{After}$
Step 2	Identify the appropriate test statistic.	$F = \dfrac{s^2_{Before}}{s^2_{After}}$
Step 3	Specify the level of significance.	5%
Step 4	State the decision rule.	With 418 − 1 = 417 and 418 − 1 = 417 degrees of freedom, the critical values are 0.82512 and 1.21194. Reject the null if the calculated *F*-statistic is less than 0.82512 or greater than 1.21194.

Excel

Left side: F.INV(0.025,417,417)

Right side: F.INV(0.975,417,417)

R qf(c(.025,.975),417,417)

Python from scipy.stats import f

Left side: f.ppf(.025,417,417)

Right side: f.ppf(.975,417,417)

Step 5 Calculate the test statistic.

$$F = \frac{4.644}{3.919} = 1.18500$$

Step 6 Make a decision.

Fail to reject the null hypothesis since the calculated *F*-statistic falls within the bounds of the two critical values. There is not sufficient evidence to indicate that the weekly variances of returns are different in the periods before and after the regulation change.

Solution to 2

Step 1 State the hypotheses.

$$H_0 : \sigma^2_{Before} \leq \sigma^2_{After} \text{ versus } H_a : \sigma^2_{Before} > \sigma^2_{After}$$

Step 2 Identify the appropriate test statistic.

$$F = \frac{s^2_{Before}}{s^2_{After}}$$

Step 3 Specify the level of significance.

5%

Step 4 State the decision rule.

With 418 − 1 = 417 and 418 − 1 = 417 degrees of freedom, the critical value is 1.17502.

We reject the null hypothesis if the calculated *F*-statistic is greater than 1.17502.

Excel F.INV(0.95,417,417)

R qf(.95,417,417)

Python from scipy.stats import f

f.ppf(.95,417,417)

Step 5 Calculate the test statistic.

$$F = \frac{4.644}{3.919} = 1.18500$$

Step 6 Make a decision.

Reject the null hypothesis since the calculated *F*-statistic is greater than 1.17502. There is sufficient evidence to indicate that the weekly variances of returns before the regulation change are greater than the variances after the regulation change.

EXAMPLE 14

The Volatility of Derivatives Expiration Days

You are interested in investigating whether quadruple witching days—that is, the occurrence of stock option, index option, index futures, and single stock futures expirations on the same day—exhibit greater volatility than normal trading days. Exhibit 22 presents the daily standard deviation of returns for normal trading days and quadruple witching days during a four-year period.

Exhibit 22 Standard Deviation of Returns: Normal Trading Days and Derivatives Expiration Days

Period	Type of Day	*n*	Standard Deviation (%)
1	Normal trading days	138	0.821
2	Quadruple witching days	16	1.217

Test to determine whether the variance of returns for quadruple witching days is greater than the variance for non-expiration, normal trading days. Use a 5% level of significance.

Solution

Step 1	State the hypotheses.	$H_0: \sigma^2_{Period2} \leq \sigma^2_{Period1}$ versus $H_a: \sigma^2_{Period2} > \sigma^2_{Period1}$
Step 2	Identify the appropriate test statistic.	$F = \frac{s^2_{Period2}}{s^2_{Period1}}$
Step 3	Specify the level of significance.	5%
Step 4	State the decision rule.	With 16 − 1 = 15 and 138 − 1 = 137 degrees of freedom, the critical value is 1.73997. We reject the null hypothesis if the calculated *F*-statistic is greater than 1.73997. **Excel** F.INV(0.95,15,137) **R** qf(.95,15,137) **Python** from scipy.stats import f f.ppf(.95,15,137)
Step 5	Calculate the test statistic.	$F = \frac{1.48109}{0.67404} = 2.19733$
Step 6	Make a decision.	Reject the null hypothesis since the calculated *F*-statistic is greater than 1.73997. There is sufficient evidence to indicate that the variance of returns for quadruple witching days is greater than the variance for normal trading days.

13 PARAMETRIC VS. NONPARAMETRIC TESTS

k compare and contrast parametric and nonparametric tests, and describe situations where each is the more appropriate type of test

The hypothesis-testing procedures we have discussed up to this point have two characteristics in common. First, they are concerned with parameters, and second, their validity depends on a definite set of assumptions. Mean and variance, for example, are two parameters, or defining quantities, of a normal distribution. The tests also make specific assumptions—in particular, assumptions about the distribution of the population producing the sample. Any test or procedure with either of these two characteristics is a **parametric test** or procedure. In some cases, however, we are concerned about quantities other than parameters of distributions. In other cases, we may believe that the assumptions of parametric tests do not hold. In cases where we are examining quantities other than population parameters or where assumptions of the parameters are not satisfied, a nonparametric test or procedure can be useful.

A **nonparametric test** is a test that is not concerned with a parameter or a test that makes minimal assumptions about the population from which the sample comes. In Exhibit 23, we give examples of nonparametric alternatives to the parametric, *t*-distributed tests concerning means.

Exhibit 23 Nonparametric Alternatives to Parametric Tests Concerning Means

	Parametric	Nonparametric
Tests concerning a single mean	*t*-distributed test *z*-distributed test	Wilcoxon signed-rank test
Tests concerning differences between means	*t*-distributed test	Mann–Whitney *U* test (Wilcoxon rank sum test)
Tests concerning mean differences (paired comparisons tests)	*t*-distributed test	Wilcoxon signed-rank test Sign test

13.1 Uses of Nonparametric Tests

We primarily use nonparametric procedures in four situations: (1) when the data we use do not meet distributional assumptions, (2) when there are outliers, (3) when the data are given in ranks or use an ordinal scale, or (4) when the hypotheses we are addressing do not concern a parameter.

The first situation occurs when the data available for analysis suggest that the distributional assumptions of the parametric test are not satisfied. For example, we may want to test a hypothesis concerning the mean of a population but believe that neither *t*- nor *z*-distributed tests are appropriate because the sample is small and may come from a markedly non-normally distributed population. In that case, we may use a nonparametric test. The nonparametric test will frequently involve the conversion of observations (or a function of observations) into ranks according to magnitude, and sometimes it will involve working with only "greater than" or "less than" relationships (using the + and − signs to denote those relationships). Characteristically, one must refer to specialized statistical tables to determine the rejection points of the test statistic, at least for small samples. Such tests, then, typically interpret the null hypothesis as a hypothesis about ranks or signs.

Second, whereas the underlying distribution of the population may be normal, there may be extreme values or outliers that influence the parametric statistics but not the nonparametric statistics. For example, we may want to use a nonparametric test of the median, in the case of outliers, instead of a test of the mean.

Third, we may have a sample in which observations are ranked. In those cases, we also use nonparametric tests because parametric tests generally require a stronger measurement scale than ranks. For example, if our data were the rankings of investment managers, we would use nonparametric procedures to test the hypotheses concerning those rankings.

A fourth situation in which we use nonparametric procedures occurs when our question does not concern a parameter. For example, if the question concerns whether a sample is random or not, we use the appropriate nonparametric test (a "runs test"). The nonparametric runs test is used to test whether stock price changes can be used to forecast future stock price changes—in other words, a test of the random-walk theory. Another type of question that nonparametric methods can address is whether a sample came from a population following a particular probability distribution.

13.2 Nonparametric Inference: Summary

Nonparametric statistical procedures extend the reach of inference because they make few assumptions, can be used on ranked data, and may address questions unrelated to parameters. Quite frequently, nonparametric tests are reported alongside parametric tests; the user can then assess how sensitive the statistical conclusion is to the assumptions underlying the parametric test. However, if the assumptions of the parametric test are met, the parametric test (where available) is generally preferred over the nonparametric test because the parametric test may have more power—that is, a greater ability to reject a false null hypothesis.

EXAMPLE 15

The Use of Nonparametric Tests

A nonparametric test is most appropriate when the:

A data consist of ranked values.

B validity of the test depends on many assumptions.

C sample sizes are large but are drawn from a population that may be non-normal.

Solution

A is correct. When the samples consist of ranked values, parametric tests are not appropriate. In such cases, nonparametric tests are most appropriate.

14 TESTS CONCERNING CORRELATION

l. explain parametric and nonparametric tests of the hypothesis that the population correlation coefficient equals zero, and determine whether the hypothesis is rejected at a given level of significance

In many contexts in investments, we want to assess the strength of the linear relationship between two variables; that is, we want to evaluate the **correlation** between them. A significance test of a correlation coefficient allows us to assess whether the relationship between two random variables is the result of chance. If we decide that the relationship does not result from chance, then we are inclined to use this information in modeling or forecasting.

If the correlation between two variables is zero, we conclude that there is no linear relation between the two variables. We use a test of significance to assess whether the correlation is different from zero. After we estimate a correlation coefficient, we need to ask whether the estimated correlation is significantly different from zero.

A correlation may be positive (that is, the two variables tend to move in the same direction at the same time) or negative (that is, the two variables tend to move in different directions at the same time). The correlation coefficient is a number between −1 and +1, where −1 denotes a perfect negative or inverse, straight-line relationship between the two variables; +1 denotes a perfect positive, straight-line relationship; and 0 represents the absence of any straight-line relationship (that is, no correlation).

The most common hypotheses concerning correlation occur when comparing the population correlation coefficient with zero because we are often asking whether there is a relationship, which implies a null of the correlation coefficient equal to zero (that is, no relationship). Hypotheses concerning the population correlation coefficient may be two or one sided, as we have seen in other tests. Let ρ represent the population correlation coefficient. The possible hypotheses are as follows:

Two sided: H_0: $\rho = 0$ versus H_a: $\rho \neq 0$

One sided (right side): H_0: $\rho \leq 0$ versus H_a: $\rho > 0$

One sided (left side): H_0: $\rho \geq 0$ versus H_a: $\rho < 0$

We use the sample correlation to test these hypotheses on the population correlation.

14.1 Parametric Test of a Correlation

The parametric pairwise correlation coefficient is often referred to as the **Pearson correlation**, the **bivariate correlation**, or simply the correlation. Our focus is on the testing of the correlation and not the actual calculation of this statistic, but it helps distinguish this correlation from the nonparametric correlation if we look at the formula for the sample correlation. Consider two variables, X and Y. The sample correlation, r_{XY}, is

$$r_{XY} = \frac{s_{XY}}{s_X s_Y},$$

where s_{XY} is the sample covariance between the X and Y variables, s_X is the standard deviation of the X variable, and s_Y is the standard deviation of the Y variable. We often drop the subscript to represent the correlation as simply r.

Therefore, you can see from this formula that each observation is compared with its respective variable mean and that, because of the covariance, it matters how much each observation differs from its respective variable mean. Note that the covariance drives the sign of the correlation.

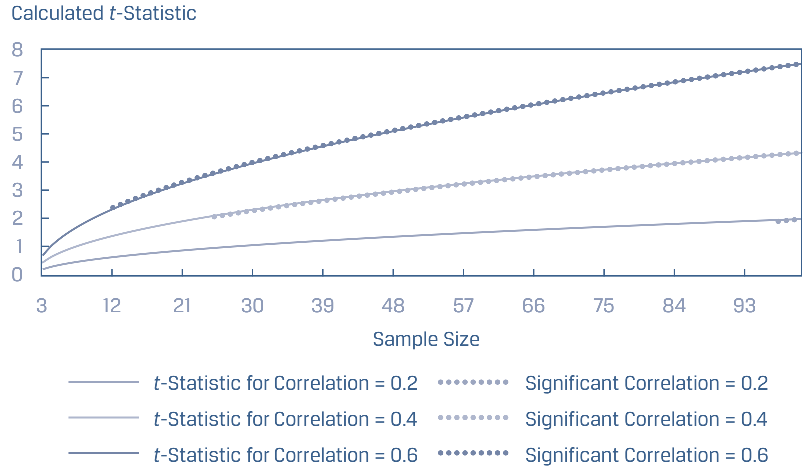
If the two variables are normally distributed, we can test to determine whether the null hypothesis (H_0: $\rho = 0$) should be rejected using the sample correlation, r. The formula for the t-test is

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}. \tag{10}$$

This test statistic is t-distributed with $n - 2$ degrees of freedom. One practical observation concerning Equation 10 is that the magnitude of r needed to reject the null hypothesis decreases as sample size n increases, for two reasons. First, as n increases, the number of degrees of freedom increases and the absolute value of the critical value of the t-statistic decreases. Second, the absolute value of the numerator increases with larger n, resulting in a larger magnitude of the calculated t-statistic. For example, with sample size $n = 12$, $r = 0.35$ results in a t-statistic of 1.182, which is not different from zero at the 0.05 level ($t_{\alpha/2} = \pm 2.228$). With a sample size of $n = 32$, the same sample correlation, $r = 0.35$, yields a t-statistic of 2.046, which is just significant at the 0.05 level ($t_{\alpha/2} = \pm 2.042$).

Another way to make this point is that when sampling from the same population, a false null hypothesis is more likely to be rejected (that is, the power of the test increases) as we increase the sample size, all else equal, because a higher number of observations increases the numerator of the test statistic. We show this in Exhibit 24 for three different sample correlation coefficients, with the corresponding calculated t-statistics and significance at the 5% level for a two-sided alternative hypothesis. As the sample size increases, significance is more likely to be indicated, but the rate of achieving this significance depends on the sample correlation coefficient; the higher the sample correlation, the faster significance is achieved when increasing the sample size. As the sample sizes increase as ever-larger datasets are examined, the null hypothesis is almost always rejected and other tools of data analysis must be applied.

Exhibit 24 **Calculated Test Statistics for Different Sample Sizes and Sample Correlations with a 5% Level of Significance**



EXAMPLE 16

Examining the Relationship between Returns on Investment One and Investment Two

An analyst is examining the annual returns for Investment One and Investment Two, as displayed in Exhibit 1. Although this time series plot provides some useful information, the analyst is most interested in quantifying how the returns of these two series are related, so she calculates the correlation coefficient, equal to 0.43051, between these series.

Is there a significant positive correlation between these two return series if she uses a 1% level of significance?

Solution

Step 1	State the hypotheses.	$H_0 : \rho \leq 0$ versus $H_a : \rho > 0$
Step 2	Identify the appropriate test statistic.	$t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$
Step 3	Specify the level of significance.	1%
Step 4	State the decision rule	With 33 – 2 = 31 degrees of freedom and a one-sided test with a 1% level of significance, the critical value is 2.45282. We reject the null hypothesis if the calculated *t*-statistic is greater than 2.45282.
Step 5	Calculate the test statistic.	$t = \dfrac{0.43051\sqrt{33-2}}{\sqrt{1-0.18534}} = 2.65568$
Step 6	Make a decision	Reject the null hypothesis since the calculated *t*-statistic is greater than 2.45282. There is sufficient evidence to reject the H_0 in favor of H_a, that the correlation between the annual returns of these two investments is positive.

14.2 Tests Concerning Correlation: The Spearman Rank Correlation Coefficient

l. explain parametric and nonparametric tests of the hypothesis that the population correlation coefficient equals zero, and determine whether the hypothesis is rejected at a given level of significance

When we believe that the population under consideration meaningfully departs from normality, we can use a test based on the **Spearman rank correlation coefficient**, r_S. The Spearman rank correlation coefficient is essentially equivalent to the usual correlation coefficient but is calculated on the *ranks* of the two variables (say, X and Y) within their respective samples. The calculation of r_S requires the following steps:

1. Rank the observations on X from largest to smallest. Assign the number 1 to the observation with the largest value, the number 2 to the observation with second largest value, and so on. In case of ties, assign to each tied observation the average of the ranks that they jointly occupy. For example, if the third and fourth largest values are tied, we assign both observations the rank of 3.5 (the average of 3 and 4). Perform the same procedure for the observations on Y.
2. Calculate the difference, d_i, between the ranks for each pair of observations on X and Y, and then calculate d_i^2 (the squared difference in ranks).
3. With n as the sample size, the Spearman rank correlation is given by

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}. \tag{11}$$

Suppose an analyst is examining the relationship between returns for two investment funds, A and B, of similar risk over 35 years. She is concerned that the assumptions for the parametric correlation may not be met, so she decides to test Spearman rank correlations. Her hypotheses are H_0: $r_S = 0$ and H_a: $r_S \neq 0$. She gathers the returns, ranks the returns for each fund, and calculates the difference in ranks and the squared differences. A partial table is provided in Exhibit 25.

Exhibit 25 Differences and Squared Differences in Ranks for Fund A and Fund B over 35 Years

Year	Fund A	Fund B	Rank of A	Rank of B	*d*	*d*²
1	2.453	1.382	27	31	−4	16
2	3.017	3.110	24	24	0	0
3	4.495	6.587	19	7	12	144
4	3.627	3.300	23	23	0	0
.						
.						
.						
30	2.269	0.025	28	35	−7	49
31	6.354	4.428	10	19	−9	81
32	6.793	4.165	8	20	−12	144
33	7.300	7.623	5	5	0	0
34	6.266	4.527	11	18	−7	49

(continued)

Exhibit 25 (Continued)

Year	Fund A	Fund B	Rank of A	Rank of B	d	d^2
35	1.257	4.704	34	16	18	324
					Sum =	2,202

The Spearman rank correlation is:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} = 1 - \frac{6(2,202)}{35(1,225 - 1)} = 0.6916.$$

The test of hypothesis for the Spearman rank correlation depends on whether the sample is small or large ($n > 30$). For small samples, the researcher requires a specialized table of critical values, but for large samples, we can conduct a t-test using the test statistic in Equation 10, which is t-distributed with $n - 2$ degrees of freedom.

In this example, for a two-tailed test with a 5% significance level, the critical values for $n - 2 = 35 - 2 = 33$ degrees of freedom are ±2.0345. For the sample information in Exhibit 24, the calculated test statistic is

$$t = \frac{0.6916\sqrt{33}}{\sqrt{1 - (0.6916^2)}} = 5.5005.$$

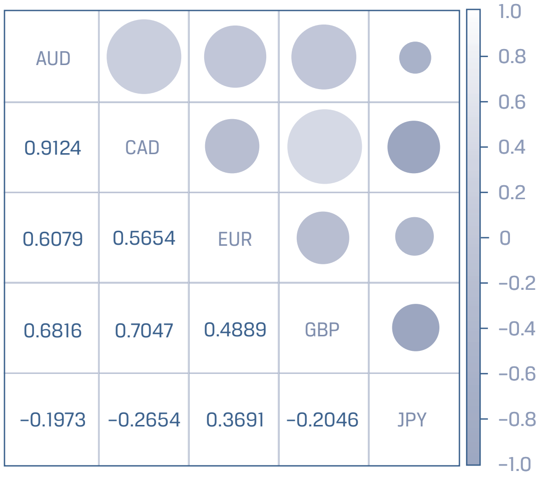
Accordingly, we reject the null hypothesis (H_0: $r_S = 0$), concluding that there is sufficient evidence to indicate that the correlation between the returns of Fund A and Fund B is different from zero.

EXAMPLE 17

Testing the Exchange Rate Correlation

An analyst gathers exchange rate data for five currencies relative to the US dollar. Upon inspection of the distribution of these exchange rates, she observes a departure from normality, especially with negative skewness for four of the series and positive skewness for the fifth. Therefore, she decides to examine the relationships among these currencies using Spearman rank correlations. She calculates these correlations between the currencies over 180 days, which are shown in the correlogram in Exhibit 26. In this correlogram, the lower triangle reports the pairwise correlations and the upper triangle provides a visualization of the magnitude of the correlations, with larger circles indicating larger absolute value of the correlations and darker circles indicating correlations that are negative.

Exhibit 26 **Spearman Rank Correlations between Exchanges Rates Relative to the US Dollar**



AUD				
0.9124	CAD			
0.6079	0.5654	EUR		
0.6816	0.7047	0.4889	GBP	
−0.1973	−0.2654	0.3691	−0.2046	JPY

For any of these pairwise Spearman rank correlations, can we reject the null hypothesis of no correlation (H_0: $r_S = 0$ and H_a: $r_S \neq 0$) at the 5% level of significance?

Solution

The critical t-values for 2.5% in each tail of the distribution are ±1.97338.

There are five exchange rates, so there are ${}_5C_2$, or 10, unique correlation pairs. Therefore, we need to calculate 10 t-statistics. For example, the correlation between EUR/USD and AUD/USD is 0.6079. The calculated t-statistic is $\frac{0.6079\sqrt{180-2}}{\sqrt{1-0.6079^2}} = \frac{8.11040}{0.79401} = 10.2144$. Repeating this t-statistic calculation for each pair of exchange rates yields the test statistics shown in Exhibit 27.

Exhibit 27 **Calculated Test Statistics for Test of Spearman Rank Correlations**

	AUD/USD	CAD/USD	EUR/USD	GBP/USD
CAD/USD	29.7409			
EUR/USD	10.2144	9.1455		
GBP/USD	12.4277	13.2513	7.4773	
JPY/USD	−2.6851	−3.6726	5.2985	−2.7887

The analyst should reject all 10 null hypotheses, because the calculated t-statistics for all exchange rate pairs fall outside the bounds of the two critical values. She should conclude that all the exchange rate pair correlations are different from zero at the 5% level.

15 TEST OF INDEPENDENCE USING CONTINGENCY TABLE DATA

m explain tests of independence based on contingency table data

When faced with categorical or discrete data, we cannot use the methods that we have discussed up to this point to test whether the classifications of such data are independent. Suppose we observe the following **frequency table** of 1,594 exchange-traded funds (ETFs) based on two classifications: size (that is, market capitalization) and investment type (value, growth, or blend), as shown in Exhibit 28. The classification of the investment type is discrete, so we cannot use correlation to assess the relationship between size and investment type.

Exhibit 28 Size and Investment Type Classifications of 1,594 ETFs

	Size Based on Market Capitalization			
Investment Type	**Small**	**Medium**	**Large**	**Total**
Value	50	110	343	503
Growth	42	122	202	366
Blend	56	149	520	725
Total	148	381	1,065	1,594

This table is referred to as a **contingency table** or a **two-way table** (because there are two classifications, or classes—size and investment type).

If we want to test whether there is a relationship between the size and investment type, we can perform a test of independence using a nonparametric test statistic that is chi-square distributed:

$$\chi^2 = \sum_{i=1}^{m} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}, \quad (12)$$

where

m = the number of cells in the table, which is the number of groups in the first class multiplied by the number of groups in the second class

O_{ij} = the number of observations in each cell of row i and column j (i.e., observed frequency)

E_{ij} = the expected number of observations in each cell of row i and column j, assuming independence (i.e., expected frequency)

This test statistic has $(r - 1)(c - 1)$ degrees of freedom, where r is the number of rows and c is the number of columns.

In Exhibit 28, size class has three groups (small, medium, and large) and investment type class has three groups (value, growth, and blend), so m is 9 (= 3 × 3). The number of ETFs in each cell (O_{ij}), the observed frequency, is given, so to calculate the chi-square test statistic, we need to estimate E_{ij}, the expected frequency, which is the number of ETFs we would expect to be in each cell if size and investment type are completely independent. The expected number of ETFs (E_{ij}) is calculated using

$$E_{ij} = \frac{(\text{Total row } i) \times (\text{Total column } j)}{\text{Overall total}}. \quad (13)$$

Consider one combination of size and investment type, small-cap value:

$$E_{ij} = \frac{503 \times 148}{1{,}594} = 46.703.$$

We repeat this calculation for each combination of size and investment type (i.e., m = 9 pairs) to arrive at the expected frequencies, shown in Panel A of Exhibit 29.

Next, we calculate $\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$, the squared difference between observed and expected frequencies scaled by expected frequency, for each cell as shown in Panel B of Exhibit 29. Finally, by summing the values of $\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$ for each of the m cells, we calculate the chi-square statistic as 32.08025.

Exhibit 29 Inputs to Chi-Square Test Statistic Calculation for 1,594 ETFs Assuming Independence of Size and Investment Type

A. Expected Frequency of ETFs by Size and Investment Type

	Size Based on Market Capitalization		
Investment Type	**Small**	**Medium**	**Large**
Value	46.703	120.228	336.070
Growth	33.982	87.482	244.536
Blend	67.315	173.290	484.395
Total	148.000	381.000	1,065.000

B. Scaled Squared Deviation for Each Combination of Size and Investment Type

	Size Based on Market Capitalization		
Investment Type	**Small**	**Medium**	**Large**
Value	0.233	0.870	0.143
Growth	1.892	13.620	7.399
Blend	1.902	3.405	2.617

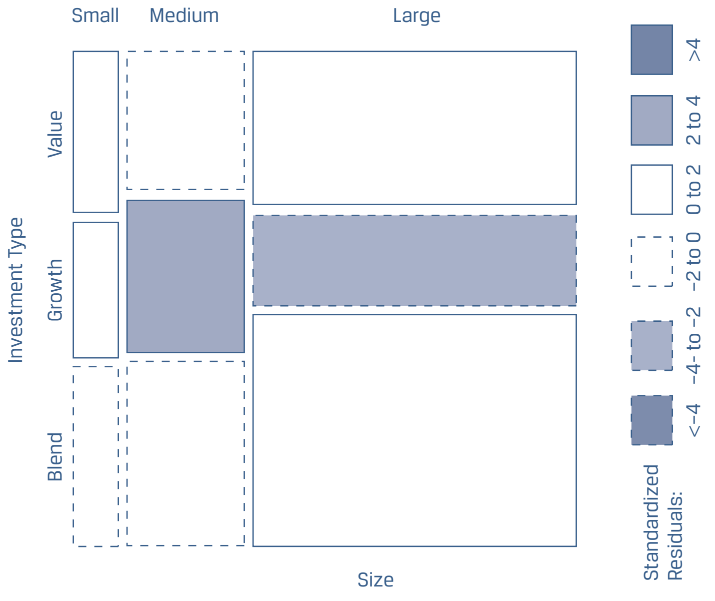
In our ETF example, we test the null hypothesis of independence between the two classes (i.e., no relationship between size and investment type) versus the alternative hypothesis of dependence (i.e., a relationship between size and investment type) using a 5% level of significance, as shown in Exhibit 30. If, on the one hand, the observed values are equal to the expected values, the calculated test statistic would be zero. If, on the other hand, there are differences between the observed and expected values, these differences are squared, so the calculated chi-square statistic will be positive. Therefore, for the test of independence using a contingency table, there is only one rejection region, on the right side.

Exhibit 30 Test of Independence of Size and Investment Type for 1,594 ETFs

Step 1	State the hypotheses.	H_0: ETF size and investment type are not related, so these classifications are independent; H_a : ETF size and investment type are related, so these classifications are not independent.
Step 2	Identify the appropriate test statistic.	$\chi^2 = \sum_{i=1}^{m} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$
Step 3	Specify the level of significance.	5%
Step 4	State the decision rule.	With (3 – 1) × (3 – 1) = 4 degrees of freedom and a one-sided test with a 5% level of significance, the critical value is 9.4877. We reject the null hypothesis if the calculated χ^2 statistic is greater than 9.4877. **Excel** CHISQ.INV(0.95,4) **R** qchisq(.95,4) **Python** from scipy.stats import chi2 chi2.ppf(.95,4)
Step 5	Calculate the test statistic.	$\chi^2 = 32.08025$
Step 6	Make a decision.	Reject the null hypothesis of independence because the calculated χ^2 test statistic is greater than 9.4877. There is sufficient evidence to conclude that ETF size and investment type are related (i.e., not independent).

We can visualize the contingency table in a graphic referred to as a mosaic. In a mosaic, a grid reflects the comparison between the observed and expected frequencies. Consider Exhibit 31, which represents the ETF contingency table.

Exhibit 31 Mosaic of the ETF Contingency Table



The width of the rectangles in Exhibit 31 reflect the proportion of ETFs that are small, medium, and large, whereas the height reflects the proportion that are value, growth, and blend. The darker shading indicates whether there are more observations than expected under the null hypothesis of independence, whereas the lighter shading indicates that there are fewer observations than expected, with "more" and "fewer" determined by reference to the standardized residual boxes. The standardized residual, also referred to as a Pearson residual, is

$$\text{Standardized residual} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}}}. \quad (14)$$

The interpretation for this ETF example is that there are more medium-size growth ETFs (standardized residual of 3.69) and fewer large-size growth ETFs (standardized residual of –2.72) than would be expected if size and investment type were independent.

EXAMPLE 18

Using Contingency Tables to Test for Independence

Consider the contingency table in Exhibit 32, which classifies 500 randomly selected companies on the basis of two environmental, social, and governance (ESG) rating dimensions: environmental rating and governance rating.

Exhibit 32 Classification of 500 Randomly Selected Companies Based on Environmental and Governance Ratings

	Governance Rating			
Environmental Rating	**Progressive**	**Average**	**Poor**	**Total**
Progressive	35	40	5	80
Average	80	130	50	260
Poor	40	60	60	160
Total	155	230	115	500

1 What are the expected frequencies for these two ESG rating dimensions if these categories are independent?

2 Using a 5% level of significance, determine whether these two ESG rating dimensions are independent of one another.

Solution to 1

The expected frequencies based on independence of the governance rating and the environmental rating are shown in Panel A of Exhibit 33. For example, using Equation 12, the expected frequency for the combination of progressive governance and progressive environmental ratings is

$$E_{ij} = \frac{155 \times 80}{500} = 24.80.$$

Exhibit 33 Inputs to Chi-Square Test Statistic Calculation Assuming Independence of Environmental and Governance Ratings

A. Expected Frequencies of Environmental and Governance Ratings Assuming Independence

	Governance Rating		
Environmental Rating	**Progressive**	**Average**	**Poor**
Progressive	24.8	36.8	18.4
Average	80.6	119.6	59.8
Poor	49.6	73.6	36.8

B. Scaled Squared Deviation for Each Combination of Environmental and Governance Ratings

	Governance Rating		
Environmental Rating	**Progressive**	**Average**	**Poor**
Progressive	4.195	0.278	9.759
Average	0.004	0.904	1.606
Poor	1.858	2.513	14.626

Solution to 2

Step 1 State the hypotheses.

H_0: Governance and environmental ratings are not related, so these ratings are independent;
H_a: Governance and environmental ratings are related, so these ratings are not independent.

Step 2 Identify the appropriate test statistic.

$$\chi^2 = \sum_{i=1}^{m} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

Step 3 Specify the level of significance.

5%

Step 4 State the decision rule.

With (3 − 1) × (3 − 1) = 4 degrees of freedom and a one-sided test with a 5% level of significance, the critical value is 9.487729.

We reject the null hypothesis if the calculated χ^2 statistic is greater than 9.487729.

Excel CHISQ.INV(0.95,4)

R qchisq(.95,4)

Python from scipy.stats import chi2
chi2.ppf(.95,4)

Step 5 Calculate the test statistic.

$\chi^2 = 35.74415$

To calculate the test statistic, we first calculate the squared difference between observed and expected frequencies scaled by expected frequency for each cell, as shown in Panel B of Exhibit 33. Then, summing the values in each of the *m* cells (see Equation 11), we calculate the chi-square statistic as 35.74415.

Step 6 Make a decision.

Reject the null hypothesis because the calculated χ^2 test statistic is greater than 9.487729. There is sufficient evidence to indicate that the environmental and governance ratings are related, so they are not independent.

SUMMARY

In this reading, we have presented the concepts and methods of statistical inference and hypothesis testing.

- A hypothesis is a statement about one or more populations.
- The steps in testing a hypothesis are as follows:
  1. State the hypotheses.
  2. Identify the appropriate test statistic and its probability distribution.
  3. Specify the significance level.
  4. State the decision rule.
  5. Collect the data and calculate the test statistic.
  6. Make a decision.
- We state two hypotheses: The null hypothesis is the hypothesis to be tested; the alternative hypothesis is the hypothesis accepted if the null hypothesis is rejected.
- There are three ways to formulate hypotheses. Let θ indicate the population parameters:
  1. Two-sided alternative: H_0: $\theta = \theta_0$ versus H_a: $\theta \neq \theta_0$
  2. One-sided alternative (right side): H_0: $\theta \leq \theta_0$ versus H_a: $\theta > \theta_0$
  3. One-sided alternative (left side): H_0: $\theta \geq \theta_0$ versus H_a: $\theta < \theta_0$

  where θ_0 is a hypothesized value of the population parameter and θ is the true value of the population parameter.
- When we have a "suspected" or "hoped for" condition for which we want to find supportive evidence, we frequently set up that condition as the alternative hypothesis and use a one-sided test. However, the researcher may select a "not equal to" alternative hypothesis and conduct a two-sided test to emphasize a neutral attitude.
- A test statistic is a quantity, calculated using a sample, whose value is the basis for deciding whether to reject or not reject the null hypothesis. We compare the computed value of the test statistic to a critical value for the same test statistic to decide whether to reject or not reject the null hypothesis.
- In reaching a statistical decision, we can make two possible errors: We may reject a true null hypothesis (a Type I error, or false positive), or we may fail to reject a false null hypothesis (a Type II error, or false negative).
- The level of significance of a test is the probability of a Type I error that we accept in conducting a hypothesis test. The standard approach to hypothesis testing involves specifying only a level of significance (that is, the probability of a Type I error). The complement of the level of significance is the confidence level.
- The power of a test is the probability of correctly rejecting the null (rejecting the null when it is false). The complement of the power of the test is the probability of a Type II error.
- A decision rule consists of determining the critical values with which to compare the test statistic to decide whether to reject or not reject the null hypothesis. When we reject the null hypothesis, the result is said to be statistically significant.
- The $(1 - \alpha)$ confidence interval represents the range of values of the test statistic for which the null hypothesis is not be rejected.

- The statistical decision consists of rejecting or not rejecting the null hypothesis. The economic decision takes into consideration all economic issues pertinent to the decision.
- The *p*-value is the smallest level of significance at which the null hypothesis can be rejected. The smaller the *p*-value, the stronger the evidence against the null hypothesis and in favor of the alternative hypothesis. The *p*-value approach to hypothesis testing involves computing a *p*-value for the test statistic and allowing the user of the research to interpret the implications for the null hypothesis.
- For hypothesis tests concerning the population mean of a normally distributed population with unknown variance, the theoretically correct test statistic is the *t*-statistic.
- When we want to test whether the observed difference between two means is statistically significant, we must first decide whether the samples are independent or dependent (related). If the samples are independent, we conduct a test concerning differences between means. If the samples are dependent, we conduct a test of mean differences (paired comparisons test).
- When we conduct a test of the difference between two population means from normally distributed populations with unknown but equal variances, we use a *t*-test based on pooling the observations of the two samples to estimate the common but unknown variance. This test is based on an assumption of independent samples.
- In tests concerning two means based on two samples that are not independent, we often can arrange the data in paired observations and conduct a test of mean differences (a paired comparisons test). When the samples are from normally distributed populations with unknown variances, the appropriate test statistic is *t*-distributed.
- In tests concerning the variance of a single normally distributed population, the test statistic is chi-square with $n - 1$ degrees of freedom, where n is sample size.
- For tests concerning differences between the variances of two normally distributed populations based on two random, independent samples, the appropriate test statistic is based on an *F*-test (the ratio of the sample variances). The degrees of freedom for this *F*-test are $n_1 - 1$ and $n_2 - 1$, where n_1 corresponds to the number of observations in the calculation of the numerator and n_2 is the number of observations in the calculation of the denominator of the *F*-statistic.
- A parametric test is a hypothesis test concerning a population parameter or a hypothesis test based on specific distributional assumptions. In contrast, a nonparametric test either is not concerned with a parameter or makes minimal assumptions about the population from which the sample comes.
- A nonparametric test is primarily used when data do not meet distributional assumptions, when there are outliers, when data are given in ranks, or when the hypothesis we are addressing does not concern a parameter.
- In tests concerning correlation, we use a *t*-statistic to test whether a population correlation coefficient is different from zero. If we have n observations for two variables, this test statistic has a *t*-distribution with $n - 2$ degrees of freedom.
- The Spearman rank correlation coefficient is calculated on the ranks of two variables within their respective samples.
- A chi-square distributed test statistic is used to test for independence of two categorical variables. This nonparametric test compares actual frequencies with those expected on the basis of independence. This test statistic has degrees of freedom of $(r - 1)(c - 2)$, where r is the number of categories for the first variable and c is the number of categories of the second variable.

REFERENCES

Benjamini, Y., and Y. Hochberg. 1995. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.” *Journal of the Royal Statistical Society. Series B. Methodological* 57:289–300.

PRACTICE PROBLEMS

1 Which of the following statements about hypothesis testing is correct?

A The null hypothesis is the condition a researcher hopes to support.

B The alternative hypothesis is the proposition considered true without conclusive evidence to the contrary.

C The alternative hypothesis exhausts all potential parameter values not accounted for by the null hypothesis.

2 Identify the appropriate test statistic or statistics for conducting the following hypothesis tests. (Clearly identify the test statistic and, if applicable, the number of degrees of freedom. For example, "We conduct the test using an *x*-statistic with *y* degrees of freedom.")

A H_0: $\mu = 0$ versus H_a: $\mu \neq 0$, where μ is the mean of a normally distributed population with unknown variance. The test is based on a sample of 15 observations.

B H_0: $\mu = 5$ versus H_a: $\mu \neq 5$, where μ is the mean of a normally distributed population with unknown variance. The test is based on a sample of 40 observations.

C H_0: $\mu \leq 0$ versus H_a: $\mu > 0$, where μ is the mean of a normally distributed population with known variance σ^2. The sample size is 45.

D H_0: $\sigma^2 = 200$ versus H_a: $\sigma^2 \neq 200$, where σ^2 is the variance of a normally distributed population. The sample size is 50.

E $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_a : \sigma_1^2 \neq \sigma_2^2$, where σ_1^2 is the variance of one normally distributed population and σ_2^2 is the variance of a second normally distributed population. The test is based on two independent samples, with the first sample of size 30 and the second sample of size 40.

F H_0: $\mu_1 - \mu_2 = 0$ versus H_a: $\mu_1 - \mu_2 \neq 0$, where the samples are drawn from normally distributed populations with unknown but assumed equal variances. The observations in the two samples (of size 25 and 30, respectively) are independent.

3 For each of the following hypothesis tests concerning the population mean, μ, state the conclusion regarding the test of the hypotheses.

A H_0: $\mu = 10$ versus H_a: $\mu \neq 10$, with a calculated *t*-statistic of 2.05 and critical *t*-values of ±1.984.

B H_0: $\mu \leq 10$ versus H_a: $\mu > 10$, with a calculated *t*-statistic of 2.35 and a critical *t*-value of +1.679

C H_0: $\mu = 10$ versus H_a: $\mu \neq 10$, with a calculated *t*-statistic of 2.05, a *p*-value of 4.6352%, and a level of significance of 5%.

D H_0: $\mu \leq 10$ versus H_a: $\mu > 10$, with a 2% level of significance and a calculated test statistic with a *p*-value of 3%.

4 For each of the following hypothesis tests concerning the population mean, state the conclusion.

A H_0: $\sigma^2 = 0.10$ versus H_a: $\sigma^2 \neq 0.10$, with a calculated chi-square test statistic of 45.8 and critical chi-square values of 42.950 and 86.830.

B H_0: $\sigma^2 = 0.10$ versus H_a: $\sigma^2 \neq 0.10$, with a 5% level of significance and a *p*-value for this calculated chi-square test statistic of 4.463%.

© 2021 CFA Institute. All rights reserved.

C H_0: $\sigma_1{}^2 = \sigma_2{}^2$ versus H_a: $\sigma_1{}^2 \neq \sigma_2{}^2$, with a calculated F-statistic of 2.3. With 40 and 30 degrees of freedom, the critical F-values are 0.498 and 1.943.

D H_0: $\sigma^2 \leq 10$ versus H_a: $\mu\sigma^2 > 10$, with a calculated test statistic of 32 and a critical chi-square value of 26.296.

5 Willco is a manufacturer in a mature cyclical industry. During the most recent industry cycle, its net income averaged \$30 million per year with a standard deviation of \$10 million ($n = 6$ observations). Management claims that Willco's performance during the most recent cycle results from new approaches and that Willco's profitability will exceed the \$24 million per year observed in prior cycles.

A With μ as the population value of mean annual net income, formulate null and alternative hypotheses consistent with testing Willco management's claim.

B Assuming that Willco's net income is at least approximately normally distributed, identify the appropriate test statistic and calculate the degrees of freedom.

C Based on critical value of 2.015, determine whether to reject the null hypothesis.

The following information relates to Questions 6–7

Performance in Forecasting Quarterly Earnings per Share

	Number of Forecasts	Mean Forecast Error (Predicted – Actual)	Standard Deviation of Forecast Errors
Analyst A	10	0.05	0.10
Analyst B	15	0.02	0.09

Critical t-values:

	Area in the Right-Side Rejection Area	
Degrees of Freedom	$p = 0.05$	$p = 0.025$
8	1.860	2.306
9	1.833	2.262
10	1.812	2.228
11	1.796	2.201
12	1.782	2.179
13	1.771	2.160
14	1.761	2.145
15	1.753	2.131
16	1.746	2.120
17	1.740	2.110
18	1.734	2.101
19	1.729	2.093
20	1.725	2.086
21	1.721	2.080

(continued)

	Area in the Right-Side Rejection Area	
Degrees of Freedom	***p* = 0.05**	***p* = 0.025**
22	1.717	2.074
23	1.714	2.069
24	1.711	2.064
25	1.708	2.060
26	1.706	2.056
27	1.703	2.052

6 Investment analysts often use earnings per share (EPS) forecasts. One test of forecasting quality is the zero-mean test, which states that optimal forecasts should have a mean forecasting error of zero. The forecasting error is the difference between the predicted value of a variable and the actual value of the variable.

You have collected data (shown in the previous table) for two analysts who cover two different industries: Analyst A covers the telecom industry; Analyst B covers automotive parts and suppliers.

A With μ as the population mean forecasting error, formulate null and alternative hypotheses for a zero-mean test of forecasting quality.

B For Analyst A, determine whether to reject the null at the 0.05 level of significance.

C For Analyst B, determine whether to reject the null at the 0.05 level of significance.

7 Reviewing the EPS forecasting performance data for Analysts A and B, you want to investigate whether the larger average forecast errors of Analyst A relative to Analyst B are due to chance or to a higher underlying mean value for Analyst A. Assume that the forecast errors of both analysts are normally distributed and that the samples are independent.

A Formulate null and alternative hypotheses consistent with determining whether the population mean value of Analyst A's forecast errors (μ_1) are larger than Analyst B's (μ_2).

B Identify the test statistic for conducting a test of the null hypothesis formulated in Part A.

C Identify the rejection point or points for the hypotheses tested in Part A at the 0.05 level of significance.

D Determine whether to reject the null hypothesis at the 0.05 level of significance.

---

8 The following table gives data on the monthly returns on the S&P 500 Index and small-cap stocks for a 40-year period and provides statistics relating to their mean differences. Further, the entire sample period is split into two subperiods of 20 years each, and the return data for these subperiods is also given in the table.

Measure	S&P 500 Return (%)	Small-Cap Stock Return (%)	Differences (S&P 500 – Small-Cap Stock)
Entire sample period, 480 months			
Mean	1.0542	1.3117	–0.258
Standard deviation	4.2185	5.9570	3.752
First subperiod, 240 months			
Mean	0.6345	1.2741	–0.640
Standard deviation	4.0807	6.5829	4.096
Second subperiod, 240 months			
Mean	1.4739	1.3492	0.125
Standard deviation	4.3197	5.2709	3.339

Use a significance level of 0.05 and assume that mean differences are approximately normally distributed.

A Formulate null and alternative hypotheses consistent with testing whether any difference exists between the mean returns on the S&P 500 and small-cap stocks.

B Determine whether to reject the null hypothesis for the entire sample period if the critical values are ±1.96.

C Determine whether to reject the null hypothesis for the first subperiod if the critical values are ±1.96.

D Determine whether to reject the null hypothesis for the second subperiod if the critical values are ±1.96.

9 During a 10-year period, the standard deviation of annual returns on a portfolio you are analyzing was 15% a year. You want to see whether this record is sufficient evidence to support the conclusion that the portfolio's underlying variance of return was less than 400, the return variance of the portfolio's benchmark.

A Formulate null and alternative hypotheses consistent with your objective.

B Identify the test statistic for conducting a test of the hypotheses in Part A, and calculate the degrees of freedom.

C Determine whether the null hypothesis is rejected or not rejected at the 0.05 level of significance using a critical value of 3.325.

10 You are investigating whether the population variance of returns on an index changed subsequent to a market disruption. You gather the following data for 120 months of returns before the disruption and for 120 months of returns after the disruption. You have specified a 0.05 level of significance.

Time Period	*n*	Mean Monthly Return (%)	Variance of Returns
Before disruption	120	1.416	22.367
After disruption	120	1.436	15.795

A Formulate null and alternative hypotheses consistent with the research goal.

B Identify the test statistic for conducting a test of the hypotheses in Part A, and calculate the degrees of freedom.

C Determine whether to reject the null hypothesis at the 0.05 level of significance if the critical values are 0.6969 and 1.4349.

11 The following table shows the sample correlations between the monthly returns for four different mutual funds and the S&P 500. The correlations are based on 36 monthly observations. The funds are as follows:

Fund 1	Large-cap fund
Fund 2	Mid-cap fund
Fund 3	Large-cap value fund
Fund 4	Emerging market fund
S&P 500	US domestic stock index

	Fund 1	**Fund 2**	**Fund 3**	**Fund 4**	**S&P 500**
Fund 1	1				
Fund 2	0.9231	1			
Fund 3	0.4771	0.4156	1		
Fund 4	0.7111	0.7238	0.3102	1	
S&P 500	0.8277	0.8223	0.5791	0.7515	1

Test the null hypothesis that each of these correlations, individually, is equal to zero against the alternative hypothesis that it is not equal to zero. Use a 5% significance level and critical t-values of ±2.032.

12 In the step "stating a decision rule" in testing a hypothesis, which of the following elements must be specified?

A Critical value

B Power of a test

C Value of a test statistic

13 Which of the following statements is correct with respect to the null hypothesis?

A It can be stated as "not equal to" provided the alternative hypothesis is stated as "equal to."

B Along with the alternative hypothesis, it considers all possible values of the population parameter.

C In a two-tailed test, it is rejected when evidence supports equality between the hypothesized value and the population parameter.

14 An analyst is examining a large sample with an unknown population variance. Which of the following is the *most* appropriate test to test the hypothesis that the historical average return on an index is less than or equal to 6%?

A One-sided t-test

B Two-sided t-test

C One-sided chi-square test

15 A hypothesis test for a normally distributed population at a 0.05 significance level implies a:

A 95% probability of rejecting a true null hypothesis.

B 95% probability of a Type I error for a two-tailed test.

C 5% critical value rejection region in a tail of the distribution for a one-tailed test.

16 Which of the following statements regarding a one-tailed hypothesis test is correct?

- **A** The rejection region increases in size as the level of significance becomes smaller.
- **B** A one-tailed test more strongly reflects the beliefs of the researcher than a two-tailed test.
- **C** The absolute value of the rejection point is larger than that of a two-tailed test at the same level of significance.

17 The value of a test statistic is *best* described as the basis for deciding whether to:

- **A** reject the null hypothesis.
- **B** accept the null hypothesis.
- **C** reject the alternative hypothesis.

18 Which of the following is a Type I error?

- **A** Rejecting a true null hypothesis
- **B** Rejecting a false null hypothesis
- **C** Failing to reject a false null hypothesis

19 A Type II error is *best* described as:

- **A** rejecting a true null hypothesis.
- **B** failing to reject a false null hypothesis.
- **C** failing to reject a false alternative hypothesis.

20 The level of significance of a hypothesis test is *best* used to:

- **A** calculate the test statistic.
- **B** define the test's rejection points.
- **C** specify the probability of a Type II error.

21 You are interested in whether excess risk-adjusted return (alpha) is correlated with mutual fund expense ratios for US large-cap growth funds. The following table presents the sample.

Mutual Fund	Alpha	Expense Ratio
1	−0.52	1.34
2	−0.13	0.40
3	−0.50	1.90
4	−1.01	1.50
5	−0.26	1.35
6	−0.89	0.50
7	−0.42	1.00
8	−0.23	1.50
9	−0.60	1.45

- **A** Formulate null and alternative hypotheses consistent with the verbal description of the research goal.
- **B** Identify and justify the test statistic for conducting a test of the hypotheses in Part A.
- **C** Determine whether to reject the null hypothesis at the 0.05 level of significance if the critical values are ±2.306.

22 All else equal, is specifying a smaller significance level in a hypothesis test likely to increase the probability of a:

	Type I error?	Type II error?
A.	No	No
B.	No	Yes
C.	Yes	No

23 The probability of correctly rejecting the null hypothesis is the:

A *p*-value.

B power of a test.

C level of significance.

24 The power of a hypothesis test is:

A equivalent to the level of significance.

B the probability of not making a Type II error.

C unchanged by increasing a small sample size.

25 When making a decision about investments involving a statistically significant result, the:

A economic result should be presumed to be meaningful.

B statistical result should take priority over economic considerations.

C economic logic for the future relevance of the result should be further explored.

26 An analyst tests the profitability of a trading strategy with the null hypothesis that the average abnormal return before trading costs equals zero. The calculated *t*-statistic is 2.802, with critical values of ±2.756 at significance level $\alpha = 0.01$. After considering trading costs, the strategy's return is near zero. The results are *most likely*:

A statistically but not economically significant.

B economically but not statistically significant.

C neither statistically nor economically significant.

27 Which of the following statements is correct with respect to the *p*-value?

A It is a less precise measure of test evidence than rejection points.

B It is the largest level of significance at which the null hypothesis is rejected.

C It can be compared directly with the level of significance in reaching test conclusions.

28 Which of the following represents a correct statement about the *p*-value?

A The *p*-value offers less precise information than does the rejection points approach.

B A larger *p*-value provides stronger evidence in support of the alternative hypothesis.

C A *p*-value less than the specified level of significance leads to rejection of the null hypothesis.

29 Which of the following statements on *p*-value is correct?

A The *p*-value indicates the probability of making a Type II error.

B The lower the *p*-value, the weaker the evidence for rejecting the H_0.

C The *p*-value is the smallest level of significance at which H_0 can be rejected.

30 The following table shows the significance level (α) and the *p*-value for two hypothesis tests.

	α	*p*-Value
Test 1	0.02	0.05
Test 2	0.05	0.02

In which test should we reject the null hypothesis?

A Test 1 only

B Test 2 only

C Both Test 1 and Test 2

31 Which of the following tests of a hypothesis concerning the population mean is *most* appropriate?

A A z-test if the population variance is unknown and the sample is small

B A z-test if the population is normally distributed with a known variance

C A t-test if the population is non-normally distributed with unknown variance and a small sample

32 For a small sample from a normally distributed population with unknown variance, the *most* appropriate test statistic for the mean is the:

A z-statistic.

B t-statistic.

C χ^2 statistic.

33 An investment consultant conducts two independent random samples of five-year performance data for US and European absolute return hedge funds. Noting a return advantage of 50 bps for US managers, the consultant decides to test whether the two means are different from one another at a 0.05 level of significance. The two populations are assumed to be normally distributed with unknown but equal variances. Results of the hypothesis test are contained in the following tables.

	Sample Size	Mean Return (%)	Standard Deviation
US managers	50	4.7	5.4
European managers	50	4.2	4.8

Null and alternative hypotheses	H_0: $\mu_{US} - \mu_E = 0$; H_a: $\mu_{US} - \mu_E \neq 0$
Calculated test statistic	0.4893
Critical value rejection points	±1.984

The mean return for US funds is μ_{US}, and μ_E is the mean return for European funds.

The results of the hypothesis test indicate that the:

A null hypothesis is not rejected.

B alternative hypothesis is statistically confirmed.

C difference in mean returns is statistically different from zero.

34 A pooled estimator is used when testing a hypothesis concerning the:

A equality of the variances of two normally distributed populations.

B difference between the means of two at least approximately normally distributed populations with unknown but assumed equal variances.

C difference between the means of two at least approximately normally distributed populations with unknown and assumed unequal variances.

35 When evaluating mean differences between two dependent samples, the *most* appropriate test is a:

A z-test.

B chi-square test.

C paired comparisons test.

36 A chi-square test is *most* appropriate for tests concerning:

A a single variance.

B differences between two population means with variances assumed to be equal.

C differences between two population means with variances assumed to not be equal.

37 Which of the following should be used to test the difference between the variances of two normally distributed populations?

A t-test

B F-test

C Paired comparisons test

38 Jill Batten is analyzing how the returns on the stock of Stellar Energy Corp. are related with the previous month's percentage change in the US Consumer Price Index for Energy (CPIENG). Based on 248 observations, she has computed the sample correlation between the Stellar and CPIENG variables to be –0.1452. She also wants to determine whether the sample correlation is significantly different from zero. The critical value for the test statistic at the 0.05 level of significance is approximately 1.96. Batten should conclude that the statistical relationship between Stellar and CPIENG is:

A significant, because the calculated test statistic is outside the bounds of the critical values for the test statistic.

B significant, because the calculated test statistic has a lower absolute value than the critical value for the test statistic.

C insignificant, because the calculated test statistic is outside the bounds of the critical values for the test statistic.

39 In which of the following situations would a nonparametric test of a hypothesis *most likely* be used?

A The sample data are ranked according to magnitude.

B The sample data come from a normally distributed population.

C The test validity depends on many assumptions about the nature of the population.

40 An analyst is examining the monthly returns for two funds over one year. Both funds' returns are non-normally distributed. To test whether the mean return of one fund is greater than the mean return of the other fund, the analyst can use:

A a parametric test only.

B a nonparametric test only.

C both parametric and nonparametric tests.

41 An analyst group follows 250 firms and classifies them in two dimensions. First, they use dividend payment history and earnings forecasts to classify firms into one of three groups, with 1 indicating the dividend stars and 3 the dividend laggards. Second, they classify firms on the basis of financial leverage, using debt ratios, debt features, and corporate governance to classify the firms into three groups, with 1 indicating the least risky firms based on financial leverage and 3 indicating the riskiest. The classification of the 250 firms is as follows:

Financial Leverage Group	Dividend Group		
	1	2	3
1	40	40	40
2	30	10	20
3	10	50	10

A What are the null and alternative hypotheses to test whether the dividend and financial leverage groups are independent of one another?

B What is the appropriate test statistic to use in this type of test?

C If the critical value for the 0.05 level of significance is 9.4877, what is your conclusion?

42 Which of the following statements is correct regarding the chi-square test of independence?

A The test has a one-sided rejection region.

B The null hypothesis is that the two groups are dependent.

C If there are two categories, each with three levels or groups, there are six degrees of freedom.

SOLUTIONS

1 C is correct. Together, the null and alternative hypotheses account for all possible values of the parameter. Any possible values of the parameter not covered by the null must be covered by the alternative hypothesis (e.g., H_0: $\mu \le 5$ versus H_a: $\mu > 5$).

2 **A** The appropriate test statistic is a t-statistic, $t = \dfrac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$, with $n - 1 = 15 - 1 = 14$ degrees of freedom. A t-statistic is correct when the sample comes from an approximately normally distributed population with unknown variance.

B The appropriate test statistic is a t-statistic, $t = \dfrac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$, with $40 - 1 = 39$ degrees of freedom. A t-statistic is theoretically correct when the sample comes from a normally distributed population with unknown variance.

C The appropriate test statistic is a z-statistic, $z = \dfrac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$, because the sample comes from a normally distributed population with a known variance.

D The appropriate test statistic is chi-square, $x^2 = \dfrac{s^2(n-1)}{\sigma_0^2}$, with $50 - 1 = 49$ degrees of freedom.

E The appropriate test statistic is the F-statistic, $F = \sigma_1^2 \big/ \sigma_2^2$, with 29 and 39 degrees of freedom.

F The appropriate test statistic is a t-statistic using a pooled estimate of the population variance: $t = \dfrac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}}$, where $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$. The t-statistic has $25 + 30 - 2 = 53$ degrees of freedom. This statistic is appropriate because the populations are normally distributed with unknown variances; because the variances are assumed to be equal, the observations can be pooled to estimate the common variance. The requirement of independent samples for using this statistic has been met.

3 We make the decision either by comparing the calculated test statistic with the critical values or by comparing the p-value for the calculated test statistic with the level of significance.

A Reject the null hypothesis because the calculated test statistic is outside the bounds of the critical values.

B The calculated t-statistic is in the rejection region that is defined by +1.679, so we reject the null hypothesis.

C The p-value corresponding to the calculated test statistic is less than the level of significance, so we reject the null hypothesis.

D We fail to reject because the *p*-value for the calculated test statistic is greater than what is tolerated with a 2% level of significance.

4 We make the decision either by comparing the calculated test statistic with the critical values or by comparing the *p*-value for the calculated test statistic with the level of significance.

A The calculated chi-square falls between the two critical values, so we fail to reject the null hypothesis.

B The *p*-value for the calculated test statistic is less than the level of significance (the 5%), so we reject the null hypothesis.

C The calculated *F*-statistic falls outside the bounds of the critical *F*-values, so we reject the null hypothesis.

D The calculated chi-square exceeds the critical value for this right-side test, so we reject the null hypothesis.

5 **A** As stated in the text, we often set up the "hoped for" or "suspected" condition as the alternative hypothesis. Here, that condition is that the population value of Willco's mean annual net income exceeds $24 million. Thus, we have H_0: $\mu \leq 24$ versus H_a: $\mu > 24$.

B Given that net income is normally distributed with unknown variance, the appropriate test statistic is $t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} = 1.469694$ with $n - 1 = 6 - 1 = 5$ degrees of freedom.

C We reject the null if the calculated *t*-statistic is greater than 2.015. The calculated *t*-statistic is $t = \dfrac{30 - 24}{10/\sqrt{6}} = 1.469694$. Because the calculated test statistic does not exceed 2.015, we fail to reject the null hypothesis. There is not sufficient evidence to indicate that the mean net income is greater than $24 million.

6 **A** H_0: $\mu = 0$ versus H_a: $\mu \neq 0$.

B The *t*-test is based on $t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$ with $n - 1 = 10 - 1 = 9$ degrees of freedom. At the 0.05 significance level, we reject the null if the calculated *t*-statistic is outside the bounds of ±2.262 (from the table for 9 degrees of freedom and 0.025 in the right side of the distribution). For Analyst A, we have a calculated test statistic of $t = \dfrac{0.05 - 0}{0.10/\sqrt{10}} = 1.58114$. We, therefore, fail to reject the null hypothesis at the 0.05 level.

C For Analyst B, the *t*-test is based on *t* with 15 − 1 = 14 degrees of freedom. At the 0.05 significance level, we reject the null if the calculated *t*-statistic is outside the bounds of ±2.145 (from the table for 14 degrees of freedom). The calculated test statistic is $t = \dfrac{0.02 - 0}{0.09/\sqrt{10}} = 0.86066$. Because 0.86066 is within the range of ±2.145, we fail to reject the null at the 0.05 level.

7 **A** Stating the suspected condition as the alternative hypothesis, we have

$$H_0: \mu_A - \mu_B \leq 0 \text{ versus } H_a: \mu_A - \mu_B > 0,$$

where

μ_A = the population mean value of Analyst A's forecast errors
μ_B = the population mean value of Analyst B's forecast errors

B We have two normally distributed populations with unknown variances. Based on the samples, it is reasonable to assume that the population variances are equal. The samples are assumed to be independent; this assumption is reasonable because the analysts cover different industries. The appropriate test statistic is t using a pooled estimate of the common variance: $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$, where $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$. The number of degrees of freedom is $n_A + n_B - 2 = 10 + 15 - 2 = 23$.

C For df = 23, according to the table, the rejection point for a one-sided (right side) test at the 0.05 significance level is 1.714.

D We first calculate the pooled estimate of variance:

$$s_p^2 = \frac{(10-1)0.01 + (15-1)0.0081}{10 + 15 - 2} = 0.0088435.$$

We then calculate the t-distributed test statistic:

$$t = \frac{(0.05 - 0.02) - 0}{\sqrt{\frac{0.0088435}{10} + \frac{0.0088435}{15}}} = \frac{0.03}{0.0383916} = 0.78142.$$

Because 0.78142 < 1.714, we fail to reject the null hypothesis. There is not sufficient evidence to indicate that the mean for Analyst A exceeds that for Analyst B.

8 **A** We test H_0: $\mu_d = 0$ versus H_a: $\mu_d \neq 0$, where μ_d is the population mean difference.

B This is a paired comparisons t-test with $n - 1 = 480 - 1 = 479$ degrees of freedom. At the 0.05 significance level, we reject the null hypothesis if the calculated t is less than −1.96 or greater than 1.96.

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}} = \frac{-0.258 - 0}{3.752/\sqrt{480}} = \frac{-0.258}{0.171255} = -1.506529, \text{ or } -1.51.$$

Because the calculate t-statistic is between ±1.96, we do not reject the null hypothesis that the mean difference between the returns on the S&P 500 and small-cap stocks during the entire sample period was zero.

C This t-test now has $n - 1 = 240 - 1 = 239$ degrees of freedom. At the 0.05 significance level, we reject the null hypothesis if the calculated t is less than −1.96 or greater than 1.96.

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}} = \frac{-0.640 - 0}{4.096/\sqrt{240}} = \frac{-0.640}{0.264396} = -2.420615, \text{ or } -2.42.$$

Because −2.42 < −1.96, we reject the null hypothesis at the 0.05 significance level. We conclude that during this subperiod, small-cap stocks significantly outperformed the S&P 500.

D This t-test has $n - 1 = 240 - 1 = 239$ degrees of freedom. At the 0.05 significance level, we reject the null hypothesis if the calculated t-statistic is less than −1.96 or greater than 1.96. The calculated test statistic is

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}} = \frac{0.125 - 0}{3.339/\sqrt{240}} = \frac{0.125}{0.215532} = 0.579962, \text{ or } 0.58.$$

At the 0.05 significance level, because the calculated test statistic of 0.58 is between ±1.96, we fail to reject the null hypothesis for the second subperiod.

9 **A** We have a "less than" alternative hypothesis, where σ^2 is the variance of return on the portfolio. The hypotheses are H_0: $\sigma^2 \geq 400$ versus H_a: $\sigma^2 < 400$, where 400 is the hypothesized value of variance, σ_0^2. This means that the rejection region is on the left side of the distribution.

B The test statistic is chi-square distributed with $10 - 1 = 9$ degrees of freedom: $Ç^2 = \frac{(n-1)s^2}{Ã_0^2}$.

C The test statistic is calculated as

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{9 \times 15^2}{400} = \frac{2{,}025}{400} = 5.0625, \text{ or } 5.06.$$

Because 5.06 is not less than 3.325, we do not reject the null hypothesis; the calculated test statistic falls to the right of the critical value, where the critical value separates the left-side rejection region from the region where we fail to reject.

We can determine the critical value for this test using software:

Excel [CHISQ.INV(0.05,9)]
R [qchisq(.05,9)]
Python [from scipy.stats import chi2 and chi2.ppf(.05,9)]

We can determine the p-value for the calculated test statistic of 17.0953 using software:

Excel [CHISQ.DIST(5.06,9,TRUE)]
R [pchisq(5.06,9,lower.tail=TRUE)]
Python [from scipy.stats import chi2 and chi2.cdf(5.06,9)]

10 **A** We have a "not equal to" alternative hypothesis:

$$H_0\text{: } \sigma^2_{Before} = \sigma^2_{After} \text{ versus } H_a\text{: } \sigma^2_{Before} \neq \sigma^2_{After}$$

B To test a null hypothesis of the equality of two variances, we use an F-test:

$$F = \frac{s_1^2}{s_2^2}.$$

$F = 22.367/15.795 = 1.416$, with $120 - 1 = 119$ numerator and $120 - 1 = 119$ denominator degrees of freedom. Because this is a two-tailed test, we use critical values for the $0.05/2 = 0.025$ level. The calculated test statistic falls within the bounds of the critical values (that is, between 0.6969 and 1.4349), so we fail to reject the null hypothesis; there is not enough evidence to indicate that the variances are different before and after the disruption. Note that we could also have formed the F-statistic as 15.796/22.367 = 0.706 and draw the same conclusion.

We could also use software to calculate the critical values:

Excel [F.INV(0.025,119,119) and F.INV(0.975,119,119)]

R [qf(c(.025,.975),119,119)]

Python from scipy.stats import f and f.ppf
[(.025,119,119) and
f.ppf(.975,119,119)]

Additionally, we could use software to calculate the p-value of the calculated test statistic, which is 5.896% and then compare it with the level of significance:

Excel [(1-F.DIST(22.367/15.796,119,119,TRUE))*2 or
F.DIST(15.796/22.367,119,119,TRUE)*2]

R [(1-pf(22.367/15.796,119,119))*2 or
pf(15.796/22.367,119,119)*2]

Python from scipy.stats import f and f.cdf
[(15.796/22.367,119,119)*2 or
(1-f.cdf(22.367/15.796,119,119))*2]

11 The hypotheses are H_0: ρ = 0 and H_a: ρ ≠ 0. The calculated test statistics are based on the formula $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$. For example, the calculated t-statistic for the correlation of Fund 3 and Fund 4 is

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.3102\sqrt{36-2}}{\sqrt{1-0.3102^2}} = 1.903.$$

Repeating this calculation for the entire matrix of correlations gives the following:

	Calculated *t*-Statistics for Correlations				
	Fund 1	**Fund 2**	**Fund 3**	**Fund 4**	**S&P 500**
Fund 1					
Fund 2	13.997				
Fund 3	3.165	2.664			
Fund 4	5.897	6.116	1.903		
S&P 500	8.600	8.426	4.142	6.642	

With critical values of ±2.032, with the exception of the correlation between Fund 3 and Fund 4 returns, we reject the null hypothesis for these correlations. In other words, there is sufficient evidence to indicate that the correlations are different from zero, with the exception of the correlation of returns between Fund 3 and Fund 4.

We could use software to determine the critical values:

Excel [T.INV(0.025,34) and T.INV(0.975,34)]

R [qt(c(.025,.975),34)]

Python [from scipy.stats import t and t.ppf(.025,34) and t.ppf(.975,34)]

We could also use software to determine the *p*-value for the calculated test statistic to enable a comparison with the level of significance. For example, for $t = 2.664$, the *p*-value is 0.01172:

Excel [(1-T.DIST(2.664,34,TRUE))*2]

R [(1-pt(2.664,34))*2]

Python [from scipy.stats import t and (1-t.cdf(2.664,34))*2]

12 B is correct. The critical value in a decision rule is the rejection point for the test. It is the point with which the test statistic is compared to determine whether to reject the null hypothesis, which is part of the fourth step in hypothesis testing.

13 A is correct. The null hypothesis and the alternative hypothesis are complements of one another and together are exhaustive; that is, the null and alternative hypotheses combined consider all the possible values of the population parameter.

14 A is correct. If the population sampled has unknown variance and the sample is large, a *z*-test may be used. Hypotheses involving "greater than" or "less than" postulations are one sided (one tailed). In this situation, the null and alternative hypotheses are stated as H_0: $\mu \leq 6\%$ and H_a: $\mu > 6\%$, respectively. A one-tailed *t*-test is also acceptable in this case, and the rejection region is on the right side of the distribution.

15 C is correct. For a one-tailed hypothesis test, there is a 5% rejection region in one tail of the distribution.

16 B is correct. One-tailed tests in which the alternative is "greater than" or "less than" represent the beliefs of the researcher more firmly than a "not equal to" alternative hypothesis.

17 A is correct. Calculated using a sample, a test statistic is a quantity whose value is the basis for deciding whether to reject the null hypothesis.

18 A is correct. The definition of a Type I error is when a true null hypothesis is rejected.

19 B is correct. A Type II error occurs when a false null hypothesis is not rejected.

20 B is correct. The level of significance is used to establish the rejection points of the hypothesis test.

21 **A** We have a "not equal to" alternative hypothesis:

$$H_0: \rho = 0 \text{ versus } H_a: \rho \neq 0$$

B Mutual fund expense ratios are bounded from above and below; in practice, there is at least a lower bound on alpha (as any return cannot be less than −100%), and expense ratios cannot be negative. These variables may not be normally distributed, and the assumptions of a parametric test are not likely to be fulfilled. Thus, a nonparametric test appears to be appropriate.

We would use the nonparametric Spearman rank correlation coefficient to conduct the test: $r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}$ with the t-distributed test statistic of $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$.

C The calculation of the Spearman rank correlation coefficient is given in the following table.

Mutual Fund	Alpha	Expense Ratio	Rank by Alpha	Rank by Expense Ratio	Difference in Rank	Difference Squared
1	−0.52	1.34	6	6	0	0
2	−0.13	0.40	1	9	−8	64
3	−0.50	1.90	5	1	4	16
4	−1.01	1.50	9	2	7	49
5	−0.26	1.35	3	5	−2	4
6	−0.89	0.50	8	8	0	0
7	−0.42	1.00	4	7	−3	9
8	−0.23	1.50	2	2	0	0
9	−0.60	1.45	7	4	3	9
						151

$$r_s = 1 - \frac{6(151)}{9(80)} = -0.25833.$$

The calculated test statistic, using the t-distributed test statistic $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ is $t = \frac{-0.25833\sqrt{7}}{\sqrt{1 - 0.066736}} = \frac{-0.683486}{0.9332638} = -0.7075$. On the basis of this value falling within the range of ±2.306, we fail to reject the null hypothesis that the Spearman rank correlation coefficient is zero.

22 B is correct. Specifying a smaller significance level decreases the probability of a Type I error (rejecting a true null hypothesis) but increases the probability of a Type II error (not rejecting a false null hypothesis). As the level of significance decreases, the null hypothesis is less frequently rejected.

23 B is correct. The power of a test is the probability of rejecting the null hypothesis when it is false.

24 B is correct. The power of a hypothesis test is the probability of correctly rejecting the null when it is false. Failing to reject the null when it is false is a Type II error. Thus, the power of a hypothesis test is the probability of not committing a Type II error.

25 C is correct. When a statistically significant result is also economically meaningful, one should further explore the logic of why the result might work in the future.

26 A is correct. The hypothesis is a two-tailed formulation. The t-statistic of 2.802 falls outside the critical rejection points of less than −2.756 and greater than 2.756. Therefore, the null hypothesis is rejected; the result is statistically significant. However, despite the statistical results, trying to profit on the strategy is not likely to be economically meaningful because the return is near zero after transaction costs.

27 C is correct. When directly comparing the *p*-value with the level of significance, it can be used as an alternative to using rejection points to reach conclusions on hypothesis tests. If the *p*-value is smaller than the specified level of significance, the null hypothesis is rejected. Otherwise, the null hypothesis is not rejected.

28 C is correct. The *p*-value is the smallest level of significance at which the null hypothesis can be rejected for a given value of the test statistic. The null hypothesis is rejected when the *p*-value is less than the specified significance level.

29 C is correct. The *p*-value is the smallest level of significance (α) at which the null hypothesis can be rejected.

30 B is correct. The *p*-value is the smallest level of significance (α) at which the null hypothesis can be rejected. If the *p*-value is less than α, the null is rejected. In Test 1, the *p*-value exceeds the level of significance, whereas in Test 2, the *p*-value is less than the level of significance.

31 B is correct. The *z*-test is theoretically the correct test to use in those limited cases when testing the population mean of a normally distributed population with known variance.

32 B is correct. A *t*-statistic is the most appropriate for hypothesis tests of the population mean when the variance is unknown and the sample is small but the population is normally distributed.

33 A is correct. The calculated *t*-statistic value of 0.4893 falls within the bounds of the critical *t*-values of ±1.984. Thus, H_0 cannot be rejected; the result is not statistically significant at the 0.05 level.

34 B is correct. The assumption that the variances are equal allows for the combining of both samples to obtain a pooled estimate of the common variance.

35 C is correct. A paired comparisons test is appropriate to test the mean differences of two samples believed to be dependent.

36 A is correct. A chi-square test is used for tests concerning the variance of a single normally distributed population.

37 B is correct. An *F*-test is used to conduct tests concerning the difference between the variances of two normally distributed populations with random independent samples.

38 A is correct. The calculated test statistic is

$$\begin{aligned} t &= \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \\ &= \frac{-0.1452\sqrt{248-2}}{\sqrt{1-\left(-0.1452\right)^2}} = -2.30177. \end{aligned}$$

Because the value of $t = -2.30177$ is outside the bounds of ±1.96, we reject the null hypothesis of no correlation and conclude that there is enough evidence to indicate that the correlation is different from zero.

39 A is correct. A nonparametric test is used when the data are given in ranks.

40 B is correct. There are only 12 (monthly) observations over the one year of the sample and thus the samples are small. Additionally, the funds' returns are non-normally distributed. Therefore, the samples do not meet the distributional assumptions for a parametric test. The Mann–Whitney U test (a nonparametric test) could be used to test the differences between population means.

41 **A** The hypotheses are as follows:

H_0: Dividend and financial leverage ratings are not related, so these groupings are independent.

H_a: Dividend and financial leverage ratings are related, so these groupings are not independent.

B The appropriate test statistic is $\chi^2 = \sum_{i=1}^{m} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$, where O_{ij} represents the observed frequency for the i and j group and E_{ij} represents the expected frequency for the i and j group if the groupings are independent.

C The expected frequencies based on independence are as follows:

Financial Leverage Group	Dividend Group 1	2	3	Sum
1	38.4	48	33.6	120
2	19.2	24	16.8	60
3	22.4	28	19.6	70
Sum	80	100	70	250

The scaled squared deviations for each combination of financial leverage and dividend grouping are:

Financial Leverage Group	Dividend Group 1	2	3
1	0.06667	1.33333	1.21905
2	6.07500	8.16667	0.60952
3	6.86429	17.28571	4.70204

The sum of these scaled squared deviations is the calculated chi-square statistic of 46.3223. Because this calculated value exceeds the critical value of 9.4877, we reject the null hypothesis that these groupings are independent.

42 A is correct. The test statistic comprises squared differences between the observed and expected values, so the test involves only one side, the right side. B is incorrect because the null hypothesis is that the groups are independent, and C is incorrect because with three levels of groups for the two categorical variables, there are four degrees of freedom.

READING

7

Introduction to Linear Regression

by Pamela Peterson Drake, PhD, CFA

Pamela Peterson Drake, PhD, CFA, is at James Madison University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe a simple linear regression model and the roles of the dependent and independent variables in the model;
☐	**b.** describe the least squares criterion, how it is used to estimate regression coefficients, and their interpretation;
☐	**c.** explain the assumptions underlying the simple linear regression model, and describe how residuals and residual plots indicate if these assumptions may have been violated;
☐	**d.** calculate and interpret the coefficient of determination and the *F*-statistic in a simple linear regression;
☐	**e.** describe the use of analysis of variance (ANOVA) in regression analysis, interpret ANOVA results, and calculate and interpret the standard error of estimate in a simple linear regression;
☐	**f.** formulate a null and an alternative hypothesis about a population value of a regression coefficient, and determine whether the null hypothesis is rejected at a given level of significance;
☐	**g.** calculate and interpret the predicted value for the dependent variable, and a prediction interval for it, given an estimated linear regression model and a value for the independent variable;
☐	**h.** describe different functional forms of simple linear regressions.

SIMPLE LINEAR REGRESSION

1

a describe a simple linear regression model and the roles of the dependent and independent variables in the model

© 2021 CFA Institute. All rights reserved.

Financial analysts often need to examine whether a variable is useful for explaining another variable. For example, the analyst may want to know whether earnings growth, or perhaps cash flow growth, helps explain the company's value in the marketplace. **Regression analysis** is a tool for examining this type of issue.

Suppose an analyst is examining the return on assets (ROA) for an industry and observes the ROA for the six companies shown in Exhibit 1. The average of these ROAs is 12.5%, but the range is from 4% to 20%.

Exhibit 1 Return on Assets of Selected Companies

Company	ROA (%)
A	6
B	4
C	15
D	20
E	10
F	20

In trying to understand why the ROAs differ among these companies, we could look at why the ROA of Company A differs from that of Company B, why the ROA of Company A differs from that of Company D, why the ROA of Company F differs from that of Company C, and so on, comparing each pair of ROAs. A way to make this a simpler exploration is to try to understand why each company's ROA differs from the mean ROA of 12.5%. We look at the sum of squared deviations of the observations from the mean to capture variations in ROA from their mean. Let Y represent the variable that we would like to explain, which in this case is the return on assets. Let Y_i represent an observation of a company's ROA, and let $\overline{Y}$ represent the mean ROA for the sample of size n. We can describe the variation of the ROAs as

$$\text{Variation of } Y = \sum\nolimits_{i=1}^{n}(Y_i - \overline{Y})^2. \tag{1}$$

Our goal is to understand what drives these returns on assets or, in other words, what explains the variation of Y. The variation of Y is often referred to as the **sum of squares total (SST)**, or the total sum of squares.

We now ask whether it is possible to explain the variation of the ROA using another variable that also varies among the companies; note that if this other variable is constant or random, it would not serve to explain why the ROAs differ from one another. Suppose the analyst believes that the capital expenditures in the previous period, scaled by the prior period's beginning property, plant, and equipment, are a driver for the ROA variable. Let us represent this scaled capital expenditures variable as CAPEX, as we show in Exhibit 2.

Exhibit 2 Return on Assets and Scaled Capital Expenditures

Company	ROA (%)	CAPEX (%)
A	6.0	0.7
B	4.0	0.4
C	15.0	5.0

Exhibit 2 (Continued)

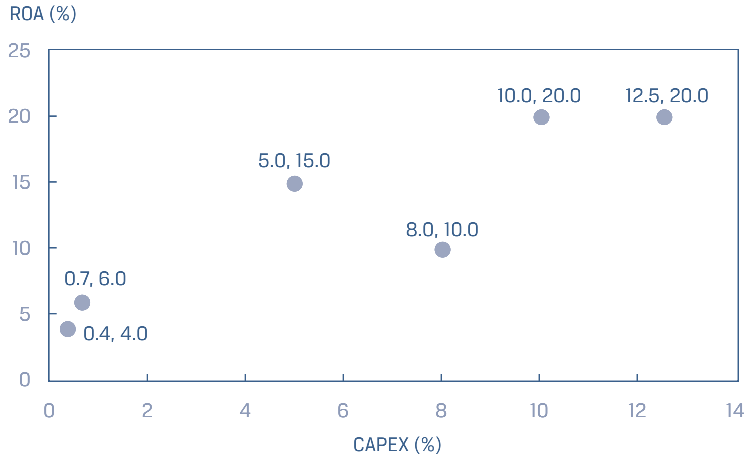
Company	ROA (%)	CAPEX (%)
D	20.0	10.0
E	10.0	8.0
F	20.0	12.5
Arithmetic mean	12.50	6.10

The variation of X, in this case CAPEX, is calculated as

$$\text{Variation of } X = \sum_{i=1}^{n}(X_i - \bar{X})^2. \tag{2}$$

We can see the relation between ROA and CAPEX in the **scatter plot** (or scattergram) in Exhibit 3, which represents the two variables in two dimensions. Typically, we present the variable whose variation we want to explain along the vertical axis and the variable whose variation we want to use to explain that variation along the horizontal axis. Each point in this scatter plot represents a paired observation that consists of CAPEX and ROA. From a casual visual inspection, there appears to be a positive relation between ROA and CAPEX: Companies with higher CAPEX tend to have a higher ROA.

Exhibit 3 Scatter Plot of ROA and CAPEX



In the ROA example, we use the capital expenditures to explain the returns on assets. We refer to the variable whose variation is being explained as the **dependent variable**, or the explained variable; it is typically denoted by Y. We refer to the variable whose variation is being used to explain the variation of the dependent variable as the **independent variable**, or the explanatory variable; it is typically denoted by X. Therefore, in our example, the ROA is the dependent variable (Y) and CAPEX is the independent variable (X).

A common method for relating the dependent and independent variables is through the estimation of a linear relationship, which implies describing the relation between the two variables as represented by a straight line. If we have only one independent variable, we refer to the method as **simple linear regression (SLR)**; if we have more than one independent variable, we refer to the method as multiple regression.

Linear regression allows us to test hypotheses about the relationship between two variables, by quantifying the strength of the relationship between the two variables, and to use one variable to make predictions about the other variable. Our focus is on linear regression with a single independent variable—that is, simple linear regression.

EXAMPLE 1

Identifying the Dependent and Independent Variables in a Regression

An analyst is researching the relationship between corporate earnings growth and stock returns. Specifically, she is interested in whether earnings revisions affect stock price returns in the same period. She collects five years of monthly data on "Wall Street" EPS revisions for a sample of 100 companies and on their monthly stock price returns over the five-year period.

What are the dependent and independent variables in her model?

Solution

The dependent variable is monthly stock price returns, and the independent variable is Wall Street EPS revisions, since in the analyst's model, the variation in monthly stock price returns is being explained by the variation in EPS revisions.

2 ESTIMATING THE PARAMETERS OF A SIMPLE LINEAR REGRESSION

b describe the least squares criterion, how it is used to estimate regression coefficients, and their interpretation

2.1 The Basics of Simple Linear Regression

Regression analysis begins with the dependent variable, the variable whose variation you are seeking to explain. The independent variable is the variable whose variation you are using to explain changes in the dependent variable. For example, you might try to explain small-stock returns (the dependent variable) using returns to the S&P 500 Index (the independent variable). Or you might try to explain a country's inflation rate (the dependent variable) as a function of growth in its money supply (the independent variable).

As the name implies, **linear regression** assumes a linear relationship between the dependent and the independent variables. The goal is to fit a line to the observations on Y and X to minimize the squared deviations from the line; this is the least squares criterion—hence, the name least squares regression. Because of its common use, linear regression is often referred to as ordinary least squares (OLS) regression.

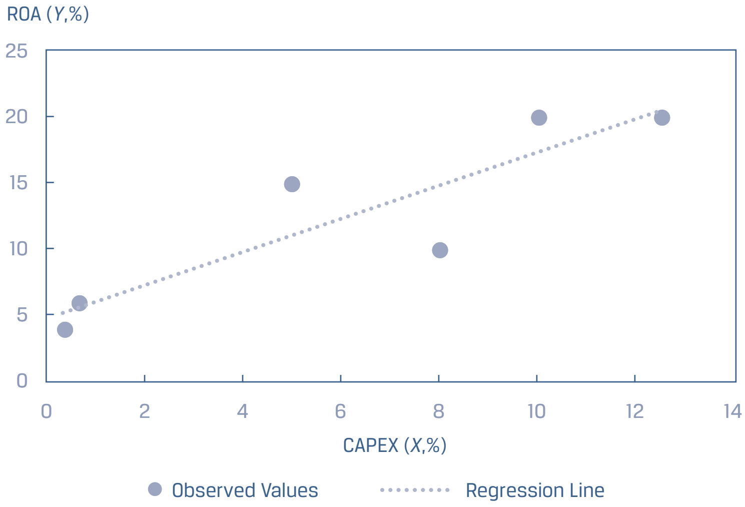
Using notation, the linear relation between the dependent and independent variables is described as

$$Y_i = b_0 + b_1 X_i + \varepsilon_i,\ i = 1, \ldots, n. \quad (3)$$

Equation 3 is a model that does not require that every (Y, X) pair for an observation fall on the regression line. This equation states that the dependent variable, Y, is equal to the **intercept**, b_0, plus a **slope coefficient**, b_1, multiplied by the independent variable, X, plus an **error term**, ε. The error term, or simply the error, represents the difference between the observed value of Y and that expected from the true underlying population relation between Y and X. We refer to the intercept, b_0, and the slope coefficient, b_1, as the **regression coefficients**. A way that we often describe this simple linear regression relation is that Y is regressed on X.

Consider the ROA and CAPEX scatter diagram from Exhibit 3, which we elaborate on in Exhibit 4 by including the fitted regression line. This line represents the average relationship between ROA and CAPEX; not every observation falls on the line, but the line describes the mean relation between ROA and CAPEX.

Exhibit 4 Fitted Regression Line of ROA and CAPEX



2.2 Estimating the Regression Line

We cannot observe the population parameter values b_0 and b_1 in a regression model. Instead, we observe only $\hat{b}_0$ and $\hat{b}_1$, which are estimates (as indicated by the "hats" above the coefficients) of the population parameters based on the sample. Thus, predictions must be based on the parameters' estimated values, and testing is based on estimated values in relation to the hypothesized population values.

We estimate the regression line as the line that best fits the observations. *In simple linear regression, the estimated intercept, $\hat{b}_0$, and slope, $\hat{b}_1$, are such that the sum of the squared vertical distances from the observations to the fitted line is minimized.* The focus is on the sum of the squared differences between the observations on Y_i and the corresponding estimated value, $\hat{Y}_i$, on the regression line.

We represent the value of the dependent variable for the ith observation that falls on the line as $\hat{Y}_i$, which is equal to $\hat{b}_0 + \hat{b}_1 X_i$. The $\hat{Y}_i$ is what the estimated value of the Y variable would be for the ith observation based on the mean relationship between

Y and X. The **residual** for the ith observation, e_i, is how much the observed value of Y_i differs from the $\hat{Y}_i$ estimated using the regression line: $e_i = Y_i - \hat{Y}_i$. Note the subtle difference between the error term and the residual: The error term refers to the true underlying population relationship, whereas the residual refers to the fitted linear relation based on the sample.

Fitting the line requires minimizing the sum of the squared residuals, the **sum of squares error (SSE)**, also known as the residual sum of squares:

$$\text{Sum of squares error} = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

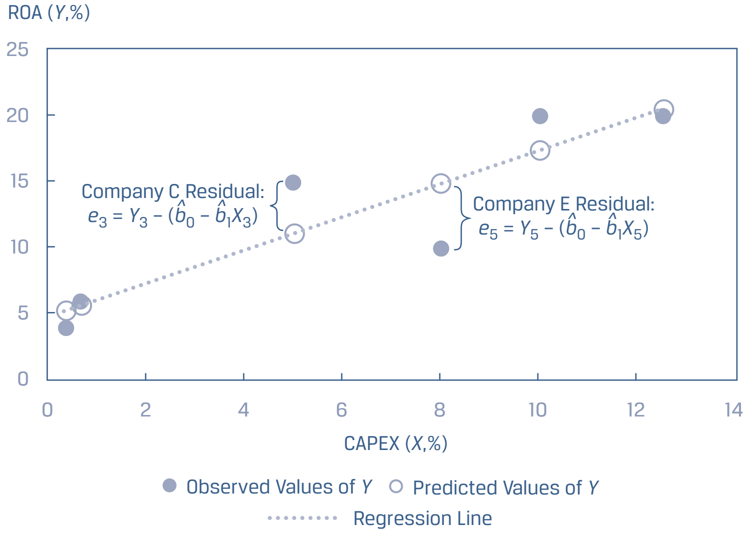
$$= \sum_{i=1}^{n}\left[Y_i - \left(\hat{b}_0 + \hat{b}_1 X_i\right)\right]^2 \qquad (4)$$

$$= \sum_{i=1}^{n} e_i^2.$$

Using least squares regression to estimate the values of the population parameters of b_0 and b_1, we can fit a line through the observations of X and Y that explains the value that Y takes for any particular value of X.

As seen in Exhibit 5, the residuals are represented by the vertical distances from the fitted line (see the third and fifth observations, Companies C and E, respectively) and are, therefore, in the units of measurement represented by the dependent variable. The residual is in the same unit of measurement as the dependent variable: If the dependent variable is in euros, the error term is in euros, and if the dependent variable is in growth rates, the error term is in growth rates.

Exhibit 5 Residuals of the Linear Regression



How do we calculate the intercept ($\hat{b}_0$) and the slope ($\hat{b}_1$) for a given sample of (Y, X) pairs of observations? The slope is the ratio of the covariance between Y and X to the variance of X, where $\overline{Y}$ is the mean of the Y variable and $\overline{X}$ is the mean of X variable:

$$\hat{b}_1 = \frac{\text{Covariance of } Y \text{ and } X}{\text{Variance of } X} = \frac{\dfrac{\sum_{i=1}^{n}(Y_i - \overline{Y})(X_i - \overline{X})}{n-1}}{\dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}.$$

Simplifying,

$$\hat{b}_1 = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})(X_i - \overline{X})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}. \tag{5}$$

Once we estimate the slope, we can then estimate the intercept using the mean of Y and the mean of X:

$$\hat{b}_0 = \overline{Y} - \hat{b}_1\overline{X}. \tag{6}$$

We show the calculation of the slope and the intercept in Exhibit 6.

Exhibit 6 Estimating Slope and Intercept for the ROA Model

Company	ROA (Y_i)	CAPEX (X_i)	$(Y_i - \overline{Y})^2$	$(X_i - \overline{X})^2$	$(Y_i - \overline{Y})(X_i - \overline{X})$
A	6.0	0.7	42.25	29.16	35.10
B	4.0	0.4	72.25	32.49	48.45
C	15.0	5.0	6.25	1.21	−2.75
D	20.0	10.0	56.25	15.21	29.25
E	10.0	8.0	6.25	3.61	−4.75
F	20.0	12.5	56.25	40.96	48.00
Sum	75.0	36.6	239.50	122.64	153.30
Arithmetic mean	12.5	6.100			

Slope coefficient: $\hat{b}_1 = \dfrac{153.30}{122.64} = 1.25.$

Intercept: $\hat{b}_0 = 12.5 - (1.25 \times 6.10) = 4.875$

ROA regression model: $\hat{Y}_i = 4.875 + 1.25X_i + \varepsilon_i.$

Notice the similarity of the formula for the slope coefficient and that of the pairwise correlation. The sample correlation, r, is the ratio of the covariance to the product of the standard deviations:

$$r = \frac{\text{Covariance of } Y \text{ and } X}{\left(\begin{array}{c}\text{Standard deviation}\\ \text{of } Y\end{array}\right)\left(\begin{array}{c}\text{Standard deviation}\\ \text{of } X\end{array}\right)}.$$

The subtle difference between the slope and the correlation formulas is in the denominator: For the slope, this is the variance of the independent variable, but for the correlation, the denominator is the product of the standard deviations. For our ROA and CAPEX analysis,

$$\text{Covariance of } Y \text{ and } X\text{: } \operatorname{cov}_{XY} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{n-1} = \frac{153.30}{5} = 30.66. \tag{7a}$$

Standard deviation of Y and X:

$$\begin{aligned} S_Y &= \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}} = \sqrt{\frac{239.50}{5}} = 6.9210; \\ S_X &= \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{122.64}{5}} = 4.9526. \end{aligned} \tag{7b}$$

$$r = \frac{30.66}{(6.9210)(4.9526)} = 0.89458945.$$

Because the denominators of both the slope and the correlation are positive, the sign of the slope and the correlation are driven by the numerator: If the covariance is positive, both the slope and the correlation are positive, and if the covariance is negative, both the slope and the correlation are negative.

HOW DO ANALYSTS PERFORM SIMPLE LINEAR REGRESSION?

Typically, an analyst will use the data analysis functions on a spreadsheet, such as Microsoft Excel, or a statistical package in the R or Python programming languages to perform linear regression analysis. The following are some of the more common choices in practice.

Simple Linear Regression: Intercept and Slope

- *Excel*: Use the INTERCEPT, SLOPE functions.
- *R*: Use the lm function.
- *Python*: Use the sm.OLS function in the statsmodels package.

Correlations

- *Excel*: Use the CORREL function.
- *R*: Use the cor function in the stats library.
- *Python*: Use the corrcoef function in the numpy library.

Note that in R and Python, there are many choices for regression and correlation analysis.

2.3 Interpreting the Regression Coefficients

What is the meaning of the regression coefficients? The intercept is the value of the dependent variable if the value of the independent variable is zero. Importantly, this does not make sense in some contexts, especially if it is unrealistic that the independent variable would be zero. For example, if we have a model where money supply explains GDP growth, the intercept has no meaning because, practically speaking, zero money supply is not possible. If the independent variable were money supply growth, however, the intercept is meaningful. The slope is the change in the dependent variable for a one-unit change in the independent variable. If the slope is positive, then the change in the independent variable and that of the dependent variable will be in the same direction; if the slope is negative, the change in the independent variable and that of the dependent variable will be in opposite directions.

INTERPRETING POSITIVE AND NEGATIVE SLOPES

Suppose the dependent variable (Y) is in millions of euros and the independent variable (*X*) is in millions of US dollars.

If the slope is positive 1.2, then
↑ USD1 million → ↑ EUR1.2 million
↓ USD1 million → ↓ EUR1.2 million
If the slope is negative 1.2, then
↑ USD1 million → ↓ EUR1.2 million
↓ USD1 million → ↑ EUR1.2 million

Using the ROA regression model from Exhibit 6, we would interpret the estimated coefficients as follows:

- The return on assets for a company is 4.875% if the company makes no capital expenditures.
- If CAPEX increases by one unit—say, from 4% to 5%—ROA increases by 1.25%.

Using the estimated regression coefficients, we can determine the values of the dependent variable if they follow the average relationship between the dependent and independent variables. A result of the mathematics of the least squares fitting of the regression line is that the expected value of the residual term is zero: E(ε) = 0.

We show the calculation of the predicted dependent variable and residual term for each observation in the ROA example in Exhibit 7. Note that the sum and average of Y_i and $\hat{Y}_1$ are the same, and the sum of the residuals is zero.

Exhibit 7 Calculation of the Dependent Variable and Residuals for the ROA and CAPEX Model

	(1)	(2)	(3)	(4)
Company	**ROA (Y_i)**	**CAPEX (X_i)**	**Predicted ROA ($\hat{Y}_i$)**	**(1) – (3) Residual (e_i)**
A	6.0	0.7	5.750	0.250
B	4.0	0.4	5.375	–1.375
C	15.0	5.0	11.125	3.875
D	20.0	10.0	17.375	2.625
E	10.0	8.0	14.875	–4.875
F	20.0	12.5	20.500	–0.500

(continued)

Exhibit 7 (Continued)

Company	(1) ROA (Y_i)	(2) CAPEX (X_i)	(3) Predicted ROA ($\hat{Y}_i$)	(4) (1) – (3) Residual (e_i)
Sum	75.0	36.6	75.000	0.000
Average	12.5	6.1	12.5	0.000

For Company C (i = 3),

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i + \varepsilon_i = 4.875 + 1.25X_i + \varepsilon_i$$

$$\hat{Y}_i = 4.875 + (1.25 \times 5.0) = 4.875 + 6.25 = 11.125$$

$Y_i - \hat{Y}_i = e_i = 15.0 - 11.125 = 3.875$, the vertical distance in Exhibit 5.

Whereas the sum of the residuals must equal zero by design, the focus of fitting the regression line in a simple linear regression is minimizing the sum of the squared residual terms.

2.4 Cross-Sectional vs. Time-Series Regressions

Regression analysis uses two principal types of data: cross sectional and time series. A cross-sectional regression involves many observations of X and Y for the same time period. These observations could come from different companies, asset classes, investment funds, countries, or other entities, depending on the regression model. For example, a cross-sectional model might use data from many companies to test whether predicted EPS growth explains differences in price-to-earnings ratios during a specific time period. Note that if we use cross-sectional observations in a regression, we usually denote the observations as $i = 1, 2, \ldots, n$.

Time-series data use many observations from different time periods for the same company, asset class, investment fund, country, or other entity, depending on the regression model. For example, a time-series model might use monthly data from many years to test whether a country's inflation rate determines its short-term interest rates. If we use time-series data in a regression, we usually denote the observations as $t = 1, 2, \ldots, T$. Note that in the sections that follow, we primarily use the notation $i = 1, 2, \ldots, n$, even for time series.

EXAMPLE 2

Estimating a Simple Linear Regression Model

An analyst is exploring the relationship between a company's net profit margin and research and development expenditures. He collects data for an industry and calculates the ratio of research and development expenditures to revenues (RDR) and the net profit margin (NPM) for eight companies. Specifically, he wants to explain the variation that he observes in the net profit margin by using the variation he observes in the companies' research and development spending. He reports the data in Exhibit 8.

Exhibit 8 Observations on NPM and RDR for Eight Companies

Company	NPM (%)	RDR (%)
1	4	8
2	5	10
3	10	6
4	9	5
5	5	7
6	6	9
7	12	5
8	3	10

1. What is the slope coefficient for this simple linear regression model?
2. What is the intercept for this regression model?
3. How is this estimated linear regression model represented?
4. What is the pairwise correlation between NPM and RDR?

Solutions

1. The slope coefficient for the regression model is −1.3, and the details for the inputs to this calculation are in Exhibit 9.

Exhibit 9 Details of Calculation of Slope of NPM Regressed on RDR

Company	NPM (%) (Y_i)	RDR (%) (X_i)	$Y_i - \bar{Y}$	$X_i - \bar{X}$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})(X_i - \bar{X})$
1	4	8	−2.8	0.5	7.5625	0.25	−1.375
2	5	10	−1.8	2.5	3.0625	6.25	−4.375
3	10	6	3.3	−1.5	10.5625	2.25	−4.875
4	9	5	2.3	−2.5	5.0625	6.25	−5.625
5	5	7	−1.8	−0.5	3.0625	0.25	0.875
6	6	9	−0.8	1.5	0.5625	2.25	−1.125
7	12	5	5.3	−2.5	27.5625	6.25	−13.125
8	3	10	−3.8	2.5	14.0625	6.25	−9.375
Sum	54.0	60.0	0.0	0.0	71.5000	30.00	−39.0
Average	6.75	7.5					

Slope coefficient: $\hat{b}_1 = \frac{-39}{30} = -1.3.$

2. The intercept of the regression model is 16.5:

$$\text{Intercept: } \hat{b}_0 = 6.75 - (-1.3 \times 7.5) = 6.75 + 9.75 = 16.5$$

3 The regression model is represented by $\hat{Y}_i = 16.5 - 1.3X_i + \varepsilon_i$.

4 The pairwise correlation is –0.8421:

$$r = \frac{-39/7}{\sqrt{71.5/7}\sqrt{30/7}} = \frac{-5.5714}{(3.1960)(2.0702)} = -0.8421.$$

EXAMPLE 3

Interpreting Regression Coefficients

An analyst has estimated a model that regresses a company's return on equity (ROE) against its growth opportunities (GO), defined as the company's three-year compounded annual growth rate in sales, over 20 years and produces the following estimated simple linear regression:

$$\text{ROE}_i = 4 + 1.8\ \text{GO}_i + \varepsilon_i.$$

Both variables are stated in percentages, so a GO observation of 5% is included as 5.

1 The predicted value of the company's ROE if its GO is 10% is closest to:
- A 1.8%.
- B 15.8%.
- C 22.0%.

2 The change in ROE for a change in GO from 5% to 6% is closest to:
- A 1.8%.
- B 4.0%.
- C 5.8%.

3 The residual in the case of a GO of 8% and an observed ROE of 21% is closest to:
- A –1.8%.
- B 2.6%.
- C 12.0%.

Solutions

1 C is correct. The predicted value of ROE = 4 + (1.8 × 10) = 22.

2 A is correct. The slope coefficient of 1.8 is the expected change in the dependent variable (ROE) for a one-unit change in the independent variable (GO).

3 B is correct. The predicted value is ROE = 4 + (1.8 × 8) = 18.4. The observed value of ROE is 21, so the residual is 2.6 = 21.0 – 18.4.

ASSUMPTIONS OF THE SIMPLE LINEAR REGRESSION MODEL

3

c explain the assumptions underlying the simple linear regression model, and describe how residuals and residual plots indicate if these assumptions may have been violated

We have discussed how to interpret the coefficients in a simple linear regression model. Now we turn to the statistical assumptions underlying this model. Suppose that we have *n* observations of both the dependent variable, *Y*, and the independent variable, *X*, and we want to estimate the simple linear regression of *Y* regressed on *X*. We need to make the following four key assumptions to be able to draw valid conclusions from a simple linear regression model:

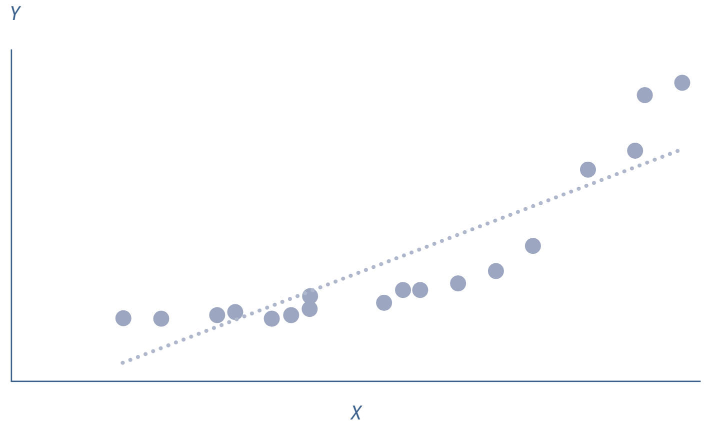
1. Linearity: The relationship between the dependent variable, *Y*, and the independent variable, *X*, is linear.
2. Homoskedasticity: The variance of the regression residuals is the same for all observations.
3. Independence: The observations, pairs of *Y*s and *X*s, are independent of one another. This implies the regression residuals are uncorrelated across observations.
4. Normality: The regression residuals are normally distributed.

Now we take a closer look at each of these assumptions and introduce the "best practice" of examining residual plots of regression results to identify potential violations of these key assumptions.

3.1 Assumption 1: Linearity

We are fitting a linear model, so we must assume that the true underlying relationship between the dependent and independent variables is linear. If the relationship between the independent and dependent variables is nonlinear in the parameters, estimating that relation with a simple linear regression model will produce invalid results: The model will be biased, because it will under- and overestimate the dependent variable at certain points. For example, $Y_i = b_0 e^{b_1 X_i} + \varepsilon_i$ is nonlinear in b_1, so we should not apply the linear regression model to it. Exhibit 10 shows an example of this exponential model, with a regression line indicated. You can see that this line does not fit this relationship well: For lower and higher values of *X*, the linear model underestimates the *Y*, whereas for the middle values, the linear model overestimates *Y*.

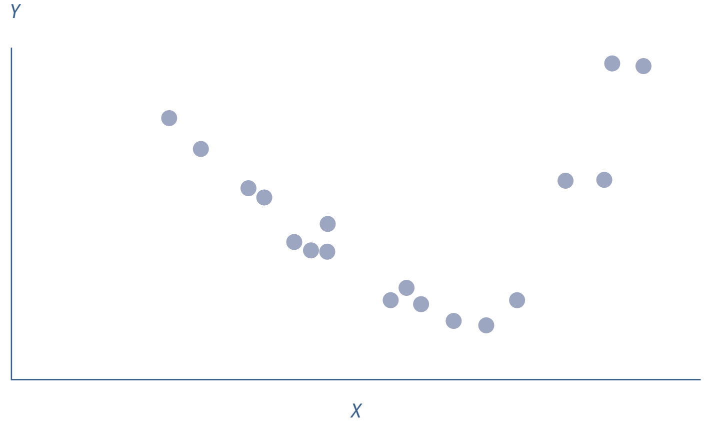
Exhibit 10 **Illustration of Nonlinear Relationship Estimated as a Linear Relationship**



Another implication of this assumption is that the independent variable, X, must not be random; that is, it is non-stochastic. If the independent variable is random, there would be no linear relation between the dependent and independent variables. Although we may initially assume that the independent variable in the regression model is not random, that assumption may not always be true.

When we look at the residuals of a model, what we would like to see is that the residuals are random. The residuals should not exhibit a pattern when plotted against the independent variable. As we show in Exhibit 11, the residuals from the Exhibit 10 linear regression do not appear to be random but, rather, exhibit a relationship with the independent variable, X, falling for some range of X and rising in another.

Exhibit 11 **Illustration of Residuals in a Nonlinear Relationship Estimated as a Linear Relationship**



3.2 Assumption 2: Homoskedasticity

Assumption 2, that the variance of the residuals is the same for all observations, is known as the **homoskedasticity** assumption. In terms of notation, this assumption relates to the squared residuals:

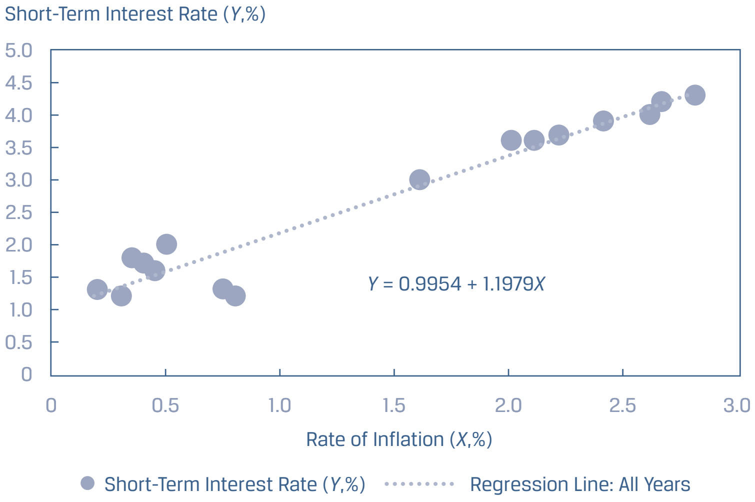
$$\mathrm{E}\left(\varepsilon_i^2\right) = \sigma_\varepsilon^2,\ i = 1, \ldots, n. \tag{8}$$

If the residuals are not homoscedastic, that is, if the variance of residuals differs across observations, then we refer to this as **heteroskedasticity**.

Suppose you are examining a time series of short-term interest rates as the dependent variable and inflation rates as the independent variable over 16 years. We may believe that short-term interest rates (Y) and inflation rates (X) should be related (that is, interest rates are higher with higher rates of inflation. If this time series spans many years, with different central bank actions that force short-term interest rates to be (artificially) low for the last eight years of the series, then it is likely that the residuals in this estimated model will appear to come from two different models. We will refer to the first eight years as Regime 1 (normal rates) and the second eight years as Regime 2 (low rates). If the model fits differently in the two regimes, the residuals and their variances will be different.

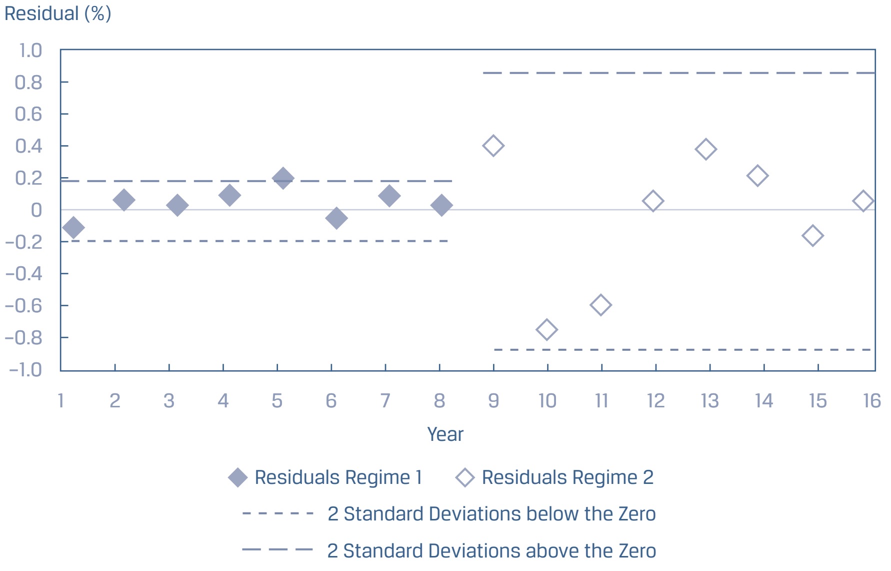
You can see this situation in Exhibit 12, which shows a scatter plot with an estimated regression line. The slope of the regression line over all 16 years is 1.1979.

Exhibit 12 Scatter Plot of Interest Rates (*Y*) and Inflation Rates (*X*)



We plot the residuals of this model in Exhibit 13 against the years. In this plot, we indicate the distance that is two standard deviations from zero (the mean of the residuals) for the first eight years' residuals and then do the same for the second eight years. As you can see, the residuals appear different for the two regimes: the variation in the residuals for the first eight years is much smaller than the variation for the second eight years.

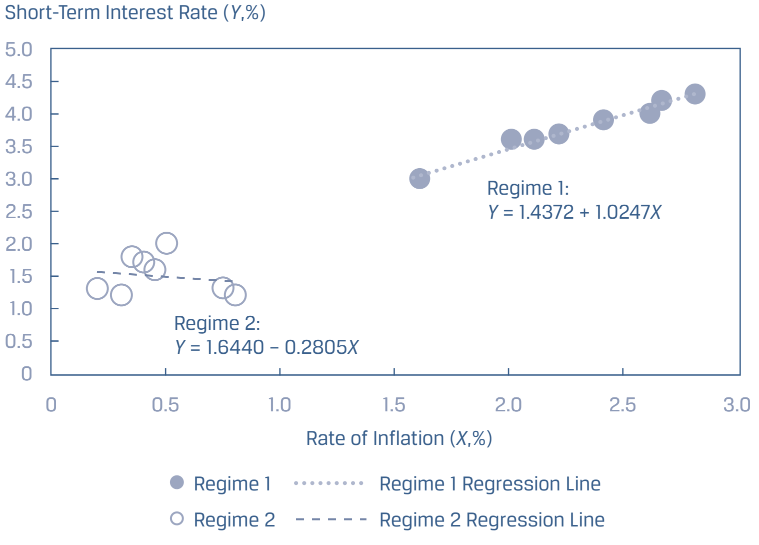
Exhibit 13 Residual Plot for Interest Rates (*Y*) vs. Inflation Rates (*X*) Model



Why does this happen? The model seems appropriate, but when we examine the residuals (Exhibit 13), an important step in assessing the model fit, we see that the model fits better in some years compared with others. The difference in variance of residuals between the two regimes is apparent from the much wider band around residuals for Regime 2 (the low-rate period). This indicates a clear violation of the homoskedasticity assumption.

If we estimate a regression line for each regime, we can see that the model for the two regimes is quite different, as we show in Exhibit 14. In the case of Regime 1 (normal rates), the slope is 1.0247, whereas in Regime 2 (low rates) the slope is −0.2805. In sum, the clustering of residuals in two groups with much different variances clearly indicates the existence of distinct regimes for the relationship between short-term interest rates and the inflation rate.

Exhibit 14 Fitted Regression Lines for the Two Regimes

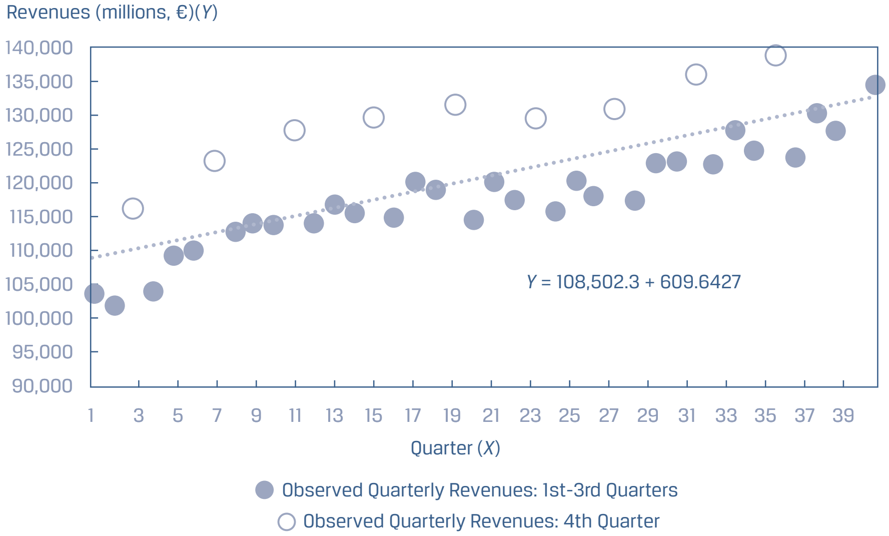


3.3 Assumption 3: Independence

We assume that the observations (Y and X pairs) are uncorrelated with one another, meaning they are independent. If there is correlation between observations (that is, autocorrelation), they are not independent and the residuals will be correlated. The assumption that the residuals are uncorrelated across observations is also necessary for correctly estimating the variances of the **estimated parameters** of b_0 and b_1 (i.e., $\hat{b}_0$ and $\hat{b}_1$) that we use in hypothesis tests of the intercept and slope, respectively. It is important to examine whether the residuals exhibit a pattern, suggesting a violation of this assumption. Therefore, we need to visually and statistically examine the residuals for a regression model.

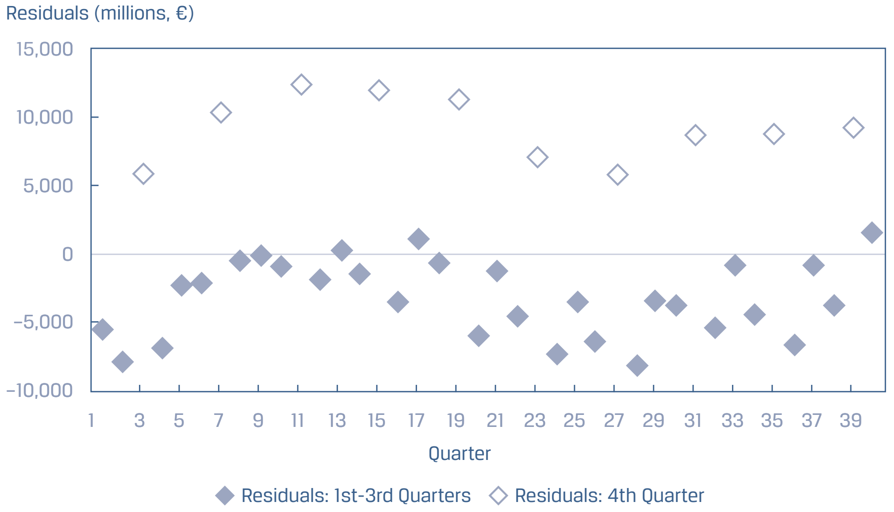
Consider the quarterly revenues of a company regressed over 40 quarters, as shown in Exhibit 15, with the regression line included. It is clear that these revenues display a seasonal pattern, an indicator of autocorrelation.

Exhibit 15 Regression of Quarterly Revenues vs. Time (40 Quarters)



In Exhibit 16, we plot the residuals from this model and see that there is a pattern. These residuals are correlated, specifically jumping up in Quarter 4 and then falling back the subsequent quarter. In sum, the patterns in both Exhibits 15 and 16 indicate a violation of the assumption of independence.

Exhibit 16 Residual Plot for Quarterly Revenues vs. Time Model



3.4 Assumption 4: Normality

The assumption of normality requires that the residuals be normally distributed. This does not mean that the dependent and independent variables must be normally distributed; it only means that the residuals from the model are normally distributed. However, in estimating any model, it is good practice to understand the distribution of the dependent and independent variables to explore for outliers. An outlier in either or both variables can substantially influence the fitted line such that the estimated model will not fit well for most of the other observations.

With normally distributed residuals, we can test a particular hypothesis about a linear regression model. For large sample sizes, we may be able to drop the assumption of normality by appealing to the central limit theorem; asymptotic theory (which deals with large samples) shows that in many cases, the test statistics produced by standard regression programs are valid even if the model's residuals are not normally distributed.

EXAMPLE 4

Assumptions of Simple Linear Regression

An analyst is investigating a company's revenues and estimates a simple linear time-series model by regressing revenues against time, where time—1, 2, . . . , 15—is measured in years. She plots the company's observed revenues and the estimated regression line, as shown in Exhibit 17. She also plots the residuals from this regression model, as shown in Exhibit 18.

Exhibit 17 Revenues vs. Time Using Simple Linear Regression

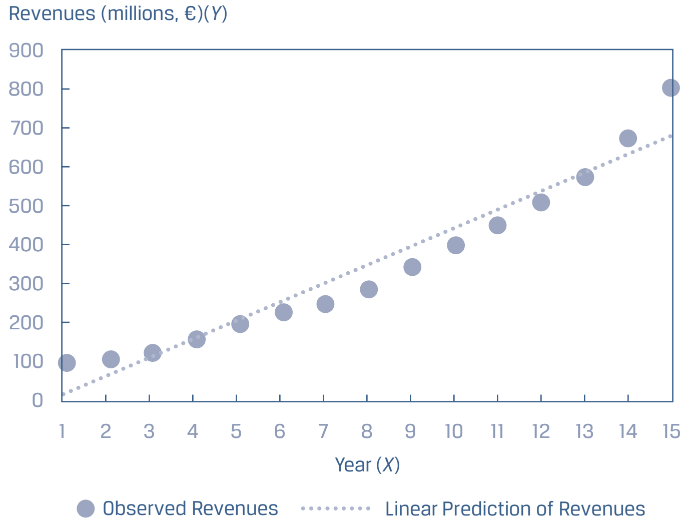
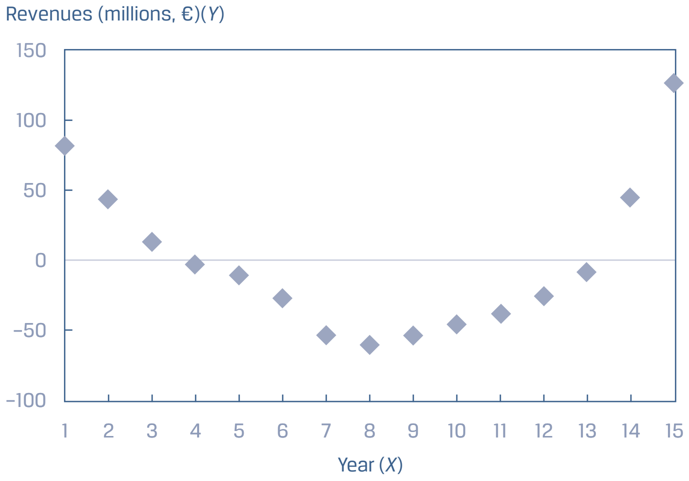


Exhibit 18 Residual Plot for Revenues vs. Time



Based on Exhibits 17 and 18, describe which assumption(s) of simple linear regression the analyst's model may be violating.

Solution

The correct model is not linear, as evident from the pattern of the revenues in Exhibit 17. In the earlier years (i.e., 1 and 2) and later years (i.e., 14 and 15), the linear model underestimates revenues, whereas for the middle years (i.e., 7–11), the linear model overestimates revenues. Moreover, the curved pattern of residuals in Exhibit 18 indicates potential heteroskedasticity (residuals have unequal variances), lack of independence of observations, and non-normality (a concern given the small sample size of $n = 15$). In sum, the analyst should be concerned that her model violates all the assumptions governing simple linear regression (linearity, homoskedasticity, independence, and normality).

4

ANALYSIS OF VARIANCE

d calculate and interpret the coefficient of determination and the *F*-statistic in a simple linear regression

The simple linear regression model sometimes describes the relationship between two variables quite well, but sometimes it does not. We must be able to distinguish between these two cases to use regression analysis effectively. Remember our goal is to explain the variation of the dependent variable. So, how well has this goal been achieved, given our choice of independent variable?

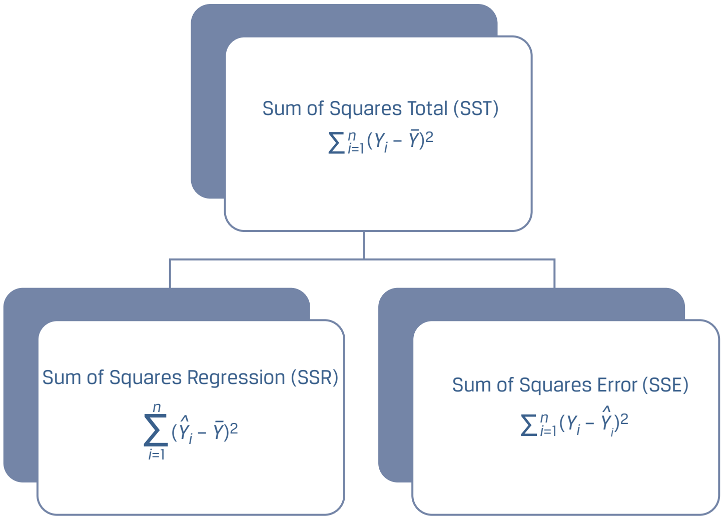
4.1 Breaking down the Sum of Squares Total into Its Components

We begin with the sum of squares total and then break it down into two parts: the sum of squares error and the **sum of squares regression (SSR)**. The sum of squares regression is the sum of the squared differences between the predicted value of the dependent variable, $\hat{Y}_i$, based on the estimated regression line, and the mean of the dependent variable, $\overline{Y}$:

$$\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2. \tag{9}$$

We have already defined the sum of squares total, which is the total variation in *Y*, and the sum of squares error, the unexplained variation in *Y*. Note that the sum of squares regression is the explained variation in *Y*. So, as illustrated in Exhibit 19, SST = SSR + SSE, meaning total variation in *Y* equals explained variation in *Y* plus unexplained variation in *Y*.

Exhibit 19 Breakdown of Variation of Dependent Variable


Sum of Squares Total (SST)
$\sum_{i=1}^{n}(Y_i - \bar{Y})^2$
Sum of Squares Regression (SSR)
$\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$
Sum of Squares Error (SSE)
$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$

We show the breakdown of the sum of squares total formula for our ROA regression example in Exhibit 20. The total variation of ROA that we want to explain (SST) is 239.50. This number comprises the variation unexplained (SSE), 47.88, and the variation explained (SSR), 191.63. These sum of squares values are important inputs into measures of the fit of the regression line.

Exhibit 20 Breakdown of Sum of Squares Total for ROA Model

Company	ROA (Y_i)	CAPEX (X_i)	Predicted ROA ($\hat{Y}$)	Variation to Be Explained $\left(Y_i - \bar{Y}\right)^2$	Variation Unexplained $\left(Y_i - \hat{Y}_i\right)^2$	Variation Explained $\left(\hat{Y}_i - \bar{Y}\right)^2$
A	6.0	0.7	5.750	42.25	0.063	45.563
B	4.0	0.4	5.375	72.25	1.891	50.766
C	15.0	5.0	11.125	6.25	15.016	1.891
D	20.0	10.0	17.375	56.25	6.891	23.766
E	10.0	8.0	14.875	6.25	23.766	5.641
F	20.0	12.5	20.500	56.25	0.250	64.000
				239.50	47.88	191.625
Mean	12.50					

Sum of squares total = 239.50.
Sum of squares error = 47.88.
Sum of squares regression = 191.63.

4.2 Measures of Goodness of Fit

There are several measures that we can use to evaluate goodness of fit—that is, how well the regression model fits the data. These include the coefficient of determination, the *F*-statistic for the test of fit, and the standard error of the regression.

The **coefficient of determination**, also referred to as the *R*-squared or R^2, is the percentage of the variation of the dependent variable that is explained by the independent variable:

$$\begin{gathered}\text{Coefficient of determination} = \frac{\text{Sum of squares regression}}{\text{Sum of squares total}} \\ \text{Coefficient of determination} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}.\end{gathered} \tag{10}$$

By construction, the coefficient of determination ranges from 0% to 100%. In our ROA example, the coefficient of determination is 191.625 ÷ 239.50, or 0.8001, so 80.01% of the variation in ROA is explained by CAPEX. In a simple linear regression, the square of the pairwise correlation is equal to the coefficient of determination:

$$r^2 = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2}{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2} = R^2.$$

In our earlier ROA regression analysis, $r = 0.8945$, so we now see that r^2 is indeed equal to the coefficient of determination (R^2), since $(0.8945)^2 = 0.8001$.

Whereas the coefficient of determination—the portion of the variation of the dependent variable explained by the independent variable—is descriptive, it is not a statistical test. To see if our regression model is likely to be statistically meaningful, we will need to construct an *F*-distributed test statistic.

In general, we use an F-distributed test statistic to compare two variances. In regression analysis, we can use an F-distributed test statistic to test whether the slopes in a regression are equal to zero, with the slopes designated as b_i, against the alternative hypothesis that at least one slope is not equal to zero:

H_0: $b_1 = b_2 = b_3 = \ldots = b_k = 0$.

H_a: At least one b_k is not equal to zero.

For simple linear regression, these hypotheses simplify to

H_0: $b_1 = 0$.

H_a: $b_1 \neq 0$.

The F-distributed test statistic is constructed by using the sum of squares regression and the sum of squares error, each adjusted for degrees of freedom; in other words, it is the ratio of two variances. We divide the sum of squares regression by the number of independent variables, represented by k. In the case of a simple linear regression, k = 1, so we arrive at the **mean square regression (MSR)**, which is the same as the sum of squares regression:

$$\text{MSR} = \frac{\text{Sum of squares regression}}{k} = \frac{\sum_{i=1}^{n}(\widehat{Y}_i - \bar{Y})^2}{1}.$$

So, for simple linear regression,

$$\text{MSR} = \sum_{i=1}^{n}(\widehat{Y}_i - \bar{Y})^2. \tag{11}$$

Next, we calculate the **mean square error (MSE)**, which is the sum of squares error divided by the degrees of freedom, which are $n - k - 1$. In simple linear regression, $n - k - 1$ becomes $n - 2$:

$$\text{MSE} = \frac{\text{Sum of squares error}}{n - k - 1}.$$

$$\text{MSE} = \frac{\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2}{n - 2}. \tag{12}$$

Therefore, the F-distributed test statistic (MSR/MSE) is

$$F = \frac{\dfrac{\text{Sum of squares regression}}{k}}{\dfrac{\text{Sum of squares error}}{n - k - 1}} = \frac{\text{MSR}}{\text{MSE}}$$

$$F = \frac{\dfrac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{1}}{\dfrac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n - 2}}, \tag{13}$$

which is distributed with 1 and $n - 2$ degrees of freedom in simple linear regression. The F-statistic in regression analysis is one sided, with the rejection region on the right side, because we are interested in whether the variation in Y explained (the numerator) is larger than the variation in Y unexplained (the denominator).

4.3 ANOVA and Standard Error of Estimate in Simple Linear Regression

e describe the use of analysis of variance (ANOVA) in regression analysis, interpret ANOVA results, and calculate and interpret the standard error of estimate in a simple linear regression

We often represent the sums of squares from a regression model in an **analysis of variance (ANOVA)** table, as shown in Exhibit 21, which presents the sums of squares, the degrees of freedom, the mean squares, and the *F*-statistic. Notice that the variance of the dependent variable is the ratio of the sum of squares total to $n - 1$.

Exhibit 21 Analysis of Variance Table for Simple Linear Regression

Source	Sum of Squares	Degrees of Freedom	Mean Square	*F*-Statistic
Regression	$\text{SSR} = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$	1	$\text{MSR} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{1}$	$F = \frac{\text{MSR}}{\text{MSE}} = \frac{\frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{1}}{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n-2}}$
Error	$\text{SSE} = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$	$n - 2$	$\text{MSE} = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n-2}$	
Total	$\text{SST} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$	$n - 1$		

From the ANOVA table, we can also calculate the **standard error of the estimate** (s_e), which is also known as the standard error of the regression or the root mean square error. The s_e is a measure of the distance between the observed values of the dependent variable and those predicted from the estimated regression; the smaller the s_e, the better the fit of the model. The s_e, along with the coefficient of determination and the *F*-statistic, is a measure of the goodness of the fit of the estimated regression line. Unlike the coefficient of determination and the *F*-statistic, which are relative measures of fit, the standard error of the estimate is an absolute measure of the distance of the observed dependent variable from the regression line. Thus, the s_e is an important statistic used to evaluate a regression model and is used in calculating prediction intervals and performing tests on the coefficients. The calculation of s_e is straightforward once we have the ANOVA table because it is the square root of the MSE:

$$\text{Standard error of the estimate } (s_e) = \sqrt{\text{MSE}} = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n-2}}. \tag{14}$$

We show the ANOVA table for our ROA regression example in Exhibit 22, using the information from Exhibit 20. For a 5% level of significance, the critical *F*-value for the test of whether the model is a good fit (that is, whether the slope coefficient is different from zero) is 7.71. We can get this critical value in the following ways:

- *Excel* [F.INV(0.95,1,4)]
- *R* [qf(.95,1,4)]
- *Python* [from scipy.stats import f *and* f.ppf(.95,1,4)]

With a calculated *F*-statistic of 16.0104 and a critical *F*-value of 7.71, we reject the null hypothesis and conclude that the slope of our simple linear regression model for ROA is different from zero.

Exhibit 22 ANOVA Table for ROA Regression Model

Source	Sum of Squares	Degrees of Freedom	Mean Square	*F*-Statistic
Regression	191.625	1	191.625	16.0104
Error	47.875	4	11.96875	
Total	239.50	5		

The calculations to derive the ANOVA table and ultimately to test the goodness of fit of the regression model can be time consuming, especially for samples with many observations. However, statistical packages, such as SAS, SPSS Statistics, and Stata, as well as software, such as Excel, R, and Python, produce the ANOVA table as part of the output for regression analysis.

EXAMPLE 5

Using ANOVA Table Results to Evaluate a Simple Linear Regression

Suppose you run a cross-sectional regression for 100 companies, where the dependent variable is the annual return on stock and the independent variable is the lagged percentage of institutional ownership (INST). The results of this simple linear regression estimation are shown in Exhibit 23. Evaluate the model by answering the questions below.

Exhibit 23 ANOVA Table for Annual Stock Return Regressed on Institutional Ownership

Source	Sum of Squares	Degrees of Freedom	Mean Square
Regression	576.1485	1	576.1485
Error	1,873.5615	98	19.1180
Total	2,449.7100		

1 What is the coefficient of determination for this regression model?

2 What is the standard error of the estimate for this regression model?

3 At a 5% level of significance, do we reject the null hypothesis of the slope coefficient equal to zero if the critical *F*-value is 3.938?

4 Based on your answers to the preceding questions, evaluate this simple linear regression model.

Solutions

1 The coefficient of determination is sum of squares regression/sum of squares total: 576.148 ÷ 2,449.71 = 0.2352, or 23.52%.

2 The standard error of the estimate is the square root of the mean square error: $\sqrt{19.1180} = 4.3724$.

3 Using a six-step process for testing hypotheses, we get the following:

Step 1	State the hypotheses.	H_0: $b_1 = 0$ versus H_a: $b_1 \neq 0$
Step 2	Identify the appropriate test statistic.	$F = \dfrac{\text{MSR}}{\text{MSE}}$ with 1 and 98 degrees of freedom.
Step 3	Specify the level of significance.	α = 5% (one tail, right side).
Step 4	State the decision rule.	Critical F-value = 3.938. Reject the null hypothesis if the calculated F-statistic is greater than 3.938.
Step 5	Calculate the test statistic.	$F = \dfrac{576.1485}{19.1180} = 30.1364$
Step 6	Make a decision.	Reject the null hypothesis because the calculated F-statistic is greater than the critical F-value. There is sufficient evidence to indicate that the slope coefficient is different from 0.0.

4 The coefficient of determination indicates that variation in the independent variable explains 23.52% of the variation in the dependent variable. Also, the F-statistic test confirms that the model's slope coefficient is different from 0 at the 5% level of significance. In sum, the model seems to fit the data reasonably well.

HYPOTHESIS TESTING OF LINEAR REGRESSION COEFFICIENTS

5

f formulate a null and an alternative hypothesis about a population value of a regression coefficient, and determine whether the null hypothesis is rejected at a given level of significance

5.1 Hypothesis Tests of the Slope Coefficient

We can use the F-statistic to test for the significance of the slope coefficient (that is, whether it is significantly different from zero), but we also may want to perform other hypothesis tests for the slope coefficient—for example, testing whether the population slope is different from a specific value or whether the slope is positive. We can use a t-distributed test statistic to test such hypotheses about a regression coefficient.

Suppose we want to check a stock's valuation using the market model; we hypothesize that the stock has an average systematic risk (i.e., risk similar to that of the market), as represented by the coefficient on the market returns variable. Or we may want to test the hypothesis that economists' forecasts of the inflation rate are unbiased (that is, on average, not overestimating or underestimating actual inflation rates). In each case, does the evidence support the hypothesis? Such questions as these can be addressed with hypothesis tests on the regression slope. To test a hypothesis about a slope, we calculate the test statistic by subtracting the hypothesized population slope (B_1) from the estimated slope coefficient ($\hat{b}_1$) and then dividing this difference by the standard error of the slope coefficient, $s_{\hat{b}_1}$:

$$t = \frac{\hat{b}_1 - B_1}{s_{\hat{b}_1}}. \tag{15}$$

This test statistic is t-distributed with $n - k - 1$ or $n - 2$ degrees of freedom because two parameters (an intercept and a slope) were estimated in the regression.

The **standard error of the slope coefficient** ($s_{\hat{b}_1}$) for a simple linear regression is the ratio of the model's standard error of the estimate (s_e) to the square root of the variation of the independent variable:

$$s_{\hat{b}_1} = \frac{s_e}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}}. \tag{16}$$

We compare the calculated t-statistic with the critical values to test hypotheses. Note that the greater the variability of the independent variable, the lower the standard error of the slope (Equation 16) and hence the greater the calculated t-statistic (Equation 15). If the calculated t-statistic is outside the bounds of the critical t-values, we reject the null hypothesis, but if the calculated t-statistic is within the bounds of the critical values, we fail to reject the null hypothesis. Similar to tests of the mean, the alternative hypothesis can be two sided or one sided.

Consider our previous simple linear regression example with ROA as the dependent variable and CAPEX as the independent variable. Suppose we want to test whether the slope coefficient of CAPEX is different from zero to confirm our intuition of a significant relationship between ROA and CAPEX. We can test the hypothesis concerning the slope using the six-step process, as we show in Exhibit 24. As a result of this test, we conclude that the slope is different from zero; that is, CAPEX is a significant explanatory variable of ROA.

Exhibit 24 Test of the Slope for the Regression of ROA on CAPEX

Step 1	State the hypotheses.	H_0: $b_1 = 0$ versus H_a: $b_1 \neq 0$
Step 2	Identify the appropriate test statistic.	$t = \frac{\hat{b}_1 - B_1}{s_{\hat{b}_1}}$ with 6 − 2 = 4 degrees of freedom.
Step 3	Specify the level of significance.	α = 5%.

Exhibit 24 (Continued)

Step 4	State the decision rule.	Critical t-values = ±2.776. We can determine this from **Excel** *Lower:* T.INV(0.025,4) *Upper:* T.INV(0.975,4) **R** qt(c(.025,.975),4) **Python** from scipy.stats import t *Lower:* t.ppf(.025,4) *Upper:* t.ppf(.975,4) We reject the null hypothesis if the calculated t-statistic is less than −2.776 or greater than +2.776.
Step 5	Calculate the test statistic.	The slope coefficient is 1.25 (Exhibit 6). The mean square error is 11.96875 (Exhibit 22). The variation of CAPEX is 122.640 (Exhibit 6). $s_e = \sqrt{11.96875} = 3.459588.$ $s_{\hat{b}_1} = \dfrac{3.459588}{\sqrt{122.640}} = 0.312398.$ $t = \dfrac{1.25 - 0}{0.312398} = 4.00131.$
Step 6	Make a decision.	Reject the null hypothesis of a zero slope. There is sufficient evidence to indicate that the slope is different from zero.

A feature of simple linear regression is that the t-statistic used to test whether the slope coefficient is equal to zero and the t-statistic to test whether the pairwise correlation is zero (that is, H_0: $\rho = 0$ versus H_a: $\rho \neq 0$) are the same value. Just as with a test of a slope, both two-sided and one-sided alternatives are possible for a test of a correlation—for example, H_0: $\rho \leq 0$ versus H_a: $\rho > 0$. The test-statistic to test whether the correlation is equal to zero is

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

In our example of ROA regressed on CAPEX, the correlation (r) is 0.8945. To test whether this correlation is different from zero, we perform a test of hypothesis, shown in Exhibit 25. As you can see, we draw a conclusion similar to that for our test of the slope, but it is phrased in terms of the correlation between ROA and CAPEX: There is a significant correlation between ROA and CAPEX.

Exhibit 25 Test of the Correlation between ROA and CAPEX

Step 1	State the hypotheses.	H_0: $\rho = 0$ versus H_a: $\rho \neq 0$
Step 2	Identify the appropriate test statistic.	$t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$ with 6 − 2 = 4 degrees of freedom.
Step 3	Specify the level of significance.	$\alpha = 5\%$.
Step 4	State the decision rule.	Critical t-values = ±2.776. Reject the null if the calculated t-statistic is less than −2.776 or greater than +2.776.

(continued)

Exhibit 25 (Continued)

Step 5	Calculate the test statistic.	$t = \dfrac{0.8945\sqrt{4}}{\sqrt{1 - 0.8001}} = 4.00131.$
Step 6	Make a decision.	Reject the null hypothesis of no correlation. There is sufficient evidence to indicate that the correlation between ROA and CAPEX is different from zero.

Another interesting feature of simple linear regression is that the test-statistic used to test the fit of the model (that is, the *F*-distributed test statistic) is related to the calculated *t*-statistic used to test whether the slope coefficient is equal to zero: $t^2 = F$; therefore, $4.00131^2 = 16.0104$.

What if instead we want to test whether there is a one-to-one relationship between ROA and CAPEX, implying a slope coefficient of 1.0. The hypotheses become H_0: $b_1 = 1$ and H_a: $b_1 \neq 1$. The calculated *t*-statistic is

$$t = \frac{1.25 - 1}{0.312398} = 0.80026.$$

This calculated test statistic falls within the bounds of the critical values, ±2.776, so we fail to reject the null hypothesis: There is not sufficient evidence to indicate that the slope is different from 1.0.

What if instead we want to test whether there is a positive slope or positive correlation, as our intuition suggests? In this case, all the steps are the same as in Exhibits 24 and 25 except the critical values because the tests are one sided. For a test of a positive slope or positive correlation, the critical value for a 5% level of significance is +2.132. We show the test of hypotheses for a positive slope and a positive correlation in Exhibit 26. Our conclusion is that there is sufficient evidence supporting both a positive slope and a positive correlation.

Exhibit 26 One-Sided Tests for the Slope and Correlation

		Test of the Slope	**Test of the Correlation**
Step 1	State the hypotheses.	H_0: $b_1 \leq 0$ versus H_a: $b_1 > 0$	H_0: $\rho \leq 0$ versus H_a: $\rho > 0$
Step 2	Identify the appropriate test statistic.	$t = \dfrac{\hat{b}_1 - B_1}{s_{\hat{b}_1}}$ with 6 − 2 = 4 degrees of freedom.	$t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$ with 6 − 2 = 4 degrees of freedom.
Step 3	Specify the level of significance.	$\alpha = 5\%$.	$\alpha = 5\%$.
Step 4	State the decision rule.	Critical *t*-value = 2.132. Reject the null if the calculated *t*-statistic is greater than 2.132.	Critical *t*-value = 2.132. Reject the null if the calculated *t*-statistic is greater than 2.132.
Step 5	Calculate the test statistic.	$t = \dfrac{1.25 - 0}{0.312398} = 4.00131$	$t = \dfrac{0.8945\sqrt{4}}{\sqrt{1 - 0.8001}} = 4.00131$
Step 6	Make a decision.	Reject the null hypothesis. There is sufficient evidence to indicate that the slope is greater than zero.	Reject the null hypothesis. There is sufficient evidence to indicate that the correlation is greater than zero.

5.2 Hypothesis Tests of the Intercept

There are occasions when we want to test whether the population intercept is a specific value. As a reminder on how to interpret the intercept, consider the simple linear regression with a company's revenue growth rate as the dependent variable (Y) and the GDP growth rate of its home country as the independent variable (X). The intercept is the company's revenue growth rate if the GDP growth rate is 0%.

The equation for the standard error of the intercept, $s_{\hat{b}_0}$, is

$$s_{\hat{b}_0} = \sqrt{\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}. \qquad (17)$$

We can test whether the intercept is different from the hypothesized value, B_0, by comparing the estimated intercept ($\hat{b}_0$) with the hypothesized intercept and then dividing the difference by the standard error of the intercept:

$$t_{intercept} = \frac{\hat{b}_0 - B_0}{s_{\hat{b}_0}} = \frac{\hat{b}_0 - B_0}{\sqrt{\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}}.$$

In the ROA regression example, the intercept is 4.875%. Suppose we want to test whether the intercept is greater than 3%. The one-sided hypothesis test is shown in Exhibit 27. As you can see, we reject the null hypothesis. In other words, there is sufficient evidence that if there are no capital expenditures (CAPEX = 0), ROA is greater than 3%.

Exhibit 27 Test of Hypothesis for Intercept for Regression of ROA on CAPEX

Step 1	State the hypotheses.	H_0: $b_0 \le 3\%$ versus H_a: $b_0 > 3\%$
Step 2	Identify the appropriate test statistic.	$t_{intercept} = \frac{\hat{b}_0 - B_0}{s_{\hat{b}_0}}$ with 6 – 2 = 4 degrees of freedom.
Step 3	Specify the level of significance.	α = 5%.
Step 4	State the decision rule.	Critical *t*-value = 2.132. Reject the null if the calculated *t*-statistic is greater than 2.132.
Step 5	Calculate the test statistic.	$t_{intercept} = \frac{4.875 - 3.0}{\sqrt{\frac{1}{6} + \frac{6.1^2}{122.64}}} = \frac{1.875}{0.68562} = 2.73475$
Step 6	Make a decision.	Reject the null hypothesis. There is sufficient evidence to indicate that the intercept is greater than 3%.

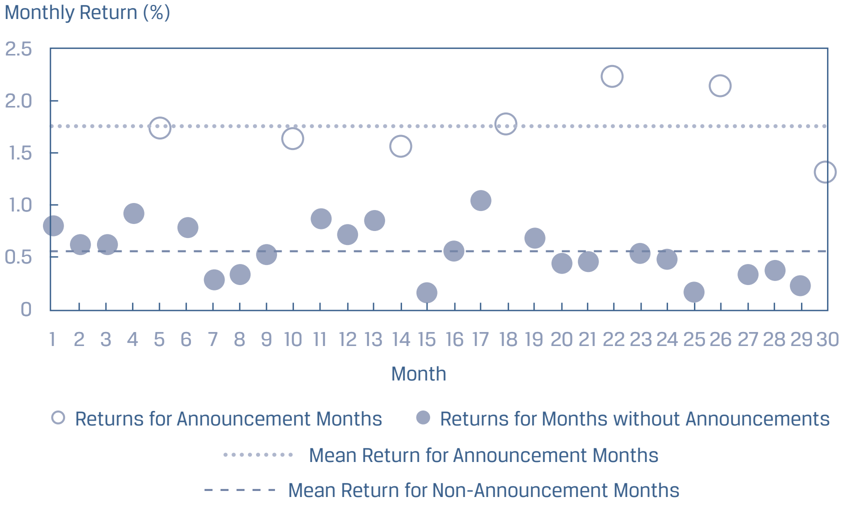
5.3 Hypothesis Tests of Slope When Independent Variable Is an Indicator Variable

Suppose we want to examine whether a company's quarterly earnings announcements influence its monthly stock returns. In this case, we could use an **indicator variable**, or dummy variable, that takes on only the values 0 or 1 as the independent variable. Consider the case of a company's monthly stock returns over a 30-month period. A simple linear regression model for investigating this question would be monthly returns, RET, regressed on the indicator variable, EARN, that takes on a value of 0 if there is no earnings announcement that month and 1 if there is an earnings announcement:

$$\text{RET}_i = b_0 + b_1\text{EARN}_i + \varepsilon_i.$$

This regression setup allows us to test whether there are different returns for earnings-announcement months versus non-earnings-announcement months. The observations and regression results are shown graphically in Exhibit 28.

Exhibit 28 Earnings Announcements, Dummy Variable, and Stock Returns



Clearly there are some months in which the returns are different from other months, and these correspond to months in which there was an earnings announcement. We estimate the simple linear regression model and perform hypothesis testing in the same manner as if the independent variable were a continuous variable. In a simple linear regression, the interpretation of the intercept is the predicted value of the dependent variable if the indicator variable is zero. Moreover, the slope, when the indicator variable is 1, is the difference in the means if we grouped the observations by the indicator variable. The results of the regression are given in Panel A of Exhibit 29.

Exhibit 29 Regression and Test of Differences Using an Indicator Variable

A. Regression Estimation Results

	Estimated Coefficients	Standard Error of Coefficients	Calculated Test Statistic
Intercept	0.5629	0.0560	10.0596
EARN	1.2098	0.1158	10.4435

Degrees of freedom = 28.
Critical *t*-values = +2.0484 (5% significance).

B. Test of Differences in Means

	RET for Earnings-Announcement Months	RET for Non-Earnings-Announcement Months	Difference in Means
Mean	1.7727	0.5629	1.2098
Variance	0.1052	0.0630	
Observations	7	23	
Pooled variance			0.07202
Calculated test statistic			10.4435

Degrees of freedom = 28.
Critical *t*-values = +2.0484 (5% significance).

We can see the following from Panel A of Exhibit 29:

- The intercept (0.5629) is the mean of the returns for non-earnings-announcement months.
- The slope coefficient (1.2098) is the difference in means of returns between earnings-announcement and non-announcement months.
- We reject the null hypothesis that the slope coefficient on EARN is equal to zero. We also reject the null hypothesis that the intercept is zero. The reason is that in both cases, the calculated test statistic exceeds the critical *t*-value.

We could also test whether the mean monthly return is the same for both the non-earnings-announcement months and the earnings-announcement months by testing the following:

$$H_0 : \mu_{RETearnings} = \mu_{RETNon-earnings} \text{ and } H_a : \mu_{RETearnings} \neq \mu_{RETNon-earnings}$$

The results of this hypothesis test are gleaned from Panel B of Exhibit 29. As you can see, we reject the null hypothesis that there is no difference in the mean RET for the earnings-announcement and non-earnings-announcements months at the 5% level of significance, since the calculated test statistic (10.4435) exceeds the critical value (2.0484).

5.4 Test of Hypotheses: Level of Significance and *p*-Values

The choice of significance level in hypothesis testing is always a matter of judgment. Analysts often choose the 0.05 level of significance, which indicates a 5% chance of rejecting the null hypothesis when, in fact, it is true (a Type I error, or false positive). Of course, decreasing the level of significance from 0.05 to 0.01 decreases the probability of Type I error, but it also increases the probability of Type II error—failing to reject the null hypothesis when, in fact, it is false (that is, a false negative).

The *p*-value is the smallest level of significance at which the null hypothesis can be rejected. The smaller the *p*-value, the smaller the chance of making a Type I error (i.e., rejecting a true null hypothesis), so the greater the likelihood the regression model is valid. For example, if the *p*-value is 0.005, we reject the null hypothesis that the true parameter is equal to zero at the 0.5% significance level (99.5% confidence). In most software packages, the *p*-values provided for regression coefficients are for a test of null hypothesis that the true parameter is equal to zero against the alternative that the parameter is not equal to zero.

In our ROA regression example, the calculated t-statistic for the test of whether the slope coefficient is zero is 4.00131. The p-value corresponding to this test statistic is 0.008, which means there is just a 0.8% chance of rejecting the null hypotheses when it is true. Comparing this p-value with the level of significance of 5% (and critical values of ±2.776) leads us to easily reject the null hypothesis of H_0: $b_1 = 0$.

How do we determine the p-values? Since this is the area in the distribution outside the calculated test statistic, we need to resort to software tools. For the p-value corresponding to the t = 4.00131 from the ROA regression example, we could use the following:

- **Excel** 1-T.DIST(4.00131,4,TRUE))*2
- **R** (1-pt(4.00131,4))*2
- **Python** from scipy.stats import t *and* (1 - t.cdf(4.00131,4))*2

EXAMPLE 6

Hypothesis Testing of Simple Linear Regression Results

An analyst is interested in interpreting the results of and performing tests of hypotheses for the market model estimation that regresses the daily return on ABC stock on the daily return on the fictitious Europe–Asia–Africa (EAA) Equity Index, his proxy for the stock market. He has generated the regression results presented in Exhibit 30.

Exhibit 30 Selected Results of Estimation of Market Model for ABC Stock

Standard error of the estimate (s_e)	1.26
Standard deviation of ABC stock returns	0.80
Standard deviation of EAA Equity Index returns	0.70
Number of observations	1,200
	Coefficients
Intercept	0.010
Slope of EAA Equity Index returns	0.982

1 If the critical t-values are ±1.96 (at the 5% significance level), is the slope coefficient different from zero?

2 If the critical t-values are ±1.96 (at the 5% significance level), is the slope coefficient different from 1.0?

Solutions

1 First, we calculate the variation of the independent variable using the standard deviation of the independent variable:

$$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1} \times (n-1).$$

So,

$$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = 0.70^2 \times 1{,}199 = 587.51.$$

Next, the standard error of the estimated slope coefficient is

$$s_{\hat{b}_1} = \frac{s_e}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}} = \frac{1.26}{\sqrt{587.51}} = 0.051983,$$

and the test statistic is

$$t = \frac{\hat{b}_1 - B_1}{s_{\hat{b}_1}} = \frac{0.982 - 0}{0.051983} = 18.89079$$

The calculated test statistic is outside the bounds of ±1.96, so we reject the null hypothesis of a slope coefficient equal to zero.

2 The calculated test statistic for the test of whether the slope coefficient is equal to 1.0 is

$$t = \frac{0.982 - 1}{0.051983} = -0.3463.$$

The calculated test statistic is within the bounds of ±1.96, so we fail to reject the null hypothesis of a slope coefficient equal to 1.0, which is evidence that the true population slope may be 1.0.

PREDICTION USING SIMPLE LINEAR REGRESSION AND PREDICTION INTERVALS

6

g calculate and interpret the predicted value for the dependent variable, and a prediction interval for it, given an estimated linear regression model and a value for the independent variable

Financial analysts often want to use regression results to make predictions about a dependent variable. For example, we might ask, “How fast will the sales of XYZ Corporation grow this year if real GDP grows by 4%?” But we are not merely interested in making these forecasts; we also want to know how certain we can be about the forecasts’ results. A forecasted value of the dependent variable, $\hat{Y}_f$, is determined using the estimated intercept and slope, as well as the expected or forecasted independent variable, X_f:

$$\hat{Y}_f = \hat{b}_0 + \hat{b}_1 X_f \tag{18}$$

In our ROA regression model, if we forecast a company’s CAPEX to be 6%, the forecasted ROA based on our estimated equation is 12.375%:

$$\hat{Y}_f = 4.875 + \left(1.25 \times 6\right) = 12.375$$

However, we need to consider that the estimated regression line does not describe the relation between the dependent and independent variables perfectly; it is an average of the relation between the two variables. This is evident because the residuals are not all zero.

Therefore, an interval estimate of the forecast is needed to reflect this uncertainty. The estimated variance of the prediction error, s_f^2, of Y, given X, is

$$s_f^2 = s_e^2\left[1 + \frac{1}{n} + \frac{\left(X_f - \bar{X}\right)^2}{(n-1)s_X^2}\right] = s_e^2\left[1 + \frac{1}{n} + \frac{\left(X_f - \bar{X}\right)^2}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\right],$$

and the **standard error of the forecast** is

$$s_f = s_e\sqrt{1 + \frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}. \tag{19}$$

The standard error of the forecast depends on

- the standard error of the estimate, s_e;
- the number of observations, n;
- the forecasted value of the independent variable, X_f, used to predict the dependent variable and its deviation from the estimated mean, $\bar{X}$; and
- the variation of the independent variable.

We can see the following from the equation for the standard error of the forecast:

1 The better the fit of the regression model, the smaller the standard error of the estimate (s_e) and, therefore, the smaller standard error of the forecast.

2 The larger the sample size (n) in the regression estimation, the smaller the standard error of the forecast.

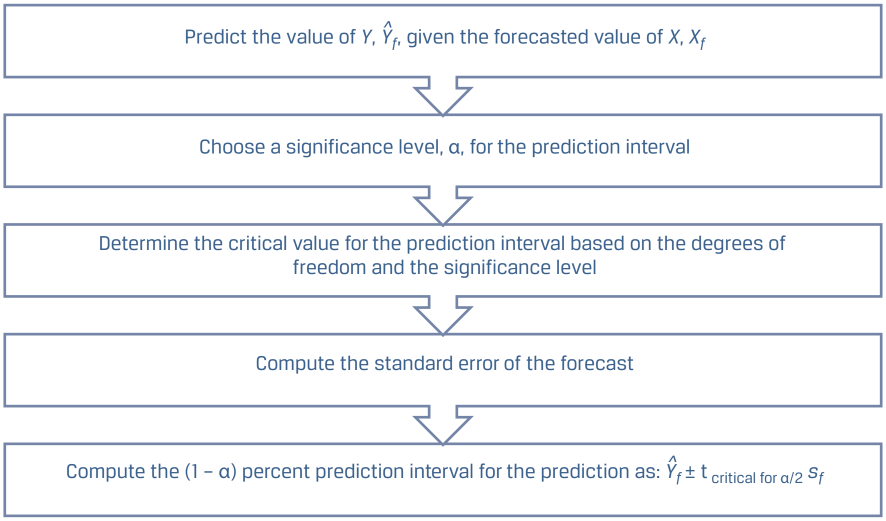
3 The closer the forecasted independent variable (X_f) is to the mean of the independent variable ($\bar{X}$) used in the regression estimation, the smaller the standard error of the forecast.

Once we have this estimate of the standard error of the forecast, determining a prediction interval around the predicted value of the dependent variable ($\hat{Y}_f$) is very similar to estimating a confidence interval around an estimated parameter. The prediction interval is

$$\hat{Y}_f \pm t_{critical\ for\ \alpha/2} s_f. \tag{20}$$

We outline the steps for developing the prediction interval in Exhibit 31.

Exhibit 31 Creating a Prediction Interval around the Predicted Dependent Variable


Predict the value of Y, $\hat{Y}_f$, given the forecasted value of X, X_f
Choose a significance level, α, for the prediction interval
Determine the critical value for the prediction interval based on the degrees of freedom and the significance level
Compute the standard error of the forecast
Compute the (1 – α) percent prediction interval for the prediction as: $\hat{Y}_f \pm t_{\text{critical for } \alpha/2}\, s_f$

For our ROA regression model, given that the forecasted value of CAPEX is 6.0, the predicted value of Y is 12.375:

$$\hat{Y}_f = 4.875 + 1.25X_f = 4.875 + (1.25 \times 6.0) = 12.375.$$

Assuming a 5% significance level (α), two sided, with n – 2 degrees of freedom (so, df = 4), the critical values for the prediction interval are ±2.776.

The standard error of the forecast is

$$s_f = 3.459588\sqrt{1 + \frac{1}{6} + \frac{(6 - 6.1)^2}{122.640}} = 3.459588\sqrt{1.166748} = 3.736912.$$

The 95% prediction interval then becomes

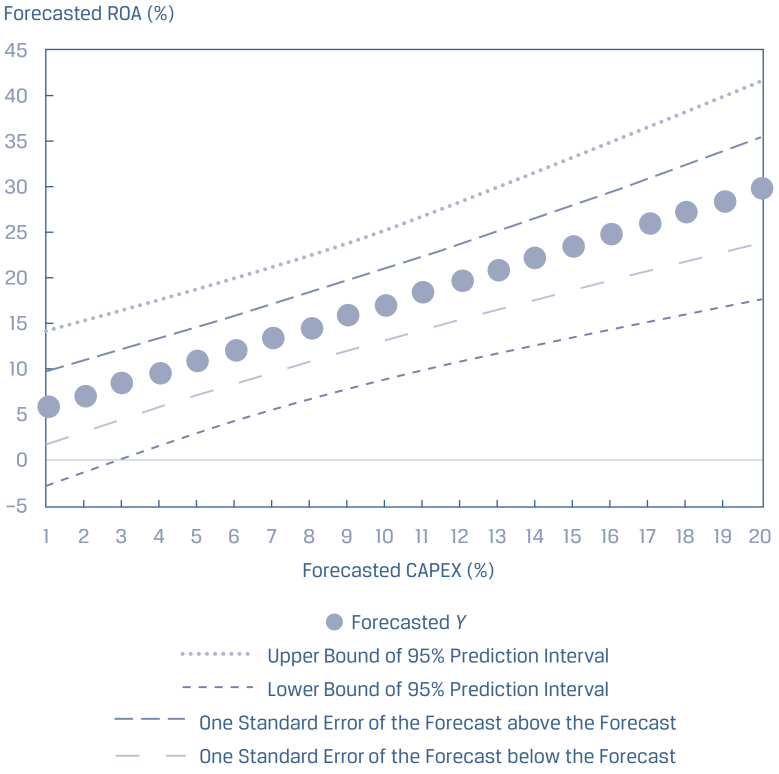
$$12.375 \pm 2.776(3.736912)$$

$$12.375 \pm 10.3737$$

$$\{2.0013 < \hat{Y}_f < 22.7487\}$$

For our ROA regression example, we can see how the standard error of the forecast (s_f) changes as our forecasted value of the independent variable gets farther from the mean of the independent variable $(X_f - \overline{X})$ in Exhibit 32. The mean of CAPEX is 6.1%, and the band that represents one standard error of the forecast, above and below the forecast, is minimized at that point and increases as the independent variable gets farther from $\overline{X}$.

Exhibit 32 ROA Forecasts and Standard Error of the Forecast



EXAMPLE 7

Predicting Net Profit Margin Using R&D Spending

Suppose we want to forecast a company's net profit margin (NPM) based on its research and development expenditures scaled by revenues (RDR), using the model estimated in Example 2 and the details provided in Exhibit 8. The regression model was estimated using data on eight companies as

$$\hat{Y}_f = 16.5 - 1.3X_f,$$

with a standard error of the estimate (s_e) of 1.8618987 and variance of RDR, $\dfrac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}$, of 4.285714, as given.

1 What is the predicted value of NPM if the forecasted value of RDR is 5?

2 What is the standard error of the forecast (s_f) if the forecasted value of RDR is 5?

3 What is the 95% prediction interval for the predicted value of NPM using critical *t*-values (df = 6) of ±2.447?

4 What is the predicted value of NPM if the forecasted value of RDR is 15?

5 What is the standard error of the forecast if the forecasted value of RDR is 15?

6 What is the 95% prediction interval for the predicted value of NPM using critical *t*-values (df = 6) of ±2.447?

Solutions

1 The predicted value of NPM is 10: 16.5 − (1.3 × 5) = 10.

2 To derive the standard error of the forecast (s_f), we first have to calculate the variation of RDR. Then, we have the all the pieces to calculate s_f:

$$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = 4.285714 \times 7 = 30.$$

$$s_f = 1.8618987\sqrt{1 + \frac{1}{8} + \frac{(5 - 7.5)^2}{30}} = 2.1499.$$

3 The 95% prediction interval for the predicted value of NPM is

$$\left\{10 \pm 2.447(2.1499)\right\}$$

$$\left\{4.7392 < \hat{Y}_f < 15.2608\right\}$$

4 The predicted value of NPM is −3: 16.5 − (1.3 × 15) = −3.

5 To derive the standard error of the forecast, we first must calculate the variation of RDR. Then, we can calculate s_f:

$$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = 4.285714 \times 7 = 30.$$

$$s_f = 1.8618987\sqrt{1 + \frac{1}{8} + \frac{(15 - 7.5)^2}{30}} = 3.2249.$$

6 The 95% prediction interval for the predicted value of NPM is

$$\left\{-3 \pm 2.447(3.2249)\right\}$$

$$\left\{-10.8913 < \hat{Y}_f < 4.8913\right\}$$

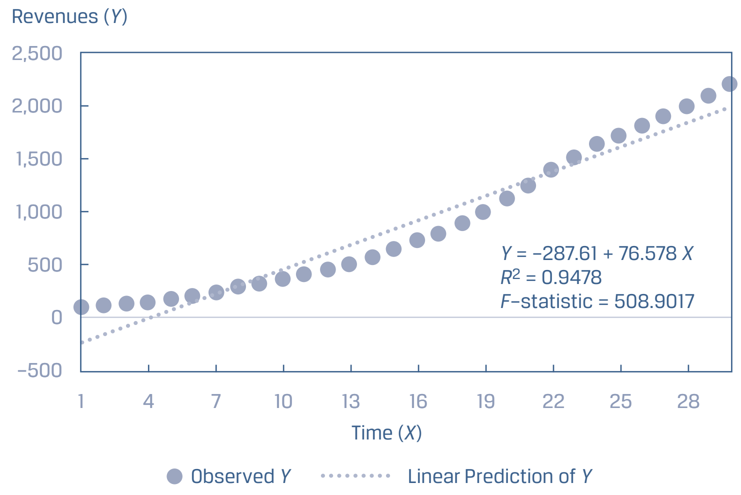
FUNCTIONAL FORMS FOR SIMPLE LINEAR REGRESSION

7

h describe different functional forms of simple linear regressions

Not every set of independent and dependent variables has a linear relation. In fact, we often see non-linear relationships in economic and financial data. Consider the revenues of a company over time illustrated in Exhibit 33, with revenues as the dependent (Y) variable and time as the independent (X) variable. Revenues grow at a rate of 15% per year for several years, but then the growth rate eventually declines to just 5% per year. Estimating this relationship as a simple linear model would understate the dependent variable, revenues, for some ranges of the independent variable, time, and would overstate it for other ranges of the independent variable.

Exhibit 33 Company Revenues over Time



We can still use the simple linear regression model, but we need to modify either the dependent or the independent variables to make it work well. This is the case with many different financial or economic data that you might use as dependent and independent variables in your regression analysis.

There are several different functional forms that can be used to potentially transform the data to enable their use in linear regression. These transformations include using the log (i.e., natural logarithm) of the dependent variable, the log of the independent variable, the reciprocal of the independent variable, the square of the independent variable, or the differencing of the independent variable. We illustrate and discuss three often-used functional forms, each of which involves log transformation:

1. the **log-lin model**, in which the dependent variable is logarithmic but the independent variable is linear;
2. the **lin-log model**, in which the dependent variable is linear but the independent variable is logarithmic; and
3. the **log-log model**, where both the dependent and independent variables are in logarithmic form.

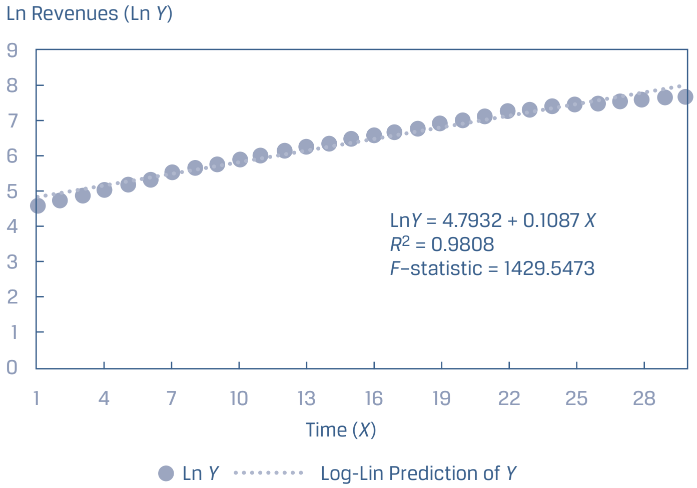
7.1 The Log-Lin Model

In the log-lin model, the dependent variable is in logarithmic form and the independent variable is not, as follows:

$$\ln Y_i = b_0 + b_1 X_i. \quad (21)$$

The slope coefficient in this model is the relative change in the dependent variable for an absolute change in the independent variable. We can transform the Y variable (revenues) in Exhibit 33 into its natural log (ln) and then fit the regression line, as we show in Exhibit 34. From this chart, we see that the log-lin model is a better fitting model than the simple linear regression model.

Exhibit 34 Log-Lin Model Applied to Company Revenues over Time



It is important to note that in working with a log-lin model, you must take care when making a forecast. For example, suppose the estimated regression model is ln $Y = -7 + 2X$. If X is 2.5%, then the forecasted value of ln Y is −2. In this case, the predicted value of Y is the antilog of −2, or $e^{-2} = 0.135335$. Another caution is that you cannot directly compare a log-lin model with a lin-lin model (that is, the regression of Y on X without any transformation) because the dependent variables are not in the same form – we would have to transform the R^2 and F-statistic to enable a comparison. However, looking at the residuals is helpful.

7.2 The Lin-Log Model

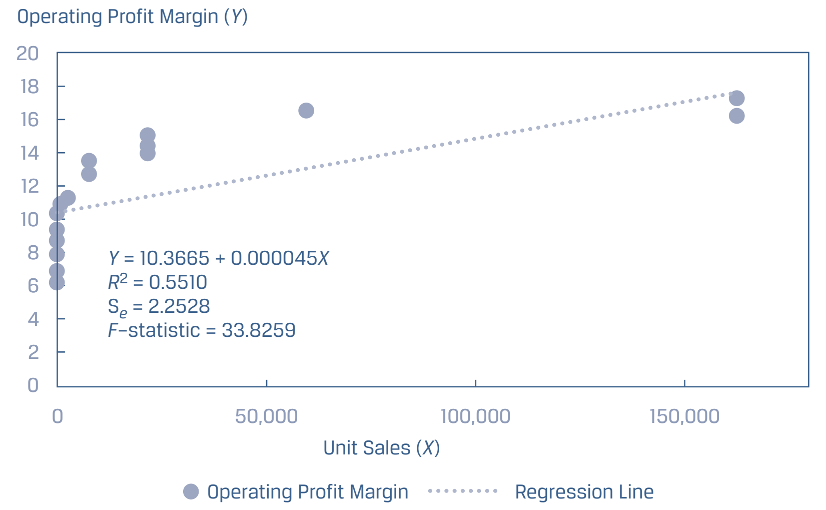
The lin-log model is similar to the log-lin model, but only the independent variable is in logarithmic form:

$$Y_i = b_0 + b_1 \ln X_i, \tag{22}$$

The slope coefficient in this regression model provides the absolute change in the dependent variable for a relative change in the independent variable.

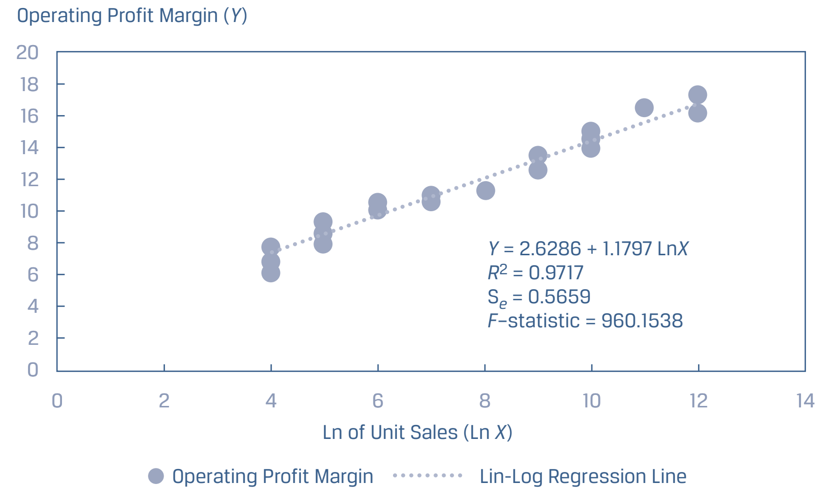
Suppose an analyst is examining the cross-sectional relationship between operating profit margin, the dependent variable (Y), and unit sales, the independent variable (X), and gathers data on a sample of 30 companies. The scatter plot and regression line for these observations are shown in Exhibit 35. Although the slope is different from zero at the 5% level (the calculated t-statistic on the slope is 5.8616, compared with critical t-values of ±2.048), given the R^2 of 55.10%, the issue is whether we can get a better fit by using a different functional form.

Exhibit 35 Relationship between Operating Profit Margin and Unit Sales



If instead we use the natural log of the unit sales as the independent variable in our model, we get a very different picture, as shown in Exhibit 36. The R^2 for the model of operating profit margin regressed on the natural log of unit sales jumps to 97.17%. Since the dependent variable is the same in both the original and transformed models, we can compare the standard error of the estimate: 2.2528 with the original independent variable and a much lower 0.5629 with the transformed independent variable. Clearly the log-transformed explanatory variable has resulted in a better fitting model.

Exhibit 36 Relationship Between Operating Profit Margin and Natural Logarithm of Unit Sales



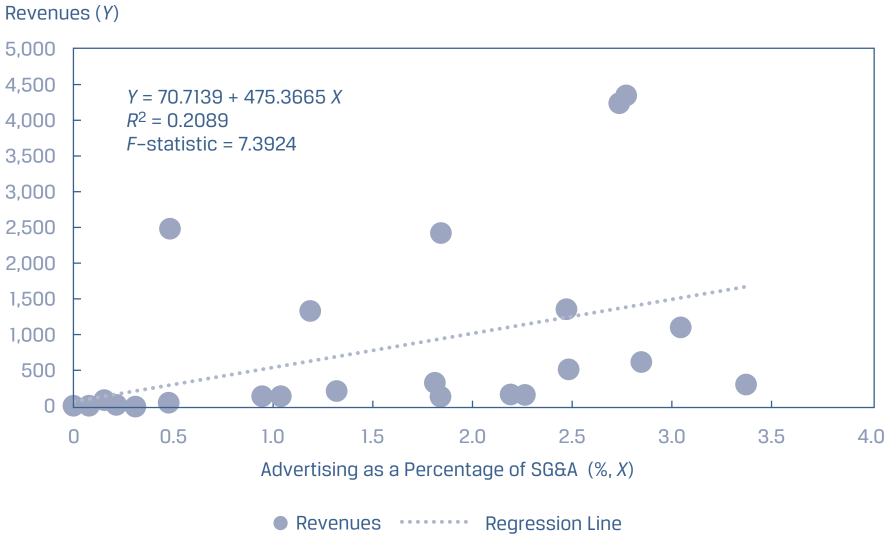
7.3 The Log-Log Model

The log-log model, in which both the dependent variable and the independent variable are linear in their logarithmic forms, is also referred to as the double-log model.

$$\ln Y_i = b_0 + b_1 \ln X_i. \quad (23)$$

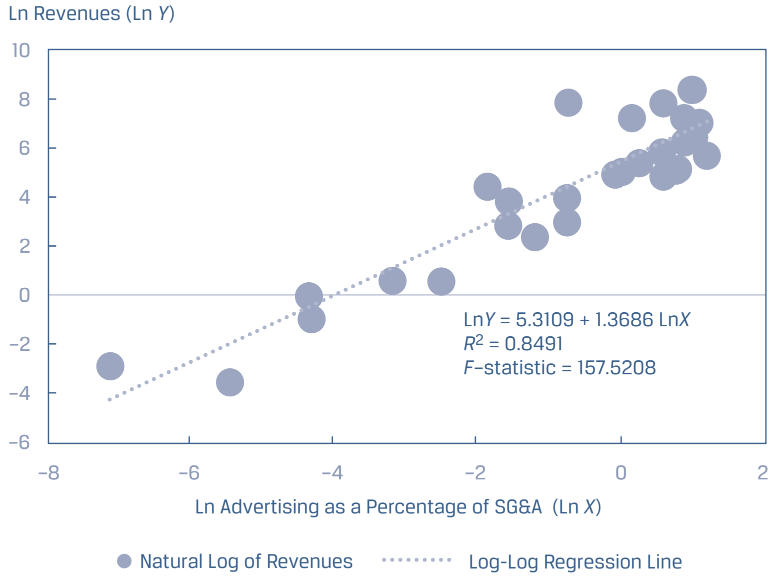
This model is useful in calculating elasticities because the slope coefficient is the relative change in the dependent variable for a relative change in the independent variable. Consider a cross-sectional model of company revenues (the *Y* variable) regressed on advertising spending as a percentage of selling, general, and administrative expenses, ADVERT (the *X* variable). As shown in Exhibit 37, a simple linear regression model results in a shallow regression line, with a coefficient of determination of just 20.89%.

Exhibit 37 Fitting a Linear Relation Between Revenues and Advertising Spending



However, if instead we use the natural logarithms of both the revenues and ADVERT, we get a much different picture of this relationship. As shown in Exhibit 38, the estimated regression line has a significant positive slope; the log-log model's R^2 increases by more than four times, from 20.89% to 84.91%; and the *F*-statistic jumps from 7.39 to 157.52. So, using the log-log transformation dramatically improves the regression model fit relative to our data.

Exhibit 38 Fitting a Log-Log Model of Revenues and Advertising Spending



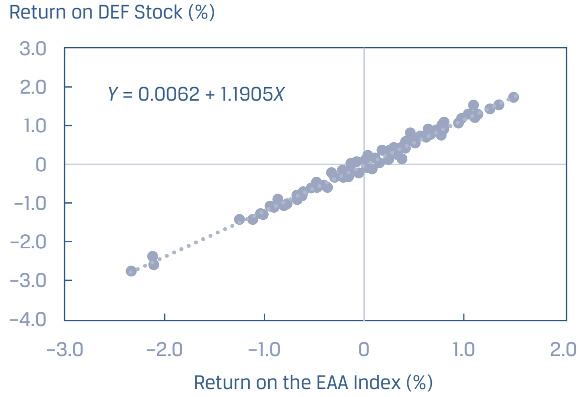
7.4 Selecting the Correct Functional Form

The key to fitting the appropriate functional form of a simple linear regression is examining the goodness of fit measures—the coefficient of determination (R^2), the F-statistic, and the standard error of the estimate (s_e)—as well as examining whether there are patterns in the residuals. In addition to fit statistics, most statistical packages provide plots of residuals as part of the regression output, which enables you to visually inspect the residuals. To reiterate an important point, what you want to see in these plots is random residuals.

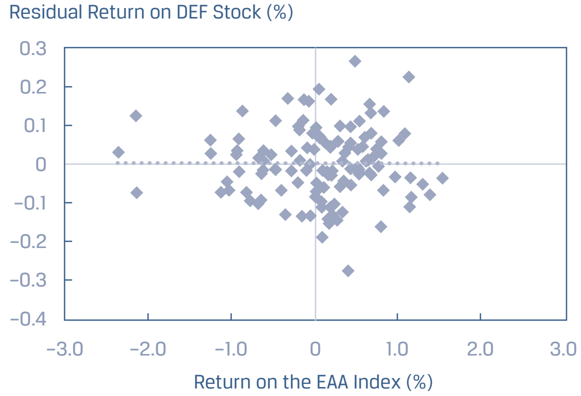
As an example, consider the relationship between the monthly returns on DEF stock and the monthly returns of the EAA Equity Index, as depicted in Panel A of Exhibit 39, with the regression line indicated. Using the equation for this regression line, we calculate the residuals and plot them against the EAA Equity Index, as shown in Panel B of Exhibit 39. The residuals appear to be random, bearing no relation to the independent variable. The distribution of the residuals, shown in Panel C of Exhibit 39, shows that the residuals are approximately normal. Using statistical software, we can investigate further by examining the distribution of the residuals, including using a normal probability plot or statistics to test for normality of the residuals.

Exhibit 39 Monthly Returns on DEF Stock Regressed on Returns on the EAA Index

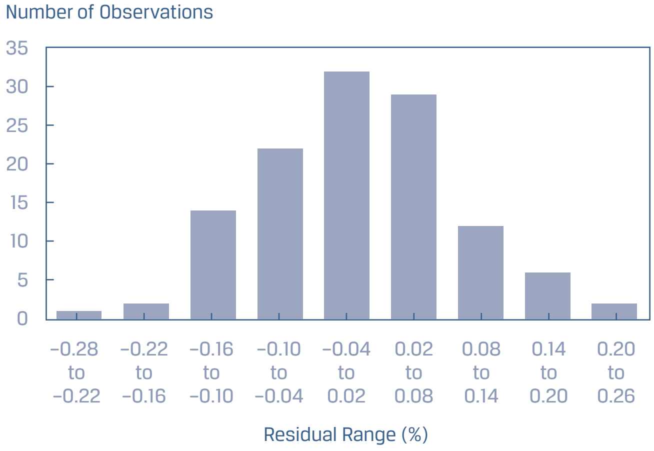
A. Scatterplot of Returns on DEF Stock and Return on the EAA Index



B. Scatterplot of Residuals and the Returns on the EAA Index



C. Histogram of Residuals



EXAMPLE 8

Comparing Functional Forms

An analyst is investigating the relationship between the annual growth in consumer spending (CONS) in a country and the annual growth in the country's GDP (GGDP). The analyst estimates the following two models:

	Model 1	Model 2
	$GGDP_i = b_0 + b_1 CONS_i + \varepsilon_i$	$GGDP_i = b_0 + b_1 \ln(CONS_i) + \varepsilon_i$
Intercept	1.040	1.006
Slope	0.669	1.994
R^2	0.788	0.867
Standard error of the estimate	0.404	0.320
F-statistic	141.558	247.040

1 Identify the functional form used in these models.

2 Explain which model has better goodness-of-fit with the sample data.

Solution

1 Model 1 is the simple linear regression with no variable transformation, whereas Model 2 is a lin-log model with the natural log of the variable CONS as the independent variable.

2 The lin-log model, Model 2, fits the data better. Since the dependent variable is the same for the two models, we can compare the fit of the models using either the relative measures (R^2 or F-statistic) or the absolute measure of fit, the standard error of the estimate. The standard error of the estimate is lower for Model 2, whereas the R^2 and F-statistic are higher for Model 2 compared with Model 1.

SUMMARY

- The dependent variable in a linear regression is the variable whose variability the regression model tries to explain. The independent variable is the variable whose variation the researcher uses to explain the variation of the dependent variable.
- If there is one independent variable in a linear regression and there are n observations of the dependent and independent variables, the regression model is $Y_i = b_0 + b_1 X_i + \varepsilon_i$, $i = 1, \ldots, n$, where Y_i is the dependent variable, X_i is the independent variable, and ε_i is the error term. In this model, the coefficients b_0 and b_1 are the population intercept and slope, respectively.

- The intercept is the expected value of the dependent variable when the independent variable has a value of zero. The slope coefficient is the estimate of the population slope of the regression line and is the expected change in the dependent variable for a one-unit change in the independent variable.
- The assumptions of the classic simple linear regression model are as follows:
  - Linearity: A linear relation exists between the dependent variable and the independent variable.
  - Homoskedasticity: The variance of the error term is the same for all observations.
  - Independence: The error term is uncorrelated across observations.
  - Normality: The error term is normally distributed.
- The estimated parameters in a simple linear regression model minimize the sum of the squared errors.
- The coefficient of determination, or R^2, measures the percentage of the total variation in the dependent variable explained by the independent variable.
- To test the fit of the simple linear regression, we can calculate an F-distributed test statistic and test the hypotheses H_0: $b_1 = 0$ versus H_a: $b_1 \neq 0$, with 1 and $n - 2$ degrees of freedom.
- The standard error of the estimate is an absolute measure of the fit of the model calculated as the square root of the mean square error.
- We can evaluate a regression model by testing whether the population value of a regression coefficient is equal to a particular hypothesized value. We do this by calculating a t-distributed test statistic that compares the estimated parameter with the hypothesized parameter, dividing this difference by the standard error of the coefficient.
- An indicator (or dummy) variable takes on only the values 0 or 1 and can be used as the independent variable in a simple linear regression. In such a model, the interpretation of the intercept is the predicted value of the dependent variable if the indicator variable is 0, and when the indicator variable is 1, the slope is the difference in the means if we grouped the observations by the indicator variable.
- We calculate a prediction interval for a regression coefficient using the estimated coefficient, the standard error of the estimated coefficient, and the critical value for the t-distributed test statistic based on the level of significance and the appropriate degrees of freedom, which are $n - 2$ for simple regression.
- We can make predictions for the dependent variable using an estimated linear regression by inserting the forecasted value of the independent variable into the estimated model.
- The standard error of the forecast is the product of the standard error of the estimate and a term that reflects the sample size of the regression, the variation of the independent variable, and the deviation between the forecasted value of the independent variable and the mean of the independent variable in the regression.

- The prediction interval for a particular forecasted value of the dependent variable is formed by using the forecasted value of the dependent variable and extending above and below this value a quantity that reflects the critical *t*-value corresponding to the degrees of freedom, the level of significance, and the standard error of the forecast.
- If the relationship between the independent variable and the dependent variable is not linear, we can often transform one or both of these variables to convert this relation to a linear form, which then allows the use of simple linear regression.

PRACTICE PROBLEMS

1 Julie Moon is an energy analyst examining electricity, oil, and natural gas consumption in different regions over different seasons. She ran a simple regression explaining the variation in energy consumption as a function of temperature. The total variation of the dependent variable was 140.58, and the explained variation was 60.16. She had 60 monthly observations.

A Calculate the coefficient of determination.

B Calculate the *F*-statistic to test the fit of the model.

C Calculate the standard error of the estimate of the regression estimation.

D Calculate the sample standard deviation of monthly energy consumption.

2 Homoskedasticity is best described as the situation in which the variance of the residuals of a regression is:

A zero.

B normally distributed.

C constant across observations.

The following information relates to Questions 3–6

An analyst is examining the annual growth of the money supply for a country over the past 30 years. This country experienced a central bank policy shift 15 years ago, which altered the approach to the management of the money supply. The analyst estimated a model using the annual growth rate in the money supply regressed on the variable (SHIFT) that takes on a value of 0 before the policy shift and 1 after. She estimated the following:

	Coefficients	Standard Error	*t*-Stat.
Intercept	5.767264	0.445229	12.95348
SHIFT	−5.13912	0.629649	−8.16188

Critical *t*-values, level of significance of 0.05:

One-sided, left side: −1.701

One-sided, right side: +1.701

Two-sided: ±2.048

3 The variable SHIFT is best described as:

A an indicator variable.

B a dependent variable.

C a continuous variable.

4 The interpretation of the intercept is the mean of the annual growth rate of the money supply:

A over the enter entire period.

© 2019 CFA Institute. All rights reserved.

- **B** after the shift in policy.
- **C** before the shift in policy.

5 The interpretation of the slope is the:

- **A** change in the annual growth rate of the money supply per year.
- **B** average annual growth rate of the money supply after the shift in policy.
- **C** difference in the average annual growth rate of the money supply from before to after the shift in policy.

6 Testing whether there is a change in the money supply growth after the shift in policy, using a 0.05 level of significance, we conclude that there is:

- **A** sufficient evidence that the money supply growth changed.
- **B** not enough evidence that the money supply growth is different from zero.
- **C** not enough evidence to indicate that the money supply growth changed.

---

7 You are examining the results of a regression estimation that attempts to explain the unit sales growth of a business you are researching. The analysis of variance output for the regression is given in the following table. The regression was based on five observations (n = 5).

Source	df	Sum of Squares	Mean Square	*F*	*p*-Value
Regression	1	88.0	88.0	36.667	0.00904
Residual	3	7.2	2.4		
Total	4	95.2			

- **A** Calculate the sample variance of the dependent variable using information in the table.
- **B** Calculate the coefficient of determination for this estimated model.
- **C** What hypothesis does the *F*-statistic test?
- **D** Is the *F*-test significant at the 0.05 significance level?
- **E** Calculate the standard error of the estimate.

8 An economist collected the monthly returns for KDL's portfolio and a diversified stock index. The data collected are shown in the following table:

Month	Portfolio Return (%)	Index Return (%)
1	1.11	−0.59
2	72.10	64.90
3	5.12	4.81
4	1.01	1.68
5	−1.72	−4.97
6	4.06	−2.06

The economist calculated the correlation between the two returns and found it to be 0.996. The regression results with the KDL return as the dependent variable and the index return as the independent variable are given as follows:

Regression Statistics	
R^2	0.9921
Standard error	2.8619
Observations	6

Source	df	Sum of Squares	Mean Square	*F*	*p*-Value
Regression	1	4,101.6205	4,101.6205	500.7921	0.0000
Residual	4	32.7611	8.1903		
Total	5	4,134.3815			

	Coefficients	Standard Error	*t*-Statistic	*p*-Value
Intercept	2.2521	1.2739	1.7679	0.1518
Index return (%)	1.0690	0.0478	22.3784	0.0000

When reviewing the results, Andrea Fusilier suspected that they were unreliable. She found that the returns for Month 2 should have been 7.21% and 6.49%, instead of the large values shown in the first table. Correcting these values resulted in a revised correlation of 0.824 and the following revised regression results:

Regression Statistics	
R^2	0.6784
Standard error	2.0624
Observations	6

Source	df	Sum of Squares	Mean Square	*F*	*p*-Value
Regression	1	35.8950	35.8950	8.4391	0.044
Residual	4	17.0137	4.2534		
Total	5	52.91			

	Coefficients	Standard Error	*t*-Statistic	*p*-Value
Intercept	2.2421	0.8635	2.5966	0.060
Slope	0.6217	0.2143	2.9050	0.044

Explain how the bad data affected the results.

The following information relates to Questions 9–12

Kenneth McCoin, CFA, is a challenging interviewer. Last year, he handed each job applicant a sheet of paper with the information in the following table, and he then asked several questions about regression analysis. Some of McCoin's questions, along with a sample of the answers he received to each, are given below. McCoin told the applicants that the independent variable is the ratio of net income to sales for restaurants with a market cap of more than $100 million and the dependent variable is the ratio of cash flow from operations to sales for those restaurants. Which of the choices provided is the best answer to each of McCoin's questions?

Regression Statistics	
R^2	0.7436
Standard error	0.0213
Observations	24

Source	**df**	**Sum of Squares**	**Mean Square**	***F***	***p*-Value**
Regression	1	0.029	0.029000	63.81	0
Residual	22	0.010	0.000455		
Total	23	0.040			

	Coefficients	**Standard Error**	***t*-Statistic**	***p*-Value**
Intercept	0.077	0.007	11.328	0
Net income to sales (%)	0.826	0.103	7.988	0

9 The coefficient of determination is *closest* to:

A 0.7436.

B 0.8261.

C 0.8623.

10 The correlation between X and Y is *closest* to:

A −0.7436.

B 0.7436.

C 0.8623.

11 If the ratio of net income to sales for a restaurant is 5%, the predicted ratio of cash flow from operations (CFO) to sales is *closest* to:

A −4.054.

B 0.524.

C 4.207.

12 Is the relationship between the ratio of cash flow to operations and the ratio of net income to sales significant at the 0.05 level?

A No, because the R^2 is greater than 0.05

B No, because the p-values of the intercept and slope are less than 0.05

C Yes, because the p-values for F and t for the slope coefficient are less than 0.05

The following information relates to Questions 13–17

Howard Golub, CFA, is preparing to write a research report on Stellar Energy Corp. common stock. One of the world's largest companies, Stellar is in the business of refining and marketing oil. As part of his analysis, Golub wants to evaluate the sensitivity of the stock's returns to various economic factors. For example, a client recently asked Golub whether the price of Stellar Energy Corp. stock has tended to rise following increases in retail energy prices. Golub believes the association between the two variables is negative, but he does not know the strength of the association.

Golub directs his assistant, Jill Batten, to study the relationships between (1) Stellar monthly common stock returns and the previous month's percentage change in the US Consumer Price Index for Energy (CPIENG) and (2) Stellar monthly common stock returns and the previous month's percentage change in the US Producer Price Index for Crude Energy Materials (PPICEM). Golub wants Batten to run both a correlation and a linear regression analysis. In response, Batten compiles the summary statistics shown in Exhibit 1 for 248 months. All the data are in decimal form, where 0.01 indicates a 1% return. Batten also runs a regression analysis using Stellar monthly returns as the dependent variable and the monthly change in CPIENG as the independent variable. Exhibit 2 displays the results of this regression model.

Exhibit 1 Descriptive Statistics

	Stellar Common Stock Monthly Return	Lagged Monthly Change	
		CPIENG	PPICEM
Mean	0.0123	0.0023	0.0042
Standard deviation	0.0717	0.0160	0.0534
Covariance, Stellar vs. CPIENG	−0.00017		
Covariance, Stellar vs. PPICEM	−0.00048		
Covariance, CPIENG vs. PPICEM	0.00044		
Correlation, Stellar vs. CPIENG	−0.1452		

Exhibit 2 Regression Analysis with CPIENG

Regression Statistics	
R^2	0.0211
Standard error of the estimate	0.0710
Observations	248

(continued)

Exhibit 2 (Continued)

Regression Statistics

	Coefficients	Standard Error	*t*-Statistic
Intercept	0.0138	0.0046	3.0275
CPIENG (%)	−0.6486	0.2818	−2.3014

Critical *t*-values

One-sided, left side: −1.651

One-sided, right side: +1.651

Two-sided: ±1.967

13 Which of the following best describes Batten's regression?

A Time-series regression

B Cross-sectional regression

C Time-series and cross-sectional regression

14 Based on the regression, if the CPIENG *decreases* by 1.0%, the expected return on Stellar common stock during the next period is *closest* to:

A 0.0073 (0.73%).

B 0.0138 (1.38%).

C 0.0203 (2.03%).

15 Based on Batten's regression model, the coefficient of determination indicates that:

A Stellar's returns explain 2.11% of the variability in CPIENG.

B Stellar's returns explain 14.52% of the variability in CPIENG.

C changes in CPIENG explain 2.11% of the variability in Stellar's returns.

16 For Batten's regression model, 0.0710 is the standard deviation of:

A the dependent variable.

B the residuals from the regression.

C the predicted dependent variable from the regression.

17 For the analysis run by Batten, which of the following is an *incorrect* conclusion from the regression output?

A The estimated intercept from Batten's regression is statistically different from zero at the 0.05 level of significance.

B In the month after the CPIENG declines, Stellar's common stock is expected to exhibit a positive return.

C Viewed in combination, the slope and intercept coefficients from Batten's regression are not statistically different from zero at the 0.05 level of significance.

The following information relates to Questions 18–26

Anh Liu is an analyst researching whether a company's debt burden affects investors' decision to short the company's stock. She calculates the short interest ratio (the ratio of short interest to average daily share volume, expressed in days) for 50 companies as of the end of 2016 and compares this ratio with the companies' debt ratio (the ratio of total liabilities to total assets, expressed in decimal form).

Liu provides a number of statistics in Exhibit 1. She also estimates a simple regression to investigate the effect of the debt ratio on a company's short interest ratio. The results of this simple regression, including the analysis of variance (ANOVA), are shown in Exhibit 2.

In addition to estimating a regression equation, Liu graphs the 50 observations using a scatter plot, with the short interest ratio on the vertical axis and the debt ratio on the horizontal axis.

Exhibit 1 Summary Statistics

Statistic	Debt Ratio X_i	Short Interest Ratio Y_i
Sum	19.8550	192.3000
Sum of squared deviations from the mean	$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = 2.2225.$	$\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2 = 412.2042.$
Sum of cross-products of deviations from the mean	$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right) = -9.2430.$	

Exhibit 2 Regression of the Short Interest Ratio on the Debt Ratio

ANOVA	Degrees of Freedom (df)	Sum of Squares	Mean Square
Regression	1	38.4404	38.4404
Residual	48	373.7638	7.7867
Total	49	412.2042	

Regression Statistics	
R^2	0.0933
Standard error of estimate	2.7905
Observations	50

	Coefficients	Standard Error	*t*-Statistic
Intercept	5.4975	0.8416	6.5322
Debt ratio (%)	–4.1589	1.8718	–2.2219

Critical t-values for a 0.05 level of significance:

One-sided, left side: −1.677

One-sided, right side: +1.677

Two-sided: ±2.011

Liu is considering three interpretations of these results for her report on the relationship between debt ratios and short interest ratios:

Interpretation 1	Companies' higher debt ratios cause lower short interest ratios.
Interpretation 2	Companies' higher short interest ratios cause higher debt ratios.
Interpretation 3	Companies with higher debt ratios tend to have lower short interest ratios.

She is especially interested in using her estimation results to predict the short interest ratio for MQD Corporation, which has a debt ratio of 0.40.

18 Based on Exhibits 1 and 2, if Liu were to graph the 50 observations, the scatter plot summarizing this relation would be *best* described as:

A horizontal.

B upward sloping.

C downward sloping.

19 Based on Exhibit 1, the sample covariance is *closest to*:

A −9.2430.

B −0.1886.

C 8.4123.

20 Based on Exhibits 1 and 2, the correlation between the debt ratio and the short interest ratio is *closest to*:

A −0.3054.

B 0.0933.

C 0.3054.

21 Which of the interpretations *best* describes Liu's findings?

A Interpretation 1

B Interpretation 2

C Interpretation 3

22 The dependent variable in Liu's regression analysis is the:

A intercept.

B debt ratio.

C short interest ratio.

23 Based on Exhibit 2, the degrees of freedom for the t-test of the slope coefficient in this regression are:

A 48.

B 49.

C 50.

24 Which of the following should Liu conclude from the results shown in Exhibit 2?

A The average short interest ratio is 5.4975.

B The estimated slope coefficient is different from zero at the 0.05 level of significance.

C The debt ratio explains 30.54% of the variation in the short interest ratio.

25 Based on Exhibit 2, the short interest ratio expected for MQD Corporation is *closest* to:

A 3.8339.

B 5.4975.

C 6.2462.

26 Based on Liu's regression results in Exhibit 2, the *F*-statistic for testing whether the slope coefficient is equal to zero is *closest* to:

A −2.2219.

B 3.5036.

C 4.9367.

The following information relates to Questions 27–31

Elena Vasileva recently joined EnergyInvest as a junior portfolio analyst. Vasileva's supervisor asks her to evaluate a potential investment opportunity in Amtex, a multinational oil and gas corporation based in the United States. Vasileva's supervisor suggests using regression analysis to examine the relation between Amtex shares and returns on crude oil.

Vasileva notes the following assumptions of regression analysis:

Assumption 1	The error term is uncorrelated across observations.
Assumption 2	The variance of the error term is the same for all observations.
Assumption 3	The dependent variable is normally distributed.

Vasileva runs a regression of Amtex share returns on crude oil returns using the monthly data she collected. Selected data used in the regression are presented in Exhibit 1, and selected regression output is presented in Exhibit 2. She uses a 1% level of significance in all her tests.

Exhibit 1 Selected Data for Crude Oil Returns and Amtex Share Returns

	Oil Return (X_i)	Amtex Return (Y_i)	Cross-Product $(X_i - \bar{X})(Y_i - \bar{Y})$	Predicted Amtex Return $\hat{Y}_i$	Regression Residual $Y_i - \hat{Y}_i$	Squared Residual $(Y_i - \hat{Y}_i)^2$
Month 1	−0.032000	0.033145	−0.000388	0.002011	−0.031134	0.000969
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Month 36	0.028636	0.062334	0.002663	0.016282	−0.046053	0.002121
Sum			0.085598			0.071475
Average	−0.018056	0.005293				

Exhibit 2 Selected Regression Output, Dependent Variable: Amtex Share Return

	Coefficient	Standard Error
Intercept	0.0095	0.0078
Oil return	0.2354	0.0760

Critical t-values for a 1% level of significance:

One-sided, left side: −2.441

One-sided, right side: +2.441

Two-sided: ±2.728

Vasileva expects the crude oil return next month, Month 37, to be −0.01. She computes the standard error of the forecast to be 0.0469.

27 Which of Vasileva's assumptions regarding regression analysis is *incorrect*?

A Assumption 1

B Assumption 2

C Assumption 3

28 Based on Exhibit 1, the standard error of the estimate is *closest* to:

A 0.04456.

B 0.04585.

C 0.05018.

29 Based on Exhibit 2, Vasileva should reject the null hypothesis that:

A the slope is less than or equal to 0.15.

B the intercept is less than or equal to zero.

C crude oil returns do not explain Amtex share returns.

30 Based on Exhibit 2 and Vasileva's prediction of the crude oil return for Month 37, the estimate of Amtex share return for Month 37 is *closest* to:

A −0.0024.

B 0.0071.

C 0.0119.

31 Using information from Exhibit 2, the 99% prediction interval for Amtex share return for Month 37 is *best* described as:

A $\hat{Y}_f \pm 0.0053$.

B $\hat{Y}_f \pm 0.0469$.

C $\hat{Y}_f \pm 0.1279$.

The following information relates to Questions 32–34

Doug Abitbol is a portfolio manager for Polyi Investments, a hedge fund that trades in the United States. Abitbol manages the hedge fund with the help of Robert Olabudo, a junior portfolio manager.

Abitbol looks at economists' inflation forecasts and would like to examine the relationship between the US Consumer Price Index (US CPI) consensus forecast and the actual US CPI using regression analysis. Olabudo estimates regression coefficients to test whether the consensus forecast is unbiased. If the consensus forecasts are unbiased, the intercept should be 0.0 and the slope will be equal to 1.0. Regression results are presented in Exhibit 1. Additionally, Olabudo calculates the 95% prediction interval of the actual CPI using a US CPI consensus forecast of 2.8.

Exhibit 1 Regression Output: Estimating US CPI

Regression Statistics

R^2	0.9859		
Standard error of estimate	0.0009		
Observations	60		
	Coefficients	**Standard Error**	***t*-Statistic**
Intercept	0.0001	0.0002	0.5000
US CPI consensus forecast	0.9830	0.0155	63.4194

Notes:

1 The absolute value of the critical value for the *t*-statistic is 2.002 at the 5% level of significance.

2 The standard deviation of the US CPI consensus forecast is s_x = 0.7539.

3 The mean of the US CPI consensus forecast is $\overline{X}$ = 1.3350.

Finally, Abitbol and Olabudo discuss the forecast and forecast interval:

Observation 1	For a given confidence level, the forecast interval is the same no matter the US CPI consensus forecast.
Observation 2	A larger standard error of the estimate will result in a wider confidence interval.

32 Based on Exhibit 1, Olabudo should:

- A conclude that the inflation predictions are unbiased.
- B reject the null hypothesis that the slope coefficient equals one.
- C reject the null hypothesis that the intercept coefficient equals zero.

33 Based on Exhibit 1, Olabudo should calculate a prediction interval for the actual US CPI *closest* to:

- A 2.7506 to 2.7544.
- B 2.7521 to 2.7529.
- C 2.7981 to 2.8019.

34 Which of Olabudo's observations of forecasting is correct?

A Only Observation 1

B Only Observation 2

C Both Observation 1 and Observations 2

The following information relates to Questions 35–38

Espey Jones is examining the relation between the net profit margin (NPM) of companies, in percent, and their fixed asset turnover (FATO). He collected a sample of 35 companies for the most recent fiscal year and fit several different functional forms, settling on the following model:

$$\ln\text{NPM}_i = b_0 + b_1\text{FATO}_i.$$

The results of this estimation are provided in Exhibit 1.

Exhibit 1 Results of Regressing NPM on FATO

Source	df	Sum of Squares	Mean Square	*F*	*p*-Value
Regression	1	102.9152	102.9152	1,486.7079	0.0000
Residual	32	2.2152	0.0692		
Total	33	105.1303			

	Coefficients	Standard Error	*t*- Statistic	*p*-Value
Intercept	0.5987	0.0561	10.6749	0.0000
FATO	0.2951	0.0077	38.5579	0.0000

35 The coefficient of determination is *closest* to:

A 0.0211.

B 0.9789.

C 0.9894.

36 The standard error of the estimate is *closest* to:

A 0.2631.

B 1.7849.

C 38.5579.

37 At a 0.01 level of significance, Jones should conclude that:

A the mean net profit margin is 0.5987%.

B the variation of the fixed asset turnover explains the variation of the natural log of the net profit margin.

C a change in the fixed asset turnover from 3 to 4 times is likely to result in a change in the net profit margin of 0.5987%.

38 The predicted net profit margin for a company with a fixed asset turnover of 2 times is *closest* to:

A 1.1889%.

B 1.8043%.

C 3.2835%

SOLUTIONS

1 **A** The coefficient of determination is 0.4279:

$$\frac{\text{Explained variation}}{\text{Total variation}} = \frac{60.16}{140.58} = 0.4279.$$

B $$F = \frac{60.16 \big/ 1}{(140.58 - 60.16) \big/ (60 - 2)} = \frac{60.16}{1.3866} = 43.3882.$$

C Begin with the sum of squares error of 140.58 − 60.16 = 80.42. Then calculate the mean square error of 80.42 ÷ (60 − 2) = 1.38655. The standard error of the estimate is the square root of the mean square error: $s_e = \sqrt{1.38655} = 1.1775$.

D The sample variance of the dependent variable uses the total variation of the dependent variable and divides it by the number of observations less one:

$$\sum_{i=1}^{n} \frac{\left(Y_i - \overline{Y}\right)^2}{n - 1} = \frac{\text{Total variation}}{n - 1} = \frac{140.58}{60 - 1} = 2.3827.$$

The sample standard deviation of the dependent variable is the square root of the variance, or $\sqrt{2.3827} = 1.544$.

2 C is correct. Homoskedasticity is the situation in which the variance of the residuals is constant across the observations.

3 A is correct. SHIFT is an indicator or dummy variable because it takes on only the values 0 and 1.

4 C is correct. In a simple regression with a single indicator variable, the intercept is the mean of the dependent variable when the indicator variable takes on a value of zero, which is before the shift in policy in this case.

5 C is correct. Whereas the intercept is the average of the dependent variable when the indicator variable is zero (that is, before the shift in policy), the slope is the difference in the mean of the dependent variable from before to after the change in policy.

6 A is correct. The null hypothesis of no difference in the annual growth rate is rejected at the 0.05 level: The calculated test statistic of −8.16188 is outside the bounds of ±2.048.

7 **A** The sample variance of the dependent variable is the sum of squares total divided by its degrees of freedom ($n - 1 = 5 - 1 = 4$, as given). Thus, the sample variance of the dependent variable is 95.2 ÷ 4 = 23.8.

B The coefficient of determination = 88.0 ÷ 95.2 = 0.92437.

C The F-statistic tests whether all the slope coefficients in a linear regression are equal to zero.

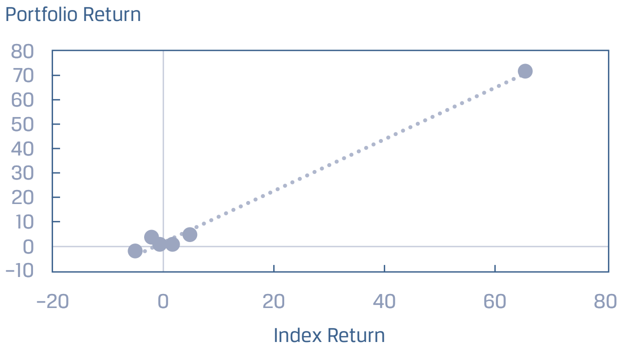
D The calculated value of the F-statistic is 36.667, as shown in the table. The corresponding p-value is less than 0.05, so you reject the null hypothesis of a slope equal to zero.

E The standard error of the estimate is the square root of the mean square error: $s_e = \sqrt{2.4} = 1.54919$.

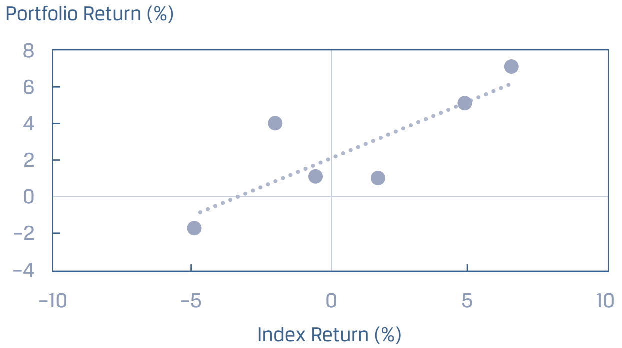
8 The Month 2 data point is an outlier, lying far away from the other data values. Because this outlier was caused by a data entry error, correcting the outlier improves the validity and reliability of the regression. In this case, revised R^2 is lower (from 0.9921 to 0.6784). The outliers created the illusion of a better fit from the higher R^2; the outliers altered the estimate of the slope. The standard error of the estimate is lower when the data error is corrected (from 2.8619 to 2.0624), as a result of the lower mean square error. However, at a 0.05 level of significance, both models fit well. The difference in the fit is illustrated in Exhibit 1.

Exhibit 1 The Fit of the Model with and without Data Errors

A. Before the Data Errors Are Corrected



B. After the Data Errors Are Corrected



9 A is correct. The coefficient of determination is the same as R^2, which is 0.7436 in the table.

10 C is correct. Because the slope is positive, the correlation between X and Y is simply the square root of the coefficient of determination: $\sqrt{0.7436} = 0.8623$.

11 C is correct. To make a prediction using the regression model, multiply the slope coefficient by the forecast of the independent variable and add the result to the intercept. Expected value of CFO to sales = 0.077 + (0.826 × 5) = 4.207.

12 C is correct. The p-value is the smallest level of significance at which the null hypotheses concerning the slope coefficient can be rejected. In this case, the p-value is less than 0.05, and thus the regression of the ratio of cash flow from operations to sales on the ratio of net income to sales is significant at the 5% level.

13 A is correct. The data are observations over time.

14 C is correct. From the regression equation, Expected return = 0.0138 + (−0.6486 × −0.01) = 0.0138 + 0.006486 = 0.0203, or 2.03%.

15 C is correct. R^2 is the coefficient of determination. In this case, it shows that 2.11% of the variability in Stellar's returns is explained by changes in CPIENG.

16 B is correct. The standard error of the estimate is the standard deviation of the regression residuals.

17 C is the correct response because it is a false statement. The slope and intercept are both statistically different from zero at the 0.05 level of significance.

18 C is correct. The slope coefficient (shown in Exhibit 2) is negative. We could also determine this by looking at the cross-product (Exhibit 1), which is negative.

19 B is correct. The sample covariance is calculated as

$$\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{n-1} = -9.2430 \div 49 = -0.1886$$

20 A is correct. In simple regression, the R^2 is the square of the pairwise correlation. Because the slope coefficient is negative, the correlation is the negative of the square root of 0.0933, or −0.3054.

21 C is correct. Conclusions cannot be drawn regarding causation; they can be drawn only about association; therefore, Interpretations 1 and 2 are incorrect.

22 C is correct. Liu explains the variation of the short interest ratio using the variation of the debt ratio.

23 A is correct. The degrees of freedom are the number of observations minus the number of parameters estimated, which equals 2 in this case (the intercept and the slope coefficient). The number of degrees of freedom is 50 − 2 = 48.

24 B is correct. The t-statistic is −2.2219, which is outside the bounds created by the critical t-values of ±2.011 for a two-tailed test with a 5% significance level. The value of 2.011 is the critical t-value for the 5% level of significance (2.5% in one tail) for 48 degrees of freedom. A is incorrect because the mean of the short interest ratio is 192.3 ÷ 50 = 3.846. C is incorrect because the debt ratio explains 9.33% of the variation of the short interest ratio.

25 A is correct. The predicted value of the short interest ratio = 5.4975 + (−4.1589 × 0.40) = 5.4975 − 1.6636 = 3.8339.

26 C is correct because $F = \dfrac{\text{Mean square regression}}{\text{Mean square error}} = \dfrac{38.4404}{7.7867} = 4.9367$.

27 C is correct. The assumptions of the linear regression model are that (1) the relationship between the dependent variable and the independent variable is linear in the parameters b_0 and b_1, (2) the residuals are independent of one another, (3) the variance of the error term is the same for all observations, and (4) the error term is normally distributed. Assumption 3 is incorrect because the dependent variable need not be normally distributed.

28 B is correct. The standard error of the estimate for a linear regression model with one independent variable is calculated as the square root of the mean square error:

$$s_e = \sqrt{\frac{0.071475}{34}} = 0.04585.$$

29 C is correct. Crude oil returns explain the Amtex share returns if the slope coefficient is statistically different from zero. The slope coefficient is 0.2354, and the calculated t-statistic is

$$t = \frac{0.2354 - 0.0000}{0.0760} = 3.0974,$$

which is outside the bounds of the critical values of ±2.728.

Therefore, Vasileva should reject the null hypothesis that crude oil returns do not explain Amtex share returns, because the slope coefficient is statistically different from zero.

A is incorrect because the calculated t-statistic for testing the slope against 0.15 is $t = \frac{0.2354 - 0.1500}{0.0760} = 1.1237$, which is less than the critical value of +2.441.

B is incorrect because the calculated t-statistic is $t = \frac{0.0095 - 0.0000}{0.0078} = 1.2179$, which is less than the critical value of +2.441.

30 B is correct. The predicted value of the dependent variable, Amtex share return, given the value of the independent variable, crude oil return, −0.01, is calculated as $\hat{Y} = \hat{b}_0 + \hat{b}_1 X_i = 0.0095 + \left[0.2354 \times (-0.01)\right] = 0.0071$.

31 C is correct. The predicted share return is 0.0095 + [0.2354 × (−0.01)] = 0.0071. The lower limit for the prediction interval is 0.0071 − (2.728 × 0.0469) = −0.1208, and the upper limit for the prediction interval is 0.0071 + (2.728 × 0.0469) = 0.1350.

A is incorrect because the bounds of the interval should be based on the standard error of the forecast and the critical t-value, not on the mean of the dependent variable.

B is incorrect because bounds of the interval are based on the product of the standard error of the forecast *and* the critical t-value, not simply the standard error of the forecast.

32 A is correct. We fail to reject the null hypothesis of a slope equal to one, and we fail to reject the null hypothesis of an intercept equal to zero. The test of the slope equal to 1.0 is $t = \frac{0.9830 - 1.000}{0.0155} = -1.09677$. The test of the intercept equal to 0.0 is $t = \frac{0.0001 - 0.0000}{.00002} = 0.5000$. Therefore, we conclude that the forecasts are unbiased.

33 A is correct. The forecast interval for inflation is calculated in three steps:

Step 1. Make the prediction given the US CPI forecast of 2.8:

$$\begin{aligned}\widehat{Y} &= b_0 + b_1 X \\ &= 0.0001 + (0.9830 \times 2.8) \\ &= 2.7525.\end{aligned}$$

Step 2. Compute the variance of the prediction error:

$$s_f^2 = s_e^2\left\{1 + (1/n) + \left[\left(X_f - \bar{X}\right)^2\right]\Big/\left[(n-1) \times s_x^2\right]\right\}.$$

$$s_f^2 = 0.0009^2\left\{1 + (1/60) + \left[(2.8 - 1.3350)^2\right]\Big/\left[(60 - 1) \times 0.7539^2\right]\right\}.$$

$$s_f^2 = 0.00000088.$$

$$s_f = 0.0009.$$

Step 3. Compute the prediction interval:

$$\hat{Y} \pm t_c \times s_f$$

2.7525 ± (2.0 × 0.0009)

Lower bound: 2.7525 – (2.0 × 0.0009) = 2.7506.

Upper bound: 2.7525 + (2.0 × 0.0009) = 2.7544.

So, given the US CPI forecast of 2.8, the 95% prediction interval is 2.7506 to 2.7544.

34 B is correct. The confidence level influences the width of the forecast interval through the critical *t*-value that is used to calculate the distance from the forecasted value: The larger the confidence level, the wider the interval. Therefore, Observation 1 is not correct.

Observation 2 is correct. The greater the standard error of the estimate, the greater the standard error of the forecast.

35 B is correct. The coefficient of determination is 102.9152 ÷ 105.1303 = 0.9789.

36 A is correct. The standard error is the square root of the mean square error, or $\sqrt{0.0692} = 0.2631$.

37 B is correct. The *p*-value corresponding to the slope is less than 0.01, so we reject the null hypothesis of a zero slope, concluding that the fixed asset turnover explains the natural log of the net profit margin.

38 C is correct. The predicted natural log of the net profit margin is 0.5987 + (2 × 0.2951) = 1.1889. The predicted net profit margin is $e^{1.1889} = 3.2835\%$.

APPENDICES

Appendix A	Cumulative Probabilities for a Standard Normal Distribution
Appendix B	Table of the Student's *t*-Distribution (One-Tailed Probabilities)
Appendix C	Values of X^2 (Degrees of Freedom, Level of Significance)
Appendix D	Table of the *F*-Distribution
Appendix E	Critical Values for the Durbin-Watson Statistic ($\alpha = .05$)

© 2021 CFA Institute. All rights reserved.

Appendix A
Cumulative Probabilities for a Standard Normal Distribution
$P(Z \le x) = N(x)$ for $x \ge 0$ or $P(Z \le z) = N(z)$ for $z \ge 0$

x or z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.60	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.70	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.80	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.90	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.00	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.10	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.20	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.30	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.40	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.50	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.60	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.70	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.90	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.00	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.10	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.20	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.30	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.40	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.50	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.60	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.70	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.80	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.90	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.00	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.10	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.20	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.30	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.40	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.50	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.60	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.70	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.80	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.90	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.00	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

For example, to find the *z*-value leaving 2.5 percent of the area/probability in the upper tail, find the element 0.9750 in the body of the table. Read 1.90 at the left end of the element's row and 0.06 at the top of the element's column, to give 1.90 + 0.06 = 1.96. *Table generated with Excel.*

Quantitative Methods for Investment Analysis, Second Edition, by Richard A. DeFusco, CFA, Dennis W. McLeavey, CFA, Jerald E. Pinto, CFA, and David E. Runkle, CFA. Copyright © 2004 by CFA Institute.

Appendix A (continued)

Cumulative Probabilities for a Standard Normal Distribution

$P(Z \leq x) = N(x)$ for $x \leq 0$ or $P(Z \leq z) = N(z)$ for $z \leq 0$

x or z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
−0.10	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.20	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.30	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.40	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.50	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.60	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.70	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.80	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.90	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−1.00	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−1.10	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.20	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.30	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.40	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.50	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.60	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.70	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.80	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.90	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−2.00	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−2.10	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.20	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.30	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.40	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.50	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.60	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.70	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.80	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.90	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−3.00	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−3.10	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.20	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.30	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.40	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.50	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
−3.60	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
−3.70	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
−3.80	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
−3.90	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
−4.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

For example, to find the *z*-value leaving 2.5 percent of the area/probability in the lower tail, find the element 0.0250 in the body of the table. Read −1.90 at the left end of the element's row and 0.06 at the top of the element's column, to give −1.90 − 0.06 = −1.96. *Table generated with Excel.*

Appendix B
Table of the Student's *t*-Distribution (One-Tailed Probabilities)

df	$p = 0.10$	$p = 0.05$	$p = 0.025$	$p = 0.01$	$p = 0.005$	df	$p = 0.10$	$p = 0.05$	$p = 0.025$	$p = 0.01$	$p = 0.005$
1	3.078	6.314	12.706	31.821	63.657	**31**	1.309	1.696	2.040	2.453	2.744
2	1.886	2.920	4.303	6.965	9.925	**32**	1.309	1.694	2.037	2.449	2.738
3	1.638	2.353	3.182	4.541	5.841	**33**	1.308	1.692	2.035	2.445	2.733
4	1.533	2.132	2.776	3.747	4.604	**34**	1.307	1.691	2.032	2.441	2.728
5	1.476	2.015	2.571	3.365	4.032	**35**	1.306	1.690	2.030	2.438	2.724
6	1.440	1.943	2.447	3.143	3.707	**36**	1.306	1.688	2.028	2.434	2.719
7	1.415	1.895	2.365	2.998	3.499	**37**	1.305	1.687	2.026	2.431	2.715
8	1.397	1.860	2.306	2.896	3.355	**38**	1.304	1.686	2.024	2.429	2.712
9	1.383	1.833	2.262	2.821	3.250	**39**	1.304	1.685	2.023	2.426	2.708
10	1.372	1.812	2.228	2.764	3.169	**40**	1.303	1.684	2.021	2.423	2.704
11	1.363	1.796	2.201	2.718	3.106	**41**	1.303	1.683	2.020	2.421	2.701
12	1.356	1.782	2.179	2.681	3.055	**42**	1.302	1.682	2.018	2.418	2.698
13	1.350	1.771	2.160	2.650	3.012	**43**	1.302	1.681	2.017	2.416	2.695
14	1.345	1.761	2.145	2.624	2.977	**44**	1.301	1.680	2.015	2.414	2.692
15	1.341	1.753	2.131	2.602	2.947	**45**	1.301	1.679	2.014	2.412	2.690
16	1.337	1.746	2.120	2.583	2.921	**46**	1.300	1.679	2.013	2.410	2.687
17	1.333	1.740	2.110	2.567	2.898	**47**	1.300	1.678	2.012	2.408	2.685
18	1.330	1.734	2.101	2.552	2.878	**48**	1.299	1.677	2.011	2.407	2.682
19	1.328	1.729	2.093	2.539	2.861	**49**	1.299	1.677	2.010	2.405	2.680
20	1.325	1.725	2.086	2.528	2.845	**50**	1.299	1.676	2.009	2.403	2.678
21	1.323	1.721	2.080	2.518	2.831	**60**	1.296	1.671	2.000	2.390	2.660
22	1.321	1.717	2.074	2.508	2.819	**70**	1.294	1.667	1.994	2.381	2.648
23	1.319	1.714	2.069	2.500	2.807	**80**	1.292	1.664	1.990	2.374	2.639
24	1.318	1.711	2.064	2.492	2.797	**90**	1.291	1.662	1.987	2.368	2.632
25	1.316	1.708	2.060	2.485	2.787	**100**	1.290	1.660	1.984	2.364	2.626
26	1.315	1.706	2.056	2.479	2.779	**110**	1.289	1.659	1.982	2.361	2.621
27	1.314	1.703	2.052	2.473	2.771	**120**	1.289	1.658	1.980	2.358	2.617
28	1.313	1.701	2.048	2.467	2.763	**200**	1.286	1.653	1.972	2.345	2.601
29	1.311	1.699	2.045	2.462	2.756	∞	1.282	1.645	1.960	2.326	2.576
30	1.310	1.697	2.042	2.457	2.750						

To find a critical t-value, enter the table with df and a specified value for α, the significance level. For example, with 5 df, $\alpha = 0.05$ and a one-tailed test, the desired probability in the tail would be $p = 0.05$ and the critical t-value would be $t(5, 0.05) = 2.015$. With $\alpha = 0.05$ and a two-tailed test, the desired probability in each tail would be $p = 0.025 = \alpha/2$, giving $t(0.025) = 2.571$. *Table generated using Excel.*

Quantitative Methods for Investment Analysis, Second Edition, by Richard A. DeFusco, CFA, Dennis W. McLeavey, CFA, Jerald E. Pinto, CFA, and David E. Runkle, CFA. Copyright © 2004 by CFA Institute.

Appendix C
Values of χ^2 (Degrees of Freedom, Level of Significance)

Degrees of Freedom	Probability in Right Tail								
	0.99	**0.975**	**0.95**	**0.9**	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**
1	0.000157	0.000982	0.003932	0.0158	2.706	3.841	5.024	6.635	7.879
2	0.020100	0.050636	0.102586	0.2107	4.605	5.991	7.378	9.210	10.597
3	0.1148	0.2158	0.3518	0.5844	6.251	7.815	9.348	11.345	12.838
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.086	16.750
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688	29.819
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	12.878	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.335
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
50	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
80	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
100	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.170

To have a probability of 0.05 in the right tail when df = 5, the tabled value is $\chi^2(5, 0.05) = 11.070$.

Quantitative Methods for Investment Analysis, Second Edition, by Richard A. DeFusco, CFA, Dennis W. McLeavey, CFA, Jerald E. Pinto, CFA, and David E. Runkle, CFA. Copyright © 2004 by CFA Institute.

Appendix D
Table of the *F*-Distribution

Panel A. Critical values for right-hand tail area equal to 0.05 **Numerator: df$_1$ and Denominator: df$_2$**

	df1:1	2	3	4	5	6	7	8	9	10	11	12	15	20	21	22	23	24	25	30	40	60	120	∞
df2: 1	161	200	216	225	230	234	237	239	241	242	243	244	246	248	248	249	249	249	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.70	8.66	8.65	8.65	8.64	8.64	8.63	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.86	5.80	5.79	5.79	5.78	5.77	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.62	4.56	4.55	4.54	4.53	4.53	4.52	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.94	3.87	3.86	3.86	3.85	3.84	3.83	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.51	3.44	3.43	3.43	3.42	3.41	3.40	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.22	3.15	3.14	3.13	3.12	3.12	3.11	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.01	2.94	2.93	2.92	2.91	2.90	2.89	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.85	2.77	2.76	2.75	2.75	2.74	2.73	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.72	2.65	2.64	2.63	2.62	2.61	2.60	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.62	2.54	2.53	2.52	2.51	2.51	2.50	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.53	2.46	2.45	2.44	2.43	2.42	2.41	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.46	2.39	2.38	2.37	2.36	2.35	2.34	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.40	2.33	2.32	2.31	2.30	2.29	2.28	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.35	2.28	2.26	2.25	2.24	2.24	2.23	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.31	2.23	2.22	2.21	2.20	2.19	2.18	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.27	2.19	2.18	2.17	2.16	2.15	2.14	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.23	2.16	2.14	2.13	2.12	2.11	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.20	2.12	2.11	2.10	2.09	2.08	2.07	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.18	2.10	2.08	2.07	2.06	2.05	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.15	2.07	2.06	2.05	2.04	2.03	2.02	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.24	2.20	2.13	2.05	2.04	2.02	2.01	2.01	2.00	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.11	2.03	2.01	2.00	1.99	1.98	1.97	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.09	2.01	2.00	1.98	1.97	1.96	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.01	1.93	1.92	1.91	1.90	1.89	1.88	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.92	1.84	1.83	1.81	1.80	1.79	1.78	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.84	1.75	1.73	1.72	1.71	1.70	1.69	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.87	1.83	1.75	1.66	1.64	1.63	1.62	1.61	1.60	1.55	1.50	1.43	1.35	1.25
Infinity	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75	1.67	1.57	1.56	1.54	1.53	1.52	1.51	1.46	1.39	1.32	1.22	1.00

Appendix D (continued)
Table of the *F*-Distribution

Panel B. Critical values for right-hand tail area equal to 0.025 **Numerator: df_1 and Denominator: df_2**

	df1: 1	2	3	4	5	6	7	8	9	10	11	12	15	20	21	22	23	24	25	30	40	60	120	∞
df2: 1	648	799	864	900	922	937	948	957	963	969	973	977	985	993	994	995	996	997	998	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.41	39.43	39.45	39.45	39.45	39.45	39.46	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.37	14.34	14.25	14.17	14.16	14.14	14.13	14.12	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.79	8.75	8.66	8.56	8.55	8.53	8.52	8.51	8.50	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.57	6.52	6.43	6.33	6.31	6.30	6.29	6.28	6.27	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.41	5.37	5.27	5.17	5.15	5.14	5.13	5.12	5.11	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.71	4.67	4.57	4.47	4.45	4.44	4.43	4.41	4.40	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.24	4.20	4.10	4.00	3.98	3.97	3.96	3.95	3.94	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.91	3.87	3.77	3.67	3.65	3.64	3.63	3.61	3.60	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.66	3.62	3.52	3.42	3.40	3.39	3.38	3.37	3.35	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.47	3.43	3.33	3.23	3.21	3.20	3.18	3.17	3.16	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.32	3.28	3.18	3.07	3.06	3.04	3.03	3.02	3.01	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.20	3.15	3.05	2.95	2.93	2.92	2.91	2.89	2.88	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.09	3.05	2.95	2.84	2.83	2.81	2.80	2.79	2.78	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	3.01	2.96	2.86	2.76	2.74	2.73	2.71	2.70	2.69	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.93	2.89	2.79	2.68	2.67	2.65	2.64	2.63	2.61	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.87	2.82	2.72	2.62	2.60	2.59	2.57	2.56	2.55	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.81	2.77	2.67	2.56	2.54	2.53	2.52	2.50	2.49	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.76	2.72	2.62	2.51	2.49	2.48	2.46	2.45	2.44	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.72	2.68	2.57	2.46	2.45	2.43	2.42	2.41	2.40	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.68	2.64	2.53	2.42	2.41	2.39	2.38	2.37	2.36	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.65	2.60	2.50	2.39	2.37	2.36	2.34	2.33	2.32	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.62	2.57	2.47	2.36	2.34	2.33	2.31	2.30	2.29	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.59	2.54	2.44	2.33	2.31	2.30	2.28	2.27	2.26	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.56	2.51	2.41	2.30	2.28	2.27	2.26	2.24	2.23	2.18	2.12	2.05	1.98	1.91
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.46	2.41	2.31	2.20	2.18	2.16	2.15	2.14	2.12	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.33	2.29	2.18	2.07	2.05	2.03	2.02	2.01	1.99	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.22	2.17	2.06	1.94	1.93	1.91	1.90	1.88	1.87	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.10	2.05	1.94	1.82	1.81	1.79	1.77	1.76	1.75	1.69	1.61	1.53	1.43	1.31
Infinity	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.99	1.94	1.83	1.71	1.69	1.67	1.66	1.64	1.63	1.57	1.48	1.39	1.27	1.00

Appendix D (continued)
Table of the *F*-Distribution

Panel C. Critical values for right-hand tail area equal to 0.01 **Numerator: df_1 and Denominator: df_2**

	df1: 1	2	3	4	5	6	7	8	9	10	11	12	15	20	21	22	23	24	25	30	40	60	120	∞
df2: 1	4052	5000	5403	5625	5764	5859	5928	5982	6023	6056	6083	6106	6157	6209	6216	6223	6229	6235	6240	6261	6287	6313	6339	6366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	27.1	26.9	26.7	26.7	26.6	26.6	26.6	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.5	14.4	14.2	14.0	14.0	14.0	13.9	13.9	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	10.0	9.89	9.72	9.55	9.53	9.51	9.49	9.47	9.45	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.56	7.40	7.37	7.35	7.33	7.31	7.30	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.54	6.47	6.31	6.16	6.13	6.11	6.09	6.07	6.06	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.73	5.67	5.52	5.36	5.34	5.32	5.30	5.28	5.26	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.18	5.11	4.96	4.81	4.79	4.77	4.75	4.73	4.71	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71	4.56	4.41	4.38	4.36	4.34	4.33	4.31	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40	4.25	4.10	4.08	4.06	4.04	4.02	4.01	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16	4.01	3.86	3.84	3.82	3.80	3.78	3.76	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.82	3.66	3.64	3.62	3.60	3.59	3.57	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.86	3.80	3.66	3.51	3.48	3.46	3.44	3.43	3.41	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	3.52	3.37	3.35	3.33	3.31	3.29	3.28	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.62	3.55	3.41	3.26	3.24	3.22	3.20	3.18	3.16	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.46	3.31	3.16	3.14	3.12	3.10	3.08	3.07	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.43	3.37	3.23	3.08	3.05	3.03	3.02	3.00	2.98	2.92	2.84	2.75	2.66	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30	3.15	3.00	2.98	2.96	2.94	2.92	2.91	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23	3.09	2.94	2.92	2.90	2.88	2.86	2.84	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.24	3.17	3.03	2.88	2.86	2.84	2.82	2.80	2.79	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12	2.98	2.83	2.81	2.78	2.77	2.75	2.73	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.14	3.07	2.93	2.78	2.76	2.74	2.72	2.70	2.69	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03	2.89	2.74	2.72	2.70	2.68	2.66	2.64	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	3.06	2.99	2.85	2.70	2.68	2.66	2.64	2.62	2.60	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84	2.70	2.55	2.53	2.51	2.49	2.47	2.45	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66	2.52	2.37	2.35	2.33	2.31	2.29	2.27	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50	2.35	2.20	2.17	2.15	2.13	2.12	2.10	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.40	2.34	2.19	2.03	2.01	1.99	1.97	1.95	1.93	1.86	1.76	1.66	1.53	1.38
Infinity	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18	2.04	1.88	1.85	1.83	1.81	1.79	1.77	1.70	1.59	1.47	1.32	1.00

Appendix D (continued)
Table of the *F*-Distribution

Panel D. Critical values for right-hand tail area equal to 0.005 **Numerator: df_1 and Denominator: df_2**

	df1: 1	2	3	4	5	6	7	8	9	10	11	12	15	20	21	22	23	24	25	30	40	60	120	∞
df2: 1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24222	24334	24426	24630	24836	24863	24892	24915	24940	24959	25044	25146	25253	25359	25464
2	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	200
3	55.55	49.80	47.47	46.20	45.39	44.84	44.43	44.13	43.88	43.68	43.52	43.39	43.08	42.78	42.73	42.69	42.66	42.62	42.59	42.47	42.31	42.15	41.99	41.83
4	31.33	26.28	24.26	23.15	22.46	21.98	21.62	21.35	21.14	20.97	20.82	20.70	20.44	20.17	20.13	20.09	20.06	20.03	20.00	19.89	19.75	19.61	19.47	19.32
5	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.49	13.38	13.15	12.90	12.87	12.84	12.81	12.78	12.76	12.66	12.53	12.40	12.27	12.14
6	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.13	10.03	9.81	9.59	9.56	9.53	9.50	9.47	9.45	9.36	9.24	9.12	9.00	8.88
7	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.27	8.18	7.97	7.75	7.72	7.69	7.67	7.64	7.62	7.53	7.42	7.31	7.19	7.08
8	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.10	7.01	6.81	6.61	6.58	6.55	6.53	6.50	6.48	6.40	6.29	6.18	6.06	5.95
9	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.31	6.23	6.03	5.83	5.80	5.78	5.75	5.73	5.71	5.62	5.52	5.41	5.30	5.19
10	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.75	5.66	5.47	5.27	5.25	5.22	5.20	5.17	5.15	5.07	4.97	4.86	4.75	4.64
11	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54	5.42	5.32	5.24	5.05	4.86	4.83	4.80	4.78	4.76	4.74	4.65	4.55	4.45	4.34	4.23
12	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.99	4.91	4.72	4.53	4.50	4.48	4.45	4.43	4.41	4.33	4.23	4.12	4.01	3.90
13	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94	4.82	4.72	4.64	4.46	4.27	4.24	4.22	4.19	4.17	4.15	4.07	3.97	3.87	3.76	3.65
14	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72	4.60	4.51	4.43	4.25	4.06	4.03	4.01	3.98	3.96	3.94	3.86	3.76	3.66	3.55	3.44
15	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.33	4.25	4.07	3.88	3.86	3.83	3.81	3.79	3.77	3.69	3.59	3.48	3.37	3.26
16	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38	4.27	4.18	4.10	3.92	3.73	3.71	3.68	3.66	3.64	3.62	3.54	3.44	3.33	3.22	3.11
17	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25	4.14	4.05	3.97	3.79	3.61	3.58	3.56	3.53	3.51	3.49	3.41	3.31	3.21	3.10	2.98
18	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14	4.03	3.94	3.86	3.68	3.50	3.47	3.45	3.42	3.40	3.38	3.30	3.20	3.10	2.99	2.87
19	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04	3.93	3.84	3.76	3.59	3.40	3.37	3.35	3.33	3.31	3.29	3.21	3.11	3.00	2.89	2.78
20	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.76	3.68	3.50	3.32	3.29	3.27	3.24	3.22	3.20	3.12	3.02	2.92	2.81	2.69
21	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88	3.77	3.68	3.60	3.43	3.24	3.22	3.19	3.17	3.15	3.13	3.05	2.95	2.84	2.73	2.61
22	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81	3.70	3.61	3.54	3.36	3.18	3.15	3.12	3.10	3.08	3.06	2.98	2.88	2.77	2.66	2.55
23	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75	3.64	3.55	3.47	3.30	3.12	3.09	3.06	3.04	3.02	3.00	2.92	2.82	2.71	2.60	2.48
24	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69	3.59	3.50	3.42	3.25	3.06	3.04	3.01	2.99	2.97	2.95	2.87	2.77	2.66	2.55	2.43
25	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64	3.54	3.45	3.37	3.20	3.01	2.99	2.96	2.94	2.92	2.90	2.82	2.72	2.61	2.50	2.38
30	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.25	3.18	3.01	2.82	2.80	2.77	2.75	2.73	2.71	2.63	2.52	2.42	2.30	2.18
40	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22	3.12	3.03	2.95	2.78	2.60	2.57	2.55	2.52	2.50	2.48	2.40	2.30	2.18	2.06	1.93
60	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.82	2.74	2.57	2.39	2.36	2.33	2.31	2.29	2.27	2.19	2.08	1.96	1.83	1.69
120	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.62	2.54	2.37	2.19	2.16	2.13	2.11	2.09	2.07	1.98	1.87	1.75	1.61	1.43
Infinity	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.43	2.36	2.19	2.00	1.97	1.95	1.92	1.90	1.88	1.79	1.67	1.53	1.36	1.00

With 1 degree of freedom (df) in the numerator and 3 df in the denominator, the critical *F*-value is 10.1 for a right-hand tail area equal to 0.05.

Quantitative Methods for Investment Analysis, Second Edition, by Richard A. DeFusco, CFA, Dennis W. McLeavey, CFA, Jerald E. Pinto, CFA, and David E. Runkle, CFA. Copyright © 2004 by CFA Institute.

Appendix E
Critical Values for the Durbin-Watson Statistic ($\alpha = .05$)

	K = 1		*K* = 2		*K* = 3		*K* = 4		*K* = 5	
n	d_l	d_u	d_l	d_u	d_l	d_u	d_l	d_u	d_l	d_u
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

Note: *K* = the number of slope parameters in the model.

Source: From J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression, II." *Biometrika* 38 (1951): 159–178.

Glossary

A priori probability A probability based on logical analysis rather than on observation or personal judgment.

Abandonment option The ability to terminate an investment at some future time if the financial results are disappointing.

Abnormal return The amount by which a security's actual return differs from its expected return, given the security's risk and the market's return.

Absolute advantage A country's ability to produce a good or service at a lower absolute cost than its trading partner.

Absolute dispersion The amount of variability present without comparison to any reference point or benchmark.

Absolute frequency The actual number of observations counted for each unique value of the variable (also called raw frequency).

Accelerated book build An offering of securities by an investment bank acting as principal that is accomplished in only one or two days.

Accelerated methods Depreciation methods that allocate a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Accounting costs Monetary value of economic resources used in performing an activity. These can be explicit, out-of-pocket, current payments, or an allocation of historical payments (depreciation) for resources. They do not include implicit opportunity costs.

Accounting profit Income as reported on the income statement, in accordance with prevailing accounting standards, before the provisions for income tax expense. Also called *income before taxes* or *pretax income*.

Accounts payable Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Accrued expenses Liabilities related to expenses that have been incurred but not yet paid as of the end of an accounting period—an example of an accrued expense is rent that has been incurred but not yet paid, resulting in a liability "rent payable." Also called *accrued liabilities*.

Accrued interest Interest earned but not yet paid.

Acquisition method A method of accounting for a business combination where the acquirer is required to measure each identifiable asset and liability at fair value. This method was the result of a joint project of the IASB and FASB aiming at convergence in standards for the accounting of business combinations.

Action lag Delay from policy decisions to implementation.

Active investment An approach to investing in which the investor seeks to outperform a given benchmark.

Active return The return on a portfolio minus the return on the portfolio's benchmark.

Activity ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory. Also called *asset utilization ratios* or *operating efficiency ratios*.

Add-on rates Bank certificates of deposit, repos, and indexes such as Libor and Euribor are quoted on an add-on rate basis (bond equivalent yield basis).

Addition rule for probabilities A principle stating that the probability that *A* or *B* occurs (both occur) equals the probability that *A* occurs, plus the probability that *B* occurs, minus the probability that both *A* and *B* occur.

Agency bond See *quasi-government bond*.

Agency costs Costs associated with the conflict of interest present between principals and agents when a company is managed by non-owners. Agency costs result from the inherent conflicts of interest between managers, bondholders, and equity owners.

Agency costs of debt Costs arising from conflicts of interest between managers and debtholders.

Agency costs of equity The smaller the stake managers have in the company, the less their share in bearing the cost of excessive perquisite consumption—consequently, the less their desire to give their best efforts in running the company.

Agency RMBS In the United States, securities backed by residential mortgage loans and guaranteed by a federal agency or guaranteed by either of the two GSEs (Fannie Mae and Freddie Mac).

Aggregate demand The quantity of goods and services that households, businesses, government, and non-domestic customers want to buy at any given level of prices.

Aggregate demand curve Inverse relationship between the price level and real output.

Aggregate income The value of all the payments earned by the suppliers of factors used in the production of goods and services.

Aggregate output The value of all the goods and services produced during a specified period.

Aggregate supply The quantity of goods and services producers are willing to supply at any given level of price.

Aggregate supply curve The level of domestic output that companies will produce at each price level.

All-or-nothing (AON) orders An order that includes the instruction to trade only if the trade fills the entire quantity (size) specified.

Allocationally efficient A characteristic of a market, a financial system, or an economy that promotes the allocation of resources to their highest value uses.

Alternative data Non-traditional data types generated by the use of electronic devices, social media, satellite and sensor networks, and company exhaust.

Alternative hypothesis The hypothesis that is accepted if the null hypothesis is rejected.

Alternative investment markets Market for investments other than traditional securities investments (i.e., traditional common and preferred shares and traditional fixed income instruments). The term usually encompasses direct and indirect investment in real estate (including timberland and farmland) and commodities (including precious metals); hedge funds, private equity, and other investments requiring specialized due diligence.

Alternative trading systems Trading venues that function like exchanges but that do not exercise regulatory authority over their subscribers except with respect to the conduct

of the subscribers' trading in their trading systems. Also called *electronic communications networks* or *multilateral trading facilities*.

American depository receipt A US dollar-denominated security that trades like a common share on US exchanges.

American depository share The underlying shares on which American depository receipts are based. They trade in the issuing company's domestic market.

American-style Type of option contract that can be exercised at any time up to the option's expiration date.

Amortisation The process of allocating the cost of intangible long-term assets having a finite useful life to accounting periods; the allocation of the amount of a bond premium or discount to the periods remaining until bond maturity.

Amortised cost The historical cost (initially recognised cost) of an asset, adjusted for amortisation and impairment.

Amortizing bond Bond with a payment schedule that calls for periodic payments of interest and repayments of principal.

Amortizing loans Loans with a payment schedule that calls for periodic payments of interest and repayments of principal.

Analysis of variance (ANOVA) The analysis that breaks the total variability of a dataset (such as observations on the dependent variable in a regression) into components representing different sources of variation. With reference to regression, ANOVA provides the inputs for an *F*-test of the significance of the regression as a whole, as well as the inputs for the coefficient of determination and the standard error of the estimate.

Analytical duration The use of mathematical formulas to estimate the impact of benchmark yield-to-maturity changes on bond prices.

Anchoring and adjustment bias An information-processing bias in which the use of a psychological heuristic influences the way people estimate probabilities.

Annual percentage rate The cost of borrowing expressed as a yearly rate.

Annuity A finite set of level sequential cash flows.

Annuity due An annuity having a first cash flow that is paid immediately.

Anomalies Apparent deviations from market efficiency.

Antidilutive With reference to a transaction or a security, one that would increase earnings per share (EPS) or result in EPS higher than the company's basic EPS—antidilutive securities are not included in the calculation of diluted EPS.

Arbitrage 1) The simultaneous purchase of an undervalued asset or portfolio and sale of an overvalued but equivalent asset or portfolio, in order to obtain a riskless profit on the price differential. Taking advantage of a market inefficiency in a risk-free manner. 2) The condition in a financial market in which equivalent assets or combinations of assets sell for two different prices, creating an opportunity to profit at no risk with no commitment of money. In a well-functioning financial market, few arbitrage opportunities are possible. 3) A risk-free operation that earns an expected positive net profit but requires no net investment of money.

Arbitrage-free pricing The overall process of pricing derivatives by arbitrage and risk neutrality. Also called the *principle of no arbitrage*.

Arbitrageurs Traders who engage in arbitrage. See *arbitrage*.

Arithmetic mean The sum of the observations divided by the number of observations.

Artificial intelligence Computer systems that exhibit cognitive and decision-making ability comparable (or superior) to that of humans.

Ask The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size). See *offer*.

Ask size The maximum quantity of an asset that pertains to a specific ask price from a trader. For example, if the ask for a share issue is $30 for a size of 1,000 shares, the trader is offering to sell at $30 up to 1,000 shares.

Asset allocation The process of determining how investment funds should be distributed among asset classes.

Asset-backed securities A type of bond issued by a legal entity called a *special purpose entity* (SPE) on a collection of assets that the SPE owns. Also, securities backed by receivables and loans other than mortgages.

Asset-based valuation models Valuation based on estimates of the market value of a company's assets.

Asset beta The unlevered beta; which reflects the business risk of the assets; the assets' systematic risk.

Asset class A group of assets that have similar characteristics, attributes, and risk–return relationships.

Asset swap Converts the periodic fixed coupon of a specific bond to an MRR plus or minus a spread.

Asset utilization ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Assets Resources controlled by an enterprise as a result of past events and from which future economic benefits to the enterprise are expected to flow.

Assignment of accounts receivable The use of accounts receivable as collateral for a loan.

Asymmetric information The differential of information between corporate insiders and outsiders regarding the company's performance and prospects. Managers typically have more information about the company's performance and prospects than owners and creditors.

At the money An option in which the underlying's price equals the exercise price.

Auction A type of bond issuing mechanism often used for sovereign bonds that involves bidding.

Autarkic price The price of a good or service in an autarkic economy.

Autarky A state in which a country does not trade with other countries.

Automatic stabilizer A countercyclical factor that automatically comes into play as an economy slows and unemployment rises.

Availability bias An information-processing bias in which people take a heuristic approach to estimating the probability of an outcome based on how easily the outcome comes to mind.

Available-for-sale Under US GAAP, debt securities not classified as either held-to-maturity or held-for-trading securities. The investor is willing to sell but not actively planning to sell. In general, available-for-sale debt securities are reported at fair value on the balance sheet, with unrealized gains included as a component of other comprehensive income.

Average fixed cost Total fixed cost divided by quantity produced.

Average life See *weighted average life*.

Average product Measures the productivity of inputs on average and is calculated by dividing total product by the total number of units for a given input that is used to generate that output.

Average revenue Total revenue divided by quantity sold.

Average total cost Total cost divided by quantity produced.

Average variable cost Total variable cost divided by quantity produced.

Back-testing With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backfill bias A bias introduced if a fund backfills its past performance data to the database when it first starts reporting actual performance.

Backtesting With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backup line of credit A type of credit enhancement provided by a bank to an issuer of commercial paper to ensure that the issuer will have access to sufficient liquidity to repay maturing commercial paper if issuing new paper is not a viable option.

Backwardation A condition in the futures markets in which the spot price exceeds the futures price, the forward curve is downward sloping, and the convenience yield is high.

Balance of payments A double-entry bookkeeping system that summarizes a country's economic transactions with the rest of the world for a particular period of time, typically a calendar quarter or year.

Balance of trade deficit When the domestic economy is spending more on non-domestic goods and services than non-domestic economies are spending on domestic goods and services.

Balance sheet The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet). Also called *statement of financial position* or *statement of financial condition.*

Balance sheet ratios Financial ratios involving balance sheet items only.

Balanced With respect to a government budget, one in which spending and revenues (taxes) are equal.

Balloon payment Large payment required at maturity to retire a bond's outstanding principal amount.

Bar chart A chart for plotting the frequency distribution of categorical data, where each bar represents a distinct category and each bar's height is proportional to the frequency of the corresponding category. In technical analysis, a bar chart that plots four bits of data for each time interval—the high, low, opening, and closing prices. A vertical line connects the high and low prices. A cross-hatch left indicates the opening price and a cross-hatch right indicates the closing price.

Barter economy An economy where economic agents as house-holds, corporations, and governments "pay" for goods and services with another good or service.

Base-rate neglect A type of representativeness bias in which the base rate or probability of the categorization is not adequately considered.

Base rates The reference rate on which a bank bases lending rates to all other customers.

Basic EPS Net earnings available to common shareholders (i.e., net income minus preferred dividends) divided by the weighted average number of common shares outstanding.

Basis point Used in stating yield spreads, one basis point equals one-hundredth of a percentage point, or 0.01%.

Basket of listed depository receipts An exchange-traded fund (ETF) that represents a portfolio of depository receipts.

Bayes' formula The rule for updating the probability of an event of interest—given a set of prior probabilities for the event, information, and information given the event—if you receive new information.

Bearer bonds Bonds for which ownership is not recorded; only the clearing system knows who the bond owner is.

Bearish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from above; this movement is considered bearish. A **death cross** is a bearish crossover based on 50-day and 200-day moving averages.

Behavioral finance A field of finance that examines the psychological variables that affect and often distort the investment decision making of investors, analysts, and portfolio managers.

Behind the market Said of prices specified in orders that are worse than the best current price; e.g., for a limit buy order, a limit price below the best bid.

Benchmark A comparison portfolio; a point of reference or comparison.

Benchmark issue The latest sovereign bond issue for a given maturity. It serves as a benchmark against which to compare bonds that have the same features but that are issued by another type of issuer.

Benchmark rate Typically the yield-to-maturity on a government bond having the same or close to the same time-to-maturity.

Benchmark spread The yield spread over a specific benchmark, usually measured in basis points.

Bernoulli random variable A random variable having the outcomes 0 and 1.

Bernoulli trial An experiment that can produce one of two outcomes.

Best bid The highest bid in the market.

Best effort offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-efforts offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-in-class An ESG implementation approach that seeks to identify the most favorable companies in an industry based on ESG considerations.

Best offer The lowest offer (ask price) in the market.

Beta A measure of the sensitivity of a given investment or portfolio to movements in the overall market.

Bid The price at which a dealer or trader is willing to buy an asset, typically qualified by a maximum quantity.

Bid–ask spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid–offer spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid size The maximum quantity of an asset that pertains to a specific bid price from a trader.

Big Data The vast amount of data being generated by industry, governments, individuals, and electronic devices that arises from both traditional and non-traditional data sources.

Bilateral loan A loan from a single lender to a single borrower.

Bimodal A distribution that has two most frequently occurring values.

Binomial model A model for pricing options in which the underlying price can move to only one of two possible new prices.

Binomial random variable The number of successes in n Bernoulli trials for which the probability of success is constant for all trials and the trials are independent.

Binomial tree The graphical representation of a model of asset price dynamics in which, at each period, the asset moves up with probability p or down with probability $(1 - p)$.

Bitcoin A cryptocurrency using blockchain technology that was created in 2009.

Bivariate correlation See *Pearson correlation*.

Block brokers A broker (agent) that provides brokerage services for large-size trades.

Blockchain A type of digital ledger in which information is recorded sequentially and then linked together and secured using cryptographic methods.

Blue chip Widely held large market capitalization companies that are considered financially sound and are leaders in their respective industry or local stock market.

Bollinger Bands A price-based technical analysis indicator consisting of a line representing the moving average, a higher line representing the moving average plus a set number of standard deviations from the average (for the same number of periods as was used to calculate the moving average), and a lower line representing the moving average minus the same number of standard deviations.

Bond Contractual agreement between the issuer and the bondholders.

Bond equivalent yield A calculation of yield that is annualized using the ratio of 365 to the number of days to maturity. Bond equivalent yield allows for the restatement and comparison of securities with different compounding periods.

Bond indenture The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *trust deed*.

Bond market vigilantes Bond market participants who might reduce their demand for long-term bonds, thus pushing up their yields.

Bond yield plus risk premium approach An estimate of the cost of common equity that is produced by summing the before-tax cost of debt and a risk premium that captures the additional yield on a company's stock relative to its bonds. The additional yield is often estimated using historical spreads between bond yields and stock yields.

Bonus issue of shares A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Book building Investment bankers' process of compiling a "book" or list of indications of interest to buy part of an offering.

Book value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. Also called *carrying value*.

Boom An expansionary phase characterized by economic growth "testing the limits" of the economy.

Bootstrap A resampling method that repeatedly draws samples with replacement of the selected elements from the original observed sample. Bootstrap is usually conducted by using computer simulation and is often used to find standard error or construct confidence intervals of population parameters.

Bottom-up analysis An investment selection approach that focuses on company-specific circumstances rather than emphasizing economic cycles or industry analysis.

Box and whisker plot A graphic for visualizing the dispersion of data across quartiles. It consists of a "box" with "whiskers" connected to the box.

Breakdown A breakdown occurs when the price of an asset moves below a support level.

Breakeven point The number of units produced and sold at which the company's net income is zero (Revenues = Total cost); in the case of perfect competition, the quantity at which price, average revenue, and marginal revenue equal average total cost.

Breakout A breakout occurs when the price of an asset moves above a resistance level.

Bridge financing Interim financing that provides funds until permanent financing can be arranged.

Broad money Encompasses narrow money plus the entire range of liquid assets that can be used to make purchases.

Broker 1) An agent who executes orders to buy or sell securities on behalf of a client in exchange for a commission. 2) See *futures commission merchants*.

Broker–dealer A financial intermediary (often a company) that may function as a principal (dealer) or as an agent (broker) depending on the type of trade.

Brokered market A market in which brokers arrange trades among their clients.

Brownfield investment Investing in *existing* infrastructure assets.

Bubble line chart A line chart that uses varying-sized bubbles to represent a third dimension of the data. The bubbles are sometimes color-coded to present additional information.

Budget surplus/deficit The difference between government revenue and expenditure for a stated fixed period of time.

Bullet bond Bond in which the principal repayment is made entirely at maturity.

Bullish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from below; this movement is considered bullish. A **golden cross** is a bullish crossover based on 50-day and 200-day moving averages.

Business cycles Are recurrent expansions and contractions in economic activity affecting broad segments of the economy.

Business risk The risk associated with operating earnings. Operating earnings are uncertain because total revenues and many of the expenditures contributed to produce those revenues are uncertain.

Buy-side firm An investment management company or other investor that uses the services of brokers or dealers (i.e., the client of the sell side firms).

Buyback A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Call An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call market A market in which trades occur only at a particular time and place (i.e., when the market is called).

Call money rate The interest rate that buyers pay for their margin loan.

Call option An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call protection The time during which the issuer of the bond is not allowed to exercise the call option.

Callable bond A bond containing an embedded call option that gives the issuer the right to buy the bond back from the investor at specified prices on pre-determined dates.

Calmar ratio A comparison of the average annual compounded return to the maximum drawdown risk.

Candlestick chart A price chart with four bits of data for each time interval. A candle indicates the opening and closing price for the interval. The body of the candle is shaded if the opening price was higher than the closing price, and the body is white (or clear) if the opening price was lower than the closing price. Vertical lines known as wicks or shadows extend from the top and bottom of the candle to indicate, respectively, the high and low prices for the interval.

Cannibalization Cannibalization occurs when an investment takes customers and sales away from another part of the company.

Cap rate A metric by which real estate managers are often judged; the annual rent actually earned (net of any vacancies) divided by the price originally paid for the property.

Capacity The ability of the borrower to make its debt payments on time.

Capital account A component of the balance of payments account that measures transfers of capital.

Capital allocation line (CAL) A graph line that describes the combinations of expected return and standard deviation of return available to an investor from combining the optimal portfolio of risky assets with the risk-free asset.

Capital allocation process The process that companies use for decision making on capital investments—those projects with a life of one year or longer.

Capital asset pricing model (CAPM) An equation describing the expected return on any asset (or portfolio) as a linear function of its beta relative to the market portfolio.

Capital consumption allowance A measure of the wear and tear (depreciation) of the capital stock that occurs in the production of goods and services.

Capital deepening investment Increases the stock of capital relative to labor.

Capital expenditure Expenditure on physical capital (fixed assets).

Capital light Capital light businesses require little incremental investment in fixed assets or working capital to enable revenue growth.

Capital loss ratio The percentage of capital in deals that have been realized below cost, net of any recovered proceeds, divided by total invested capital.

Capital market expectations An investor's expectations concerning the risk and return prospects of asset classes.

Capital market line (CML) The line with an intercept point equal to the risk-free rate that is tangent to the efficient frontier of risky assets; represents the efficient frontier when a risk-free asset is available for investment.

Capital market securities Securities with maturities at issuance longer than one year.

Capital markets Financial markets that trade securities of longer duration, such as bonds and equities.

Capital rationing A capital rationing environment assumes that the company has a fixed amount of funds to invest.

Capital restrictions Controls placed on foreigners' ability to own domestic assets and/or domestic residents' ability to own foreign assets.

Capital stock The accumulated amount of buildings, machinery, and equipment used to produce goods and services.

Capital structure The mix of debt and equity that a company uses to finance its business; a company's specific mixture of long-term financing.

Carry The net of the costs and benefits of holding, storing, or "carrying" an asset.

Carrying amount The amount at which an asset or liability is valued according to accounting principles.

Carrying value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. For a bond, the purchase price plus (or minus) the amortized amount of the discount (or premium).

Cartel Participants in collusive agreements that are made openly and formally.

Cash collateral account Form of external credit enhancement whereby the issuer immediately borrows the credit-enhancement amount and then invests that amount, usually in highly rated short-term commercial paper.

Cash conversion cycle A financial metric that measures the length of time required for a company to convert cash invested in its operations to cash received as a result of its operations; equal to days of inventory on hand + days of sales outstanding – number of days of payables. Also called *net operating cycle*.

Cash flow additivity principle The principle that dollar amounts indexed at the same point in time are additive.

Cash flow from operating activities The net amount of cash provided from operating activities.

Cash flow from operations The net amount of cash provided from operating activities.

Cash flow yield The internal rate of return on a series of cash flows.

Cash market securities Money market securities settled on a "same day" or "cash settlement" basis.

Cash markets See *spot markets*.

Cash prices See *spot prices*.

Cash-settled forwards See *non-deliverable forwards*.

Catch-up clause A clause in an agreement that favors the GP. For a GP who earns a 20% performance fee, a catch-up clause allows the GP to receive 100% of the distributions above the hurdle rate *until* she receives 20% of the profits generated, and then every excess dollar is split 80/20 between the LPs and GP.

Categorical data Values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize (also called **qualitative data**).

CBOE Volatility Index (VIX) A measure of near-term market volatility as conveyed by S&P 500 stock index option prices.

Central bank funds market The market in which deposit-taking banks that have an excess reserve with their national central bank can lend money to banks that need funds for maturities ranging from overnight to one year. Called the Federal or Fed funds market in the United States.

Central bank funds rate Interest rate at which central bank funds are bought (borrowed) and sold (lent) for maturities ranging from overnight to one year. Called Federal or Fed funds rate in the United States.

Central banks The dominant bank in a country, usually with official or semi-official governmental status.

Central limit theorem The theorem that states the sum (and the mean) of a set of independent, identically distributed random variables with finite variances is normally distributed, whatever distribution the random variables follow.

Certificate of deposit An instrument that represents a specified amount of funds on deposit with a bank for a specified maturity and interest rate. CDs are issued in various denominations and can be negotiable or non-negotiable.

Change in polarity principle A tenet of technical analysis that states that once a support level is breached, it becomes a resistance level. The same holds true for resistance levels: Once breached, they become support levels.

Change of control put A covenant giving bondholders the right to require the issuer to buy back their debt, often at par or at some small premium to par value, in the event that the borrower is acquired.

Character The quality of a debt issuer's management.

Chartist An individual who uses charts or graphs of a security's historical prices or levels to forecast its future trends.

Chi-square test of independence A statistical test for detecting a potential association between categorical variables.

Classical cycle Refers to fluctuations in the level of economic activity when measured by GDP in volume terms.

Classified balance sheet A balance sheet organized so as to group together the various assets and liabilities into subcategories (e.g., current and noncurrent).

Clawback A requirement that the general partner return any funds distributed as incentive fees until the limited partners have received back their initial investment and a percentage of the total profit.

Clearing The process by which the exchange verifies the execution of a transaction and records the participants' identities.

Clearing instructions Instructions that indicate how to arrange the final settlement ("clearing") of a trade.

Clearinghouse An entity associated with a futures market that acts as middleman between the contracting parties and guarantees to each party the performance of the other.

Closed economy An economy that does not trade with other countries; an *autarkic economy*.

Closed-end fund A mutual fund in which no new investment money is accepted. New investors invest by buying existing shares, and investors in the fund liquidate by selling their shares to other investors.

Cluster sampling A procedure that divides a population into subpopulation groups (clusters) representative of the population and then randomly draws certain clusters to form a sample.

Clustered bar chart See *grouped bar chart*.

Co-investing In co-investing, the investor invests in assets *indirectly* through the fund but also possesses rights (known as co-investment rights) to invest *directly* in the same assets. Through co-investing, an investor is able to make an investment *alongside* a fund when the fund identifies deals.

Code of ethics An established guide that communicates an organization's values and overall expectations regarding member behavior. A code of ethics serves as a general guide for how community members should act.

Coefficient of determination The percentage of the variation of the dependent variable that is explained by the independent variable. Also referred to as the "R-squared" or "R^2."

Coefficient of variation The ratio of a set of observations' standard deviation to the observations' mean value.

Cognitive cost The effort involved in processing new information and updating beliefs.

Cognitive dissonance The mental discomfort that occurs when new information conflicts with previously held beliefs or cognitions.

Cognitive errors Behavioral biases resulting from faulty reasoning; cognitive errors stem from basic statistical, information-processing, or memory errors.

Coincident economic indicators Turning points that are usually close to those of the overall economy; they are believed to have value for identifying the economy's present state.

Collateral Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Collateral manager Buys and sells debt obligations for and from the CDO's portfolio of assets (i.e., the collateral) to generate sufficient cash flows to meet the obligations to the CDO bondholders.

Collateral trust bonds Bonds secured by securities, such as common shares, other bonds, or other financial assets.

Collateralized bond obligations A structured asset-backed security that is collateralized by a pool of bonds.

Collateralized debt obligation Generic term used to describe a security backed by a diversified pool of one or more debt obligations.

Collateralized loan obligations A structured asset-backed security that is collateralized by a pool of loans.

Collateralized mortgage obligation A security created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collateralized mortgage obligations Securities created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collaterals Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Combination A listing in which the order of the listed items does not matter.

Combination formula (binomial formula) The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is ${}_nC_r = \binom{n}{r} = \dfrac{n!}{(n-r)!r!}$.

Commercial paper A short-term, negotiable, unsecured promissory note that represents a debt obligation of the issuer.

Committed (Regular) lines of credit A bank commitment to extend credit up to a pre-specified amount; the commitment is considered a short-term liability and is usually in effect for 364 days (one day short of a full year).

Committed capital The amount that the limited partners have agreed to provide to the private equity fund.

Commodity swap A swap in which the underlying is a commodity such as oil, gold, or an agricultural product.

Common market Level of economic integration that incorporates all aspects of the customs union and extends it by allowing free movement of factors of production among members.

Common shares A type of security that represent an ownership interest in a company.

Common-size analysis The restatement of financial statement items using a common denominator or reference item that allows one to identify trends and major differences; an example is an income statement in which all items are expressed as a percent of revenue.

Common stock See *common shares.*

Company analysis Analysis of an individual company.

Comparable company A company that has similar business risk, usually in the same industry and preferably with a single line of business.

Comparative advantage A country's ability to produce a good or service at a lower relative cost, or opportunity cost, than its trading partner.

Competitive strategy A company's plans for responding to the threats and opportunities presented by the external environment.

Complement The event not-S, written S^C, given the event S. Note that $P(S) + P(S^C) = 1$.

Complements Goods that tend to be used together; technically, two goods whose cross-price elasticity of demand is negative.

Complete markets Informally, markets in which the variety of distinct securities traded is so broad that any desired payoff in a future state-of-the-world is achievable.

Component cost of capital The rate of return required by suppliers of capital for an individual source of a company's funding, such as debt or equity.

Compounding The process of accumulating interest on interest.

Comprehensive income The change in equity of a business enterprise during a period from nonowner sources; includes all changes in equity during a period except those resulting from investments by owners and distributions to owners; comprehensive income equals net income plus other comprehensive income.

Conditional expected value The expected value of a stated event given that another event has occurred.

Conditional pass-through covered bonds Covered bonds that convert to pass-through securities after the original maturity date if all bond payments have not yet been made and the sponsor is in default.

Conditional probability The probability of an event given (conditioned on) another event.

Conditional variances The variance of one variable, given the outcome of another.

Confidence level The complement of the level of significance.

Confirmation bias A belief perseverance bias in which people tend to look for and notice what confirms their beliefs, to ignore or undervalue what contradicts their beliefs, and to misinterpret information as support for their beliefs.

Confusion matrix A type of contingency table used for evaluating the performance of a classification model.

Conservatism bias A belief perseverance bias in which people maintain their prior views or forecasts by inadequately incorporating new information.

Consolidation The movement of a stock's price within a well-defined range of trading levels for a period of time. The price consolidates between a support level and a resistance level.

Constant-yield price trajectory A graph that illustrates the change in the price of a fixed-income bond over time assuming no change in yield-to-maturity. The trajectory shows the "pull to par" effect on the price of a bond trading at a premium or a discount to par value.

Constituent securities With respect to an index, the individual securities within an index.

Consumer surplus The difference between the value that a consumer places on units purchased and the amount of money that was required to pay for them.

Contango A condition in the futures markets in which the spot price is lower than the futures price, the forward curve is upward sloping, and there is little or no convenience yield.

Contingency provision Clause in a legal document that allows for some action if a specific event or circumstance occurs.

Contingency table A tabular format that displays the frequency distributions of two or more categorical variables simultaneously and is used for finding patterns between the variables. A contingency table for two categorical variables is also known as a two-way table.

Contingent claim A security whose value is based on some other underlying security.

Contingent claims Derivatives in which the payoffs occur if a specific event occurs; generally referred to as options.

Contingent convertible bonds Bonds that automatically convert into equity if a specific event or circumstance occurs, such as the issuer's equity capital falling below the minimum requirement set by the regulators. Also called *CoCos.*

Continuation pattern A type of pattern used in technical analysis to predict the resumption of a market trend that was in place prior to the formation of a pattern.

Continuous data Data that can be measured and can take on any numerical value in a specified range of values.

Continuous random variable A random variable for which the range of possible outcomes is the real line (all real numbers between $-\infty$ and $+\infty$) or some subset of the real line.

Continuous trading market A market in which trades can be arranged and executed any time the market is open.

Continuously compounded return The natural logarithm of 1 plus the holding period return, or equivalently, the natural logarithm of the ending price over the beginning price.

Contra account An account that offsets another account.

Contract rate See *mortgage rate.*

Contraction The period of a business cycle after the peak and before the trough; often called a *recession* or, if exceptionally severe, called a *depression*.

Contraction risk The risk that when interest rates decline, the security will have a shorter maturity than was anticipated at the time of purchase because borrowers refinance at the new, lower interest rates.

Contractionary Tending to cause the real economy to contract.

Contractionary fiscal policy A fiscal policy that has the objective to make the real economy contract.

Contracts for differences See *non-deliverable forwards.*

Contribution margin The amount available for fixed costs and profit after paying variable costs; revenue minus variable costs.

Controlling shareholders A particular shareholder or block of shareholders holding a percentage of shares that gives them significant voting power.

Convenience sampling A procedure of selecting an element from a population on the basis of whether or not it is accessible to a researcher or how easy it is for a researcher to access the element.

Convenience yield A non-monetary advantage of holding an asset.

Conventional bond See *plain vanilla bond*.

Conventional cash flows A conventional cash flow pattern is one with an initial outflow followed by a series of inflows.

Convergence The tendency for differences in output per capita across countries to diminish over time. In technical analysis, the term describes the case when an indicator moves in the same manner as the security being analyzed.

Conversion premium The difference between the convertible bond's price and its conversion value.

Conversion price For a convertible bond, the price per share at which the bond can be converted into shares.

Conversion ratio For a convertible bond, the number of common shares that each bond can be converted into.

Conversion value For a convertible bond, the current share price multiplied by the conversion ratio.

Convertible bond Bond that gives the bondholder the right to exchange the bond for a specified number of common shares in the issuing company.

Convertible preference shares A type of equity security that entitles shareholders to convert their shares into a specified number of common shares.

Convexity adjustment For a bond, one half of the annual or approximate convexity statistic multiplied by the change in the yield-to-maturity squared.

Core inflation Refers to the inflation rate calculated based on a price index of goods and services except food and energy.

Corporate governance The system of internal controls and procedures by which individual companies are managed.

Correlation A measure of the linear relationship between two random variables.

Correlation coefficient A number between −1 and +1 that measures the consistency or tendency for two investments to act in a similar way. It is used to determine the effect on portfolio risk when two assets are combined.

Cost averaging The periodic investment of a fixed amount of money.

Cost of capital The cost of financing to a company; the rate of return that suppliers of capital require as compensation for their contribution of capital.

Cost of carry See *carry*.

Cost of debt The cost of debt financing to a company, such as when it issues a bond or takes out a bank loan; the rate of return required by a company's bondholders.

Cost of equity The rate of return required by a company's common stockholders.

Cost of preferred stock The cost to a company of issuing preferred stock; the dividend yield that a company must commit to pay preferred stockholders.

Cost structure The mix of a company's variable costs and fixed costs.

Counterparty risk The risk that the other party to a contract will fail to honor the terms of the contract.

Coupon rate The interest rate promised in a contract; this is the rate used to calculate the periodic interest payments.

Cournot assumption Assumption in which each firm determines its profit-maximizing production level assuming that the other firms' output will not change.

Covariance A measure of the co-movement (linear association) between two random variables.

Covariance matrix A matrix or square array whose entries are covariances; also known as a variance–covariance matrix.

Covenants The terms and conditions of lending agreements that the issuer must comply with; they specify the actions that an issuer is obligated to perform (affirmative covenant) or prohibited from performing (negative covenant).

Covered bond Debt obligation secured by a segregated pool of assets called the cover pool. The issuer must maintain the value of the cover pool. In the event of default, bondholders have recourse against both the issuer and the cover pool.

Covered bonds Senior debt obligations issued by a financial institution and backed by a segregated pool of assets. Covered bond investors have recourse to both the underlying assets and the financial institution.

Credit analysis The evaluation of credit risk; the evaluation of the creditworthiness of a borrower or counterparty.

Credit default swap (CDS) A type of credit derivative in which one party, the credit protection buyer who is seeking credit protection against a third party, makes a series of regularly scheduled payments to the other party, the credit protection seller. The seller makes no payments until a credit event occurs.

Credit derivatives A contract in which one party has the right to claim a payment from another party in the event that a specific credit event occurs over the life of the contract.

Credit enhancements Provisions that may be used to reduce the credit risk of a bond issue.

Credit-linked coupon bond Bond for which the coupon changes when the bond's credit rating changes.

Credit-linked note (CLN) Fixed-income security in which the holder of the security has the right to withhold payment of the full amount due at maturity if a credit event occurs.

Credit migration risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *downgrade risk*.

Credit risk The risk of loss caused by a counterparty's or debtor's failure to make a promised payment. Also called *default risk*.

Credit spread option An option on the yield spread on a bond.

Credit tranching A structure used to redistribute the credit risk associated with the collateral; a set of bond classes created to allow investors a choice in the amount of credit risk that they prefer to bear.

Creditworthiness The perceived ability of the borrower to pay its debt obligations in a timely manner; it represents the ability of a company to withstand adverse impacts on its cash flows.

Critical values Values of the test statistic at which the decision changes from fail to reject the null hypothesis to reject the null hypothesis.

Cross-default Covenant or contract clause that specifies a borrower is considered in default if they default on another debt obligation.

Cross-default provisions Provisions whereby events of default, such as non-payment of interest on one bond, trigger default on all outstanding debt; implies the same default probability for all issues.

Cross-price elasticity of demand The percentage change in quantity demanded for a given percentage change in the price of another good; the responsiveness of the demand for Product A that is associated with the change in price of Product B.

Cross-sectional analysis Analysis that involves comparisons across individuals in a group over a given time period or at a given point in time.

Cross-sectional data A list of the observations of a specific variable from multiple observational units at a given point in time. The observational units can be individuals, groups, companies, trading markets, regions, etc.

Crossing networks Trading systems that match buyers and sellers who are willing to trade at prices obtained from other markets.

Crowding out The thesis that government borrowing may divert private sector investment from taking place.

Cryptocurrency An electronic medium of exchange that lacks physical form.

Cryptography An algorithmic process to encrypt data, making the data unusable if received by unauthorized parties.

Cumulative absolute frequency Cumulates (i.e., adds up) in a frequency distribution the absolute frequencies as one moves from the first bin to the last bin.

Cumulative distribution function A function giving the probability that a random variable is less than or equal to a specified value.

Cumulative frequency distribution chart A chart that plots either the cumulative absolute frequency or the cumulative relative frequency on the y-axis against the upper limit of the interval and allows one to see the number or the percentage of the observations that lie below a certain value.

Cumulative preference shares Preference shares for which any dividends that are not paid accrue and must be paid in full before dividends on common shares can be paid.

Cumulative relative frequency A sequence of partial sums of the relative frequencies in a frequency distribution.

Cumulative voting A voting process whereby shareholders can accumulate and vote all their shares for a single candidate in an election, as opposed to having to allocate their voting rights evenly among all candidates.

Currencies Monies issued by national monetary authorities.

Currency option bonds Bonds that give bondholders the right to choose the currency in which they want to receive interest payments and principal repayments.

Currency swap A swap in which each party makes interest payments to the other in different currencies.

Current account A component of the balance of payments account that measures the flow of goods and services.

Current assets Assets that are expected to be consumed or converted into cash in the near future, typically one year or less. *Also called liquid assets.*

Current cost With reference to assets, the amount of cash or cash equivalents that would have to be paid to buy the same or an equivalent asset today; with reference to liabilities, the undiscounted amount of cash or cash equivalents that would be required to settle the obligation today.

Current government spending With respect to government expenditures, spending on goods and services that are provided on a regular, recurring basis including health, education, and defense.

Current liabilities Short-term obligations, such as accounts payable, wages payable, or accrued liabilities, that are expected to be settled in the near future, typically one year or less.

Current yield The sum of the coupon payments received over the year divided by the flat price; also called the *income* or *interest yield* or *running yield*.

Curve duration The sensitivity of the bond price (or the market value of a financial asset or liability) with respect to a benchmark yield curve.

Customs union Extends the free trade area (FTA) by not only allowing free movement of goods and services among members, but also creating a common trade policy against nonmembers.

CVaR Conditional VaR, a tail loss measure. The weighted average of all loss outcomes in the statistical distribution that exceed the VaR loss.

Cyclical A cyclical company is one whose profits are strongly correlated with the strength of the overall economy.

Cyclical companies Companies with sales and profits that regularly expand and contract with the business cycle or state of economy.

Daily settlement See *mark to market* and *marking to market.*

Dark pools Alternative trading systems that do not display the orders that their clients send to them.

Data A collection of numbers, characters, words, and text—as well as images, audio, and video—in a raw or organized format to represent facts or information.

Data mining The practice of determining a model by extensive searching through a dataset for statistically significant patterns. Also called *data snooping.*

Data science An interdisciplinary field that brings computer science, statistics, and other disciplines together to analyze and produce insights from Big Data.

Data snooping The practice of determining a model by extensive searching through a dataset for statistically significant patterns.

Data table see **two-dimensional rectangular array**.

Day order An order that is good for the day on which it is submitted. If it has not been filled by the close of business, the order expires unfilled.

Days of inventory on hand An activity ratio equal to the number of days in the period divided by inventory turnover over the period.

Dealers A financial intermediary that acts as a principal in trades.

Dealing securities Securities held by banks or other financial intermediaries for trading purposes.

Death cross See **Bearish crossover**.

Debentures Type of bond that can be secured or unsecured.

Debt rating An independent measure of the quality and safety of a company's debt based upon an analysis of the company's ability to pay the promised cash flows in full and on time. It includes an analysis of any indentures.

Debt-rating approach A method for estimating a company's before-tax cost of debt based on the yield on comparably rated bonds for maturities that closely match that of the company's existing debt.

Debt tax shield The tax benefit from interest paid on debt being tax deductible from income, equal to the marginal tax rate multiplied by the value of the debt.

Debt-to-assets ratio A solvency ratio calculated as total debt divided by total assets.

Debt-to-capital ratio A solvency ratio calculated as total debt divided by total debt plus total shareholders' equity.

Debt-to-equity ratio A solvency ratio calculated as total debt divided by total shareholders' equity.

Deciles Quantiles that divide a distribution into 10 equal parts.

Declaration date The day that the corporation issues a statement declaring a specific dividend.

Decreasing returns to scale When a production process leads to increases in output that are proportionately smaller than the increase in inputs.

Deductible temporary differences Temporary differences that result in a reduction of or deduction from taxable income in a future period when the balance sheet item is recovered or settled.

Deep learning Machine learning using neural networks with many hidden layers.

Deep learning nets Machine learning using neural networks with many hidden layers.

Default probability The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default risk*.

Default risk The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default probability*.

Default risk premium An extra return that compensates investors for the possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount.

Defensive companies Companies with sales and profits that have little sensitivity to the business cycle or state of the economy.

Defensive interval ratio A liquidity ratio that estimates the number of days that an entity could meet cash needs from liquid assets; calculated as (cash + short-term marketable investments + receivables) divided by daily cash expenditures.

Deferred coupon bond Bond that pays no coupons for its first few years but then pays a higher coupon than it otherwise normally would for the remainder of its life. Also called *split coupon bond*.

Deferred income A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred tax assets A balance sheet asset that arises when an excess amount is paid for income taxes relative to accounting profit. The taxable income is higher than accounting profit and income tax payable exceeds tax expense. The company expects to recover the difference during the course of future operations when tax expense exceeds income tax payable.

Deferred tax liabilities A balance sheet liability that arises when a deficit amount is paid for income taxes relative to accounting profit. The taxable income is less than the accounting profit and income tax payable is less than tax expense. The company expects to eliminate the liability over the course of future operations when income tax payable exceeds tax expense.

Defined benefit pension plans Plans in which the company promises to pay a certain annual amount (defined benefit) to the employee after retirement. The company bears the investment risk of the plan assets.

Defined contribution pension plans Individual accounts to which an employee and typically the employer makes contributions during their working years and expect to draw on the accumulated funds at retirement. The employee bears the investment and inflation risk of the plan assets.

Deflation Negative inflation.

Degree of confidence The probability that a confidence interval includes the unknown population parameter.

Degree of financial leverage (DFL) The ratio of the percentage change in net income to the percentage change in operating income; the sensitivity of the cash flows available to owners when operating income changes.

Degree of operating leverage (DOL) The ratio of the percentage change in operating income to the percentage change in units sold; the sensitivity of operating income to changes in units sold.

Degree of total leverage The ratio of the percentage change in net income to the percentage change in units sold; the sensitivity of the cash flows to owners to changes in the number of units produced and sold.

Degrees of freedom The number of independent variables used in defining sample statistics, such as variance, and the probability distributions they measure.

Delta The sensitivity of the derivative price to a small change in the value of the underlying asset.

Demand curve Graph of the inverse demand function. A graph showing the demand relation, either the highest quantity willingly purchased at each price or the highest price willingly paid for each quantity.

Demand function A relationship that expresses the quantity demanded of a good or service as a function of own-price and possibly other variables.

Demand shock A typically unexpected disturbance to demand, such as an unexpected interruption in trade or transportation.

Dependent With reference to events, the property that the probability of one event occurring depends on (is related to) the occurrence of another event.

Dependent variable The variable whose variation about its mean is to be explained by the regression; the left-side variable in a regression equation. Also referred to as the *explained variable*.

Depository bank A bank that raises funds from depositors and other investors and lends it to borrowers.

Depository institutions Commercial banks, savings and loan banks, credit unions, and similar institutions that raise funds from depositors and other investors and lend it to borrowers.

Depository receipt A security that trades like an ordinary share on a local exchange and represents an economic interest in a foreign company.

Depreciation The process of systematically allocating the cost of long-lived (tangible) assets to the periods during which the assets are expected to provide economic benefits.

Derivative pricing rule A pricing rule used by crossing networks in which a price is taken (derived) from the price that is current in the asset's primary market.

Derivatives A financial instrument whose value depends on the value of some underlying asset or factor (e.g., a stock price, an interest rate, or exchange rate).

Descriptive statistics Measures that summarize central tendency and spread variation in the data's distribution.

Diffuse prior The assumption of equal prior probabilities.

Diffusion index Reflects the proportion of the index's components that are moving in a pattern consistent with the overall index.

Diluted EPS The EPS that would result if all dilutive securities were converted into common shares.

Diluted shares The number of shares that would be outstanding if all potentially dilutive claims on common shares (e.g., convertible debt, convertible preferred stock, and employee stock options) were exercised.

Diminishing balance method An accelerated depreciation method, i.e., one that allocates a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Diminishing marginal productivity Describes a state in which each additional unit of input produces less output than previously.

Direct format With reference to the cash flow statement, a format for the presentation of the statement in which cash flow from operating activities is shown as operating cash receipts less operating cash disbursements. Also called *direct method*.

Direct investing Occurs when an investor makes a direct investment in an asset without the use of an intermediary.

Direct method See *direct format*.

Direct taxes Taxes levied directly on income, wealth, and corporate profits.

Direct write-off method An approach to recognizing credit losses on customer receivables in which the company waits until such time as a customer has defaulted and only then recognizes the loss.

Discount To reduce the value of a future payment in allowance for how far away it is in time; to calculate the present value of some future amount. Also, the amount by which an instrument is priced below its face (par) value.

Discount margin See *required margin*.

Discount rates In general, the interest rates used to calculate present values. In the money market, however, a discount rate is a specific type of quoted rate.

Discounted cash flow models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security.

Discouraged worker A person who has stopped looking for a job or has given up seeking employment.

Discrete data Numerical values that result from a counting process; therefore, practically speaking, the data are limited to a finite number of values.

Discrete random variable A random variable that can take on at most a countable number of possible values.

Discriminatory pricing rule A pricing rule used in continuous markets in which the limit price of the order or quote that first arrived determines the trade price.

Diseconomies of scale Increase in cost per unit resulting from increased production.

Dispersion The variability of a population or sample of observations around the central tendency.

Display size The size of an order displayed to public view.

Disposition effect As a result of loss aversion, an emotional bias whereby investors are reluctant to dispose of losers. This reluctance results in an inefficient and gradual adjustment to deterioration in fundamental value.

Distributed ledger A type of database that may be shared among entities in a network.

Distributed ledger technology Technology based on a distributed ledger.

Divergence In technical analysis, a term that describes the case when an indicator moves differently from the security being analyzed.

Diversification ratio The ratio of the standard deviation of an equally weighted portfolio to the standard deviation of a randomly selected security.

Dividend A distribution paid to shareholders based on the number of shares owned.

Dividend discount model (DDM) A present value model that estimates the intrinsic value of an equity share based on the present value of its expected future dividends.

Dividend payout ratio The ratio of cash dividends paid to earnings for a period.

Divisor A number (denominator) used to determine the value of a price return index. It is initially chosen at the inception of an index and subsequently adjusted by the index provider, as necessary, to avoid changes in the index value that are unrelated to changes in the prices of its constituent securities.

Doji In the Japanese terminology used in candlestick charting, the doji signifies that after a full day of trading, the positive price influence of buyers and the negative price influence of sellers exactly counteracted each other—with opening and closing prices that are virtually equal—which suggests that the market under analysis is in balance.

Domestic content provisions Stipulate that some percentage of the value added or components used in production should be of domestic origin.

Double bottom In technical analysis, a reversal pattern that is formed when the price reaches a low, rebounds, and then declines back to the first low level. A double bottom is used to predict a change from a downtrend to an uptrend.

Double coincidence of wants A prerequisite to barter trades, in particular that both economic agents in the transaction want what the other is selling.

Double declining balance depreciation An accelerated depreciation method that involves depreciating the asset at double the straight-line rate. This rate is multiplied by the book value of the asset at the beginning of the period (a declining balance) to calculate depreciation expense.

Double top In technical analysis, a reversal pattern that is formed when an uptrend reverses twice at roughly the same high price level. A double top is used to predict a change from an uptrend to a downtrend.

Down transition probability The probability that an asset's value moves down in a model of asset price dynamics.

Downgrade risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *credit migration risk*.

Downside risk Risk of incurring returns below a specified value.

Downtrend A pattern that occurs when the price of an asset moves lower over a period of time.

Drag on liquidity When receipts lag, creating pressure from the decreased available funds.

Drawdown A percentage peak-to-trough reduction in net asset value.

Dual-currency bonds Bonds that make coupon payments in one currency and pay the par value at maturity in another currency.

DuPont analysis An approach to decomposing return on investment, e.g., return on equity, as the product of other financial ratios.

Duration A measure of the approximate sensitivity of a security to a change in interest rates (i.e., a measure of interest rate risk).

Duration gap A bond's Macaulay duration minus the investment horizon.

Dutch Book Theorem A result in probability theory stating that inconsistent probabilities create profit opportunities.

Early repayment option See *prepayment option*.

Earnings per share The amount of income earned during a period per share of common stock.

Earnings surprise The portion of a company's earnings that is unanticipated by investors and, according to the efficient market hypothesis, merits a price adjustment.

Economic costs All the remuneration needed to keep a productive resource in its current employment or to acquire the resource for productive use; the sum of total accounting costs and implicit opportunity costs.

Economic indicator A variable that provides information on the state of the overall economy.

Economic loss The amount by which accounting profit is less than normal profit.

Economic profit Equal to accounting profit less the implicit opportunity costs not included in total accounting costs; the difference between total revenue (TR) and total cost (TC). Also called *abnormal profit* or *supernormal profit*.

Economic stabilization Reduction of the magnitude of economic fluctuations.

Economic union Incorporates all aspects of a common market and in addition requires common economic institutions and coordination of economic policies among members.

Economies of scale Reduction in cost per unit resulting from increased production.

Effective annual rate The amount by which a unit of currency will grow in a year with interest on interest included.

Effective convexity A *curve convexity* statistic that measures the secondary effect of a change in a benchmark yield curve on a bond's price.

Effective duration The sensitivity of a bond's price to a change in a benchmark yield curve.

Effective interest rate The borrowing rate or market rate that a company incurs at the time of issuance of a bond.

Efficient market A market in which asset prices reflect new information quickly and rationally.

Elastic Said of a good or service when the magnitude of elasticity is greater than one.

Elasticity The percentage change in one variable for a percentage change in another variable; a general measure of how sensitive one variable is to a change in the value of another variable.

Elasticity of demand A measure of the sensitivity of quantity demanded to a change in a product's own price: $\%\Delta Q^D/\%\Delta P$.

Elasticity of supply A measure of the sensitivity of quantity supplied to a change in price: $\%\Delta Q^S/\%\Delta P$.

Electronic communications networks See *alternative trading systems*.

Embedded option Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Embedded options Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Emotional biases Behavioral biases resulting from reasoning influenced by feelings; emotional biases stem from impulse or intuition.

Empirical duration The use of statistical methods and historical bond prices to estimate the price–yield relationship for a specific bond or portfolio of bonds.

Empirical probability The probability of an event estimated as a relative frequency of occurrence.

Employed The number of people with a job.

Endowment bias An emotional bias in which people value an asset more when they hold rights to it than when they do not.

Engagement/active ownership An ESG investment approach that uses shareholder power to influence corporate behavior through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is directed by ESG guidelines.

Enterprise risk management An overall assessment of a company's risk position. A centralized approach to risk management sometimes called firmwide risk management.

Enterprise value A measure of a company's total market value from which the value of cash and short-term investments have been subtracted.

Equal weighting An index weighting method in which an equal weight is assigned to each constituent security at inception.

Equipment trust certificates Bonds secured by specific types of equipment or physical assets.

Equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Equity risk premium The expected return on equities minus the risk-free rate; the premium that investors demand for investing in equities.

Equity swap A swap transaction in which at least one cash flow is tied to the return to an equity portfolio position, often an equity index.

Error term The difference between an observation and its expected value, where the expected value is based on the true underlying population relation between the dependent and independent variables. Also known simply as the *error*.

ESG An acronym that encompasses environmental, social, and governance.

ESG integration An ESG investment approach that focuses on systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions for the purpose of achieving the product's stated investment objectives. Used interchangeably with **ESG investing**.

ESG investing The consideration of environmental, social, and governance factors in the investment process. Used interchangeably with **ESG integration**.

Estimate The particular value calculated from sample observations using an estimator.

Estimated parameters With reference to a regression analysis, the estimated values of the population intercept and population slope coefficients in a regression.

Estimator An estimation formula; the formula used to compute the sample mean and other sample statistics are examples of estimators.

Ethical principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Ethics The study of moral principles or of making good choices. Ethics encompasses a set of moral principles and rules of conduct that provide guidance for our behavior.

Eurobonds Type of bond issued internationally, outside the jurisdiction of the country in whose currency the bond is denominated.

European-style Said of an option contract that can only be exercised on the option's expiration date.

Event Any outcome or specified set of outcomes of a random variable.

Ex-dividend date The first date that a share trades without (i.e., "ex") the dividend.

Excess kurtosis Degree of kurtosis (fatness of tails) relative to the kurtosis of the normal distribution.

Exchanges Places where traders can meet to arrange their trades.

Execution instructions Instructions that indicate how to fill an order.

Exercise The process of using an option to buy or sell the underlying.

Exercise price The fixed price at which an option holder can buy or sell the underlying. Also called *strike price*, *striking price*, or *strike*.

Exercise value The value obtained if an option is exercised based on current conditions. Also known as *intrinsic value*.

Exhaustive Covering or containing all possible outcomes.

Expansion The period of a business cycle after its lowest point and before its highest point.

Expansionary Tending to cause the real economy to grow.

Expansionary fiscal policy Fiscal policy aimed at achieving real economic growth.

Expected inflation The level of inflation that economic agents expect in the future.

Expected loss Default probability times loss severity given default.

Expected return on the portfolio ($E(R_p)$) The weighted average of the expected returns (R_1 to R_n) on the component securities using their respective weights (w_1 to w_n).

Expected value The probability-weighted average of the possible outcomes of a random variable.

Expenses Outflows of economic resources or increases in liabilities that result in decreases in equity (other than decreases because of distributions to owners); reductions in net assets associated with the creation of revenues.

Export subsidy Paid by the government to the firm when it exports a unit of a good that is being subsidized.

Exports Goods and services that an economy sells to other countries.

Extension risk The risk that when interest rates rise, fewer prepayments will occur because homeowners are reluctant to give up the benefits of a contractual interest rate that now looks low. As a result, the security becomes longer in maturity than anticipated at the time of purchase.

Externalities Spillover effects of production and consumption activities onto others who are not directly involved in a particular transaction, activity, or decision.

Externality An effect of a market transaction that is borne by parties other than those who transacted.

Extra dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Extreme value theory A branch of statistics that focuses primarily on extreme outcomes.

Face value The amount of cash payable by a company to the bondholders when the bonds mature; the promised payment at maturity separate from any coupon payment.

Factoring arrangement When a company sells its accounts receivable to a lender (known as a factor) who assumes responsibility for the credit granting and collection process.

Fair value The amount at which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's-length transaction; the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants.

False discovery approach An adjustment in the *p*-values for tests performed multiple times.

False discovery rate The rate of Type I errors in testing a null hypothesis multiple times for a given level of significance.

Fat-Tailed Describes a distribution that has fatter tails than a normal distribution (also called leptokurtic).

Fed funds rate The US interbank lending rate on overnight borrowings of reserves.

Federal funds rate The US interbank lending rate on overnight borrowings of reserves.

Fiat money Money that is not convertible into any other commodity.

Fiduciary An entity designated to represent the rights and responsibilities of a beneficiary whose assets they are managing, such as a bond trustee acting on behalf of fixed-income investors.

Fiduciary call A combination of a European call and a risk-free bond that matures on the option expiration day and has a face value equal to the exercise price of the call.

FIFO method The first in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the order in which they were placed in inventory.

Fill or kill See *immediate or cancel order*.

Finance lease From the lessee perspective, under US GAAP, a type of lease which is more akin to the purchase of an asset by the lessee. From the lessor perspective, under IFRS, a lease which "transfers substantially all the risks and rewards incidental to ownership of an underlying asset."

Financial account A component of the balance of payments account that records investment flows.

Financial distress Heightened uncertainty regarding a company's ability to meet its various obligations because of diminished earnings power or actual current losses.

Financial flexibility The ability to react and adapt to financial adversity and opportunities.

Financial leverage The extent to which a company can effect, through the use of debt, a proportional change in the return on common equity that is greater than a given proportional change in operating income; also, short for the financial leverage ratio.

Financial leverage ratio A measure of financial leverage calculated as average total assets divided by average total equity.

Financial risk The risk that environmental, social, or governance risk factors will result in significant costs or other losses to a company and its shareholders; the risk arising from a company's obligation to meet required payments under its financing agreements.

Financing activities Activities related to obtaining or repaying capital to be used in the business (e.g., equity and long-term debt).

Fintech Technological innovation in the design and delivery of financial services and products in the financial industry.

Firm commitment offering See *underwritten offering*.

First-degree price discrimination Where a monopolist is able to charge each customer the highest price the customer is willing to pay.

First lien debt Debt secured by a pledge of certain assets that could include buildings, but it may also include property and equipment, licenses, patents, brands, etc.

First mortgage debt Debt secured by a pledge of a specific property.

Fiscal multiplier The ratio of a change in national income to a change in government spending.

Fiscal policy The use of taxes and government spending to affect the level of aggregate expenditures.

Fisher effect The thesis that the real rate of interest in an economy is stable over time so that changes in nominal interest rates are the result of changes in expected inflation.

Fixed charge coverage A solvency ratio measuring the number of times interest and lease payments are covered by operating income, calculated as (EBIT + lease payments) divided by (interest payments + lease payments).

Fixed costs Costs that remain at the same level regardless of a company's level of production and sales.

Fixed-for-floating interest rate swap An interest rate swap in which one party pays a fixed rate and the other pays a floating rate, with both sets of payments in the same currency. Also called *plain vanilla swap* or *vanilla swap*.

Fixed-rate perpetual preferred stock Nonconvertible, noncallable preferred stock that has a fixed dividend rate and no maturity date.

Flag A technical analysis continuation pattern formed by parallel trendlines, typically over a short period.

Flat price The full price of a bond minus the accrued interest; also called the *quoted* or *clean* price.

Float-adjusted market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by adjusting its market capitalization for its market float.

Floaters See *floating-rate notes*.

Floating-rate notes Notes on which interest payments are not fixed but instead vary from period to period depending on the current level of a reference interest rate.

Flotation cost Fees charged to companies by investment bankers and other costs associated with raising new capital.

Foreclosure Allows the lender to take possession of a mortgaged property if the borrower defaults and then sell it to recover funds.

Foreign currency reserves Holding by the central bank of non-domestic currency deposits and non-domestic bonds.

Foreign direct investment Direct investment by a firm in one country (the source country) in productive assets in a foreign country (the host country).

Foreign exchange gains (or losses) Gains (or losses) that occur when the exchange rate changes between the investor's currency and the currency that foreign securities are denominated in.

Foreign portfolio investment Shorter-term investment by individuals, firms, and institutional investors (e.g., pension funds) in foreign financial instruments such as foreign stocks and foreign government bonds.

Forward commitments Class of derivatives that provides the ability to lock in a price to transact in the future at a previously agreed-upon price.

Forward contract An agreement between two parties in which one party, the buyer, agrees to buy from the other party, the seller, an underlying asset at a later date for a price established at the start of the contract.

Forward curve A series of forward rates, each having the same time frame.

Forward market For future delivery, beyond the usual settlement time period in the cash market.

Forward price The fixed price or rate at which the transaction scheduled to occur at the expiration of a forward contract will take place. This price is agreed on at the initiation date of the contract.

Forward rate The interest rate on a bond or money market instrument traded in a forward market. A forward rate can be interpreted as an incremental, or marginal, return for extending the time-to-maturity for an additional time period.

Forward rate agreements A forward contract calling for one party to make a fixed interest payment and the other to make an interest payment at a rate to be determined at the contract expiration.

Fractile A value at or below which a stated fraction of the data lies. Also called quantile.

Fractional reserve banking Banking in which reserves constitute a fraction of deposits.

Framing bias An information-processing bias in which a person answers a question differently based on the way in which it is asked (framed).

Free cash flow The actual cash that would be available to the company's investors after making all investments necessary to maintain the company as an ongoing enterprise (also referred to as free cash flow to the firm); the internally generated funds that can be distributed to the company's investors (e.g., shareholders and bondholders) without impairing the value of the company.

Free cash flow hypothesis The hypothesis that higher debt levels discipline managers by forcing them to make fixed debt service payments and by reducing the company's free cash flow.

Free-cash-flow-to-equity models Valuation models based on discounting expected future free cash flow to equity.

Free float The number of shares that are readily and freely tradable in the secondary market.

Free trade When there are no government restrictions on a country's ability to trade.

Free trade areas One of the most prevalent forms of regional integration, in which all barriers to the flow of goods and services among members have been eliminated.

Frequency distribution A tabular display of data constructed either by counting the observations of a variable by distinct values or groups or by tallying the values of a numerical variable into a set of numerically ordered bins (also called a one-way table).

Frequency polygon A graph of a frequency distribution obtained by drawing straight lines joining successive points representing the class frequencies.

Frequency table A representation of the frequency of occurrence of two discrete variables.

Full price The price of a security with accrued interest; also called the *invoice* or *dirty* price.

Fund investing In fund investing, the investor invests in assets indirectly by contributing capital to a fund as part of a group of investors. Fund investing is available for all major alternative investment types.

Fundamental analysis The examination of publicly available information and the formulation of forecasts to estimate the intrinsic value of assets.

Fundamental value The underlying or true value of an asset based on an analysis of its qualitative and quantitative characteristics. Also called *intrinsic value*.

Fundamental weighting An index weighting method in which the weight assigned to each constituent security is based on its underlying company's size. It attempts to address the disadvantages of market-capitalization weighting by using measures that are independent of the constituent security's price.

Funds of hedge funds Funds that hold a portfolio of hedge funds, more commonly shortened to *funds of funds*.

Fungible Freely exchangeable, interchangeable, or substitutable with other things of the same type. Money and commodities are the most common examples.

Future value (FV) The amount to which a payment or series of payments will grow by a stated future date.

Futures contract A variation of a forward contract that has essentially the same basic definition but with some additional features, such as a clearinghouse guarantee against credit losses, a daily settlement of gains and losses, and an organized electronic or floor trading facility.

Futures price The agreed-upon price of a futures contract.

FX swap The combination of a spot and a forward FX transaction.

G-spread The yield spread in basis points over an actual or interpolated government bond.

Gains Asset inflows not directly related to the ordinary activities of the business.

Game theory The set of tools decision makers use to incorporate responses by rival decision makers into their strategies.

Gamma A numerical measure of how sensitive an option's delta (the sensitivity of the derivative's price) is to a change in the value of the underlying.

Gap opening A gap is an area of a chart where a security's price either rises or falls from the previous day's close with no trading occurring in between. A gap opening is the start of a new trading session with a gap.

GDP deflator A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

General partner (GP) The partner that runs the business and theoretically bears unlimited liability for the business's debts and obligations.

Geometric mean A measure of central tendency computed by taking the nth root of the product of n non-negative values.

Giffen goods Goods that are consumed more as the price of the good rises because it is a very inferior good whose income effect overwhelms its substitution effect when price changes.

Gilts Bonds issued by the UK government.

Global depository receipt A depository receipt that is issued outside of the company's home country and outside of the United States.

Global minimum-variance portfolio The portfolio on the minimum-variance frontier with the smallest variance of return.

Global registered share A common share that is traded on different stock exchanges around the world in different currencies.

Gold standard With respect to a currency, if a currency is on the gold standard a given amount can be converted into a prespecified amount of gold.

Golden Cross See **Bullish crossover**.

Good-on-close An execution instruction specifying that an order can only be filled at the close of trading. Also called *market on close*.

Good-on-open An execution instruction specifying that an order can only be filled at the opening of trading.

Good-till-cancelled order An order specifying that it is valid until the entity placing the order has cancelled it (or, commonly, until some specified amount of time such as 60 days has elapsed, whichever comes sooner).

Goodwill An intangible asset that represents the excess of the purchase price of an acquired company over the value of the net assets acquired.

Government equivalent yield A yield that restates a yield-to-maturity based on a 30/360 day count to one based on actual/actual.

Green bonds Bonds used in green finance whereby the proceeds are earmarked toward environmental-related products.

Green finance A type of finance that addresses environmental concerns while achieving economic growth.

Green loans Any loan instruments made available exclusively to finance or re-finance, in whole or in part, new and/or existing eligible green projects. Green loans are commonly aligned in the market with the Green Loan Principles.

Greenfield investment Investing in infrastructure assets *that are to be constructed*.

Grey market The forward market for bonds about to be issued. Also called "when issued" market.

Gross domestic product The market value of all final goods and services produced within the economy during a given period (output definition) or, equivalently, the aggregate income earned by all households, all companies, and the government within the economy during a given period (income definition).

Gross margin Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit margin The ratio of gross profit to revenues.

Grouped bar chart A bar chart for showing joint frequencies for two categorical variables (also known as a **clustered bar chart**).

Grouping by function With reference to the presentation of expenses in an income statement, the grouping together of expenses serving the same function, e.g. all items that are costs of goods sold.

Grouping by nature With reference to the presentation of expenses in an income statement, the grouping together of expenses by similar nature, e.g., all depreciation expenses.

Growth cycle Refers to fluctuations in economic activity around the long-term potential trend growth level, focusing on how much actual economic activity is below or above trend growth in economic activity.

Growth cyclical A term sometimes used to describe companies that are growing rapidly on a long-term basis but that still experience above-average fluctuation in their revenues and profits over the course of a business cycle.

Growth investors With reference to equity investors, investors who seek to invest in high-earnings-growth companies.

Growth option The ability to make additional investments in a project at some future time if the financial results are strong. Also called *expansion option*.

Growth rate cycle Refers to fluctuations in the growth rate of economic activity.

Guarantee certificate A type of structured financial instrument that provides investors with capital protection. It combines a zero-coupon bond and a call option on some underlying asset.

Haircut See *repo margin*.

Halo effect An emotional bias that extends a favorable evaluation of some characteristics to other characteristics.

Hard-bullet covered bonds Covered bonds for which a bond default is triggered and bond payments are accelerated in the event of sponsor default if payments do not occur according to the original maturity schedule.

Harmonic mean A type of weighted mean computed as the reciprocal of the arithmetic average of the reciprocals.

Head and shoulders pattern In technical analysis, a reversal pattern that is formed in three parts: a left shoulder, a head, and a right shoulder. A head and shoulders pattern is used to predict a change from an uptrend to a downtrend.

Headline inflation Refers to the inflation rate calculated based on the price index that includes all goods and services in an economy.

Heat map A type of graphic that organizes and summarizes data in a tabular format and represents it using a color spectrum.

Hedge funds Private investment vehicles that typically use leverage, derivatives, and long and short investment strategies.

Hedge portfolio A hypothetical combination of the derivative and its underlying that eliminates risk.

Held-to-maturity Debt (fixed-income) securities that a company intends to hold to maturity; these are presented at their original cost, updated for any amortisation of discounts or premiums.

Herding Clustered trading that may or may not be based on information.

Heteroskedasticity A violation of the simple linear regression assumption of homoskedasticity; with heteroskedasticity, the variability of the residuals is not equal across all observations.

Hidden order An order that is exposed not to the public but only to the brokers or exchanges that receive it.

High-frequency trading A form of algorithmic trading that makes use of vast quantities of data to execute trades on ultra-high-speed networks in fractions of a second.

High-water mark The highest value, net of fees, which a fund has reached in history. It reflects the highest cumulative return used to calculate an incentive fee.

Hindsight bias A bias with selective perception and retention aspects in which people may see past events as having been predictable and reasonable to expect.

Histogram A chart that presents the distribution of numerical data by using the height of a bar or column to represent the absolute frequency of each bin or interval in the distribution.

Historical cost In reference to assets, the amount paid to purchase an asset, including any costs of acquisition and/or preparation; with reference to liabilities, the amount of proceeds received in exchange in issuing the liability.

Historical equity risk premium approach An estimate of a country's equity risk premium that is based on the historical averages of the risk-free rate and the rate of return on the market portfolio.

Holder-of-record date The date that a shareholder listed on the corporation's books will be deemed to have ownership of the shares for purposes of receiving an upcoming dividend.

Holding period return The return that an investor earns during a specified holding period; a synonym for total return.

Home bias A preference for securities listed on the exchanges of one's home country.

Homogeneity of expectations The assumption that all investors have the same economic expectations and thus have the same expectations of prices, cash flows, and other investment characteristics.

Homoskedasticity An assumption of simple linear regression such that the variability of the residuals is equal across all observations.

Horizon yield The internal rate of return between the total return (the sum of reinvested coupon payments and the sale price or redemption amount) and the purchase price of the bond.

Horizontal analysis Common-size analysis that involves comparing a specific financial statement with that statement in prior or future time periods; also, cross-sectional analysis of one company with another.

Horizontal demand schedule Implies that at a given price, the response in the quantity demanded is infinite.

Hostile takeover An attempt by one entity to acquire a company without the consent of the company's management.

Household A person or a group of people living in the same residence, taken as a basic unit in economic analysis.

Human capital The accumulated knowledge and skill that workers acquire from education, training, or life experience and the corresponding present value of future earnings to be generated by said skilled individual.

Hurdle rate The minimum rate of return on investment that a fund must reach before a GP receives carried interest.

Hypothesis A proposed explanation or theory that can be tested.

Hypothesis testing The process of testing of hypotheses about one or more populations using statistical inference.

I-spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Iceberg order An order in which the display size is less than the order's full size.

If-converted method A method for accounting for the effect of convertible securities on earnings per share (EPS) that specifies what EPS would have been if the convertible securities had been converted at the beginning of the period, taking account of the effects of conversion on net income and the weighted average number of shares outstanding.

Illusion of control bias A bias in which people tend to believe that they can control or influence outcomes when, in fact, they cannot.

Immediate or cancel order An order that is valid only upon receipt by the broker or exchange. If such an order cannot be filled in part or in whole upon receipt, it cancels immediately. Also called *fill or kill*.

Impact investing Investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.

Impact lag The lag associated with the result of actions affecting the economy with delay.

Implicit price deflator for GDP A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

Implicit selection bias One type of selection bias introduced through the presence of a threshold that filters out some unqualified members.

Implied forward rates Calculated from spot rates, an implied forward rate is a breakeven reinvestment rate that links the return on an investment in a shorter-term zero-coupon bond to the return on an investment in a longer-term zero-coupon bond.

Implied volatility The volatility that option traders use to price an option, implied by the price of the option and a particular option-pricing model.

Import license Specifies the quantity of a good that can be imported into a country.

Imports Goods and services that a domestic economy (i.e., house-holds, firms, and government) purchases from other countries.

In the money Options that, if exercised, would result in the value received being worth more than the payment required to exercise.

Income Increases in economic benefits in the form of inflows or enhancements of assets, or decreases of liabilities that result in an increase in equity (other than increases resulting from contributions by owners).

Income elasticity of demand A measure of the responsiveness of demand to changes in income, defined as the percentage change in quantity demanded divided by the percentage change in income.

Income tax paid The actual amount paid for income taxes in the period; not a provision, but the actual cash outflow.

Income tax payable The income tax owed by the company on the basis of taxable income.

Increasing marginal returns When the marginal product of a resource increases as additional units of that input are employed.

Increasing returns to scale When a production process leads to increases in output that are proportionately larger than the increase in inputs.

Incremental cash flow The cash flow that is realized because of a decision; the changes or increments to cash flows resulting from a decision or action.

Indenture Legal contract that describes the form of a bond, the obligations of the issuer, and the rights of the bondholders. Also called the *trust deed*.

Independent With reference to events, the property that the occurrence of one event does not affect the probability of another event occurring. With reference to two random variables X and Y, they are independent if and only if $P(X,Y) = P(X)P(Y)$.

Independent projects Independent projects are capital investments whose cash flows are independent of each other.

Independent variable A variable used to explain the dependent variable in a regression; a right-side variable in a regression equation. Also referred to as the *explanatory variable*.

Independently and identically distributed With respect to random variables, the property of random variables that are independent of each other but follow the identical probability distribution.

Index-linked bond Bond for which coupon payments and/or principal repayment are linked to a specified index.

Indexing An investment strategy in which an investor constructs a portfolio to mirror the performance of a specified index.

Indicator variable A variable that takes on only one of two values, 0 or 1, based on a condition. In simple linear regression, the slope is the difference in the dependent variable for the two conditions. Also referred to as a *dummy variable*.

Indifference curve A curve representing all the combinations of two goods or attributes such that the consumer is entirely indifferent among them.

Indirect format With reference to cash flow statements, a format for the presentation of the statement which, in the operating cash flow section, begins with net income then shows additions and subtractions to arrive at operating cash flow. Also called *indirect method*.

Indirect method See *indirect format*.

Indirect taxes Taxes such as taxes on spending, as opposed to direct taxes.

Industry A group of companies offering similar products and/or services.

Industry analysis The analysis of a specific branch of manufacturing, service, or trade.

Inelastic Said of a good or service when the magnitude of elasticity is less than one. Insensitive to price changes.

Inferior goods A good whose consumption decreases as income increases.

Inflation The percentage increase in the general price level from one period to the next; a sustained rise in the overall level of prices in an economy.

Inflation-linked bond Type of index-linked bond that offers investors protection against inflation by linking the bond's coupon payments and/or the principal repayment to an index of consumer prices. Also called *linkers*.

Inflation premium An extra return that compensates investors for expected inflation.

Inflation reports A type of economic publication put out by many central banks.

Inflation uncertainty The degree to which economic agents view future rates of inflation as difficult to forecast.

Information cascade The transmission of information from those participants who act first and whose decisions influence the decisions of others.

Information-motivated traders Traders that trade to profit from information that they believe allows them to predict future prices.

Informationally efficient market A market in which asset prices reflect new information quickly and rationally.

Initial coin offering An unregulated process whereby companies raise capital by selling crypto tokens to investors in exchange for fiat money or another agreed-upon cryptocurrency.

Initial margin The amount that must be deposited in a clearinghouse account when entering into a futures contract.

Initial margin requirement The margin requirement on the first day of a transaction as well as on any day in which additional margin funds must be deposited.

Initial public offering (IPO) The first issuance of common shares to the public by a formerly private corporation.

Input productivity The amount of output produced by workers in a given period of time—for example, output per hour worked; measures the efficiency of labor.

Intangible assets Assets lacking physical substance, such as patents and trademarks.

Interbank market The market of loans and deposits between banks for maturities ranging from overnight to one year.

Intercept The expected value of the dependent variable when the independent variable in a simple linear regression is equal to zero.

Interest Payment for lending funds.

Interest coverage A solvency ratio calculated as EBIT divided by interest payments.

Interest-only mortgage A loan in which no scheduled principal repayment is specified for a certain number of years.

Interest rate A rate of return that reflects the relationship between differently dated cash flows; a discount rate.

Interest rate effect The effect through which price level changes, through demand for money, impact interest rate, which in turn impacts investment and consumption.

Interest rate swap A swap in which the underlying is an interest rate. Can be viewed as a currency swap in which both currencies are the same and can be created as a combination of currency swaps.

Intermarket analysis A field within technical analysis that combines analysis of the major categories of securities—namely, equities, bonds, currencies, and commodities—to identify market trends and possible inflections in trends.

Internal rate of return The discount rate that makes net present value equal 0; the discount rate that makes the present value of an investment's costs (outflows) equal to the present value of the investment's benefits (inflows).

Internet of Things A network arrangement of structures and devices whereby the objects on the network are able to interact and share information.

Interpolated spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Interquartile range The difference between the third and first quartiles of a dataset.

Interval With reference to grouped data, a set of values within which an observation falls.

Intrinsic value See *exercise value*.

Inventory investment Net change in business inventory.

Inventory turnover An activity ratio calculated as cost of goods sold divided by average inventory.

Inverse demand function A restatement of the demand function in which price is stated as a function of quantity.

Inverse floater A type of leveraged structured financial instrument. The cash flows are adjusted periodically and move in the opposite direction of changes in the reference rate.

Inverse transformation method A method using randomly generated numbers from the continuous uniform distribution to generate random observations from any distribution.

Investing activities Activities associated with the acquisition and disposal of property, plant, and equipment; intangible assets; other long-term assets; and both long-term and short-term investments in the equity and debt (bonds and loans) issued by other companies.

Investment banks Financial intermediaries that provide advice to their mostly corporate clients and help them arrange transactions such as initial and seasoned securities offerings.

Investment policy statement (IPS) A written planning document that describes a client's investment objectives and risk tolerance over a relevant time horizon, along with constraints that apply to the client's portfolio.

Investment property Property used to earn rental income or capital appreciation (or both).

Jackknife A resampling method that repeatedly draws samples by taking the original observed data sample and leaving out one observation at a time (without replacement) from the set.

January effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January. Also called *turn-of-the-year effect*.

Joint frequencies The entry in the cells of the contingency table that represent the joining of one variable from a row and the other variable from a column to count observations.

Joint probability The probability of the joint occurrence of stated events.

Joint probability function A function giving the probability of joint occurrences of values of stated random variables.

Judgmental Sampling A procedure of selectively handpicking elements from the population based on a researcher's knowledge and professional judgment.

Key rate duration A method of measuring the interest rate sensitivities of a fixed-income instrument or portfolio to shifts in key points along the yield curve.

Keynesians Economists who believe that fiscal policy can have powerful effects on aggregate demand, output, and employment when there is substantial spare capacity in an economy.

Kurtosis The statistical measure that indicates the combined weight of the tails of a distribution relative to the rest of the distribution.

Labor force The portion of the working age population (over the age of 16) that is employed or is available for work but not working (unemployed).

Labor productivity The quantity of goods and services (real GDP) that a worker can produce in one hour of work.

Lagging economic indicators Turning points that take place later than those of the overall economy; they are believed to have value in identifying the economy's past condition.

Law of demand The principle that as the price of a good rises, buyers will choose to buy less of it, and as its price falls, they will buy more.

Law of diminishing marginal returns The observation that a variable factor's marginal product must eventually fall as more of it is added to a fixed amount of the other factors.

Law of diminishing returns The smallest output that a firm can produce such that its long run average costs are minimized.

Law of one price The condition in a financial market in which two equivalent financial instruments or combinations of financial instruments can sell for only one price. Equivalent to the principle that no arbitrage opportunities are possible.

Lead underwriter The lead investment bank in a syndicate of investment banks and broker–dealers involved in a securities underwriting.

Leading economic indicators Turning points that usually precede those of the overall economy; they are believed to have value for predicting the economy's future state, usually near-term.

Lease A contract that conveys the right to use an asset for a period of time in exchange for consideration.

Legal tender Something that must be accepted when offered in exchange for goods and services.

Lender of last resort An entity willing to lend money when no other entity is ready to do so.

Leptokurtic Describes a distribution that has fatter tails than a normal distribution (also called fat-tailed).

Lessee The party obtaining the use of an asset through a lease.

Lessor The owner of an asset that grants the right to use the asset to another party.

Letter of credit Form of external credit enhancement whereby a financial institution provides the issuer with a credit line to reimburse any cash flow shortfalls from the assets backing the issue.

Level of significance The probability of a Type I error in testing a hypothesis.

Leverage In the context of corporate finance, leverage refers to the use of fixed costs within a company's cost structure. Fixed costs that are operating costs (such as depreciation or rent) create operating leverage. Fixed costs that are financial costs (such as interest expense) create financial leverage.

Leveraged buyout A transaction whereby the target company's management team converts the target to a privately held company by using heavy borrowing to finance the purchase of the target company's outstanding shares.

Liabilities Present obligations of an enterprise arising from past events, the settlement of which is expected to result in an outflow of resources embodying economic benefits; creditors' claims on the resources of a company.

LIFO layer liquidation With respect to the application of the LIFO inventory method, the liquidation of old, relatively low-priced inventory; happens when the volume of sales rises above the volume of recent purchases so that some sales are made from relatively old, low-priced inventory. Also called *LIFO liquidation*.

LIFO method The last in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the reverse order the items were placed in inventory (i.e., inventory produced or acquired last are assumed to be sold first).

LIFO reserve The difference between the reported LIFO inventory carrying amount and the inventory amount that would have been reported if the FIFO method had been used (in other words, the FIFO inventory value less the LIFO inventory value).

Likelihood The probability of an observation, given a particular set of conditions.

Limit down A limit move in the futures market in which the price at which a transaction would be made is at or below the lower limit.

Limit order Instructions to a broker or exchange to obtain the best price immediately available when filling an order, but in no event accept a price higher than a specified (limit) price when buying or accept a price lower than a specified (limit) price when selling.

Limit order book The book or list of limit orders to buy and sell that pertains to a security.

Limit up A limit move in the futures market in which the price at which a transaction would be made is at or above the upper limit.

Limitations on liens Meant to put limits on how much secured debt an issuer can have.

Limited partners Partners with limited liability. Limited partnerships in hedge and private equity funds are typically restricted to investors who are expected to understand and to be able to assume the risks associated with the investments.

Limited partnership agreement (LPA) A legal document that outlines the rules of the partnership and establishes the framework that ultimately guides the fund's operations throughout its life.

Lin-log model A regression model in which the independent variable is in logarithmic form.

Line chart A type of graph used to visualize ordered observations. In technical analysis, a plot of price data, typically closing prices, with a line connecting the points.

Linear interpolation The estimation of an unknown value on the basis of two known values that bracket it, using a straight line between the two known values.

Linear regression Regression that models the straight-line relationship between the dependent and independent variables. Also known as *least squares regression* and *ordinary least squares regression*.

Linear scale A scale in which equal distances correspond to equal absolute amounts. Also called an *arithmetic scale*.

Linker See *inflation-linked bond*.

Liquid market Said of a market in which traders can buy or sell with low total transaction costs when they want to trade.

Liquidation To sell the assets of a company, division, or subsidiary piecemeal, typically because of bankruptcy; the form of bankruptcy that allows for the orderly satisfaction of creditors' claims after which the company ceases to exist.

Liquidity The extent to which a company is able to meet its short-term obligations using cash flows and those assets that can be readily transformed into cash.

Liquidity premium An extra return that compensates investors for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

Liquidity ratios Financial ratios measuring the company's ability to meet its short-term obligations to creditors.

Liquidity risk The risk that a financial instrument cannot be purchased or sold without a significant concession in price due to the size of the market.

Liquidity trap A condition in which the demand for money becomes infinitely elastic (horizontal demand curve) so that injections of money into the economy will not lower interest rates or affect real activity.

Load fund A mutual fund in which, in addition to the annual fee, a percentage fee is charged to invest in the fund and/or for redemptions from the fund.

Loan-to-value ratio The amount of the property's mortgage divided by the purchase price

Locked limit A condition in the futures markets in which a transaction cannot take place because the price would be beyond the limits.

Lockup period The minimum holding period before investors are allowed to make withdrawals or redeem shares from a fund.

Log-lin model A regression model in which the dependent variable is in logarithmic form.

Log-log model A regression model in which both the dependent and independent variables are in logarithmic form. Also known as the *double-log model*.

Logarithmic scale A scale in which equal distances represent equal proportional changes in the underlying quantity.

London interbank offered rate (Libor) Collective name for multiple rates at which a select set of banks believes they could borrow unsecured funds from other banks in the London interbank market for different currencies and different borrowing periods ranging from overnight to one year.

Long The buyer of a derivative contract. Also refers to the position of owning a derivative.

Long-lived assets Assets that are expected to provide economic benefits over a future period of time, typically greater than one year. Also called *long-term assets*.

Long position A position in an asset or contract in which one owns the asset or has an exercisable right under the contract.

Long-run average total cost The curve describing average total cost when no costs are considered fixed.

Look-ahead bias A bias caused by using information that was unavailable on the test date.

Loss aversion The tendency of people to dislike losses more than they like comparable gains.

Loss-aversion bias A bias in which people tend to strongly prefer avoiding losses as opposed to achieving gains.

Loss severity Portion of a bond's value (including unpaid interest) an investor loses in the event of default.

Losses Asset outflows not directly related to the ordinary activities of the business.

Lower bound The lowest possible value of an option.

M^2 A measure of what a portfolio would have returned if it had taken on the same total risk as the market index.

M^2 alpha Difference between the risk-adjusted performance of the portfolio and the performance of the benchmark.

Macaulay duration The approximate amount of time a bond would have to be held for the market discount rate at purchase to be realized if there is a single change in interest rate. It indicates the point in time when the coupon reinvestment and price effects of a change in yield-to-maturity offset each other.

Machine learning Computer based techniques that seek to extract knowledge from large amounts of data by "learning" from known examples and then generating structure or predictions. ML algorithms aim to "find the pattern, apply the pattern."

Macroeconomics The branch of economics that deals with aggregate economic quantities, such as national output and national income.

Maintenance covenants Covenants in bank loan agreements that require the borrower to satisfy certain financial ratio tests while the loan is outstanding.

Maintenance margin The minimum amount that is required by a futures clearinghouse to maintain a margin account and to protect against default. Participants whose margin balances drop below the required maintenance margin must replenish their accounts.

Maintenance margin requirement The margin requirement on any day other than the first day of a transaction.

Management buy-in Leveraged buyout in which the current management team is being replaced and the acquiring team will be involved in managing the company.

Management buyout A leveraged buyout event in which a group of investors consisting primarily of the company's existing management purchases at least controlling interest in its outstanding shares. At the extreme, they may purchase all shares and take the company private.

Management fee A fee based on assets under management or committed capital, as applicable—also called a *base fee*.

MAR ratio A variation of the Calmar ratio that uses a full investment history and the average drawdown.

Margin The amount of money that a trader deposits in a margin account. The term is derived from the stock market practice in which an investor borrows a portion of the money required to purchase a certain amount of stock. In futures markets, there is no borrowing so the margin is more of a down payment or performance bond.

Margin bond A cash deposit required by the clearinghouse from the participants to a contract to provide a credit guarantee. Also called a *performance bond*.

Margin call A request for the short to deposit additional funds to bring their balance up to the initial margin.

Margin loan Money borrowed from a broker to purchase securities.

Marginal cost The cost of producing an additional unit of a good.

Marginal frequencies The sums determined by adding joint frequencies across rows or across columns in a contingency table.

Marginal product Measures the productivity of each unit of input and is calculated by taking the difference in total product from adding another unit of input (assuming other resource quantities are held constant).

Marginal propensity to consume The proportion of an additional unit of disposable income that is consumed or spent; the change in consumption for a small change in income.

Marginal propensity to save The proportion of an additional unit of disposable income that is saved (not spent).

Marginal revenue The change in total revenue divided by the change in quantity sold; simply, the additional revenue from selling one more unit.

Marginal value curve A curve describing the highest price consumers are willing to pay for each additional unit of a good.

Mark to market The revaluation of a financial asset or liability to its current market value or fair value.

Mark-to-market Refers to the current expected fair market value for which a given security would likely be available for purchase or sale if traded in current market conditions.

Market anomaly Change in the price or return of a security that cannot directly be linked to current relevant information known in the market or to the release of new information into the market.

Market bid–ask spread The difference between the best bid and the best offer.

Market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its market capitalization by the total market capitalization (sum of the market capitalization) of all securities in the index. Also called *value weighting*.

Market discount rate The rate of return required by investors given the risk of the investment in a bond; also called the *required yield* or the *required rate of return*.

Market float The number of shares that are available to the investing public.

Market liquidity risk The risk that the price at which investors can actually transact—buying or selling—may differ from the price indicated in the market.

Market model A regression equation that specifies a linear relationship between the return on a security (or portfolio) and the return on a broad market index.

Market multiple models Valuation models based on share price multiples or enterprise value multiples.

Market-on-close An execution instruction specifying that an order can only be filled at the close of trading.

Market order Instructions to a broker or exchange to obtain the best price immediately available when filling an order.

Market-oriented investors With reference to equity investors, investors whose investment disciplines cannot be clearly categorized as value or growth.

Market rate of interest The rate demanded by purchasers of bonds, given the risks associated with future cash payment obligations of the particular bond issue.

Market reference rate (MRR) Collective name for a set of rates covering different currencies for different maturities, ranging from overnight to one year.

Market risk The risk that arises from movements in interest rates, stock prices, exchange rates, and commodity prices.

Market value The price at which an asset or security can currently be bought or sold in an open market.

Marketable limit order A buy limit order in which the limit price is placed above the best offer, or a sell limit order in which the limit price is placed below the best bid. Such orders generally will partially or completely fill right away.

Markowitz efficient frontier The graph of the set of portfolios offering the maximum expected return for their level of risk (standard deviation of return).

Matching principle The accounting principle that expenses should be recognized in the same period in which the associated revenue is recognized.

Matrix pricing Process of estimating the market discount rate and price of a bond based on the quoted or flat prices of more frequently traded comparable bonds.

Maturity premium An extra return that compensates investors for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended.

Maturity structure A factor explaining the differences in yields on similar bonds; also called *term structure*.

Mean absolute deviation With reference to a sample, the mean of the absolute values of deviations from the sample mean.

Mean square error (MSE) The sum of squares error divided by the degrees of freedom, $n - k - 1$; in a simple linear regression, $n - k - 1 = n - 2$.

Mean square regression (MSR) The sum of squares regression divided by the number of independent variables k; in a simple linear regression, $k = 1$.

Mean–variance analysis An approach to portfolio analysis using expected means, variances, and covariances of asset returns.

Measure of central tendency A quantitative measure that specifies where data are centered.

Measure of value A standard for measuring value; a function of money.

Measures of location Quantitative measures that describe the location or distribution of data. They include not only measures of central tendency but also other measures, such as percentiles.

Median The value of the middle item of a set of items that has been sorted into ascending or descending order (i.e., the 50th percentile).

Medium of exchange Any asset that can be used to purchase goods and services or to repay debts; a function of money.

Medium-term note A corporate bond offered continuously to investors by an agent of the issuer, designed to fill the funding gap between commercial paper and long-term bonds.

Mental accounting bias An information-processing bias in which people treat one sum of money differently from another equal-sized sum based on which mental account the money is assigned to.

Menu costs A cost of inflation in which businesses constantly have to incur the costs of changing the advertised prices of their goods and services.

Mesokurtic Describes a distribution with kurtosis equal to that of the normal distribution, namely, kurtosis equal to three.

Microeconomics The branch of economics that deals with markets and decision making of individual economic units, including consumers and businesses.

Minimum efficient scale The smallest output that a firm can produce such that its long-run average total cost is minimized.

Minimum-variance portfolio The portfolio with the minimum variance for each given level of expected return.

Minority shareholders Particular shareholders or a block of shareholders holding a small proportion of a company's outstanding shares, resulting in a limited ability to exercise control in voting activities.

Minsky moment Named for Hyman Minksy. A point in a business cycle when, after individuals become overextended in borrowing to finance speculative investments, people start realizing that something is likely to go wrong and a panic ensues, leading to asset sell-offs.

Modal interval With reference to grouped data, the interval containing the greatest number of observations (i.e., highest frequency).

Mode The most frequently occurring value in a distribution.

Modern portfolio theory (MPT) The analysis of rational portfolio choices based on the efficient use of risk.

Modified duration A measure of the percentage price change of a bond given a change in its yield-to-maturity.

Momentum oscillator A graphical representation of market sentiment that is constructed from price data and calculated so that it oscillates either between a low and a high or around some number.

Monetarists Economists who believe that the rate of growth of the money supply is the primary determinant of the rate of inflation.

Monetary policy Actions taken by a nation's central bank to affect aggregate output and prices through changes in bank reserves, reserve requirements, or its target interest rate.

Monetary transmission mechanism The process whereby a central bank's interest rate gets transmitted through the economy and ultimately affects the rate of increase of prices.

Monetary union An economic union in which the members adopt a common currency.

Money A generally accepted medium of exchange and unit of account.

Money convexity For a bond, the annual or approximate convexity multiplied by the full price.

Money creation The process by which changes in bank reserves translate into changes in the money supply.

Money duration A measure of the price change in units of the currency in which the bond is denominated given a change in its yield-to-maturity.

Money market The market for short-term debt instruments (one-year maturity or less).

Money market securities Fixed-income securities with maturities at issuance of one year or less.

Money multiplier Describes how a change in reserves is expected to affect the money supply; in its simplest form, 1 divided by the reserve requirement.

Money neutrality The thesis that an increase in the money supply leads in the long-run to an increase in the price level, while leaving real variables like output and employment unaffected.

Money-weighted return The internal rate of return on a portfolio, taking account of all cash flows.

Moneyness The relationship between the price of the underlying and an option's exercise price.

Monopolistic competition Highly competitive form of imperfect competition; the competitive characteristic is a notably large number of firms, while the monopoly aspect is the result of product differentiation.

Monopoly In pure monopoly markets, there are no substitutes for the given product or service. There is a single seller, which exercises considerable power over pricing and output decisions.

Monte Carlo simulation A technique involving the use of computer software to represent the operation of a complex financial system. A characteristic feature is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Moral principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Mortgage-backed securities Debt obligations that represent claims to the cash flows from pools of mortgage loans, most commonly on residential property.

Mortgage loan A loan secured by the collateral of some specified real estate property that obliges the borrower to make a predetermined series of payments to the lender.

Mortgage pass-through security A security created when one or more holders of mortgages form a pool of mortgages and sell shares or participation certificates in the pool.

Mortgage rate The interest rate on a mortgage loan; also called *contract rate* or *note rate*.

Moving average The average of the closing price of a security over a specified number of periods. With each new period, the average is recalculated.

Moving-average convergence/divergence oscillator A momentum oscillator that is based on the difference between short-term and long-term moving averages of a security's price.

Multi-factor model A model that explains a variable in terms of the values of a set of factors.

Multi-market indexes Comprised of indexes from different countries, designed to represent multiple security markets.

Multi-step format With respect to the format of the income statement, a format that presents a subtotal for gross profit (revenue minus cost of goods sold).

Multilateral trading facilities See *alternative trading systems*.

Multinational corporation A company operating in more than one country or having subsidiary firms in more than one country.

Multinomial formula (general formula for labeling problems) The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by $\dfrac{n!}{n_1!n_2!\ldots n_k!}$.

Multiple of invested capital (MOIC) A simplified calculation that measures the total value of all distributions and residual asset values relative to an initial total investment; also known as a *money multiple*.

Multiple testing problem The risk of getting statistically significant test results when performing a test multiple times.

Multiplication rule for counting If one task can be done in n_1 ways, and a second task, given the first, can be done in n_2 ways, and a third task, given the first two tasks, can be done in n_3 ways, and so on for k tasks, then the number of ways the k tasks can be done is $(n_1)(n_2)(n_3) \ldots (n_k)$.

Multiplication rule for expected value of the product of uncorrelated random variables This rule is $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated.

Multiplication rule for independent events The rule that when two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B.

Multiplication rule for probability The rule that the joint probability of events A and B equals the probability of A given B times the probability of B.

Multiplier models Valuation models based on share price multiples or enterprise value multiples.

Multivariate distribution A probability distribution that specifies the probabilities for a group of related random variables.

Multivariate normal distribution A probability distribution for a group of random variables that is completely defined by the means and variances of the variables plus all the correlations between pairs of the variables.

Muni See *municipal bond*.

Municipal bond A type of non-sovereign bond issued by a state or local government in the United States. It very often (but not always) offers income tax exemptions.

Mutual fund A comingled investment pool in which investors in the fund each have a pro-rata claim on the income and value of the fund.

Mutually exclusive Indicates that only one event can occur at a time.

Mutually exclusive projects Mutually exclusive projects compete directly with each other. For example, if Projects A and B are mutually exclusive, you can choose A or B but you cannot choose both.

***n* Factorial** For a positive integer n, the product of the first n positive integers; 0 factorial equals 1 by definition. n factorial is written as $n!$.

Narrow money The notes and coins in circulation in an economy, plus other very highly liquid deposits.

Nash equilibrium When two or more participants in a non-coop-erative game have no incentive to deviate from their respective equilibrium strategies given their opponent's strategies.

National income The income received by all factors of production used in the generation of final output. National income equals gross domestic product (or, in some countries, gross national product) minus the capital consumption allowance and a statistical discrepancy.

Natural language processing Computer programs developed to analyze and interpret human language.

Natural rate of unemployment Effective unemployment rate, below which pressure emerges in labor markets.

Negative screening An ESG investment style that focuses on the exclusion of certain sectors, companies, or practices in a fund or portfolio on the basis of specific ESG criteria.

Net book value The remaining (undepreciated) balance of an asset's purchase cost. For liabilities, the face value of a bond minus any unamortized discount, or plus any unamortized premium.

Net exports The difference between the value of a country's exports and the value of its imports (i.e., value of exports minus imports).

Net income The difference between revenue and expenses; what remains after subtracting all expenses (including depreciation, interest, and taxes) from revenue.

Net present value The present value of an investment's cash inflows (benefits) minus the present value of its cash outflows (costs).

Net profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also called *profit margin* or *return on sales*.

Net realisable value Estimated selling price in the ordinary course of business less the estimated costs necessary to make the sale.

Net revenue Revenue after adjustments (e.g., for estimated returns or for amounts unlikely to be collected).

Net tax rate The tax rate net of transfer payments.

Neural networks Computer programs based on how our own brains learn and process information.

Neutral rate of interest The rate of interest that neither spurs on nor slows down the underlying economy.

No-load fund A mutual fund in which there is no fee for investing in the fund or for redeeming fund shares, although there is an annual fee based on a percentage of the fund's net asset value.

Node Each value on a binomial tree from which successive moves or outcomes branch.

Nominal data Categorical values that are not amenable to being organized in a logical order. An example of nominal data is the classification of publicly listed stocks into sectors.

Nominal GDP The value of goods and services measured at current prices.

Nominal risk-free interest rate The sum of the real risk-free interest rate and the inflation premium.

Non-accelerating inflation rate of unemployment Effective unemployment rate below which pressure emerges in labor markets.

Non-agency RMBS In the United States, securities issued by private entities that are not guaranteed by a federal agency or a GSE.

Non-Bank lender Unlike typical banks, which make loans and take deposits, these lenders only make loans. Although they lend money to companies, they usually specialize in issuing mortgages to first-time homebuyers and parties looking to refinance an existing mortgage.

Non-cumulative preference shares Preference shares for which dividends that are not paid in the current or subsequent periods are forfeited permanently (instead of being accrued and paid at a later date).

Non-current assets Assets that are expected to benefit the company over an extended period of time (usually more than one year).

Non-current liabilities Obligations that broadly represent a probable sacrifice of economic benefits in periods generally greater than one year in the future.

Non-cyclical A company whose performance is largely independent of the business cycle. Also known as defensive.

Non-deliverable forwards Cash-settled forward contracts, used predominately with respect to foreign exchange forwards. Also called *contracts for differences*.

Non-financial risks Risks that arise from sources other than changes in the external financial markets, such as changes in accounting rules, legal environment, or tax rates.

Non-participating preference shares Preference shares that do not entitle shareholders to share in the profits of the company. Instead, shareholders are only entitled to receive a fixed dividend payment and the par value of the shares in the event of liquidation.

Non-probability sampling A sampling plan dependent on factors other than probability considerations, such as a sampler's judgment or the convenience to access data.

Non-recourse loan A loan in which the lender does not have a shortfall claim against the borrower, so the lender can look only to the property to recover the outstanding mortgage balance.

Non-renewable resources Finite resources that are depleted once they are consumed, such as oil and coal.

Non-sovereign bond A bond issued by a government below the national level, such as a province, region, state, or city.

Nonconventional cash flows In a nonconventional cash flow pattern, the initial outflow is not followed by inflows only, but the cash flows can flip from positive (inflows) to negative (outflows) again or even change signs several times.

Nonparametric test A test that is not concerned with a parameter or that makes minimal assumptions about the population from which a sample comes.

Nonsystematic risk Unique risk that is local or limited to a particular asset or industry that need not affect assets outside of that asset class.

Normal distribution A continuous, symmetric probability distribution that is completely described by its mean and its variance.

Normal goods Goods that are consumed in greater quantities as income increases.

Normal profit The level of accounting profit needed to just cover the implicit opportunity costs ignored in accounting costs.

Notching Ratings adjustment methodology where specific issues from the same borrower may be assigned different credit ratings.

Note rate See *mortgage rate*.

Notice period The length of time (typically 30 to 90 days) in advance that investors may be required to notify a fund of their intent to redeem some or all of their investment.

Notional principal An imputed principal amount.

Null hypothesis The hypothesis that is tested.

Numerical data Values that represent measured or counted quantities as a number. Also called **quantitative data**.

Objective probabilities Probabilities that generally do not vary from person to person; includes a priori and empirical probabilities.

Observation The value of a specific variable collected at a point in time or over a specified period of time.

Odds against *E* The reciprocal of odds for *E*.

Odds for *E* The probability of *E* divided by 1 minus the probability of *E*.

Off-the-run Seasoned government bonds are off-the-run securities; they are not the most recently issued or the most actively traded.

Offer The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size).

Official interest rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks. Also called *official policy rate* or *policy rate*.

Official policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Oligopoly Market structure with a relatively small number of firms supplying the market.

On-the-run The most recently issued and most actively traded sovereign securities.

One-dimensional array The simplest format for representing a collection of data of the same data type.

One-sided hypothesis test A test in which the null hypothesis is rejected only if the evidence indicates that the population parameter is greater than or less than the hypothesized parameter; occurs when the alternative hypothesis is stated either as greater than or less than the hypothesized population parameter.

Open economy An economy that trades with other countries.

Open-end fund A mutual fund that accepts new investment money and issues additional shares at a value equal to the net asset value of the fund at the time of investment.

Open interest The number of outstanding contracts in a clearinghouse at any given time. The open interest figure changes daily as some parties open up new positions, while other parties offset their old positions.

Open market operations The purchase or sale of bonds by the national central bank to implement monetary policy. The bonds traded are usually sovereign bonds issued by the national government.

Operating activities Activities that are part of the day-to-day business functioning of an entity, such as selling inventory and providing services.

Operating breakeven The number of units produced and sold at which the company's operating profit is zero (revenues = operating costs).

Operating cash flow The net amount of cash provided from operating activities.

Operating efficiency ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Operating lease A type of lease which is more akin to the rental of the underlying asset.

Operating leverage The use of fixed costs in operations.

Operating profit A company's profits on its usual business activities before deducting taxes. Also called *operating income*.

Operating profit margin A profitability ratio calculated as operating income (i.e., income before interest and taxes) divided by revenue. Also called *operating margin*.

Operating risk The risk attributed to the operating cost structure, in particular the use of fixed costs in operations; the risk arising from the mix of fixed and variable costs; the risk that a company's operations may be severely affected by environmental, social, and governance risk factors.

Operational independence A bank's ability to execute monetary policy and set interest rates in the way it thought would best meet the inflation target.

Operational risk The risk that arises from inadequate or failed people, systems, and internal policies, procedures, and processes, as well as from external events that are beyond the control of the organization but that affect its operations.

Operationally efficient Said of a market, a financial system, or an economy that has relatively low transaction costs.

Opportunity cost The value that investors forgo by choosing a particular course of action; the value of something in its best alternative use.

Optimal capital structure The capital structure at which the value of the company is maximized.

Option A financial instrument that gives one party the right, but not the obligation, to buy or sell an underlying asset from or to another party at a fixed price over a specific period of time. Also referred to as *contingent claim* or *option contract.*

Option-adjusted price The value of the embedded option plus the flat price of the bond.

Option-adjusted spread OAS = Z-spread – Option value (in basis points per year).

Option-adjusted yield The required market discount rate whereby the price is adjusted for the value of the embedded option.

Option contract See *option*.

Option premium The amount of money a buyer pays and seller receives to engage in an option transaction.

Order A specification of what instrument to trade, how much to trade, and whether to buy or sell.

Order-driven markets A market (generally an auction market) that uses rules to arrange trades based on the orders that traders submit; in their pure form, such markets do not make use of dealers.

Order precedence hierarchy With respect to the execution of orders to trade, a set of rules that determines which orders execute before other orders.

Ordinal data Categorical values that can be logically ordered or ranked.

Ordinary annuity An annuity with a first cash flow that is paid one period from the present.

Ordinary shares Equity shares that are subordinate to all other types of equity (e.g., preferred equity). Also called *common stock* or *common shares.*

Organized exchange A securities marketplace where buyers and seller can meet to arrange their trades.

Other comprehensive income Items of comprehensive income that are not reported on the income statement; comprehensive income minus net income.

Out-of-sample test A test of a strategy or model using a sample outside the period on which the strategy or model was developed.

Out of the money Options that, if exercised, would require the payment of more money than the value received and therefore would not be currently exercised.

Outcome A possible value of a random variable.

Over-the-counter (OTC) market A decentralized market where buy and sell orders initiated from various locations are matched through a communications network.

Overbought A market condition in which market sentiment is thought to be unsustainably bullish.

Overcollateralization Form of internal credit enhancement that refers to the process of posting more collateral than needed to obtain or secure financing.

Overconfidence bias A bias in which people demonstrate unwarranted faith in their own intuitive reasoning, judgments, and/or cognitive abilities.

Overfitting An undesirable result from fitting a model so closely to a dataset that it does not perform well on new data.

Oversold A market condition in which market sentiment is thought to be unsustainably bearish.

Own price The price of a good or service itself (as opposed to the price of something else).

Own-price elasticity of demand The percentage change in quantity demanded for a percentage change in good's own price, holding all other things constant.

Owners' equity The excess of assets over liabilities; the residual interest of shareholders in the assets of an entity after deducting the entity's liabilities. Also called *shareholders' equity* or *shareholders' funds*.

***p*-Value** The smallest level of significance at which the null is rejected.

Paired comparisons test See *test of the mean of the differences*.

Panel data A mix of time-series and cross-sectional data that contains observations through time on characteristics of across multiple observational units.

Par curve A sequence of yields-to-maturity such that each bond is priced at par value. The bonds are assumed to have the same currency, credit risk, liquidity, tax status, and annual yields stated for the same periodicity.

Par value The amount of principal on a bond.

Parallel shift A parallel yield curve shift implies that all rates change by the same amount in the same direction.

Parameter A descriptive measure computed from or used to describe a population of data, conventionally represented by Greek letters.

Parametric test Any test (or procedure) concerned with parameters or whose validity depends on assumptions concerning the population generating the sample.

Pari passu Covenant or contract clause that ensures a debt obligation is treated the same as the borrower's other senior debt instruments and is not subordinated to similar obligations.

Partial duration See *key rate duration*.

Participating preference shares Preference shares that entitle shareholders to receive the standard preferred dividend plus the opportunity to receive an additional dividend if the company's profits exceed a pre-specified level.

Pass-through rate The coupon rate of a mortgage pass-through security.

Passive investment A buy and hold approach in which an investor does not make portfolio changes based on short-term expectations of changing market or security performance.

Payable date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payment date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payments system The system for the transfer of money.

Pearson correlation A parametric measure of the relationship between two variables.

Pecking order theory The theory that managers consider how their actions might be interpreted by outsiders and thus order their preferences for various forms of corporate financing. Forms of financing that are least visible to outsiders (e.g., internally generated funds) are most preferable to managers, and those that are most visible (e.g., equity) are least preferable.

Peer company See *comparable company*.

Peer group A group of companies engaged in similar business activities whose economics and valuation are influenced by closely related factors.

Pennant A technical analysis continuation pattern formed by trendlines that converge to form a triangle, typically over a short period.

Per capita real GDP Real GDP divided by the size of the population, often used as a measure of a country's average standard of living.

Per unit contribution margin The amount that each unit sold contributes to covering fixed costs—that is, the difference between the price per unit and the variable cost per unit.

Percentiles Quantiles that divide a distribution into 100 equal parts that sum to 100.

Perfect capital markets Markets in which, by assumption, there are no taxes, transaction costs, or bankruptcy costs and in which all investors have equal ("symmetric") information.

Perfect competition A market structure in which the individual firm has virtually no impact on market price, because it is assumed to be a very small seller among a very large number of firms selling essentially identical products.

Perfectly elastic When the quantity demanded or supplied of a given good is infinitely sensitive to a change in the value of a specified variable (e.g., price).

Perfectly inelastic When the quantity demanded or supplied of a given good is completely insensitive to a change in the value of a specified variable (e.g., price).

Performance bond See *margin bond*.

Performance evaluation The measurement and assessment of the outcomes of investment management decisions.

Performance fee Fees paid to the general partner from the limited partner(s) based on realized net profits.

Period costs Costs (e.g., executives' salaries) that cannot be directly matched with the timing of revenues and which are thus expensed immediately.

Periodicity The assumed number of periods in the year; typically matches the frequency of coupon payments.

Permanent differences Differences between tax and financial reporting of revenue (expenses) that will not be reversed at some future date. These result in a difference between the company's effective tax rate and statutory tax rate and do not result in a deferred tax item.

Permissioned networks Networks that are fully open only to select participants on a DLT network.

Permissionless networks Networks that are fully open to any user on a DLT network.

Permutation An ordered listing.

Permutation formula The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is ${}_nP_r = \dfrac{n!}{(n-r)!}$.

Perpetual bonds Bonds with no stated maturity date.

Perpetuity A perpetual annuity, or a set of never-ending level sequential cash flows, with the first cash flow occurring one period from now. A bond that does not mature.

Personal income A broad measure of household income that includes all income received by households, whether earned or unearned; measures the ability of consumers to make purchases.

Pet projects Investments in which influential managers want the corporation to invest. Often, unfortunately, pet projects are selected without undergoing normal capital allocation analysis.

Plain vanilla bond Bond that makes periodic, fixed coupon payments during the bond's life and a lump-sum payment of principal at maturity. Also called *conventional bond*.

Platykurtic Describes a distribution that has relatively less weight in the tails than the normal distribution (also called thin-tailed).

Point estimate A single numerical estimate of an unknown quantity, such as a population parameter.

Policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Population All members of a specified group.

Portfolio company In private equity, the company in which the private equity fund is investing.

Portfolio demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments.

Portfolio planning The process of creating a plan for building a portfolio that is expected to satisfy a client's investment objectives.

Position The quantity of an asset that an entity owns or owes.

Positive screening The process of including sectors or companies based on specific ESG criteria, typically ESG performance relative to industry peers

Posterior probability An updated probability that reflects or comes after new information.

Potential GDP The level of real GDP that can be produced at full employment; measures the productive capacity of the economy.

Power of a test The probability of correctly rejecting the null—that is, rejecting the null hypothesis when it is false.

Precautionary money balances Money held to provide a buffer against unforeseen events that might require money.

Preference shares A type of equity interest which ranks above common shares with respect to the payment of dividends and the distribution of the company's net assets upon liquidation. They have characteristics of both debt and equity securities. Also called *preferred stock*.

Preferred stock See *preference shares*.

Premium In the case of bonds, premium refers to the amount by which a bond is priced above its face (par) value. In the case of an option, the amount paid for the option contract.

Prepaid expense A normal operating expense that has been paid in advance of when it is due.

Prepayment option Contractual provision that entitles the borrower to prepay all or part of the outstanding mortgage principal prior to the scheduled due date when the principal must be repaid. Also called *early repayment option*.

Prepayment penalty mortgages Mortgages that stipulate a monetary penalty if a borrower prepays within a certain time period after the mortgage is originated.

Prepayment risk The uncertainty that the timing of the actual cash flows will be different from the scheduled cash flows as set forth in the loan agreement due to the borrowers' ability to alter payments, usually to take advantage of interest rate movements.

Present value (PV) The present discounted value of future cash flows: For assets, the present discounted value of the future net cash inflows that the asset is expected to generate; for liabilities, the present discounted value of the future net cash outflows that are expected to be required to settle the liabilities.

Present value models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security. Also called *discounted cash flow models*.

Pretax margin A profitability ratio calculated as earnings before taxes divided by revenue.

Price elasticity of demand Measures the percentage change in the quantity demanded, given a percentage change in the price of a given product.

Price index Represents the average prices of a basket of goods and services.

Price limits Limits imposed by a futures exchange on the price change that can occur from one day to the next.

Price multiple A ratio that compares the share price with some sort of monetary flow or value to allow evaluation of the relative worth of a company's stock.

Price priority The principle that the highest priced buy orders and the lowest priced sell orders execute first.

Price relative A ratio of an ending price over a beginning price; it is equal to 1 plus the holding period return on the asset.

Price return Measures *only* the price appreciation or percentage change in price of the securities in an index or portfolio.

Price return index An index that reflects *only* the price appreciation or percentage change in price of the constituent securities. Also called *price index*.

Price-setting option The option to adjust prices when demand varies from what is forecast.

Price stability In economics, refers to an inflation rate that is low on average and not subject to wide fluctuation.

Price takers Producers that must accept whatever price the market dictates.

Price to book value A valuation ratio calculated as price per share divided by book value per share.

Price to cash flow A valuation ratio calculated as price per share divided by cash flow per share.

Price to earnings ratio (P/E ratio or P/E) The ratio of share price to earnings per share.

Price to sales A valuation ratio calculated as price per share divided by sales per share.

Price value of a basis point A version of money duration, it is an estimate of the change in the full price of a bond given a 1 basis point change in the yield-to-maturity.

Price weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its price by the sum of all the prices of the constituent securities.

Priced risk Risk for which investors demand compensation for bearing (e.g., equity risk, company-specific factors, macroeconomic factors).

Primary bond market A market in which issuers first sell bonds to investors to raise capital.

Primary capital markets (primary markets) The market where securities are first sold and the issuers receive the proceeds.

Primary dealer Financial institution that is authorized to deal in new issues of sovereign bonds and that serves primarily as a trading counterparty of the office responsible for issuing sovereign bonds.

Primary market The market where securities are first sold and the issuers receive the proceeds.

Prime brokers Brokers that provide services that commonly include custody, administration, lending, short borrowing, and trading.

Principal The amount of funds originally invested in a project or instrument; the face value to be paid at maturity.

Principal–agent relationship A relationship in which a principal hires an agent to perform a particular task or service; also known as an *agency relationship*.

Principal amount Amount that an issuer agrees to repay the debtholders on the maturity date.

Principal business activity The business activity from which a company derives a majority of its revenues and/or earnings.

Principal value Amount that an issuer agrees to repay the debtholders on the maturity date.

Principle of no arbitrage See *arbitrage-free pricing*.

Prior probabilities Probabilities reflecting beliefs prior to the arrival of new information.

Priority of claims Priority of payment, with the most senior or highest ranking debt having the first claim on the cash flows and assets of the issuer.

Private equity fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *venture capital fund*.

Private equity funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *venture capital funds*.

Private equity securities Securities that are not listed on public exchanges and have no active secondary market. They are issued primarily to institutional investors via non-public offerings, such as private placements.

Private investment in public equity (PIPE) An investment in the equity of a publicly traded firm that is made at a discount to the market value of the firm's shares.

Private placement Typically, a non-underwritten, unregistered offering of securities that are sold only to an investor or a small group of investors. It can be accomplished directly between the issuer and the investor(s) or through an investment bank.

Probability A number between 0 and 1 describing the chance that a stated event will occur.

Probability density function A function with non-negative values such that probability can be described by areas under the curve graphing the function.

Probability distribution A distribution that specifies the probabilities of a random variable's possible outcomes.

Probability function A function that specifies the probability that the random variable takes on a specific value.

Probability sampling A sampling plan that allows every member of the population to have an equal chance of being selected.

Probability tree diagram A diagram with branches emanating from nodes representing either mutually exclusive chance events or mutually exclusive decisions.

Production-flexibility options The options to alter production when demand varies from what is forecast.

Production function Provides the quantitative link between the levels of output that the economy can produce and the inputs used in the production process.

Productivity The amount of output produced by workers during a given period—for example, output per hour worked measures the efficiency of labor.

Profession An occupational group that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition.

Profit The return that owners of a company receive for the use of their capital and the assumption of financial risk when making their investments.

Profit and loss (P&L) statement A financial statement that provides information about a company's profitability over a stated period of time. Also called the *income statement*.

Profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses.

Profitability ratios Ratios that measure a company's ability to generate profitable sales from its resources (assets).

Project sequencing To defer the decision to invest in a future project until the outcome of some or all of a current investment is known. Investments are sequenced over time, so that making an investment creates the option to invest in future projects.

Promissory note A written promise to pay a certain amount of money on demand.

Property, plant, and equipment Tangible assets that are expected to be used for more than one period in either the production or supply of goods or services, or for administrative purposes.

Prospectus The document that describes the terms of a new bond issue and helps investors perform their analysis on the issue.

Protective put An option strategy in which a long position in an asset is combined with a long position in a put.

Proxy contest Corporate takeover mechanism in which shareholders are persuaded to vote for a group seeking a controlling position on a company's board of directors.

Proxy voting A process that enables shareholders who are unable to attend a meeting to authorize another individual to vote on their behalf.

Public offer See *public offering*.

Public offering An offering of securities in which any member of the public may buy the securities. Also called *public offer*.

Public–private partnership (PPP) An agreement between the public sector and the private sector to finance, build, and operate public infrastructure, such as hospitals and toll roads.

Pull on liquidity When disbursements are paid too quickly or trade credit availability is limited, requiring companies to expend funds before they receive funds from sales that could cover the liability.

Pure discount bonds See *zero-coupon bond*.

Put An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Put–call–forward parity The relationship among puts, calls, and forward contracts.

Put–call parity An equation expressing the equivalence (parity) of a portfolio of a call and a bond with a portfolio of a put and the underlying, which leads to the relationship between put and call prices.

Put/call ratio A technical analysis indicator that evaluates market sentiment based on the volume of put options traded divided by the volume of call options traded for a particular financial instrument.

Put option An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Putable bonds Bonds that give the bondholder the right to sell the bond back to the issuer at a predetermined price on specified dates.

Qualitative data see **categorical data**.

Quantile A value at or below which a stated fraction of the data lies. Also referred to as a fractile.

Quantitative data see **numerical data**.

Quantitative easing An expansionary monetary policy based on aggressive open market purchase operations.

Quantity equation of exchange An expression that over a given period, the amount of money used to purchase all goods and services in an economy, $M \times V$, is equal to monetary value of this output, $P \times Y$.

Quantity theory of money Asserts that total spending (in money terms) is proportional to the quantity of money.

Quartiles Quantiles that divide a distribution into four equal parts.

Quasi-fixed cost A cost that stays the same over a range of production but can change to another constant level when production moves outside of that range.

Quasi-government bond A bond issued by an entity that is either owned or sponsored by a national government. Also called *agency bond*.

Quintiles Quantiles that divide a distribution into five equal parts.

Quota rents Profits that foreign producers can earn by raising the price of their goods higher than they would without a quota.

Quotas Government policies that restrict the quantity of a good that can be imported into a country, generally for a specified period of time.

Quote-driven market A market in which dealers acting as principals facilitate trading.

Quoted interest rate A quoted interest rate that does not account for compounding within the year. Also called *stated annual interest rate*.

Quoted margin The specified yield spread over the reference rate, used to compensate an investor for the difference in the credit risk of the issuer and that implied by the reference rate.

Random number An observation drawn from a uniform distribution.

Random number generator An algorithm that produces uniformly distributed random numbers between 0 and 1.

Random variable A quantity whose future outcomes are uncertain.

Range The difference between the maximum and minimum values in a dataset.

Raw data Data available in their original form as collected.

Real exchange rate effect The effect through which changing price level impacts real exchange rate which in turn impacts net exports and aggregate demand.

Real GDP The value of goods and services produced, measured at base year prices.

Real income Income adjusted for the effect of inflation on the purchasing power of money. Also known as the *purchasing power of income*. If income remains constant and a good's price falls, real income is said to rise, even though the number of monetary units (e.g., dollars) remains unchanged.

Real interest rate Nominal interest rate minus the expected rate of inflation.

Real risk-free interest rate The single-period interest rate for a completely risk-free security if no inflation were expected.

Realizable (settlement) value With reference to assets, the amount of cash or cash equivalents that could currently be obtained by selling the asset in an orderly disposal; with reference to liabilities, the undiscounted amount of cash or cash equivalents expected to be paid to satisfy the liabilities in the normal course of business.

Rebalancing Adjusting the weights of the constituent securities in an index.

Rebalancing policy The set of rules that guide the process of restoring a portfolio's asset class weights to those specified in the strategic asset allocation.

Recession A period during which real GDP decreases (i.e., negative growth) for at least two successive quarters, or a period of significant decline in total output, income, employment, and sales usually lasting from six months to a year.

Recognition lag The lag in government response to an economic problem resulting from the delay in confirming a change in the state of the economy.

Recourse loan A loan in which the lender has a claim against the borrower for any shortfall between the outstanding mortgage balance and the proceeds received from the sale of the property.

Redemption yield See *yield-to-maturity*.

Redemptions Withdrawals of funds by investors, as allowed by the notice period and other terms in the partnership agreement.

Refinancing rate A type of central bank policy rate.

Registered bonds Bonds for which ownership is recorded by either name or serial number.

Regression analysis A tool for examining whether a variable is useful for explaining another variable.

Regression coefficients The intercept and slope coefficient of a regression.

Regret The feeling that an opportunity has been missed; typically, an expression of *hindsight bias*.

Regret-aversion bias An emotional bias in which people tend to avoid making decisions that will result in action out of fear that the decision will turn out poorly.

Relative dispersion The amount of dispersion relative to a reference value or benchmark.

Relative frequency The absolute frequency of each unique value of the variable divided by the total number of observations of the variable.

Relative price The price of a specific good or service in comparison with those of other goods and services.

Relative strength analysis A comparison of the performance of one asset with the performance of another asset or a benchmark, based on changes in the ratio of the two assets' prices over time.

Relative strength index (RSI) A technical analysis momentum oscillator that compares a security's gains with its losses over a set period.

Renewable resources Resources that can be replenished, such as a forest.

Rent Payment for the use of property.

Reorganization Agreements made by a company in bankruptcy under which a company's capital structure is altered and/or alternative arrangements are made for debt repayment; US Chapter 11 bankruptcy. The company emerges from bankruptcy as a going concern.

Replication The creation of an asset or portfolio from another asset, portfolio, and/or derivative.

Repo A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy back the same security from the purchaser at an agreed-on price and future date. The party who sells the

security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repo margin The difference between the market value of the security used as collateral and the value of the loan. Also called *haircut*.

Repo rate The interest rate on a repurchase agreement.

Representativeness bias A belief perseverance bias in which people tend to classify new information based on past experiences and classifications.

Repurchase agreement A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy the same security back from the purchaser at an agreed-on price and future date. The party who sells the security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repurchase date The date when the party who sold the security at the inception of a repurchase agreement buys back the security from the cash lending counterparty.

Repurchase price The price at which the party who sold the security at the inception of the repurchase agreement buys back the security from the cash lending counterparty.

Required margin The yield spread over or under the reference rate such that an FRN is priced at par value on a rate reset date.

Required rate of return See *market discount rate*.

Required yield See *market discount rate*.

Required yield spread The difference between the yield-to-maturity on a new bond and the benchmark rate; additional compensation required by investors for the difference in risk and tax status of a bond relative to a government bond. Sometimes called the *spread over the benchmark*.

Resampling A statistical method that repeatedly draws samples from the original observed data sample for the statistical inference of population parameters.

Reserve accounts Form of internal credit enhancement that relies on creating accounts and depositing in these accounts cash that can be used to absorb losses. Also called *reserve funds*.

Reserve funds See *reserve accounts*.

Reserve requirement The requirement for banks to hold reserves in proportion to the size of deposits.

Residual The difference between an observation and its predicted value, where the predicted value is based on the estimated linear relation between the dependent and independent variables using sample data.

Resistance In technical analysis, a price range in which selling activity is sufficient to stop the rise in the price of a security.

Responsible investing An overall (umbrella) term for several investment strategies—ESG investing, SRI investing, thematic investing, and impact investing—that incorporate ESG analysis in their investment processes.

Restricted payments A bond covenant meant to protect creditors by limiting how much cash can be paid out to shareholders over time.

Retracement In technical analysis, a reversal in the movement of a security's price such that it is counter to the prevailing longer-term price trend.

Return-generating model A model that can provide an estimate of the expected return of a security given certain parameters and estimates of the values of the independent variables in the model.

Return on assets (ROA) A profitability ratio calculated as net income divided by average total assets; indicates a company's net profit generated per dollar invested in total assets.

Return on equity (ROE) A profitability ratio calculated as net income divided by average shareholders' equity.

Return on invested capital A measure of the profitability of a company relative to the amount of capital invested by the equity- and debtholders.

Return on sales An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also referred to as *net profit margin*.

Return on total capital A profitability ratio calculated as EBIT divided by the sum of short- and long-term debt and equity.

Revaluation model Under IFRS, the process of valuing long-lived assets at fair value, rather than at cost less accumulated depreciation. Any resulting profit or loss is either reported on the income statement and/or through equity under revaluation surplus.

Revenue The amount charged for the delivery of goods or services in the ordinary activities of a business over a stated period; the inflows of economic resources to a company over a stated period.

Reversal pattern A type of pattern used in technical analysis to predict the end of a trend and a change in the direction of a security's price.

Reverse repo A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse repurchase agreement A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse stock split A reduction in the number of shares outstanding with a corresponding increase in share price, but no change to the company's underlying fundamentals.

Revolving credit agreements (also known as revolvers) The most reliable form of short-term bank borrowing facilities; they are in effect for multiple years (e.g., three to five years) and can have optional medium-term loan features.

Rho The sensitivity of the option price to the risk-free rate.

Ricardian equivalence An economic theory that implies that it makes no difference whether a government finances a deficit by increasing taxes or issuing debt.

Risk Exposure to uncertainty. The chance of a loss or adverse outcome as a result of an action, inaction, or external event.

Risk averse The assumption that an investor will choose the least risky alternative.

Risk aversion The degree of an investor's inability and unwillingness to take risk.

Risk budgeting The establishment of objectives for individuals, groups, or divisions of an organization that takes into account the allocation of an acceptable level of risk.

Risk exposure The state of being exposed or vulnerable to a risk. The extent to which an organization is sensitive to underlying risks.

Risk governance The top-down process and guidance that directs risk management activities to align with and support the overall enterprise.

Risk management The process of identifying the level of risk an organization wants, measuring the level of risk the organization currently has, taking actions that bring the actual

level of risk to the desired level of risk, and monitoring the new actual level of risk so that it continues to be aligned with the desired level of risk.

Risk management framework The infrastructure, process, and analytics needed to support effective risk management in an organization.

Risk-neutral pricing Sometimes said of derivatives pricing, uses the fact that arbitrage opportunities guarantee that a risk-free portfolio consisting of the underlying and the derivative must earn the risk-free rate.

Risk-neutral probabilities Weights that are used to compute a binomial option price. They are the probabilities that would apply if a risk-neutral investor valued an option.

Risk premium An extra return expected by investors for bearing some specified risk.

Risk shifting Actions to change the distribution of risk outcomes.

Risk tolerance The amount of risk an investor is willing and able to bear to achieve an investment goal.

Risk transfer Actions to pass on a risk to another party, often, but not always, in the form of an insurance policy.

Robo-adviser A machine-based analytical tool or service that provides technology-driven investment solutions through online platforms.

Rule of 72 The principle that the approximate number of years necessary for an investment to double is 72 divided by the stated interest rate.

Running yield See *current yield.*

Safety-first rules Rules for portfolio selection that focus on the risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Sales Generally, a synonym for revenue; “sales” is generally understood to refer to the sale of goods, whereas “revenue” is understood to include the sale of goods or services.

Sales risk Uncertainty with respect to the quantity of goods and services that a company is able to sell and the price it is able to achieve; the risk related to the uncertainty of revenues.

Sample A subset of a population.

Sample correlation coefficient A standardized measure of how two variables in a sample move together. It is the ratio of the sample covariance to the product of the two variables’ standard deviations.

Sample covariance A measure of how two variables in a sample move together.

Sample excess kurtosis A sample measure of the degree of a distribution’s kurtosis in excess of the normal distribution’s kurtosis.

Sample mean The sum of the sample observations divided by the sample size.

Sample selection bias Bias introduced by systematically excluding some members of the population according to a particular attribute—for example, the bias introduced when data availability leads to certain observations being excluded from the analysis.

Sample-size neglect A type of representativeness bias in which financial market participants incorrectly assume that small sample sizes are representative of populations (or “real” data).

Sample skewness A sample measure of the degree of asymmetry of a distribution.

Sample standard deviation The positive square root of the sample variance.

Sample statistic A quantity computed from or used to describe a sample.

Sample variance The sum of squared deviations around the mean divided by the degrees of freedom.

Sampling The process of obtaining a sample.

Sampling distribution The distribution of all distinct possible values that a statistic can assume when computed from samples of the same size randomly drawn from the same population.

Sampling error The difference between the observed value of a statistic and the estimate resulting from using subsets of the population.

Sampling plan The set of rules used to select a sample.

Say on pay A process whereby shareholders may vote on executive remuneration (compensation) matters.

Scatter plot A chart in which two variables are plotted along the axis and points on the chart represent pairs of the two variables. In regression, the dependent variable is plotted on the vertical axis and the independent variable is plotted along the horizontal axis. Also known as a scattergram and a *scatter diagram.*

Scatter plot matrix A tool for organizing scatter plots between pairs of variables, making it easy to inspect all pairwise relationships in one combined visual.

Scenario analysis Analysis that shows the changes in key financial quantities that result from given (economic) events, such as the loss of customers, the loss of a supply source, or a catastrophic event; a risk management technique involving examination of the performance of a portfolio under specified situations. Closely related to stress testing.

Screening The application of a set of criteria to reduce a set of potential investments to a smaller set having certain desired characteristics.

Seasoned offering An offering in which an issuer sells additional units of a previously issued security.

Second-degree price discrimination When the monopolist charges different per-unit prices using the quantity purchased as an indicator of how highly the customer values the product.

Second lien A secured interest in the pledged assets that ranks below first lien debt in both collateral protection and priority of payment.

Secondary bond markets Markets in which existing bonds are traded among investors.

Secondary market The market where securities are traded among investors.

Secondary precedence rules Rules that determine how to rank orders placed at the same time.

Sector A group of related industries (GICS definition).

Sector indexes Indexes that represent and track different economic sectors—such as consumer goods, energy, finance, health care, and technology—on either a national, regional, or global basis.

Secured (“asset-based”) loan A loan that is backed by specific, secured company assets.

Secured bonds Bonds secured by assets or financial guarantees pledged to ensure debt repayment in case of default.

Secured debt Debt in which the debtholder has a direct claim—a pledge from the issuer—on certain assets and their associated cash flows.

Securitization A process that involves moving assets into a special legal entity, which then uses the assets as guarantees to secure a bond issue.

Securitized assets Assets that are typically used to create asset-backed bonds; for example, when a bank securitizes a pool of loans, the loans are said to be securitized.

Security characteristic line A plot of the excess return of a security on the excess return of the market.

Security market index A portfolio of securities representing a given security market, market segment, or asset class.

Security market line (SML) The graph of the capital asset pricing model.

Security selection The process of selecting individual securities; typically, security selection has the objective of generating superior risk-adjusted returns relative to a portfolio's benchmark.

Self-attribution bias A bias in which people take too much credit for successes (*self-enhancing*) and assign responsibility to others for failures (*self-protecting*).

Self-control bias A bias in which people fail to act in pursuit of their long-term, overarching goals because of a lack of self-discipline.

Self-investment limits With respect to investment limitations applying to pension plans, restrictions on the percentage of assets that can be invested in securities issued by the pension plan sponsor.

Sell-side firm A broker/dealer that sells securities and provides independent investment research and recommendations to their clients (i.e., buy-side firms).

Semi-strong-form efficient market A market in which security prices reflect all publicly known and available information.

Semiannual bond basis yield An annual rate having a periodicity of two; also known as a *semiannual bond equivalent yield*.

Semiannual bond equivalent yield See *semiannual bond basis yield*.

Seniority ranking Priority of payment of various debt obligations.

Sensitivity analysis Analysis that shows the range of possible outcomes as specific assumptions are changed.

Separately managed account (SMA) An investment portfolio managed exclusively for the benefit of an individual or institution.

Serial maturity structure Structure for a bond issue in which the maturity dates are spread out during the bond's life; a stated number of bonds mature and are paid off each year before final maturity.

Settlement The process that occurs after a trade is completed, the securities are passed to the buyer, and payment is received by the seller.

Settlement date Date when the buyer makes cash payment and the seller delivers the security.

Settlement price The official price, designated by the clearinghouse, from which daily gains and losses will be determined and marked to market.

Share repurchase A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Shareholder activism Strategies used by shareholders to attempt to compel a company to act in a desired manner.

Shareholder engagement The process whereby companies engage with their shareholders.

Shareholders' equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Sharpe ratio The average return in excess of the risk-free rate divided by the standard deviation of return; a measure of the average excess return earned per unit of standard deviation of return.

Shelf registration A type of public offering that allows the issuer to file a single, all-encompassing offering circular that covers a series of bond issues.

Short The seller of an asset or derivative contract. Also refers to the position of being short an asset or derivative contract.

Short position A position in an asset or contract in which one has sold an asset one does not own, or in which a right under a contract can be exercised against oneself.

Short-run average total cost The curve describing average total cost when some costs are considered fixed.

Short selling A transaction in which borrowed securities are sold with the intention to repurchase them at a lower price at a later date and return them to the lender.

Shortfall risk The risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Shutdown point The point at which average revenue is equal to the firm's average variable cost.

Side letters Side agreements created between the GP and specific LPs. These agreements exist *outside* the LPA. These agreements provide additional terms and conditions related to the investment agreement.

Simple interest The interest earned each period on the original investment; interest calculated on the principal only.

Simple linear regression (SLR) A regression that summarizes the relation between the dependent variable and a single independent variable.

Simple random sample A subset of a larger population created in such a way that each element of the population has an equal probability of being selected to the subset.

Simple random sampling The procedure of drawing a sample to satisfy the definition of a simple random sample.

Simple yield The sum of the coupon payments plus the straight-line amortized share of the gain or loss, divided by the flat price.

Simulation Computer-generated sensitivity or scenario analysis that is based on probability models for the factors that drive outcomes.

Simulation trial A complete pass through the steps of a simulation.

Single-step format With respect to the format of the income statement, a format that does not subtotal for gross profit (revenue minus cost of goods sold).

Sinking fund arrangement Provision that reduces the credit risk of a bond issue by requiring the issuer to retire a portion of the bond's principal outstanding each year.

Situational influences External factors, such as environmental or cultural elements, that shape our behavior.

Skewed Not symmetrical.

Skewness A quantitative measure of skew (lack of symmetry); a synonym of skew. It is computed as the average cubed deviation from the mean standardized by dividing by the standard deviation cubed.

Slope coefficient The coefficient of an independent variable that represents the average change in the dependent variable for a one-unit change in the independent variable.

Small country A country that is a price taker in the world market for a product and cannot influence the world market price.

Smart beta Involves the use of simple, transparent, rules-based strategies as a basis for investment decisions.

Smart contract A computer program that is designed to self-execute on the basis of pre-specified terms and conditions agreed to by parties to a contract.

Socially responsible investing An investment approach that excludes investments in companies or industries that deviate from an organization's beliefs and sometimes includes investments with favorable environmental or social profiles.

Soft-bullet covered bonds Covered bonds for which bond default and payment acceleration of bond cash flows may be delayed upon sponsor default until a new final maturity date is reached.

Solvency With respect to financial statement analysis, the ability of a company to fulfill its long-term obligations.

Solvency ratios Ratios that measure a company's ability to meet its long-term obligations.

Solvency risk The risk that an organization does not survive or succeed because it runs out of cash, even though it might otherwise be solvent.

Sovereign bond A bond issued by a national government. Also called "Sovereign."

Spearman rank correlation coefficient A measure of correlation applied to ranked data.

Special dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Special purpose entity A non-operating entity created to carry out a specified purpose, such as leasing assets or securitizing receivables; can be a corporation, partnership, trust, or limited liability partnership formed to facilitate a specific type of business activity. Also called *special purpose vehicle*, *special purpose company*, or *variable interest entity*.

Special purpose vehicle See *special purpose entity*.

Specific identification method An inventory accounting method that identifies which specific inventory items were sold and which remained in inventory to be carried over to later periods.

Speculative demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments. Also called *portfolio demand for money*.

Speculative money balances Monies held in anticipation that other assets will decline in value.

Split coupon bond See *deferred coupon bond*.

Sponsored A type of depository receipt in which the foreign company whose shares are held by the depository has a direct involvement in the issuance of the receipts.

Spot curve A sequence of yields-to-maturity on zero-coupon bonds. Sometimes called *zero* or *strip curve* (because coupon payments are "stripped" off the bonds).

Spot markets Markets in which assets are traded for immediate delivery.

Spot prices The price of an asset for immediately delivery.

Spot rates A sequence of market discount rates that correspond to the cash flow dates; yields-to-maturity on zero-coupon bonds maturing at the date of each cash flow.

Spread In general, the difference in yield between different fixed-income securities. Often used to refer to the difference between the yield-to-maturity and the benchmark.

Spread over the benchmark See *required yield spread*.

Spread risk Bond price risk arising from changes in the yield spread on credit-risky bonds; reflects changes in the market's assessment and/or pricing of credit migration (or downgrade) risk and market liquidity risk.

Spurious correlation Refers to: 1) correlation between two variables that reflects chance relationships in a particular dataset; 2) correlation induced by a calculation that mixes each of two variables with a third variable; and 3) correlation between two variables arising not from a direct relation between them but from their relation to a third variable.

Stacked bar chart An alternative form for presenting the frequency distribution of two categorical variables, where bars representing the sub-groups are placed on top of each other to form a single bar. Each sub-section is shown in a different color to represent the contribution of each sub-group, and the overall height of the stacked bar represents the marginal frequency for the category.

Stackelberg model A prominent model of strategic decision making in which firms are assumed to make their decisions sequentially.

Stagflation The combination of a high inflation rate with a high level of unemployment and a slowdown of the economy.

Staggered boards Election process whereby directors are typically divided into multiple classes that are elected separately in consecutive years—that is, one class every year.

Stakeholder management The identification, prioritization, and understanding of the interests of stakeholder groups and managing the company's relationships with these groups.

Stakeholders Individuals or groups of individuals who may be affected either directly or indirectly by a decision and thus have an interest, or stake, in the decision.

Standard deviation The positive square root of the variance; a measure of dispersion in the same units as the original data.

Standard error of the estimate A measure of the fit of a regression line, calculated as the square root of the mean square error. Also known as the *standard error of the regression* and the *root mean square error*.

Standard error of the forecast A measure of the uncertainty associated with a forecasted value of the dependent variable that depends on the standard error of the estimate, the variability of the independent variable, the deviation of the forecasted independent variable from the mean in the regression, and the number of observations.

Standard error of the slope coefficient The standard error of the slope, which in a simple linear regression is the ratio of the model's standard error of the estimate (s_e) to the square root of the variation of the independent variable.

Standard normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Standardizing A transformation that involves subtracting the mean and dividing the result by the standard deviation.

Standards of conduct Behaviors required by a group; established benchmarks that clarify or enhance a group's code of ethics.

Standing limit orders A limit order at a price below market and which therefore is waiting to trade.

Stated annual interest rate A quoted interest rate that does not account for compounding within the year. Also called *quoted interest rate*.

Statement of changes in equity (statement of owners' equity) A financial statement that reconciles the beginning-of-period and end-of-period balance sheet values of

shareholders' equity; provides information about all factors affecting shareholders' equity. Also called *statement of owners' equity*.

Statement of financial condition The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of financial position The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of operations A financial statement that provides information about a company's profitability over a stated period of time.

Static trade-off theory of capital structure A theory pertaining to a company's optimal capital structure. The optimal level of debt is found at the point where additional debt would cause the costs of financial distress to increase by a greater amount than the benefit of the additional tax shield.

Statistic A summary measure of a sample of observations.

Statistically significant A result indicating that the null hypothesis can be rejected; with reference to an estimated regression coefficient, frequently understood to mean a result indicating that the corresponding population regression coefficient is different from zero.

Status quo bias An emotional bias in which people do nothing (i.e., maintain the status quo) instead of making a change.

Statutory voting A common method of voting where each share represents one vote.

Step-up coupon bond Bond for which the coupon, fixed or floating, increases by specified margins at specified dates.

Stochastic oscillator A momentum indicator that compares a particular closing price of a security to a range of the security's prices over a certain period of time.

Stock dividend A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Stock split An increase in the number of shares outstanding with a consequent decrease in share price, but no change to the company's underlying fundamentals.

Stop-loss order See *stop order*.

Stop order An order in which a trader has specified a stop price condition. Also called *stop-loss order*.

Store of value The quality of tending to preserve value.

Store of wealth Goods that depend on the fact that they do not perish physically over time, and on the belief that others would always value the good.

Straight-line method A depreciation method that allocates evenly the cost of a long-lived asset less its estimated residual value over the estimated useful life of the asset.

Straight voting A shareholder voting process in which shareholders receive one vote for each share owned.

Strategic analysis Analysis of the competitive environment with an emphasis on the implications of the environment for corporate strategy.

Strategic asset allocation The set of exposures to IPS-permissible asset classes that is expected to achieve the client's long-term objectives given the client's investment constraints.

Stratified random sampling A procedure that first divides a population into subpopulations (strata) based on classification criteria and then randomly draws samples from each stratum in sizes proportional to that of each stratum in the population.

Street convention A yield measure that neglects weekends and holidays; the internal rate of return on cash flows assuming payments are made on the scheduled dates, even when the scheduled date falls on a weekend or holiday.

Stress testing A specific type of scenario analysis that estimates losses in rare and extremely unfavorable combinations of events or scenarios.

Strong-form efficient market A market in which security prices reflect all public and private information.

Structural (or cyclically adjusted) budget deficit The deficit that would exist if the economy was at full employment (or full potential output).

Structural subordination Arises in a holding company structure when the debt of operating subsidiaries is serviced by the cash flow and assets of the subsidiaries before funds can be passed to the holding company to service debt at the parent level.

Structured data Data that are highly organized in a predefined manner, usually with repeating patterns.

Structured financial instrument A financial instrument that shares the common attribute of repackaging risks. Structured financial instruments include asset-backed securities, collateralized debt obligations, and other structured financial instruments such as capital protected, yield enhancement, participation, and leveraged instruments.

Subjective probability A probability drawing on personal or subjective judgment.

Subordinated debt A class of unsecured debt that ranks below a firm's senior unsecured obligations.

Subordination Form of internal credit enhancement that relies on creating more than one bond tranche and ordering the claim priorities for ownership or interest in an asset between the tranches. The ordering of the claim priorities is called a senior/subordinated structure, where the tranches of highest seniority are called senior followed by subordinated or junior tranches. Also called *credit tranching*.

Substitutes Said of two goods or services such that if the price of one increases the demand for the other tends to increase, holding all other things equal (e.g., butter and margarine).

Sum of squares error (SSE) The sum of the squared deviations of (1) the value of the dependent variable and (2) the value of the dependent variable based on the estimated regression line. Also referred to as the *residual sum of squares*.

Sum of squares regression (SSR) The sum of the squared deviations of (1) the value of the dependent variable based on the estimated regression line and (2) the mean of the dependent variable.

Sum of squares total (SST) The sum of the squared deviations of the dependent variable from its mean; the variation of the dependent variable. Also referred to as the *total sum of squares*.

Sunk cost A cost that has already been incurred.

Supervised learning A machine learning approach that makes use of labeled training data.

Supply shock A typically unexpected disturbance to supply.

Support In technical analysis, a price range in which buying activity is sufficient to stop the decline in the price of a security.

Support tranches Classes or tranches in CMOs that protect PAC tranches from prepayment risk.

Supranational bond A bond issued by a supranational agency such as the World Bank.

Surety bond Form of external credit enhancement whereby a rated and regulated insurance company guarantees to reimburse bondholders for any losses incurred up to a maximum amount if the issuer defaults.

Survey approach An estimate of the equity risk premium that is based on estimates provided by a panel of finance experts.

Survivorship bias Relates to the inclusion of only current investment funds in a database. As such, the returns of funds that are no longer available in the marketplace (have been liquidated) are excluded from the database. In addition, *backfill bias* is another problem, whereby certain surviving hedge funds may be added to databases and various hedge fund indexes only after they are initially successful and start to report their returns.

Sustainability linked loans These are any types of loan instruments and/or contingent facilities (such as bonding lines, guarantee lines, or letters of credit) that incentivize the borrower's achievement of ambitious, pre-determined sustainability performance objectives.

Sustainable growth rate The rate of dividend (and earnings) growth that can be sustained over time for a given level of return on equity, keeping the capital structure constant and without issuing additional common stock.

Sustainable investing A term used in a similar context to responsible investing, but its key focus is on factoring in sustainability issues while investing.

Sustainable rate of economic growth The rate of increase in the economy's productive capacity or potential GDP.

Swap contract An agreement between two parties to exchange a series of future cash flows.

Syndicated loan A loan from a group of lenders to a single borrower.

Syndicated offering A bond issue underwritten by a group of investment banks.

Systematic risk Risk that affects the entire market or economy; it cannot be avoided and is inherent in the overall market. Systematic risk is also known as non-diversifiable or market risk.

Systematic sampling A procedure of selecting every *k*th member until reaching a sample of the desired size. The sample that results from this procedure should be approximately random.

Tactical asset allocation The decision to deliberately deviate from the strategic asset allocation in an attempt to add value based on forecasts of the near-term relative performance of asset classes.

Tactical asset allocation (TAA) A portfolio strategy that shifts the percentages of assets held in various asset classes (or categories) to take advantage of market opportunities. Allocation shifts can occur within an asset class or across asset classes.

Tag cloud see **word cloud**.

Target capital structure A company's chosen proportions of debt and equity.

Target independent A bank's ability to determine the definition of inflation that they target, the rate of inflation that they target, and the horizon over which the target is to be achieved.

Target semideviation A measure of downside risk, calculated as the square root of the average of the squared deviations of observations below the target (also called target downside deviation).

Tariffs Taxes that a government levies on imported goods.

Tax base The amount at which an asset or liability is valued for tax purposes.

Tax expense An aggregate of an entity's income tax payable (or recoverable in the case of a tax benefit) and any changes in deferred tax assets and liabilities. It is essentially the income tax payable or recoverable if these had been determined based on accounting profit rather than taxable income.

Tax loss carry forward A taxable loss in the current period that may be used to reduce future taxable income.

Taxable income The portion of an entity's income that is subject to income taxes under the tax laws of its jurisdiction.

Taxable temporary differences Temporary differences that result in a taxable amount in a future period when determining the taxable profit as the balance sheet item is recovered or settled.

Technical analysis A form of security analysis that uses price and volume data, often displayed graphically, in decision making.

Technology The process a company uses to transform inputs into outputs.

Tender offer Corporate takeover mechanism that involves shareholders selling their interests directly to the group seeking to gain control.

Tenor The time-to-maturity for a bond or derivative contract. Also called *term to maturity*.

Term maturity structure Structure for a bond issue in which the bond's notional principal is paid off in a lump sum at maturity.

Term structure See *maturity structure*.

Term structure of credit spreads The relationship between the spreads over the "risk-free" (or benchmark) rates and times-to-maturity.

Term structure of yield volatility The relationship between the volatility of bond yields-to-maturity and times-to-maturity.

Terminal stock value The expected value of a share at the end of the investment horizon—in effect, the expected selling price. Also called *terminal value*.

Terminal value The expected value of a share at the end of the investment horizon—in effect, the expected selling price.

Terms of trade The ratio of the price of exports to the price of imports, representing those prices by export and import price indexes, respectively.

Test of the mean of the differences A statistical test for differences based on paired observations drawn from samples that are dependent on each other.

Text analytics The use of computer programs to analyze and derive meaning from typically large, unstructured text- or voice-based datasets.

Thematic investing An ESG investment approach that focuses on investing in themes or assets specifically relating to ESG factors, such as clean energy, green technology, or sustainable agriculture.

Thin-Tailed Describes a distribution that has relatively less weight in the tails than the normal distribution (also called platykurtic).

Third-degree price discrimination When the monopolist segregates customers into groups based on demographic or other characteristics and offers different pricing to each group.

Time-period bias The possibility that when we use a time-series sample, our statistical conclusion may be sensitive to the starting and ending dates of the sample.

Time-series data A sequence of observations for a single observational unit of a specific variable collected over time and at discrete and typically equally spaced intervals of time (such as daily, weekly, monthly, annually, or quarterly).

Time tranching The creation of classes or tranches in an ABS/MBS that possess different (expected) maturities.

Time value The difference between the market price of the option and its intrinsic value.

Time value decay Said of an option when, at expiration, no time value remains and the option is worth only its exercise value.

Time value of money The principles governing equivalence relationships between cash flows with different dates.

Time-weighted rate of return The compound rate of growth of one unit of currency invested in a portfolio during a stated measurement period; a measure of investment performance that is not sensitive to the timing and amount of withdrawals or additions to the portfolio.

Tokenization The process of representing ownership rights to physical assets on a blockchain or distributed ledger.

Top-down analysis An investment selection approach that begins with consideration of macroeconomic conditions and then evaluates markets and industries based upon such conditions.

Total comprehensive income The change in equity during a period resulting from transaction and other events, other than those changes resulting from transactions with owners in their capacity as owners.

Total cost The summation of all costs, for which costs are classified as fixed or variable.

Total factor productivity A scale factor that reflects the portion of growth unaccounted for by explicit factor inputs (e.g., capital and labor).

Total fixed cost The summation of all expenses that do not change as the level of production varies.

Total invested capital The sum of market value of common equity, book value of preferred equity, and face value of debt.

Total probability rule A rule explaining the unconditional probability of an event in terms of probabilities of the event conditional on mutually exclusive and exhaustive scenarios.

Total probability rule for expected value A rule explaining the expected value of a random variable in terms of expected values of the random variable conditional on mutually exclusive and exhaustive scenarios.

Total return Measures the price appreciation, or percentage change in price of the securities in an index or portfolio, plus any income received over the period.

Total return index An index that reflects the price appreciation or percentage change in price of the constituent securities plus any income received since inception.

Total return swap A swap in which one party agrees to pay the total return on a security. Often used as a credit derivative, in which the underlying is a bond.

Total variable cost The summation of all variable expenses.

Tracking error The standard deviation of the differences between a portfolio's returns and its benchmark's returns; a synonym of active risk.

Tracking risk The standard deviation of the differences between a portfolio's returns and its benchmarks returns. Also called *tracking error*.

Trade creation When regional integration results in the replacement of higher cost domestic production by lower cost imports from other members.

Trade credit A spontaneous form of credit in which a purchaser of the goods or service is financing its purchase by delaying the date on which payment is made.

Trade diversion When regional integration results in lower-cost imports from non-member countries being replaced with higher-cost imports from members.

Trade payables Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Trade protection Government policies that impose restrictions on trade, such as tariffs and quotas.

Trade surplus (deficit) When the value of exports is greater (less) than the value of imports.

Trading securities Under US GAAP, a category of debt securities held by a company with the intent to trade them. Also called *held-for-trading securities*.

Traditional investment markets Markets for traditional investments, which include all publicly traded debts and equities and shares in pooled investment vehicles that hold publicly traded debts and/or equities.

Transactions money balances Money balances that are held to finance transactions.

Transfer payments Welfare payments made through the social security system that exist to provide a basic minimum level of income for low-income households.

Transparency Said of something (e.g., a market) in which information is fully disclosed to the public and/or regulators.

Treasury stock method A method for accounting for the effect of options (and warrants) on earnings per share (EPS) that specifies what EPS would have been if the options and warrants had been exercised and the company had used the proceeds to repurchase common stock.

Tree-Map Another graphical tool for displaying categorical data. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group.

Trend A long-term pattern of movement in a particular direction.

Treynor ratio A measure of risk-adjusted performance that relates a portfolio's excess returns to the portfolio's beta.

Triangle pattern In technical analysis, a continuation chart pattern that forms as the range between high and low prices narrows, visually forming a triangle.

Trimmed mean A mean computed after excluding a stated small percentage of the lowest and highest observations.

Trimodal A distribution that has the three most frequently occurring values.

Triple bottom In technical analysis, a reversal pattern that results when the price forms three troughs at roughly the same price level. A triple bottom is used to predict a change from a downtrend to an uptrend.

Triple top In technical analysis, a reversal pattern that results when the price forms three peaks at roughly the same price level. A triple top is used to predict a change from an uptrend to a downtrend.

True yield The internal rate of return on cash flows using the actual calendar, including weekends and bank holidays.

Trust deed The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *bond indenture*.

Turn-of-the-year effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January.

Two-dimensional rectangular array A popular form for organizing data for processing by computers or for presenting data visually. It is comprised of columns and rows to hold multiple variables and multiple observations, respectively (also called a **data table**).

Two-fund separation theorem The theory that all investors regardless of taste, risk preferences, and initial wealth will hold a combination of two portfolios or funds: a risk-free asset and an optimal portfolio of risky assets.

Two-sided hypothesis test A test in which the null hypothesis is rejected in favor of the alternative hypothesis if the evidence indicates that the population parameter is either smaller or larger than a hypothesized value; occurs when the alternative hypothesis is stated as not equal to the hypothesized population parameters.

Two-way table See *contingency table*.

Two-week repo rate The interest rate on a two-week repurchase agreement; may be used as a policy rate by a central bank.

Type I error The error of rejecting a true null hypothesis; a false positive.

Type II error The error of not rejecting a false null hypothesis; false negative.

Unanticipated (unexpected) inflation The component of inflation that is a surprise.

Uncommitted lines of credit The least reliable form of bank borrowing in which a bank offers, without formal commitment, a line of credit for an extended period of time but reserves the right to refuse any request for its use.

Unconditional probability The probability of an event *not* conditioned on another event.

Underemployed A person who has a job but has the qualifications to work a significantly higher-paying job.

Underlying An asset that trades in a market in which buyers and sellers meet, decide on a price, and the seller then delivers the asset to the buyer and receives payment. The underlying is the asset or other derivative on which a particular derivative is based. The market for the underlying is also referred to as the *spot market*.

Underwriter A firm, usually an investment bank, that takes the risk of buying the newly issued securities from the issuer and then reselling them to investors or to dealers, thus guaranteeing the sale of the securities at the offering price negotiated with the issuer.

Underwritten offering A type of securities issue mechanism in which the investment bank guarantees the sale of the securities at an offering price that is negotiated with the issuer. Also known as *firm commitment offering*.

Unearned revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service. Also called *deferred revenue* or *deferred income*.

Unemployed People who are actively seeking employment but are currently without a job.

Unemployment rate The ratio of unemployed to the labor force.

Unexpected inflation The component of inflation that is a surprise.

Unimodal A distribution with a single value that is most frequently occurring.

Unit elastic An elasticity with a magnitude of negative one. Also called *unitary elastic*.

Unit labor cost The average labor cost to produce one unit of output.

Unit normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Unitranche debt Consists of a hybrid or blended loan structure that combines different tranches of secured and unsecured debt into a single loan with a single, blended interest rate.

Units-of-production method A depreciation method that allocates the cost of a long-lived asset based on actual usage during the period.

Univariate distribution A distribution that specifies the probabilities for a single random variable.

Universal owners Long-term investors, such as pension funds, that have significant assets invested in globally diversified portfolios.

Unlimited funds An unlimited funds environment assumes that the company can raise the funds it wants for all profitable investments simply by paying the required rate of return.

Unsecured debt Debt that gives the debtholder only a general claim on an issuer's assets and cash flow.

Unsponsored A type of depository receipt in which the foreign company whose shares are held by the depository has no involvement in the issuance of the receipts.

Unstructured data Data that do not follow any conventionally organized forms.

Unsupervised learning A machine learning approach that does not make use of labeled training data.

Up transition probability The probability that an asset's value moves up.

Uptrend A pattern that occurs when the price of an asset moves higher over a period of time.

Validity instructions Instructions which indicate when the order may be filled.

Valuation allowance A reserve created against deferred tax assets, based on the likelihood of realizing the deferred tax assets in future accounting periods.

Valuation ratios Ratios that measure the quantity of an asset or flow (e.g., earnings) in relation to the price associated with a specified claim (e.g., a share or ownership of the enterprise).

Value at risk A money measure of the minimum value of losses expected during a specified time period at a given level of probability.

Value investors With reference to equity investors, investors who are focused on paying a relatively low share price in relation to earnings or assets per share.

VaR See *value at risk*.

Variable A characteristic or quantity that can be measured, counted, or categorized and that is subject to change (also called a field, an attribute, or a feature).

Variable costs Costs that fluctuate with the level of production and sales.

Variance The expected value (the probability-weighted average) of squared deviations from a random variable's expected value.

Variation margin Additional margin that must be deposited in an amount sufficient to bring the balance up to the initial margin requirement.

Veblen goods Goods that increase in desirability with increasing price.

Vega A measure of the sensitivity of an option's price to changes in the underlying's volatility.

Venture capital Investments that provide "seed" or startup capital, early-stage financing, or later-stage financing (including mezzanine-stage financing) to companies that are in early development stages and require additional capital for expansion or preparation for an initial public offering.

Venture capital fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *private equity fund*.

Venture capital funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *private equity funds.*

Vertical analysis Common-size analysis using only one reporting period or one base financial statement; for example, an income statement in which all items are stated as percentages of sales.

Vertical demand schedule Implies that some fixed quantity is demanded, regardless of price.

Visual technique The most common and readily available method of initial data assessment. Experts in pattern recognition maintain that the visual (or "eyeball") technique is still the most effective way of searching for recognizable patterns.

Visualization The presentation of data in a pictorial or graphical format for the purpose of increasing understanding and for gaining insights into the data.

Volatility The standard deviation of the continuously compounded returns on the underlying asset.

Voluntarily unemployed A person voluntarily outside the labor force, such as a jobless worker refusing an available vacancy.

Voluntary export restraint A trade barrier under which the exporting country agrees to limit its exports of the good to its trading partners to a specific number of units.

Vote by proxy A mechanism that allows a designated party—such as another shareholder, a shareholder representative, or management—to vote on the shareholder's behalf.

Warrant Attached option that gives its holder the right to buy the underlying stock of the issuing company at a fixed exercise price until the expiration date.

Waterfall Represents the distribution method that defines the order in which allocations are made to LPs and GPs. There are two major types of waterfall: *deal by deal* (or *American*) and *whole of fund* (or *European*).

Weak-form efficient market hypothesis The belief that security prices fully reflect all past market data, which refers to all historical price and volume trading information.

Wealth effect An increase (decrease) in household wealth increases (decreases) consumer spending out of a given level of current income.

Web-based lender A lender that operates primarily on the internet, offering loans in relatively small amounts, typically to small businesses in need of cash.

Weighted average cost method An inventory accounting method that averages the total cost of available inventory items over the total units available for sale.

Weighted average cost of capital A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average cost of capital (WACC) A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average coupon rate Weighting the mortgage rate of each mortgage loan in the pool by the percentage of the mortgage outstanding relative to the outstanding amount of all the mortgages in the pool.

Weighted average life A measure that gives investors an indication of how long they can expect to hold the MBS before it is paid off; the convention-based average time to receipt of all principal repayments. Also called *average life*.

Weighted average maturity Weighting the remaining number of months to maturity for each mortgage loan in the pool by the amount of the outstanding mortgage balance.

Weighted mean An average in which each observation is weighted by an index of its relative importance.

Winsorized mean A mean computed after assigning a stated percentage of the lowest values equal to one specified low value and a stated percentage of the highest values equal to one specified high value.

Word cloud A visual device for representing textual data, which consists of words extracted from a source of textual data. The size of each distinct word is proportional to the frequency with which it appears in the given text (also known as **tag cloud**).

Working capital The difference between current assets and current liabilities.

Working capital management The management of a company's short-term assets (such as inventory) and short-term liabilities (such as money owed to suppliers).

World price The price prevailing in the world market.

Yield duration The sensitivity of the bond price with respect to the bond's own yield-to-maturity.

Yield to maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield to redemption* or *redemption yield*.

Yield-to-maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield-to-redemption* or *redemption yield*.

Yield-to-redemption See *yield-to-maturity*.

Yield-to-worst The lowest of the sequence of yields-to-call and the yield-to-maturity.

Zero-coupon bond A bond that does not pay interest during its life. It is issued at a discount to par value and redeemed at par. Also called *pure discount bond*.

Zero-volatility spread (Z-spread) Calculates a constant yield spread over a government (or interest rate swap) spot curve.

CFA Institute®
CFA Program

ECONOMICS AND FINANCIAL STATEMENT ANALYSIS

CFA® Program Curriculum
2022 • LEVEL I • VOLUME 2

WILEY

CONTENTS

How to Use the CFA Program Curriculum **ix**
- Background on the CBOK **ix**
- Organization of the Curriculum **x**
- Features of the Curriculum **x**
- Designing Your Personal Study Program **xi**
- CFA Institute Learning Ecosystem (LES) **xii**
- Prep Providers **xiii**
- Feedback **xiv**

Economics

Study Session 3 **Economics (1)** **3**

Reading 8 **Topics in Demand and Supply Analysis** **5**
- Introduction **5**
- Demand Concepts **6**
  - Demand Concepts **6**
- Price Elasticity of Demand **9**
  - Extremes of Price Elasticity **11**
- Predicting Demand Elasticity, Price Elasticity and Total Expenditure **12**
  - Elasticity and Total Expenditure **13**
- Income Elasticity of Demand, Cross-Price Elasticity of Demand **14**
  - Cross-Price Elasticity of Demand **15**
- Substitution and Income Effects; Normal Goods, Inferior Goods and Special Cases **18**
  - Normal and Inferior Goods **19**
- Supply Analysis: Cost, Marginal Return, and Productivity **23**
  - Marginal Returns and Productivity **23**
- Economc Profit Versus Accounting Profit **28**
  - Economic Cost vs. Accounting Cost **28**
- Marginal Revenue, Marginal Cost and Profit Maximization; Short-Run Cost Curves: Total, Variable, Fixed, and Marginal Costs **29**
  - Understanding the Interaction between Total, Variable, Fixed, and Marginal Cost and Output **31**
- Perfect and Imperfect Competition, Profit Maximization **35**
  - Profit-Maximization, Breakeven, and Shutdown Points of Production **36**
- Breakeven Analysis and Shutdown Decision **38**
  - The Shutdown Decision **39**
- Economies and Diseconomies of Scale with Short-Run and Long-Run Cost Analysis **43**
  - Short- and Long-Run Cost Curves **43**
  - Defining Economies of Scale and Diseconomies of Scale **44**
- *Summary* **48**
- *Practice Problems* **51**
- *Solutions* **58**

O indicates an optional segment

Reading 9 **The Firm and Market Structures** **63**

Introduction & Analysis of Market Structures **63**

Analysis of Market Structures **64**

Perfect Competition & Demand Analysis in Perfectly Competitive Markets **69**

Demand Analysis in Perfectly Competitive Markets **69**

Elasticity of Demand **71**

Other Factors Affecting Demand **73**

Consumer Surplus: Value Minus Expenditure **75**

Supply Analysis & Optimal Price and Optimal Output In Perfectly Competitive Markets **77**

Optimal Price and Output in Perfectly Competitive Markets **78**

Factors Affecting Long-Run Equilibrium in Perfectly Competitive Markets **82**

Supply, Demand, Optimal Pricing, and Optimal Ouput under Monopolistic Competition **85**

Demand Analysis in Monopolistically Competitive Markets **86**

Supply Analysis in Monopolistically Competitive Markets **86**

Optimal Price and Output in Monopolistically Competitive Markets **87**

Long-Run Equilibrium for Monopolistically Competitive Firm **87**

Oligopoly & Demand Analysis and Pricing Strategies in Oligopoly Markets: Pricing Independence Portion **88**

Demand Analysis and Pricing Strategies in Oligopoly Markets **89**

Oligopoly & Demand Analysis and Pricing Strategies in Oligopoly Markets: The Cournot Assumption Part **91**

Oligopoly & Demand Analysis and Pricing Strategies in Oligopoly Markets: The Nash Equilibrium part **93**

Supply Analysis & Optimal Price and Output & Long-Run Equilibrium in Oligopoly Markets **96**

Optimal Price and Output in Oligopoly Markets **97**

Factors Affecting Long-Run Equilibrium in Oligopoly Markets **98**

Monopoly & Demand & Supply & Optimal Price and Output in Monopoly Markets **99**

Demand Analysis in Monopoly Markets **100**

Supply Analysis in Monopoly Markets **101**

Optimal Price and Output in Monopoly Markets **103**

Price Discrimination and Consumer Surplus **104**

Factors Affecting Long-Run Equilibrium in Monopoly Markets **106**

Identification of Market Structure **107**

Econometric Approaches **108**

Simpler Measures **109**

Summary **110**

Practice Problems **113**

Solutions **117**

Reading 10 **Aggregate Output, Prices, and Economic Growth** **119**

Introduction **120**

Aggregate Output and Income **121**

Gross Domestic Product **122**

The Components of GDP **129**

  - GDP, National Income, Personal Income, and Personal Disposable Income 132
  - Relationship among Saving, Investment, the Fiscal Balance and the Trade Balance 138
- Aggregate Demand and Aggregate Supply 142
  - Aggregate Demand 142
  - Aggregate Supply 146
  - Shifts in the Aggregate Demand Curve 147
  - Equilibrium GDP and Prices 160
- Economic Growth and Sustainability 171
  - The Production Function and Potential GDP 173
  - Sources of Economic Growth 175
  - Measures of Sustainable Growth 180
  - Measuring Sustainable Growth 183
- *Summary* 186
- *Practice Problems* 190
- *Solutions* 195

Reading 11 **Understanding Business Cycles** **199**
- Introduction 199
- Overview of the Business Cycle 200
  - Phases of the Business Cycle 201
  - Leads and Lags in Business and Consumer Decision Making 205
  - Market Conditions and Investor Behavior 205
- Credit Cycles and Their Relationship to Business Cycles 207
  - Applications of Credit Cycles 207
  - Consequences for Policy 208
- Business Cycle Fluctuations from a Firm's Perspective 209
  - The Workforce and Company Costs 209
  - Fluctuations in Capital Spending 210
  - Fluctuations in Inventory Levels 212
- Consumer Behavior 214
  - Consumer Confidence 214
  - Measures of Consumption 214
  - Income Growth 215
  - Saving Rates 216
- Housing Sector Behavior 216
  - Available Statistics 217
  - Sensitivity to Interest Rates and Relationship to Credit Cycle 217
  - The Role of Demographics 217
  - Impact on the Economic Cycle 217
- External Trade Sector Behavior 218
  - Cyclical Fluctuations of Imports and Exports 218
  - The Role of the Exchange Rate 219
  - Overall Effect on Exports and Imports 219
- Theoretical Considerations 220
  - Historical Context 221
  - Neoclassical Economics 222
  - The Austrian School 223

Monetarism 223
Keynesianism 223
Modern Approach to Business Cycles 226
Economic Indicators 227
Types of Indicators 227
Composite Indicators 228
Leading Indicators 228
Using Economic Indicators 229
Other Composite Leading Indicators 230
Surveys 232
The Use of Big Data in Economic Indicators 232
Nowcasting 232
GDPNow 233
Unemployment 236
Unemployment 236
Inflation 240
Deflation, Hyperinflation, and Disinflation 240
Measuring Inflation: The Construction of Price Indexes 242
Price Indexes and Their Usage 243
Explaining Inflation 248
Summary 253
Practice Problems 257
Solutions 263

Study Session 4 **Economics (2)** 267

Reading 12 **Monetary and Fiscal Policy** 269
Introduction to Monetary and Fiscal Policy 270
Monetary Policy 272
Money: Functions, Creation, and Definition 272
The Functions of Money 273
Paper Money and the Money Creation Process 274
Definitions of Money 277
Money: Quantity Theory, Supply and Demand, Fisher Effect 278
The Demand for Money 279
The Supply and Demand for Money 280
The Fisher Effect 282
Roles of Central Banks & Objectives of Monetary Policy 285
The Objectives of Monetary Policy 288
The Costs of Inflation 289
Monetary Policy Tools 291
Open Market Operations 292
The Central Bank's Policy Rate 292
Reserve Requirements 293
The Transmission Mechanism 293
Inflation Targeting 295
Central Bank Independence 296
Credibility 297
Transparency 297

Exchange Rate Targeting 302
Monetary Policies: Contractionary, Expansionary, Limitations 305
    What's the Source of the Shock to the Inflation Rate? 306
    Limitations of Monetary Policy 306
Roles and Objectives of Fiscal Policy 311
    Roles and Objectives of Fiscal Policy 312
Deficits and the National Debt 316
Fiscal Policy Tools 320
    The Advantages and Disadvantages of Using the Different Tools of Fiscal Policy 323
    Modeling the Impact of Taxes and Government Spending: The Fiscal Multiplier 324
    The Balanced Budget Multiplier 325
Fiscal Policy Implementation 326
    Deficits and the Fiscal Stance 326
    Difficulties in Executing Fiscal Policy 327
The Relationship between Monetary and Fiscal Policy 330
    Factors Influencing the Mix of Fiscal and Monetary Policy 331
    Quantitative Easing and Policy Interaction 332
    The Importance of Credibility and Commitment 332
Summary 334
Practice Problems 336
Solutions 341

Reading 13 **International Trade and Capital Flows** 343
Introduction & International Trade-Basic Terminology 343
    International Trade 344
Patterns and Trends in International Trade and Capital Flows 348
Benefits and Costs of International Trade 351
Comparative Advantage and the Gains from Trade: Absolute and Comparative Advantage 354
    Gains from Trade: Absolute and Comparative Advantage 354
Ricardian and Heckscher–Ohlin Models of Comparative Advantage 360
Trade and Capital Flows: Restrictions & Agreements- Tariffs, Quotas and Export Subsidies 362
    Tariffs 363
    Quotas 365
    Export Subsidies 366
Trading Blocs, Common Markets, and Economic Unions 368
Capital Restrictions 373
Balance of Payments- Accounts and Components 376
    Balance of Payments Accounts 376
    Balance of Payment Components 378
Paired Transactions in the BOP Bookkeeping System 380
    Commercial Exports: Transactions (ia) and (ib) 380
    Commercial Imports: Transaction (ii) 381
    Loans to Borrowers Abroad: Transaction (iii) 381
    Purchases of Home-Country Currency by Foreign Central Banks: Transaction (iv) 382

    Receipts of Income from Foreign Investments: Transaction (v) 382
    Purchase of Non-financial Assets: Transaction (vi) 383
National Economic Accounts and the Balance of Payments 383
Trade Organizations 388
    International Monetary Fund 389
    World Bank Group 391
    World Trade Organization 392
Summary 395
Practice Problems 398
Solutions 402

Reading 14 **Currency Exchange Rates** 405
Introduction & The Foreign Exchange Market 406
    The Foreign Exchange Market 407
Market Functions 413
Market Participants, Size and Composition 418
    Market Size and Composition 421
Exchange Rate Quotations 424
    Exchange Rate Quotations 424
Cross-Rate Calculations 428
Forward Calculations 431
Exchange Rate Regimes- Ideals and Historical Perspective 439
    The Ideal Currency Regime 439
    Historical Perspective on Currency Regimes 440
A Taxonomy of Currency Regimes 442
    Arrangements with No Separate Legal Tender 444
    Currency Board System 445
    Fixed Parity 446
    Target Zone 446
    Active and Passive Crawling Pegs 446
    Fixed Parity with Crawling Bands 447
    Managed Float 447
    Independently Floating Rates 447
Exchange Rates and the Trade Balance: Introduction 451
Exchange Rates and the Trade Balance: The Elasticities Approach 452
Exchange Rates and the Trade Balance: The Absorption Approach 457
Summary 461
Practice Problems 464
Solutions 467

Financial Statement Analysis

Study Session 5 **Financial Statement Analysis (1)** 473

Reading 15 **Introduction to Financial Statement Analysis** 475
Introduction 476
Scope of Financial Statement Analysis 476
Major Financial Statements - Balance Sheet 483
    Financial Statements and Supplementary Information 483

Statement of Comprehensive Income **488**
Income Statement **488**
Other Comprehensive Income **491**
Statement of Changes in Equity and Cash Flow Statement **492**
Cash Flow Statement **493**
Financial Notes, Supplementary Schedules, and Management Commentary **495**
Management Commentary or Management's Discussion and Analysis **498**
Auditor's Reports **499**
Other Sources of Information **502**
Financial Statement Analysis Framework **503**
Articulate the Purpose and Context of Analysis **504**
Collect Data **505**
Process Data **506**
Analyze/Interpret the Processed Data **506**
Develop and Communicate Conclusions/Recommendations **506**
Follow-Up **507**
Summary **507**
Practice Problems **510**
Solutions **513**

Reading 16 **Financial Reporting Standards** **515**
Introduction **516**
The Objective of Financial Reporting **516**
Accounting Standards Boards **517**
Accounting Standards Boards **518**
Regulatory Authorities **519**
International Organization of Securities Commissions **520**
The Securities and Exchange Commission (US) **520**
Capital Markets Regulation in Europe **523**
The International Financial Reporting Standards Framework **524**
Qualitative Characteristics of Financial Reports **525**
Constraints on Financial Reports **526**
The Elements of Financial Statements **527**
Underlying Assumptions in Financial Statements **528**
Recognition of Financial Statement Elements **528**
Measurement of Financial Statement Elements **528**
General Requirements for Financial Statements **529**
Required Financial Statements **530**
General Features of Financial Statements **530**
Structure and Content Requirements **531**
Comparison of IFRS with Alternative Reporting Systems **532**
Monitoring Developments in Financial Reporting Standards **533**
New Products or Types of Transactions **534**
Evolving Standards and the Role of CFA Institute **534**
Summary **535**
Practice Problems **537**
Solutions **539**

Glossary **G-1**

How to Use the CFA Program Curriculum

Congratulations on your decision to enter the Chartered Financial Analyst (CFA®) Program. This exciting and rewarding program of study reflects your desire to become a serious investment professional. You are embarking on a program noted for its high ethical standards and the breadth of knowledge, skills, and abilities (competencies) it develops. Your commitment should be educationally and professionally rewarding.

The credential you seek is respected around the world as a mark of accomplishment and dedication. Each level of the program represents a distinct achievement in professional development. Successful completion of the program is rewarded with membership in a prestigious global community of investment professionals. CFA charterholders are dedicated to life-long learning and maintaining currency with the ever-changing dynamics of a challenging profession. CFA Program enrollment represents the first step toward a career-long commitment to professional education.

The CFA exam measures your mastery of the core knowledge, skills, and abilities required to succeed as an investment professional. These core competencies are the basis for the Candidate Body of Knowledge (CBOK™). The CBOK consists of four components:

- A broad outline that lists the major CFA Program topic areas (www.cfainstitute.org/programs/cfa/curriculum/cbok);
- Topic area weights that indicate the relative exam weightings of the top-level topic areas (www.cfainstitute.org/programs/cfa/curriculum);
- Learning outcome statements (LOS) that advise candidates about the specific knowledge, skills, and abilities they should acquire from readings covering a topic area (LOS are provided in candidate study sessions and at the beginning of each reading); and
- CFA Program curriculum that candidates receive upon exam registration.

Therefore, the key to your success on the CFA exams is studying and understanding the CBOK. The following sections provide background on the CBOK, the organization of the curriculum, features of the curriculum, and tips for designing an effective personal study program.

BACKGROUND ON THE CBOK

CFA Program is grounded in the practice of the investment profession. CFA Institute performs a continuous practice analysis with investment professionals around the world to determine the competencies that are relevant to the profession, beginning with the Global Body of Investment Knowledge (GBIK®). Regional expert panels and targeted surveys are conducted annually to verify and reinforce the continuous feedback about the GBIK. The practice analysis process ultimately defines the CBOK. The CBOK reflects the competencies that are generally accepted and applied by investment professionals. These competencies are used in practice in a generalist context and are expected to be demonstrated by a recently qualified CFA charterholder.

© 2021 CFA Institute. All rights reserved.

The CFA Institute staff—in conjunction with the Education Advisory Committee and Curriculum Level Advisors, who consist of practicing CFA charterholders—designs the CFA Program curriculum in order to deliver the CBOK to candidates. The exams, also written by CFA charterholders, are designed to allow you to demonstrate your mastery of the CBOK as set forth in the CFA Program curriculum. As you structure your personal study program, you should emphasize mastery of the CBOK and the practical application of that knowledge. For more information on the practice analysis, CBOK, and development of the CFA Program curriculum, please visit www.cfainstitute.org.

ORGANIZATION OF THE CURRICULUM

The Level I CFA Program curriculum is organized into 10 topic areas. Each topic area begins with a brief statement of the material and the depth of knowledge expected. It is then divided into one or more study sessions. These study sessions should form the basic structure of your reading and preparation. Each study session includes a statement of its structure and objective and is further divided into assigned readings. An outline illustrating the organization of these study sessions can be found at the front of each volume of the curriculum.

The readings are commissioned by CFA Institute and written by content experts, including investment professionals and university professors. Each reading includes LOS and the core material to be studied, often a combination of text, exhibits, and in-text examples and questions. End of Reading Questions (EORQs) followed by solutions help you understand and master the material. The LOS indicate what you should be able to accomplish after studying the material. The LOS, the core material, and the EORQs are dependent on each other, with the core material and EORQs providing context for understanding the scope of the LOS and enabling you to apply a principle or concept in a variety of scenarios.

The entire readings, including the EORQs, are the basis for all exam questions and are selected or developed specifically to teach the knowledge, skills, and abilities reflected in the CBOK.

You should use the LOS to guide and focus your study because each exam question is based on one or more LOS and the core material and practice problems associated with the LOS. As a candidate, you are responsible for the entirety of the required material in a study session.

We encourage you to review the information about the LOS on our website (www.cfainstitute.org/programs/cfa/curriculum/study-sessions), including the descriptions of LOS “command words” on the candidate resources page at www.cfainstitute.org.

FEATURES OF THE CURRICULUM

End of Reading Questions/Solutions *All End of Reading Questions (EORQs) as well as their solutions are part of the curriculum and are required material for the exam.* In addition to the in-text examples and questions, these EORQs help demonstrate practical applications and reinforce your understanding of the concepts presented. Some of these EORQs are adapted from past CFA exams and/or may serve as a basis for exam questions.

Glossary For your convenience, each volume includes a comprehensive Glossary. Throughout the curriculum, a **bolded** word in a reading denotes a term defined in the Glossary.

Note that the digital curriculum that is included in your exam registration fee is searchable for key words, including Glossary terms.

LOS Self-Check We have inserted checkboxes next to each LOS that you can use to track your progress in mastering the concepts in each reading.

Source Material The CFA Institute curriculum cites textbooks, journal articles, and other publications that provide additional context or information about topics covered in the readings. As a candidate, you are not responsible for familiarity with the original source materials cited in the curriculum.

Note that some readings may contain a web address or URL. The referenced sites were live at the time the reading was written or updated but may have been deactivated since then.

Some readings in the curriculum cite articles published in the *Financial Analysts Journal®*, which is the flagship publication of CFA Institute. Since its launch in 1945, the *Financial Analysts Journal* has established itself as the leading practitioner-oriented journal in the investment management community. Over the years, it has advanced the knowledge and understanding of the practice of investment management through the publication of peer-reviewed practitioner-relevant research from leading academics and practitioners. It has also featured thought-provoking opinion pieces that advance the common level of discourse within the investment management profession. Some of the most influential research in the area of investment management has appeared in the pages of the *Financial Analysts Journal*, and several Nobel laureates have contributed articles.

Candidates are not responsible for familiarity with *Financial Analysts Journal* articles that are cited in the curriculum. But, as your time and studies allow, we strongly encourage you to begin supplementing your understanding of key investment management issues by reading this, and other, CFA Institute practice-oriented publications through the Research & Analysis webpage (www.cfainstitute.org/en/research).

Errata The curriculum development process is rigorous and includes multiple rounds of reviews by content experts. Despite our efforts to produce a curriculum that is free of errors, there are times when we must make corrections. Curriculum errata are periodically updated and posted by exam level and test date online (www.cfainstitute.org/en/programs/submit-errata). If you believe you have found an error in the curriculum, you can submit your concerns through our curriculum errata reporting process found at the bottom of the Curriculum Errata webpage.

DESIGNING YOUR PERSONAL STUDY PROGRAM

Create a Schedule An orderly, systematic approach to exam preparation is critical. You should dedicate a consistent block of time every week to reading and studying. Complete all assigned readings and the associated problems and solutions in each study session. Review the LOS both before and after you study each reading to ensure that

you have mastered the applicable content and can demonstrate the knowledge, skills, and abilities described by the LOS and the assigned reading. Use the LOS self-check to track your progress and highlight areas of weakness for later review.

Successful candidates report an average of more than 300 hours preparing for each exam. Your preparation time will vary based on your prior education and experience, and you will probably spend more time on some study sessions than on others.

You should allow ample time for both in-depth study of all topic areas and additional concentration on those topic areas for which you feel the least prepared.

CFA INSTITUTE LEARNING ECOSYSTEM (LES)

As you prepare for your exam, we will email you important exam updates, testing policies, and study tips. Be sure to read these carefully.

Your exam registration fee includes access to the CFA Program Learning Ecosystem (LES). This digital learning platform provides access, even offline, to all of the readings and End of Reading Questions found in the print curriculum organized as a series of shorter online lessons with associated EORQs. This tool is your one-stop location for all study materials, including practice questions and mock exams.

The LES provides the following supplemental study tools:

Structured and Adaptive Study Plans The LES offers two ways to plan your study through the curriculum. The first is a structured plan that allows you to move through the material in the way that you feel best suits your learning. The second is an adaptive study plan based on the results of an assessment test that uses actual practice questions.

Regardless of your chosen study path, the LES tracks your level of proficiency in each topic area and presents you with a dashboard of where you stand in terms of proficiency so that you can allocate your study time efficiently.

Flashcards and Game Center The LES offers all the Glossary terms as Flashcards and tracks correct and incorrect answers. Flashcards can be filtered both by curriculum topic area and by action taken—for example, answered correctly, unanswered, and so on. These Flashcards provide a flexible way to study Glossary item definitions.

The Game Center provides several engaging ways to interact with the Flashcards in a game context. Each game tests your knowledge of the Glossary terms a in different way. Your results are scored and presented, along with a summary of candidates with high scores on the game, on your Dashboard.

Discussion Board The Discussion Board within the LES provides a way for you to interact with other candidates as you pursue your study plan. Discussions can happen at the level of individual lessons to raise questions about material in those lessons that you or other candidates can clarify or comment on. Discussions can also be posted at the level of topics or in the initial Welcome section to connect with other candidates in your area.

Practice Question Bank The LES offers access to a question bank of hundreds of practice questions that are in addition to the End of Reading Questions. These practice questions, only available on the LES, are intended to help you assess your mastery of individual topic areas as you progress through your studies. After each practice question, you will receive immediate feedback noting the correct response and indicating the relevant assigned reading so you can identify areas of weakness for further study.

Mock Exams The LES also includes access to three-hour Mock Exams that simulate the morning and afternoon sessions of the actual CFA exam. These Mock Exams are intended to be taken after you complete your study of the full curriculum and take practice questions so you can test your understanding of the curriculum and your readiness for the exam. If you take these Mock Exams within the LES, you will receive feedback afterward that notes the correct responses and indicates the relevant assigned readings so you can assess areas of weakness for further study. We recommend that you take Mock Exams during the final stages of your preparation for the actual CFA exam. For more information on the Mock Exams, please visit www.cfainstitute.org.

PREP PROVIDERS

You may choose to seek study support outside CFA Institute in the form of exam prep providers. After your CFA Program enrollment, you may receive numerous solicitations for exam prep courses and review materials. When considering a prep course, make sure the provider is committed to following the CFA Institute guidelines and high standards in its offerings.

Remember, however, that there are no shortcuts to success on the CFA exams; reading and studying the CFA Program curriculum *is* the key to success on the exam. The CFA Program exams reference only the CFA Institute assigned curriculum; no prep course or review course materials are consulted or referenced.

SUMMARY

Every question on the CFA exam is based on the content contained in the required readings and on one or more LOS. Frequently, an exam question is based on a specific example highlighted within a reading or on a specific practice problem and its solution. To make effective use of the CFA Program curriculum, please remember these key points:

1. All pages of the curriculum are required reading for the exam.
2. All questions, problems, and their solutions are part of the curriculum and are required study material for the exam. These questions are found at the end of the readings in the print versions of the curriculum. In the LES, these questions appear directly after the lesson with which they are associated. The LES provides immediate feedback on your answers and tracks your performance on these questions throughout your study.
3. We strongly encourage you to use the CFA Program Learning Ecosystem. In addition to providing access to all the curriculum material, including EORQs, in the form of shorter, focused lessons, the LES offers structured and adaptive study planning, a Discussion Board to communicate with other candidates, Flashcards, a Game Center for study activities, a test bank of practice questions, and online Mock Exams. Other supplemental study tools, such as eBook and PDF versions of the print curriculum, and additional candidate resources are available at www.cfainstitute.org.
4. Using the study planner, create a schedule and commit sufficient study time to cover the study sessions. You should also plan to review the materials, answer practice questions, and take Mock Exams.
5. Some of the concepts in the study sessions may be superseded by updated rulings and/or pronouncements issued after a reading was published. Candidates are expected to be familiar with the overall analytical framework contained in the assigned readings. Candidates are not responsible for changes that occur after the material was written.

FEEDBACK

At CFA Institute, we are committed to delivering a comprehensive and rigorous curriculum for the development of competent, ethically grounded investment professionals. We rely on candidate and investment professional comments and feedback as we work to improve the curriculum, supplemental study tools, and candidate resources.

Please send any comments or feedback to info@cfainstitute.org. You can be assured that we will review your suggestions carefully. Ongoing improvements in the curriculum will help you prepare for success on the upcoming exams and for a lifetime of learning as a serious investment professional.

Economics

STUDY SESSIONS

Study Session 3	Economics (1)
Study Session 4	Economics (2)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to demonstrate knowledge of microeconomic and macroeconomic principles.

The next study sessions introduce fundamental microeconomic and macroeconomic concepts relevant to financial analysis and investment management. Microeconomic factors such as a firm's competitive (or non-competitive) environment and its pricing strategy may be critical inputs for cash flow forecasting and bottom up security selection approaches. Economic output, global trade flows, monetary and fiscal policies, and the business cycle are key considerations for conducting top own investment analysis and economic forecasting.

Candidates should be familiar with the material covered in the following prerequisite economics readings available in Candidate Resources on the CFA Institute website:

- **Demand and Supply Analysis: Introduction**
- **Demand and Supply Analysis: Consumer Demand**
- **Demand and Supply Analysis: The Firm**

© 2021 CFA Institute. All rights reserved.

ECONOMICS
STUDY SESSION
3

Economics (1)

This study session begins by introducing fundamental concepts of demand and supply analysis for individual consumers and firms. Also covered are the various market structures (perfect competition, oligopoly, monopoly) in which firms operate. Key macroeconomic concepts and principles then follow, including aggregate output and income measurement, aggregate demand and supply analysis, and analysis of economic growth factors. The study session concludes with coverage of the business cycle and its effect on economic activity.

READING ASSIGNMENTS

Reading 8	Topics in Demand and Supply Analysis by Richard V. Eastin, PhD, and Gary L. Arbogast, PhD, CFA
Reading 9	The Firm and Market Structures by Richard Fritz, PhD, and Michele Gambera, PhD, CFA
Reading 10	Aggregate Output, Prices, and Economic Growth by Paul R. Kutasovic, PhD, CFA, and Richard Fritz, PhD
Reading 11	Understanding Business Cycles by Michele Gambera, PhD, CFA, Milton Ezrati, and Bolong Cao, PhD, CFA

© 2021 CFA Institute . All rights reserved.

READING

8

Topics in Demand and Supply Analysis

by Richard V. Eastin, PhD, and Gary L. Arbogast, PhD, CFA

Richard V. Eastin, PhD, is at the University of Southern California (USA). Gary L. Arbogast, PhD, CFA (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure;
☐	**b.** compare substitution and income effects;
☐	**c.** contrast normal goods with inferior goods;
☐	**d.** describe the phenomenon of diminishing marginal returns;
☐	**e.** determine and interpret breakeven and shutdown points of production;
☐	**f.** describe how economies of scale and diseconomies of scale affect costs.

INTRODUCTION

1

In a general sense, *economics* is the study of production, distribution, and consumption and can be divided into two broad areas of study: macroeconomics and microeconomics. **Macroeconomics** deals with aggregate economic quantities, such as national output and national income, and is rooted in **microeconomics**, which deals with markets and decision making of individual economic units, including consumers and businesses. Microeconomics is a logical starting point for the study of economics.

Microeconomics classifies private economic units into two groups: consumers (or households) and firms. These two groups give rise, respectively, to the theory of the consumer and the theory of the firm as two branches of study. The *theory of the consumer* deals with consumption (the demand for goods and services) by utility-maximizing individuals (i.e., individuals who make decisions that maximize the satisfaction received from present and future consumption). The *theory of the firm* deals with the supply of goods and services by profit-maximizing firms.

© 2019 CFA Institute. All rights reserved.

It is expected that candidates will be familiar with the basic concepts of demand and supply. This material is covered in detail in the recommended prerequisite readings. In this reading, we will explore how buyers and sellers interact to determine transaction prices and quantities. The reading is organized as follows: Sections 2–6 discuss the consumer or demand side of the market model, and Sections 7–12 discuss the supply side of the consumer goods market, paying particular attention to the firm's costs. Section 13 provides a summary of key points in the reading.

2 DEMAND CONCEPTS

a calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure;

The fundamental model of the private-enterprise economy is the demand and supply model of the market. In this section, we examine three important topics concerning the demand side of the model: (1) elasticities, (2) substitution and income effects, and (3) normal and inferior goods. The candidate is assumed to have a basic understanding of the demand and supply model and to understand how a market discovers the equilibrium price at which the quantity willingly demanded by consumers at that price is just equal to the quantity willingly supplied by firms. Here, we explore more deeply some of the concepts underlying the demand side of the model.

2.1 Demand Concepts

The quantity of a good that consumers are willing to buy depends on a number of different variables. Perhaps the most important of those variables is the item's own price. In general, economists believe that as the price of a good rises, buyers will choose to buy less of it, and as its price falls, they buy more. This opinion is so nearly universal that it has come to be called the **law of demand**.

Although a good's own price is important in determining consumers' willingness to purchase it, other variables also influence that decision. Consumers' incomes, their tastes and preferences, and the prices of other goods that serve as substitutes or complements are just a few of the other variables that influence consumers' demand for a product or service. Economists attempt to capture all these influences in a relationship called the **demand function**. (A function is a relationship that assigns a unique value to a dependent variable for any given set of values of a group of independent variables.)

Equation 1 is an example of a demand function. In Equation 1, we are saying, "The quantity demanded of good *X* depends on (is a function of) the price of good *X*, consumers' income, and the price of good *Y*":

$$Q_x^d = f\left(P_x, I, P_y\right) \tag{1}$$

where

Q_x^d = the quantity demanded of some good X (such as per household demand for gasoline in liters per month)

P_x = the price per unit of good X (such as € per liter)

I = consumers' income (as in €1,000s per household annually)

P_y = the price of another good, Y. (There can be many other goods, not just one, and they can be complements or substitutes.)

Often, economists use simple linear equations to approximate real-world demand and supply functions in relevant ranges. Equation 2 illustrates a hypothetical example of our function for gasoline demand:

$$Q_x^d = 84.5 - 6.39P_x + 0.25I - 2P_y \tag{2}$$

where the quantity of gasoline demanded $\left(Q_x^d\right)$ is a function of the price of a liter of gasoline (P_x), consumers' income in €1,000s (I), and the average price of an automobile in €1,000s (P_y).

The signs of the coefficients on gasoline price (negative) and consumers' income (positive) reflect the relationship between those variables and the quantity of gasoline consumed. The negative sign on average automobile price indicates that if automobiles go up in price, fewer will likely be purchased and driven; hence, less gasoline will be consumed. (As discussed later, such a relationship would indicate that gasoline and automobiles have a negative cross-price elasticity of demand and are thus complements.)

To continue our example, suppose that the price of gasoline (P_x) is €1.48 per liter, per household income (I) is €50,000, and the price of the average automobile (P_y) is €20,000. In this case, this function would predict that the per-household monthly demand for gasoline would be 47.54 liters, calculated as follows:

$$Q_x^d = 84.5 - 6.39(1.48) + 0.25(50) - 2(20) = 47.54$$

recalling that income and automobile prices are measured in thousands. Note that the sign on the "own-price" variable (P_x) is negative; thus, as the price of gasoline rises, per household consumption would decrease by 6.39 liters per month for every €1 increase in gas price. **Own price** is used by economists to underscore that the reference is to the price of a good itself and not the price of some other good.

In our example, there are three independent variables in the demand function and one dependent variable. If any one of the independent variables changes, so does the quantity demanded. It is often desirable to concentrate on the relationship between the dependent variable and just one of the independent variables at a time. To accomplish this goal, we can hold the other independent variables constant and rewrite the equation.

For example, to concentrate on the relationship between the quantity demanded of the good and its own price, P_x, we hold constant the values of income and the price of good Y. In our example, those values are 50 and 20, respectively. The equation would then be rewritten as

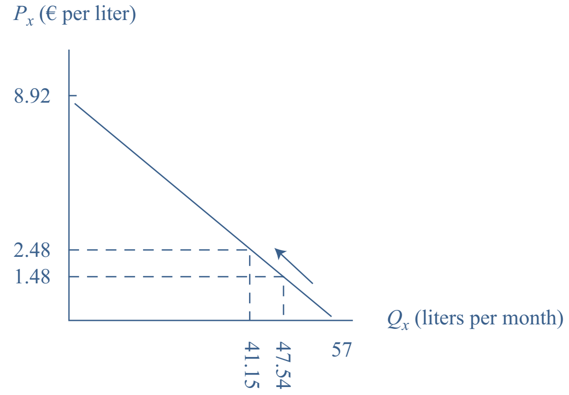
$$Q_x^d = 84.5 - 6.39P_x + 0.25(50) - 2(20) = 57 - 6.39P_x \tag{3}$$

The quantity of gasoline demanded is a function of the price of gasoline (6.39 per liter), per household income (€50,000), and the average price of an automobile (€20,000). Notice that income and the price of automobiles are not ignored; they are simply held constant, and they are "collected" in the new constant term, 57 [84.5 + (0.25)(50) – (2)(20)]. Notice also that we can solve for P_x in terms of Q_x^d by rearranging Equation 3, which gives us Equation 4:

$$P_x = 8.92 - 0.156Q_x^d \tag{4}$$

Equation 4 gives the price of gasoline as a function of the quantity of gasoline consumed per month and is referred to as the **inverse demand function**. Q_x in Equation 4 must be restricted to be less than or equal to 57 so that price is not negative. The graph of the inverse demand function is called the **demand curve** and is shown in Exhibit 1.[1]

Exhibit 1 Household Demand Curve for Gasoline



The demand curve represents the highest quantity willingly purchased at each price as well as the highest price willingly paid for each quantity. In this example, this household would be willing to purchase 47.54 liters of gasoline per month at a price of €1.48 per liter. If price were to rise to €2.48 per liter, the household would be willing to purchase only 41.15 liters per month.

This demand curve is drawn with price on the vertical axis and quantity on the horizontal axis. It can be correctly interpreted as specifying *either* the highest quantity a household would buy at a given price *or* the highest price it would be willing to pay for a given quantity. In our example, at a price of €1.48 per liter, households would each be willing to buy 47.54 liters per month. Alternatively, the highest price they would be willing to pay for 47.54 liters per month is €1.48 per liter. If the price were to rise by €1, households would reduce the quantity they each bought by 6.39 units, to 41.15 liters. The slope of the demand curve is measured as the change in price, P, divided by the change in quantity, Q ($\Delta P/\Delta Q$, where Δ stands for "the change in"). In this case, the slope of the demand curve is 1/–6.39, or –0.156.

The general model of demand and supply can be highly useful in understanding directional changes in prices and quantities that result from shifts in one curve or the other. Often, though, we need to measure how sensitive quantity demanded or supplied is to changes in the independent variables that affect them. This is the concept of **elasticity of demand** and **elasticity of supply**. Fundamentally, all elasticities are calculated in the same way: They are ratios of percentage changes. Let us begin with the sensitivity of quantity demanded to changes in the own price.

---

[1] Following usual practice, we show linear demand curves intersecting the quantity axis at a price of zero. Real-world demand functions may be non-linear in some or all parts of their domain. Thus, linear demand functions in practical cases are approximations of the true demand function that are useful for a relevant range of values.

PRICE ELASTICITY OF DEMAND

3

a calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure;

In Equation 1, we expressed the quantity demanded of some good as a function of several variables, one of which was the price of the good itself (the good's "own-price").

In Equation 3, we introduced a hypothetical household demand function for gasoline, assuming that the household's income and the price of another good (automobiles) were held constant. That function was given by the simple linear expression $Q_x^d = 57 - 6.39P_x$. Using this expression, if we were asked how sensitive the quantity of gasoline demanded is to changes in price, we might say that whenever price changes by one unit, quantity changes by 6.39 units in the opposite direction; for example, if price were to rise by €1, quantity demanded would fall by 6.39 liters per month. The coefficient on the price variable (–6.39) could be the measure of sensitivity we are seeking.

There is a drawback associated with that measure, however. It is dependent on the units in which we measured Q and P. When we want to describe the sensitivity of demand, we need to recall the specific units in which Q and P were measured—liters per month and euros per liter—in our example. This relationship cannot readily be extrapolated to other units of measure—for example, gallons and dollars. Economists, therefore, prefer to use a gauge of sensitivity that does not depend on units of measure. That metric is called **elasticity**. Elasticity is a general measure of how sensitive one variable is to any other variable, and it is expressed as the ratio of percentage changes in each variable: $\%\Delta y/\%\Delta x$. In the case of **own-price elasticity of demand**, that measure is illustrated in Equation 5:

$$E_{p_x}^d = \frac{\%\Delta Q_x^d}{\%\Delta P_x} \tag{5}$$

This equation expresses the sensitivity of the quantity demanded to a change in price. $E_{p_x}^d$ is the good's own-price elasticity and is equal to the percentage change in quantity demanded divided by the percentage change in price. This measure is independent of the units in which quantity and price are measured. If quantity demanded falls by 8% when price rises by 10%, then the elasticity of demand is simply –0.8. It does not matter whether we are measuring quantity in gallons per week or liters per day, and it does not matter whether we measure price in dollars per gallon or euros per liter; 10% is 10%, and 8% is 8%. So the ratio of the first to the second is still –0.8.

We can expand Equation 5 algebraically by noting that the percentage change in any variable x is simply the change in x (Δx) divided by the level of x. So, we can rewrite Equation 5, using a few simple steps, as

$$E_{p_x}^d = \frac{\%\Delta Q_x^d}{\%\Delta P_x} = \frac{\dfrac{\Delta Q_x^d}{Q_x^d}}{\dfrac{\Delta P_x}{P_x}} = \left(\frac{\Delta Q_x^d}{\Delta P_x}\right)\left(\frac{P_x}{Q_x^d}\right) \tag{6}$$

To get a better idea of price elasticity, it might be helpful to illustrate using our hypothetical demand function: $Q_x^d = 57 - 6.39P_x$. When the relationship between two variables is linear, $\Delta Q_x^d / \Delta P_x$ is equal to the slope coefficient on P_x in the demand function. Thus, in our example, the elasticity of demand is –6.39 multiplied by the ratio of price to quantity. We need to choose a price at which to calculate the elasticity coefficient. Using our hypothetical original price of €1.48, we can find the quantity associated with that particular price by inserting 1.48 into the demand function as given in Equation 3:

$$Q = 57 - (6.39)(1.48) = 47.54$$

and we find that $Q = 47.54$ liters per month.

The result of our calculation is that at a price of 1.48, the elasticity of our market demand function is −6.39(1.48/47.54) = −0.2. How do we interpret that value? It means, simply, that when price equals 1.48, a 1% rise in price would result in a fall in quantity demanded of 0.2%.

In our example, when the price is €1.48 per liter, demand is not very sensitive to changes in price because a 1% rise in price would reduce quantity demanded by only 0.2%. In this case, we would say that demand is **inelastic**. To be precise, when the magnitude (ignoring algebraic sign) of the own-price elasticity coefficient has a value of less than one, demand is said to be inelastic. When that magnitude is greater than one, demand is said to be **elastic**. And when the elasticity coefficient is equal to negative one, demand is said to be **unit elastic**, or unitary elastic. Note that if the law of demand holds, own-price elasticity of demand will always be negative because a rise in price will be associated with a fall in quantity demanded, but it can be either elastic (very sensitive to a change in price) or inelastic (insensitive to a change in price). In our hypothetical example, suppose the price of gasoline was very high, say, €5 per liter. In this case, the elasticity coefficient would be −1.28:

$$Q = 57 - (6.39)(5) = 25.05$$

and

$$-6.39\ (5/25.05) = -1.28$$

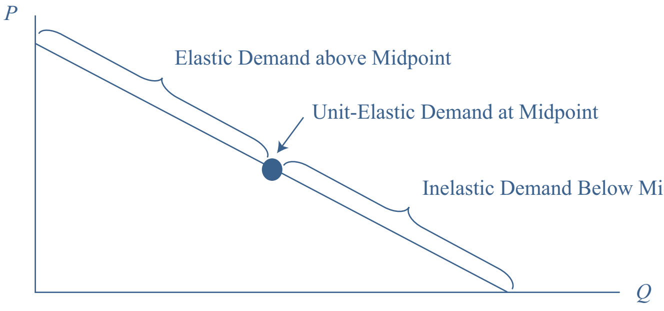
Because the magnitude of the elasticity coefficient is greater than one, we know that demand is elastic at that price.[2] In other words, at lower prices (€1.48 per liter), a slight change in the price of gasoline does not have much effect on the quantity demanded, but when gasoline is expensive (€5 per liter), consumer demand for gas is highly affected by changes in price.

By examining Equation 6 more closely, we can see that for a linear demand curve the elasticity depends on where on the curve we calculate it. The first term, $\Delta Q/\Delta P$, which is the inverse of the slope of the demand curve, remains constant along the entire demand curve. But the second term, P/Q, changes depending on where we are on the demand curve. At very low prices, P/Q is very small, so demand is inelastic. But at very high prices, Q is low and P is high, so the ratio P/Q is very high and demand is elastic. Exhibit 2 illustrates a characteristic of all negatively sloped linear demand curves. Above the midpoint of the curve, demand is elastic; below the midpoint, demand is inelastic; and at the midpoint, demand is unit elastic.

---

[2] If interested, evidence on price elasticities of demand for gasoline can be found in Molly Espey, "Explaining the Variation in Elasticity Estimates of Gasoline Demand in the United States: A Meta-analysis," *Energy Journal*, vol. 17, no. 3 (1996): 49–60. The robust estimates were about −0.26 for short-run elasticity—less than one year—and −0.58 for more than a year.

Exhibit 2 The Elasticity of a Linear Demand Curve

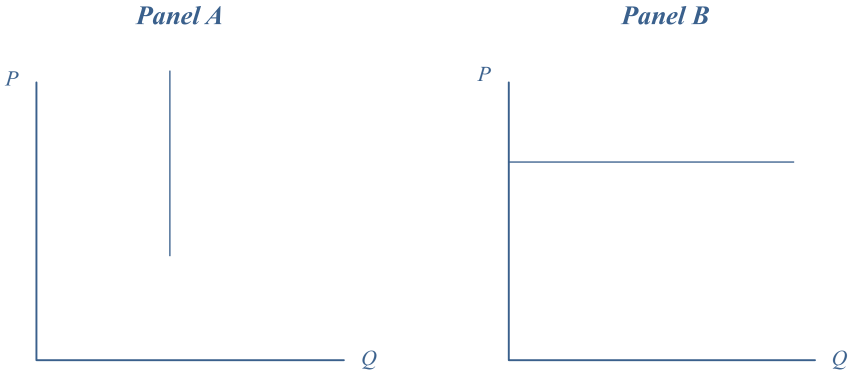


Note: For all negatively sloped, linear demand curves, elasticity varies depending on where it is calculated.

3.1 Extremes of Price Elasticity

There are two special cases in which linear demand curves have the same elasticity at all points: vertical demand curves and horizontal demand curves. Consider a vertical demand curve, as in Panel A of Exhibit 3, and a horizontal demand curve, as in Panel B. In the first case, the quantity demanded is the same, regardless of price. There is no demand curve that is perfectly vertical at all possible prices, but it is reasonable to assume that, over some range of prices, the same quantity would be purchased at a slightly higher price or a slightly lower price. Thus, in that price range, quantity demanded is not at all sensitive to price, and we would say that demand is **perfectly inelastic** in that range.

Exhibit 3 The Extremes of Price Elasticity



Note: A vertical demand has zero elasticity and is called perfectly inelastic.

Note: A horizontal demand has infinite elasticity and is called perfectly elastic.

In the second case, the demand curve is horizontal at some given price. It implies that even a minute price increase will reduce demand to zero, but at that given price, the consumer would buy some large, unknown amount. This situation is a reasonable description of the demand curve facing an individual seller in a perfectly competitive market, such as the wheat market. At the current market price of wheat, an individual farmer could sell all she has. If, however, she held out for a price above market price, it is reasonable to believe that she would not be able to sell any at all; other farmers' wheat is a perfect substitute for hers, so no one would be willing to buy any of hers at a higher price. In this case, we would say that the demand curve facing a seller under conditions of perfect competition is **perfectly elastic**.

4

PREDICTING DEMAND ELASTICITY, PRICE ELASTICITY AND TOTAL EXPENDITURE

a calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure;

Own-price elasticity of demand is a measure of how sensitive the quantity demanded is to changes in the price of a good or service, but what characteristics of a good or its market might be informative in determining whether demand is highly elastic? Perhaps the most important characteristic is whether there are close substitutes for the good in question. If there are close substitutes for the good, then if its price rises even slightly, a consumer would tend to purchase much less of this good and switch to the less costly substitute. If there are no substitutes, however, then it is likely that the demand is much less elastic. Consider a consumer's demand for some broadly defined product, such as bread. There really are no close substitutes for the entire category of bread, which includes all types from French bread to pita bread to tortillas and so on. So, if the price of all bread were to rise, perhaps a consumer would purchase a little less of it each week, but probably not a significantly smaller amount. Now, consider that the consumer's demand is for a particular baker's specialty bread instead of the category "bread" as a whole. Surely, there are close substitutes for Baker Bob's Whole Wheat Bread with Sesame Seeds than for bread in general. We would expect, then, that the demand for Baker Bob's special loaf is much more elastic than for the entire category of bread.

In addition to the degree of substitutability, other characteristics tend to be generally predictive of a good's elasticity of demand. These include the portion of the typical budget that is spent on the good, the amount of time that is allowed to respond to the change in price, the extent to which the good is seen as necessary or optional, and so on. In general, if consumers tend to spend a very small portion of their budget on a good, their demand tends to be less elastic than if they spend a very large part of their income. Most people spend only a little on toothpaste each month, for example, so it really does not matter whether the price rises 10%. They would probably still buy about the same amount. If the price of housing were to rise significantly, however, most households would try to find a way to reduce the quantity they buy, at least in the long run.

This example leads to another characteristic regarding price elasticity. For most goods and services, the long-run demand is much more elastic than the short-run demand. For example, if the price of gasoline rises, we probably would not be able to respond quickly to reduce the quantity we consume. In the short run, we tend to be locked into modes of transportation, housing and employment location, and so on. With a longer adjustment period, however, we can adjust the quantity consumed in response to the change in price by adopting a new mode of transportation or reducing the distance of our commute. Hence, for most goods, long-run elasticity of demand is greater than short-run elasticity. Durable goods, however, tend to behave in the opposite way. If the price of washing machines were to fall, people might react quickly because they have an old machine that they know will need to be replaced fairly soon anyway. So when price falls, they might decide to go ahead and make a purchase. If the price of washing machines were to stay low forever, however, it is unlikely that a typical consumer would buy more machines over a lifetime.

Knowing whether the good or service is seen to be discretionary or non-discretionary helps to understand its sensitivity to a price change. Faced with the same percentage increase in prices, consumers are much more likely to give up their Friday night restaurant meal (discretionary) than they are to cut back significantly on staples in their pantry (non-discretionary). The more a good is seen as being necessary, the less elastic its demand is likely to be.

In summary, own-price elasticity of demand is likely to be greater (i.e., more sensitive) for items that have many close substitutes, occupy a large portion of the total budget, are seen to be optional instead of necessary, or have longer adjustment times. Obviously, not all these characteristics operate in the same direction for all goods, so elasticity is likely to be a complex result of these and other characteristics. In the end, the actual elasticity of demand for a particular good turns out to be an empirical fact that can be learned only from careful observation and, often, sophisticated statistical analysis.

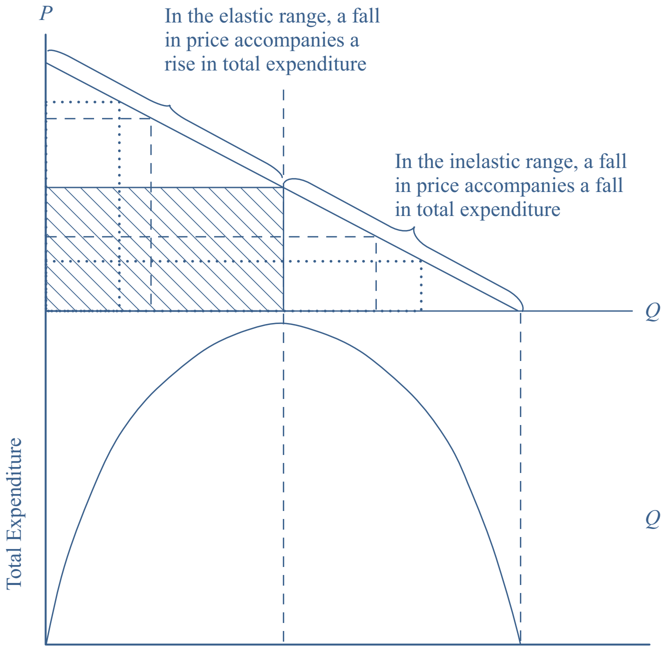
4.1 Elasticity and Total Expenditure

Because of the law of demand, an increase in price is associated with a decrease in the number of units demanded of some good or service. But what can we say about the total expenditure on that good? That is, what happens to price times quantity when price falls? Recall that elasticity is defined as the ratio of the percentage change in quantity demanded to the percentage change in price. So if demand is elastic, a decrease in price is associated with a larger percentage rise in quantity demanded. Although each unit of the good has a lower price, a sufficiently greater number of units are purchased so that total expenditure (price times quantity) would rise as price falls when demand is elastic.

If demand is inelastic, however, a given percentage decrease in price is associated with a smaller percentage rise in quantity demanded. Consequently, when demand is inelastic, a fall in price brings about a fall in total expenditure.

In summary, when demand is elastic, price and total expenditure move in *opposite* directions. When demand is inelastic, price and total expenditure move in the *same* direction. This relationship is easy to identify in the case of a linear demand curve. Recall that above the midpoint, demand is elastic, and below the midpoint, demand is inelastic. In the upper section of Exhibit 4, total expenditure ($P \times Q$) is measured as the area of a rectangle whose base is Q and height is P. Notice that as price falls, the areas of the inscribed rectangles (each outlined with their own dotted or dashed line) at first grow in size, become largest at the midpoint of the demand curve, and thereafter become smaller as price continues to fall and total expenditure declines toward zero. In the lower section of Exhibit 4, total expenditure is shown for each quantity purchased.

Exhibit 4 Elasticity and Total Expenditure



Note: Figure depicts the relationship among changes in price, changes in quantity, and changes in total expenditure. Maximum total expenditure occurs at the unit-elastic point on a linear demand curve (the cross-hatched rectangle).

The relationships just described hold for any demand curve, so it does not matter whether we are dealing with the demand curve of an individual consumer, the demand curve of the market, or the demand curve facing any given seller. For a market, the total expenditure by buyers becomes the total revenue to sellers in that market. It follows, then, that if market demand is elastic, a fall in price will result in an increase in total revenue to sellers as a whole, and if demand is inelastic, a fall in price will result in a decrease in total revenue to sellers. If the demand faced by any given seller were inelastic at the current price, that seller could increase revenue by increasing its price. But because demand is negatively sloped, the increase in price would decrease total units sold, which would almost certainly decrease total production cost. If raising price both increases revenue and decreases cost, such a move would always be profit enhancing. Faced with inelastic demand, a one-product seller would always be inclined to raise the price until the point at which demand becomes elastic.

5

INCOME ELASTICITY OF DEMAND, CROSS-PRICE ELASTICITY OF DEMAND

a calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure;

Elasticity is a measure of how sensitive one variable is to change in the value of another variable. Up to this point, we have focused on price elasticity, but the quantity demanded of a good is also a function of consumer income.

Income elasticity of demand is defined as the percentage change in quantity demanded $\left(\%\Delta Q_x^d\right)$ divided by the percentage change in income ($\%\Delta I$), holding all other things constant, as shown in Equation 7:

$$E_I^d = \frac{\%\Delta Q_x^d}{\%\Delta I} \tag{7}$$

The structure of this expression is identical to the structure of own-price elasticity given in Equation 5. (All elasticity measures that we will examine have the same general structure; the only thing that changes is the independent variable of interest.) For example, if the income elasticity of demand for some good has a value of 0.8, we would interpret that to mean that whenever income rises by 1%, the quantity demanded at each price would rise by 0.8%.

Although own-price elasticity of demand will almost always be negative, *income* elasticity of demand can be negative, positive, or zero. Positive income elasticity means that as income rises, quantity demanded also rises. Negative income elasticity of demand means that when people experience a rise in income, they buy less of these goods, and when their income falls, they buy more of the same good.

Goods with positive income elasticity are called "normal" goods. Goods with negative income elasticity are called "inferior" goods. Typical examples of inferior goods are rice, potatoes, or less expensive cuts of meat. We will discuss the concepts of normal and inferior goods in a later section.

In our discussion of the demand curve, we held all other things constant, including consumer income, to plot the relationship between price and quantity demanded. If income were to change, the entire demand curve would shift one way or the other. For normal goods, a rise in income would shift the entire demand curve upward and to the right. For inferior goods, however, a rise in income would result in a downward and leftward shift in the entire demand curve.

5.1 Cross-Price Elasticity of Demand

We previously discussed a good's own-price elasticity. However, the price of another good might also have an impact on the demand for that good or service, and we should be able to define an elasticity with respect to the other price (P_y) as well. That elasticity is called the **cross-price elasticity of demand** and takes on the same structure as own-price elasticity and income elasticity of demand, as represented in Equation 8:

$$E_{p_y}^d = \frac{\%\Delta Q_x^d}{\%\Delta P_y} \tag{8}$$

Note how similar this equation is to the equation for own-price elasticity. The only difference is that the subscript on P is now y, where y indicates some other good. This cross-price elasticity of demand measures how sensitive the demand for good X is to changes in the price of some other good, Y, holding all other things constant. For some pairs of goods, X and Y, when the price of Y rises, more of good X is demanded; the cross-price elasticity of demand is positive. Those goods are referred to as **substitutes**. In economics, if the cross-price elasticity of two goods is positive, they are substitutes, irrespective of whether someone would consider them "similar."

This concept is intuitive if you think about two goods that are seen to be close substitutes, perhaps like two brands of beer. When the price of one of your favorite brands of beer rises, you would probably buy less of that brand and more of a cheaper brand, so the cross-price elasticity of demand would be positive. For substitute goods, an increase in the price of one good would shift the demand curve for the other good upward and to the right.

Alternatively, two goods whose cross-price elasticity of demand is negative are said to be **complements**. Typically, these goods tend to be consumed together as a pair, such as gasoline and automobiles or houses and furniture. When automobile prices fall, we might expect the quantity of autos demanded to rise, and thus we might expect to see a rise in the demand for gasoline.

Whether two goods are substitutes or complements might not be immediately intuitive. For example, grocery stores often put things like coffee on sale in the hope that customers will come in for coffee and end up doing their weekly shopping there as well. In that case, coffee and, say, cabbage could very well empirically turn out to be complements even though we would not think that the price of coffee has any relation to sales of cabbage. Regardless of whether someone would see two goods as related in some fashion, if the cross-price elasticity of two goods is negative, they are complements.

Although a conceptual understanding of demand elasticities is helpful in sorting out the qualitative and directional effects among variables, using an empirically estimated demand function can yield insights into the behavior of a market. For illustration, let us return to our hypothetical individual demand function for gasoline in Equation 2, duplicated here for convenience:

$$Q_x^d = 84.5 - 6.39P_x + 0.25I - 2P_y$$

The quantity demanded of a given good $\left(Q_x^d\right)$ is a function of its own price (P_x), consumer income (I), and the price of another good (P_y).

To derive the market demand function, the individual consumers' demand functions are simply added together. If there were 1,000 individuals who represented a market and they all had identical demand functions, the market demand function would be the individual consumer's demand function multiplied by the number of consumers. Using the individual demand function given by Equation 2, the market demand function would be as shown in Equation 9:

$$Q_x^d = 84{,}500 - 6{,}390P_x + 250I - 2{,}000P_y \tag{9}$$

Earlier, when we calculated own-price elasticity of demand, we needed to choose a price at which to calculate the elasticity coefficient. Similarly, we need to choose actual values for the independent variables—P_x, I, and P_y—and insert these values into the "estimated" market demand function to find the quantity demanded. Choosing €1.48 for P_x, €50 (in thousands) for I, and €20 (in thousands) for P_y, we find that the quantity of gasoline demanded is 47,543 liters per month. We now have everything we need to calculate own-price, income, and cross-price elasticities of demand for our market. Those elasticities are expressed in Equations 10, 11, and 12. Each of those expressions has a term denoting the change in quantity divided by the change in each respective variable: own price, $\Delta Q_x/\Delta P_x$; income, $\Delta Q_x/\Delta I$, and cross price, $\Delta Q_x/\Delta P_y$.

As we stated in the discussion of own-price elasticity, when the relationship between two variables is linear, the change in quantity $\left(\Delta Q_x^d\right)$ divided by the change in own price (ΔP_x), income (ΔI), or cross price (ΔP_y) is equal to the slope coefficient on that other variable. The elasticities are calculated by inserting the slope coefficients from Equation 9 into the elasticity formulas.

Own-price elasticity:

$$E_{p_x}^d = \left(\frac{\Delta Q_x^d}{\Delta P_x}\right)\left(\frac{P_x}{Q_x^d}\right) = \left(-6{,}390\right)\left(\frac{1.48}{47{,}542.8}\right) = -0.20 \tag{10}$$

Income elasticity:

$$E_I^d = \left(\frac{\Delta Q_x^d}{\Delta I}\right)\left(\frac{I}{Q_x^d}\right) = (250)\left(\frac{50}{47{,}542.8}\right) = 0.26 \tag{11}$$

Cross-price elasticity:

$$E_{P_y}^d = \left(\frac{\Delta Q_x^d}{\Delta P_y}\right)\left(\frac{P_y}{Q_x^d}\right) = (-2000)\left(\frac{20}{47{,}542.8}\right) = -0.84 \tag{12}$$

In our example, at a price of €1.48, the own-price elasticity of demand is –0.20; a 1% increase in the price of gasoline leads to a decrease in quantity demanded of about 0.20% (Equation 10). Because the absolute value of the own-price elasticity is less than one, we characterize demand as being *inelastic* at that price; for example, an increase in price would result in an increase in total expenditure on gasoline by consumers in that market. The income elasticity of demand is 0.26 (Equation 11): A 1% increase in income would result in an increase of 0.26% in the quantity demanded of gasoline. Because that elasticity is positive (but small), we would characterize gasoline as a normal good. The cross-price elasticity of demand between gasoline and automobiles is –0.84 (Equation 12): If the price of automobiles rose by 1%, the demand for gasoline would fall by 0.84%. We would, therefore, characterize gasoline and automobiles as complements because the cross-price elasticity is negative. The magnitude is quite small, however, so we would conclude that the complementary relationship is weak.

EXAMPLE 1

Calculating Elasticities from a Given Demand Function

An individual consumer's monthly demand for downloadable e-books is given by the equation $Q_{eb}^d = 2 - 0.4P_{eb} + 0.0005I + 0.15P_{hb}$, where Q_{eb}^d equals the number of e-books demanded each month, I equals the household monthly income, P_{eb} equals the price of e-books, and P_{hb} equals the price of hardbound books. Assume that the price of e-books is €10.68, household income is €2,300, and the price of hardbound books is €21.40.

1. Determine the value of own-price elasticity of demand for e-books.
2. Determine the income elasticity of demand for e-books.
3. Determine the cross-price elasticity of demand for e-books with respect to the price of hardbound books.

Solution to 1:

The own-price elasticity of demand is given by $\left(\Delta Q_{eb}^d / \Delta P_{eb}\right)\left(P_{eb} / Q_{eb}^d\right)$. Notice from the demand function that $\Delta Q_{eb}^d / \Delta P_{eb} = -0.4$. Inserting the given variable values into the demand function yields $Q_{eb}^d = 2 - (0.4)(10.68) + (.0005)(2300) + (0.15)(21.4) = 2.088$. So at a price of €10.68, the own-price elasticity of demand equals (–0.4)(10.68/2.088) = –2.046, which is elastic because in absolute value the elasticity coefficient is greater than 1.

Solution to 2:

Recall that income elasticity of demand is given by $\left(\Delta Q_{eb}^{d}/\Delta I\right)\left(I/Q_{eb}^{d}\right)$. Notice from the demand function that $\Delta Q_{eb}^{d}/\Delta I = 0.0005$. Inserting the values for I and Q_{eb}^{d} yields income elasticity of (0.0005)(2,300/2.088) = 0.551, which is positive, so e-books are a normal good.

Solution to 3:

Recall that cross-price elasticity of demand is given by $(\Delta Q_{eb}/\Delta P_{hb})(P_{hb}/Q_{eb})$, and notice from the demand function that $\Delta Q_{eb}/\Delta P_{hb} = 0.15$. Inserting the values for P_{hb} and Q_{eb} yields a cross-price elasticity of demand for e-books of (0.15)(21.40/2.088) = 1.537, which is positive, implying that e-books and hardbound books are substitutes.

6

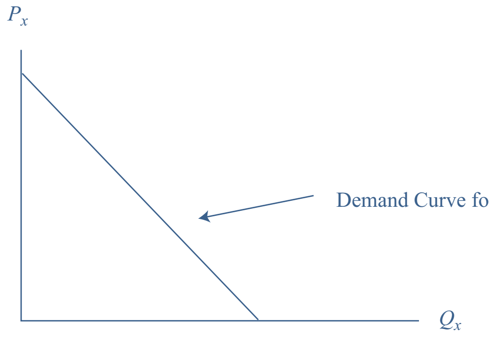
SUBSTITUTION AND INCOME EFFECTS; NORMAL GOODS, INFERIOR GOODS AND SPECIAL CASES

b compare substitution and income effects;

c contrast normal goods with inferior goods;

The law of demand states that if nothing changes other than the price of a particular good or service itself, a decrease in that good's price will tend to result in a greater quantity of that good being purchased. Simply stated, it is the assumption that a demand curve has negative slope; that is, where price per unit is measured on the vertical axis and quantity demanded per time period is measured on the horizontal axis, the demand curve is falling from left to right, as shown in Exhibit 5.

Exhibit 5 A Negatively Sloped Demand Curve—The Law of Demand



There are two reasons why a consumer would be expected to purchase more of a good when its price falls and less of a good when its price rises. These two reasons are known as the substitution effect and the income effect of a change in price. We address these two effects separately and then examine the combination of the two.

When the price of something—say, gasoline—falls, that good becomes relatively less costly compared with other goods or services a consumer might purchase. For example, gasoline is used in driving to work, so when its price falls, it is relatively cheaper to drive to work than to take public transportation. Hence, the consumer is likely to substitute a little more driving to work for a little less public transportation. When the price of beef falls, it becomes relatively cheaper than chicken. The typical consumer is, therefore, likely to purchase a little more beef and a little less chicken.

On its own, the substitution effect suggests that when the price of something falls, consumers tend to purchase more of that good. But another influence is often at work as well—the income effect. Consider a consumer spending all of her "money income" on a given combination of goods and services. (Her money income is simply the quantity of dollars or euros, or other relevant currency, that is available to her to spend in any given time period.) Now suppose the price of something she was regularly purchasing falls while her money income and the prices of all other goods remain unchanged. Economists refer to this as an increase in purchasing power or **real income**. For most goods and services, consumers tend to buy more of them when their income rises. So when the price of a good—say, beef—falls, most consumers would tend to buy more beef because of the increase in their real income. Although the consumer's money income (the number on her paycheck) is assumed not to have changed, her real income has risen because she can now buy more beef—and other goods, too—as a result of the fall in the price of that one good. So, quite apart from the substitution effect of a fall in a good's price, the income effect tends to cause consumers to purchase more of that good as well.

Substitution and income effects work the other way, too. If the price of beef were to rise, the substitution effect would cause the consumer to buy less of it and substitute more chicken for the now relatively more expensive beef. Additionally, the rise in the price of beef results in a decrease in the consumer's real income because now she can buy less goods with the same amount of money income. If beef is a good that consumers tend to buy more of when their income rises and less of when their income falls, then the rise in beef price would have an income effect that causes the consumer to buy less of it.

6.1 Normal and Inferior Goods

Economists classify goods on various dimensions, one of which relates to how consumers' purchases of a good respond to changes in consumer income. Earlier, when discussing income elasticity of demand, we introduced the concept of normal goods and inferior goods. For most goods and services, an increase in income would cause consumers to buy more; these are called **normal goods**. But that does not hold true for all goods: There are goods that consumers buy less of when their income rises and goods that they buy more of when their incomes fall. These are called **inferior goods**. This section will contrast normal goods with inferior goods.

We previously discussed income and substitution effects of a change in price. If a good is normal, a decrease in price will result in the consumer buying more of that good. Both the substitution effect and the income effect are at play here:

- A decrease in price tends to cause consumers to buy more of this good in place of other goods—the substitution effect.
- The increase in real income resulting from the decline in this good's price causes people to buy even more of this good when its price falls—the income effect.

So, we can say that for normal goods (restaurant meals, for example, as most people tend to eat out more often when their incomes rise), the substitution and the income effects reinforce one another to cause the demand curve to be negatively sloped.

For inferior goods (cheaper cuts of meat or generic beverages, for example, which most people buy less of as their incomes rise), an increase in income causes consumers to buy less, not more, and if their incomes fall, they buy more, not less. “Inferior” does not imply anything at all about the quality of the good; it is simply used to refer to a good for which an increase in income causes some people to buy less of it.

The same good could be normal for some consumers while it is inferior for others. Consider a very low-income segment of the population. For those consumers, an increase in their income might very well result in their buying more fast-food meals. They might take some of that added income and enjoy eating out at a fast-food restaurant a little more often. Now consider a high-income group. If their income rises, they might be much less inclined to eat at fast-food restaurants and instead do their dining out at a fashionable French bistro, for example. So, fast-food meals might be a normal good for some people and an inferior good for others.

Let us now consider the substitution and income effects of a change in the price of normal and inferior goods. The substitution effect says that if the price of a good falls, the consumer will substitute more of this good in the consumption bundle and buy less of some other good. The substitution effect is true for both normal and inferior goods. Next, we provide an example.

We begin with a hypothetical consumer with a certain money income (R$200,000). Given the prices for all goods, he makes a decision to buy a given amount of Good *X*, coffee. If the price of coffee falls, the consumer is better off than when the price was higher. We can assume that this consumer would have been willing to pay some amount of money each month to be able to buy coffee at the lower price. We now have two states of the world: In State 1, he spends his income on all the various goods, including his desired quantity of coffee at the original price. In State 2, he is able to buy coffee at the new lower price, but because he has paid a portion of his income to buy coffee at the lower price, he now has less money income to spend on all goods combined. If we adjusted the amount of money he would have to pay to lock in the lower price of coffee until he is just indifferent between the two states of the world, we would have exactly offset the “good” thing of the lower price with the “bad” thing of less income. This removes the income effect of the price decrease and allows us to isolate the pure substitution effect. We find that in State 2, he would buy more coffee than in State 1. The pure substitution effect is always in the direction of buying more at the lower relative price.

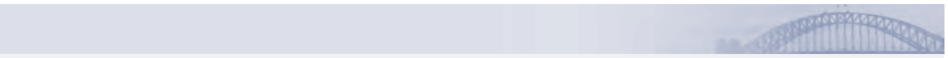
Continuing our example, assume that we give back to the consumer the amount of money he is willing to pay for the privilege of buying coffee at the lower price. Clearly, he is better off because now he can buy coffee at the lower price without having to pay for the privilege. We want to know whether, with this higher money income, he will now buy more or less coffee at the lower price. The answer depends on whether coffee is a normal or an inferior good for this consumer. Recall that for normal goods, an increase in income causes consumers to buy more, but for inferior goods, an increase in income causes consumers to buy less.

In conclusion, the substitution effect of a change in the price of a good will always be in the direction of buying more at a lower price or less at a higher price. The income effect of that same price change, however, depends on whether the good is normal or inferior. If the good is normal, the income effect reinforces the substitution effect, both leading to a negatively sloped demand curve. But if the good is inferior, the income effect and the substitution effect work in opposite directions; the income effect tends to mitigate the substitution effect.

Exhibit 6 summarizes the substitution and income effects for normal and inferior goods.

Exhibit 6 The Substitution and Income Effects of a Price Decrease on Normal and Inferior Goods

	Substitution Effect	Income Effect
Normal good	Buy more because the good is relatively cheaper than its substitutes.	Buy more because the increase in purchasing power raises the total consumption level.
Inferior good	Buy more because the good is relatively cheaper than its substitutes.	Buy less because the increase in real income prompts the consumer to buy less of the inferior good in favor of its preferred substitutes.



Exceptions to the Law of Demand

In virtually every case in the real world, the law of demand holds: A decrease in price results in an increase in quantity demanded, resulting in a negatively sloped demand curve. In a few unusual cases, however, we may find a positively sloped demand curve—a decrease (increase) in price may result in a decrease (increase) in the quantity demanded. These unusual cases are called Giffen goods and Veblen goods.

In theory, it is possible for the income effect to be so strong and so negative as to overpower the substitution effect. In such a case, more of a good would be consumed as the price rises and less would be consumed as the price falls. These goods are called **Giffen goods**, named for Robert Giffen based on his observations of the purchasing habits of the Victorian era poor. For many decades, no one really believed that a Giffen good actually existed anywhere other than in textbooks. But in recent years, studies have documented a few rare cases. One study was conducted in a poor rural community where individuals spend a very large portion of their incomes on rice. For these individuals, rice was an inferior good. Under the law of demand, the quantity of rice purchased would rise with the decline in price, but the rise in quantity would be partially offset by the income effect (a decrease in the amount of rice purchased as a result of rising incomes). What the experimenters discovered, however, was that for a certain subset of consumers, the quantity of rice purchased declined in absolute terms—the income effect actually overwhelmed the substitution effect. For consumers living at subsistence levels—incomes just barely sufficient to enable them to meet their caloric intake needs—a decline in the price of the staple enabled them to shift more of their consumption from rice to the alternate sources of calories in their diet (e.g., meat).

With some goods, the item's price tag itself might drive the consumer's preferences for it. Thorstein Veblen posited just such a circumstance in his concept of conspicuous consumption. According to this way of thinking, a consumer might derive utility out of being known by others to consume a so-called high-status good, such as a luxury automobile or a very expensive piece of jewelry. Importantly, it is the high price itself that partly imparts value to such a good. These are called **Veblen goods**, and they derive their value from the consumption of them as symbols of the purchaser's high status in society; they are certainly not inferior goods. It is argued that by increasing the price of a Veblen good, the consumer would be more inclined to purchase it, not less.

EXAMPLE 2

Income and Substitution Effects of a Decrease in Price

Monica has a monthly entertainment budget that she spends on (a) movies and (b) an assortment of other entertainment items. When the price of each movie is \$8, she spends a quarter of her budget on six movies a month and the rest of her budget on other entertainment. Monica was offered an opportunity to join a movie club at her local theater that allows her to purchase movies at half the regular price, and she can choose each month whether to join the movie club or not. There is a membership fee she must pay for each month she belongs to the club. Monica is exactly indifferent between (a) not buying the membership and, therefore, paying \$8 for movies and (b) buying the membership and paying \$4 per movie. So, she flips a coin each month to determine whether to join the club that month. In months that she does join the club, she sees eight movies. For her birthday, a friend gave her a one-month club membership as a gift, and that month she saw 12 movies.

1 If there were no club and the price of movies were to simply fall from \$8 to \$4, how many more movies would Monica buy each month?

2 Of the increased number of movies Monica would purchase if the price were to fall from \$8 to \$4, determine how much of the increase would be attributable to the substitution effect and how much to the income effect of that price decrease.

3 For Monica, are movies a normal, inferior, or Giffen good?

Solution to 1:

Six movies. When her friend gave her a club membership, she bought 12 movies instead of her usual 6. With the gift of the club membership, Monica could buy movies at a price of \$4 without paying for that privilege. This is the same as if the price of each movie fell from \$8 to \$4.

Solution to 2:

When Monica pays the club membership herself, she buys eight movies, two more than usual. Because Monica is equally well off whether she joins the club for a monthly fee and thereby pays half price or whether she does not join the club and pays full price, we can say that the income effect of the price decrease has been removed by charging her the monthly fee. So the increase from six movies to eight is the result of the substitution effect. When Monica's friend gave her the gift of a club membership, allowing her to pay half price without paying for the privilege, Monica bought 12 movies, 6 more than usual and 4 more than she would have had she paid the membership fee. The increase from 8 movies to 12 is the result of the income effect.

Solution to 3:

When the price fell from \$8 to \$4, Monica bought more movies, so clearly movies are not a Giffen good for her. Additionally, because the substitution effect and the income effect are in the same direction of buying more movies, they are a normal good for Monica. The substitution effect caused her to buy two more movies, and the income effect caused her to buy an additional four movies.

SUPPLY ANALYSIS: COST, MARGINAL RETURN, AND PRODUCTIVITY

7

d describe the phenomenon of diminishing marginal returns;

To fully comprehend the supply side of a consumer goods market, an analyst must understand the firm's costs. (As a reminder, this reading builds on the basics of the market model as covered in the recommended prerequisite reading material.)

The firm's marginal cost is the foundation of the firm's ability and willingness to offer a given quantity for sale, and its costs depend on both the productivity of its inputs and their prices. In this section, we will describe the firm's cost curves—total, average, and marginal costs in both the short run and in the long run—paying special attention to what economists call the **law of diminishing marginal returns**. We will then use this information to explore the conditions under which a firm would find it beneficial to continue operation, even if its economic profits are negative, and at what levels of production its shutdown and breakeven points occur. Long-run costs will be examined in the context of economies and diseconomies of scale.

7.1 Marginal Returns and Productivity

There is an economic phenomenon known as **increasing marginal returns**, in which **marginal product**—the productivity of each additional unit of a resource—increases as additional units of that input are employed.

Initially, a firm can experience increasing returns from adding labor to the production process because of the concepts of specialization and division of labor. At first, by having too few workers relative to total physical capital, the understaffing situation requires employees to multi-task and share duties. As more workers are added, employees can specialize, become more adept at their individual functions, and realize an increase in marginal productivity. But after a certain output level, the law of diminishing marginal returns becomes evident.

When more and more workers are added to a fixed capital base, the marginal return of the labor factor eventually decreases because the fixed input restricts the output potential of additional workers. As an illustration, consider automobile production. When an auto manufacturing plant is operating at full capacity, adding additional labor will not increase production because the physical plant is already 100% employed. More labor hours will merely add to costs without adding to output. Assuming all workers are of equal quality and motivation, the decline in marginal product occurs in the short run, where all other resources (typically, plant size, physical capital, and technology) are fixed.

Marginal returns are directly related to **input productivity**, a measure of the output per unit of input.

7.1.1 *Productivity: The Relationship between Production and Cost*

The cost of producing anything depends on the amount of *inputs*, or *factors of production* (these terms are synonymous), and the input prices. Examples of factors of production are employee hours, machine hours, raw materials, and so on. For simplicity, economists typically concentrate on only two inputs, labor and capital, although obviously there can be many inputs to a particular production process. The labor input is simply employee time, and it is measured as labor hours per time period, such as per week or per month. We denote labor hours as L. If a firm is using two laborers per week and each laborer works 35 hours per week, then L equals 70 labor hours per week. We denote hours of capital as K. If the firm is using three machines and each one is used for 12 hours per week, then K equals 36 machine hours per week. That is, the capital input is measured as machine hours used per time period. In this way, capital and labor are stated in similar terms. They represent flows of services—labor hours and machine hours—that are used to produce a flow of output per time period.

Accordingly, the respective input prices would be the wage rate per labor hour (we use w to denote wage rate) and the rental rate per machine hour (we use r to denote the rental rate per machine hour). It is helpful to think of a firm as renting the services of labor and of machines. Although the firm might own its own machines, it could in theory rent its machines out to another user, so it is forgoing the rate it could earn elsewhere when it is using its machines internally instead of renting them out. So, a firm is not using its own machines "for free." It is incurring the **opportunity cost** of not being able to rent those machines to another user.

The **total cost** of production (TC) is the number of hours of labor multiplied by the wage rate plus the number of machine hours multiplied by the rental rate of machines:

$$\text{TC} = (w)(L) + (r)(K)$$

This formula illustrates that the total cost is just the cost of all the firm's inputs. It is not a cost function, however, which is a relationship between the cost of production and the flow of output. The cost function $C = f(Q)$, where (Q) denotes the flow of output in units of production per time period, relates the production cost per time period to the number of units of output produced per time period.

Two things could cause the cost of producing any given level of output to fall: Either the price of one or both inputs could fall or the inputs themselves could become more productive and less of them would be needed (e.g., a worker is more productive when fewer hours of labor are needed to produce the same output). The reverse is true also: A rise in cost could result from either a rise in input prices or a fall in input productivity, or both.

Why is productivity important? Cost-minimization and profit-maximization behavior dictate that the firm strives to maximize productivity—for example, produce the most output per unit of input or produce any given level of output with the least amount of inputs. A firm that lags behind the industry in productivity is at a competitive disadvantage and is likely to face decreases in future earnings and shareholders' wealth. An increase in productivity lowers production costs, which leads to greater profitability and investment value. These productivity benefits can be fully or partially distributed to other stakeholders of the business, such as to consumers in the form of lower prices and to employees in the form of enhanced compensation. Transferring some or all of the productivity rewards to non-equity holders creates synergies that benefit shareholders over time.

The benefits from increased productivity are as follows:

- lower business costs, which translate into increased profitability;
- an increase in the market value of equity and shareholders' wealth resulting from an increase in profit; and
- an increase in worker rewards, which motivates further productivity increases from labor.

Undoubtedly, increases in productivity reinforce and strengthen the competitive position of the firm over the long run. A fundamental analysis of a company should examine the firm's commitment to productivity enhancements and the degree to which productivity is integrated into the competitive nature of the industry or market. In some cases, productivity is not only an important promoter of growth in firm value over the long term but is also the key factor for economic survival. A business that lags the market in terms of productivity often finds itself less competitive, while at the same time confronting profit erosion and deterioration in shareholders' wealth. Typical productivity measures for a firm are based on the concepts of total product, average product, and marginal product of labor.

7.1.2 *Total, Average, and Marginal Product of Labor*

When measuring a firm's operating efficiency, it is easier and more practical to use a single resource factor as the input variable rather than a bundle of the different resources that the firm uses in producing units of output. As discussed in the previous section, labor is typically the input that is the most identifiable and calculable for measuring productivity. However, any input that is not difficult to quantify can be used. As an example, a business that manually assembles widgets has 50 workers, one production facility, and an assortment of equipment and hand tools. The firm would like to assess its productivity when using these three input factors to produce widgets. In this example, it is most appropriate to use labor as the input factor for determining productivity because the firm uses only one (fixed) plant building and a variety of other physical capital.

We will use labor as the input variable to illustrate the concepts of total product, average product, and marginal product. Exhibit 7 provides a summary of these three concepts.

Exhibit 7 Definitions and Calculations for Total, Marginal, and Average Product of Labor

Term	Calculation
Total product	Sum of the output from all inputs during a time period; usually illustrated as the total output (Q) using labor quantity (L)
Average product	Total product divided by the quantity of a given input; measured as total product divided by the number of worker hours used at that output level (Q/L)
Marginal product	The amount of additional output resulting from using one more unit of input assuming other inputs are fixed; measured by taking the difference in total product and dividing by the change in the quantity of labor ($\Delta Q/\Delta L$)

Total product (Q) is defined as the aggregate sum of production for a firm during a time period. As a measure of productivity, total product provides superficial information about how effective and efficient a firm is in terms of producing output. For instance, three firms—Company A, Company B, and Company C—that make up an entire industry have total output levels of 100,000 units, 180,000 units, and 200,000 units, respectively. Obviously, Company C dominates the market with a 41.7% share, followed by Company B's 37.5% share and Company A's 20.8% portion of the market. However, this information says little about how efficient each firm is in generating its total output level. Total product only provides insight into a firm's production volume relative to the industry; it does not show how efficient a firm is in producing its output.

Average product of labor (AP_L) measures the productivity of an input (in this case, labor) on average and is calculated by dividing total product by the total number of units for the given input that is used to generate that output. Average product is usually measured on the basis of the labor input. It is a representative or overall measure of labor's productivity: Some workers are more productive than average, and others are less productive than average.

Exhibit 8 compares the productivity of the three firms introduced earlier. Company A employs 100 worker hours and produces 100,000 widgets per hour. Company B employs 200 worker hours and produces 180,000 widgets per hour. Company C employs 250 worker hours and produces 200,000 widgets per hour.

Exhibit 8 Comparing Productivity

	Output (Q)	Number of Worker Hours (L)	Average Product of Labor (AP_L)
Company A	100,000	100	1,000
Company B	180,000	200	900
Company C	200,000	250	800

Using this metric, it is apparent that Company A, with AP_L equal to 1,000, is the most efficient firm, despite having the lowest market share. Company C has the largest market share, but it is the least efficient of the three, with AP_L equal to 800. Assuming that Company A can maintain its productivity advantage over the long run, it will be positioned to generate the greatest return on investment through lower costs and higher profit outcomes relative to the other firms in the market.

Marginal product of labor (MP_L), also known as *marginal return*, measures the productivity of each additional unit of input and is calculated by observing the difference in total product when adding another unit of input (assuming other resource quantities are held constant). It is a gauge of the productivity of the individual additional worker hour rather than an average across all workers.

Exhibit 9 provides a numerical illustration for total, average, and marginal products of labor.

Exhibit 9 Total, Average, and Marginal Product of Labor

Labor (L)	Total Product (Q_L)	Average Product (AP_L)	Marginal Product (MP_L)
0	0	—	—
1	100	100	100
2	210	105	110
3	300	100	90
4	360	90	60
5	400	80	40
6	420	70	20
7	350	50	–70

Total product increases as the firm adds each additional hour of labor—until the seventh labor hour, at which point total production declines by 70 units. Obviously, the firm would want to avoid negative worker productivity.

At an employment level of five labor hours, AP_L is 80 units (400/5) and MP_L is 40 units [(400 – 360)/(5 – 4)]. The average productivity for all five labor hours is 80 units, but the productivity of the fifth labor hour is only 40 units.

EXAMPLE 3

Calculation and Interpretation of Total, Average, and Marginal Product

Exhibit 10 illustrates the production relationship between the number of machine hours and total product.

1. Interpret the results for total, average, and marginal product.
2. Indicate at what point increasing marginal returns change to diminishing marginal returns.

Exhibit 10

Machine Hours (K)	Total Product (Q_K)	Average Product (AP_K)	Marginal Product (MP_K)
0	0	—	—
1	1,000	1,000	1,000
2	2,500	1,250	1,500
3	4,500	1,500	2,000
4	6,400	1,600	1,900
5	7,400	1,480	1,000
6	7,500	1,250	100
7	7,000	1,000	–500

Solution to 1:

Total product increases up to six machine hours, where it tops out at 7,500. Because total product declines from Hour 6 to Hour 7, the marginal product for Machine Hour 7 is negative 500 units. Average product peaks at 1,600 units with four machine hours. Average product increases at a steady pace with the addition of Machine Hours 2 and 3. The addition of Machine Hour 4 continues to increase average product but at a decreasing rate. Beyond four machine hours, average product decreases—at an increasing rate. Marginal product peaks with Machine Hour 3 and decreases thereafter.

Solution to 2:

The marginal product, MP_K, of Machine Hour 3 is 2,000. The marginal product of each additional machine hour beyond Machine Hour 3 declines. Diminishing marginal returns are evident beyond Machine Hour 3.

A firm has a choice of using total product, average product, marginal product, or some combination of the three to measure productivity. Because total product is simply an indication of a firm's output volume and potential market share, average product and marginal product are better gauges of a firm's productivity. Both can reveal competitive advantage through production efficiency. However, individual worker productivity is not easily measurable when workers perform tasks collectively. In this case, average product is the preferred measure of productivity performance.

Referring to the total product column in Exhibit 9, output is more than twice as great (210 widgets) when two hours of labor are used as opposed to only one hour (100 widgets.) In this range of production, there is an increase in return when employee hours are added to the production process. This is the phenomenon of increasing marginal returns.

8 ECONOMC PROFIT VERSUS ACCOUNTING PROFIT

e determine and interpret breakeven and shutdown points of production;

Two important considerations of any firm are its level of profitability and whether to continue to operate in the current environment. Economists define profit differently than do accountants. **Economic profit** is defined as the difference between total revenue (TR) and total **economic costs**. **Accounting profit** is the difference between TR and total **accounting cost**. TR is the same from both an accounting standpoint and an economic standpoint; it is derived by multiplying the selling price per unit of output by the number of units: TR = $(P)(Q)$. The difference between the two measures of profit, therefore, lies in an understanding of economic cost (also called "opportunity cost," which is defined in detail in the next section).

8.1 Economic Cost vs. Accounting Cost

The opportunity cost of any particular decision, such as to produce a given level of output, can be determined by measuring the benefit forgone by not implementing the next best alternative. Suppose that a firm is currently operating with hired labor and its own plant and equipment to produce output at some level. The firm must continuously decide either to keep the level of output the same or to change it. The decision to maintain the same output requires that the firm hire the same amount of labor input and use the same level of its capital inputs as before. The labor expense is both an economic cost and an accounting cost because the money spent on labor hours could have been used for something else (opportunity cost), and it is also a current expense for the firm (accounting cost.)

Accountants typically attempt to recognize the cost of plant and equipment in the form of accounting depreciation, which is a means of distributing the historical cost of the fixed capital among the units of production for financial reporting purposes. The money spent in the past on the firm's plant and equipment is what economists call a "sunk cost." Because sunk costs cannot be altered, they cannot affect an optimal decision, which is forward looking. Sunk costs are therefore ignored, and the key management question is, Going forward, what are the opportunity costs and benefits of maintaining a given level of output?

Here is where economic depreciation comes into play. To understand the opportunity costs of using our plant and equipment—already bought and paid for—for one more period of time to produce output, we have to ask the question, What else could be done with that fixed capital if it were not used to produce our output? The answer might be that because there is no external market for our machines and buildings, we are forgoing nothing by using it to produce output. Or it might be that there is a market where we could rent out or sell our capital equipment elsewhere instead of using it to produce output. That rental rate is the economic depreciation associated with using our own equipment to produce output instead of renting or selling it elsewhere.

Economic depreciation is forward looking. It asks, What am I giving up if I use my resources to produce output in the coming period? Accounting depreciation is backward looking. It asks, How should I distribute the historical cost—that I have already paid—across units of output that I intend to produce this period? Both concepts are useful—one for making managerial decisions about output and the other as a way spreading historical costs for reporting or tax purposes—but there is not necessarily a direct relationship between the two.

MARGINAL REVENUE, MARGINAL COST AND PROFIT MAXIMIZATION; SHORT-RUN COST CURVES: TOTAL, VARIABLE, FIXED, AND MARGINAL COSTS

9

e determine and interpret breakeven and shutdown points of production;

It is assumed that any for-profit firm's management is tasked with achieving the goal of shareholder wealth maximization. Put most simply, that translates into the goal of economic profit maximization. Hereafter, when the word *profit* is used, it will be economic profit that we have in mind. Because profit is defined as TR minus TC, anything that increases revenue more than cost or decreases cost more than revenue will increase profit. Before we address profit maximization, we must introduce two important concepts: marginal revenue and marginal cost.

Marginal revenue (MR) is the additional revenue the firm realizes from the decision to increase output by one unit per time period. That is, MR = ΔTR/ΔQ. If the firm is operating in what economists call a perfectly competitive market, it is one of many sellers of identical products in an environment characterized by low or non-existent barriers to entry. Under perfect competition, the firm has no pricing power because there are many perfect substitutes for the product it sells. If it were to attempt to raise the price even by a very small amount, it would lose all of its sales to competitors. On the other hand, it can sell essentially any amount of product it wants without lowering the price below the market price.

Take the wheat market as an example of a perfectly competitive market. A seller of wheat would have no control over the market price of wheat; thus, because TR = (P)(Q), MR for this firm is simply price per unit of output. This firm is said to face a perfectly horizontal (zero-sloped), or infinitely elastic, demand curve for its product. For example, if the firm is selling 1,000 bushels of wheat per week at a price of £3 per bushel, TR is £3,000. If the firm were to increase its output by one unit, then TR would rise by exactly £3 because the firm would not have to lower its price to sell that added unit. So, for sellers in a market with perfect competition, MR = P.

In contrast, if a firm sells a product that is differentiated from other firms' products and that has a large market share, the firm is said to be operating in an environment of imperfect competition. In the extreme case of imperfect competition, there might be only one firm selling a product with no close substitutes. That firm holds a monopoly, and it is subject to the market demand curve for its product. Whether a monopoly or simply operating under imperfect competition, the firm faces a negatively sloped demand curve and must lower its price to sell another unit. Thus, MR will be lower than price.

To illustrate this concept, we will decompose MR. Recall from earlier in the reading that

TR = (P)(Q)

and

$$\text{MR} = \Delta\text{TR}/\Delta Q$$

Change in total revenue (ΔTR), the numerator of the ratio, can be written as $(P)(\Delta Q) + (Q)(\Delta P)$.

There are two competing forces affecting revenue: (1) Additional units are sold at the new price, and (2) all units must now be sold at the lower price. The firm is selling more units, but it is selling all units at a lower price than before.

To find MR, we divide the change in TR by the change in quantity:

$$\text{MR} = \frac{(P)(\Delta Q)}{\Delta Q} + \frac{(Q)(\Delta P)}{\Delta Q} = P + Q\frac{(\Delta P)}{\Delta Q}$$

In other words, MR is equal to price but with an "adjustment" equal to $(Q)(\Delta P/\Delta Q)$.

Taking this one step further, recall that earlier we said $(\Delta P/\Delta Q)$ is the slope of the demand curve. From our expression just given, $\text{MR} = P + Q(\Delta P/\Delta Q)$; so, MR is equal to price with an adjustment equal to quantity times the slope of the demand curve.

A perfectly competitive firm faces a demand curve with a slope of zero. Substituting 0 for $\Delta P/\Delta Q$ into the expression given, it becomes clear that MR is equal to price for the perfectly competitive firm—it need not lower its price to sell an additional unit. For a firm in an imperfectly competitive market, however, the demand curve is negatively sloped ($\Delta P/\Delta Q < 0$). Substituting this negative number into the expression for MR, $P + Q(\Delta P/\Delta Q)$, it becomes clear that MR for an imperfectly competitive firm is less than price.

Marginal cost (MC) is the increase to total cost resulting from the firm's decision to increase output by one additional unit per time period: $\text{MC} = \Delta\text{TC}/\Delta Q$. Economists distinguish between short-run marginal cost (SMC) and long-run marginal cost (LMC). Labor is variable over the short run, but the quantity of capital cannot be changed in the short run because there is a lead time required to build or buy new plant equipment and put it in place. In the long run, all inputs are variable.

SMC is essentially the additional cost of the variable input, labor, that must be incurred to increase the level of output by one unit. LMC is the additional cost of all inputs necessary to increase the level of output, allowing the firm the flexibility of changing both labor and capital inputs in a way that maximizes efficiency.

Understanding MC is aided by recalling that cost is *directly* related to input prices and *inversely* related to productivity. For example, if the wage rate were to rise, cost would also rise. If labor were to become more productive, cost would fall. This relationship can be captured in an expression that relates SMC to wage rate (w) and MP_L: $\text{SMC} = w/\text{MP}_L$.

This relationship between cost and productivity also holds with average variable cost. **Variable costs** are all costs that fluctuate with the level of production and sales. **Average variable cost** (AVC) is the ratio of total variable cost to total output: $\text{AVC} = \text{TVC}/Q$. Again, if labor's wage rises, AVC also rises; but if labor were to become more productive, AVC falls. This relationship is captured by the expression $\text{AVC} = w/\text{AP}_L$.

Earlier, we noted that over some range of low output, the firm might benefit from increasing marginal productivity of its labor input as workers begin to specialize. As the MP_L increases, SMCs decline. Eventually, as more and more labor is added to a fixed amount of capital, the MP_L must fall, causing SMCs to rise.

We began this section by stating that the goal of management is to maximize profit. We now address the conditions necessary for reaching that goal. Consider a firm currently producing 1,000 widgets each week and whose management is contemplating increasing that output incrementally. Would that additional unit increase profit? Clearly, profit would be increased (or losses reduced) if the additional revenue from that next unit were greater than the additional cost. So, a profit-seeking firm should increase Q if MR > MC. Conversely, if the additional unit added more to cost than to revenue, the firm should reduce output because it would save more in cost than it would lose in revenue. Only if the additional cost were exactly equal to the additional revenue would the firm be maximizing its profit.

There is another condition (called a second-order condition) necessary for profit maximization: At the level of output at which MR = MC, MC cannot be falling. This condition is fairly intuitive. If MC is falling with additional output, MP_L would be rising. (Recall that SMC = w/MP_L) If one additional hour of labor input causes MC to fall, the firm would want to add that hour and continue adding labor until SMC becomes positively sloped. We can sum up the profit-maximization decision for an operating firm as follows: Produce the level of output such that (1) MR = MC and (2) MC is not falling.

9.1 Understanding the Interaction between Total, Variable, Fixed, and Marginal Cost and Output

Exhibit 11 shows the graphical relationships between total cost, total fixed cost, and total variable cost. TC is the summation of all costs, where costs are classified on the basis of whether they are fixed or variable. **Total fixed cost** (TFC) is the summation of all expenses that do not change as the level of production varies. **Total variable cost** (TVC) is the summation of all variable expenses; TVC rises with increased production and falls with decreased production. At zero production, TC is equal to TFC because TVC at this output level is zero. The curve for TC always lies parallel to and above the TVC curve by the amount of TFC.

Exhibit 11 Total Cost, Total Variable Cost, and Total Fixed Cost

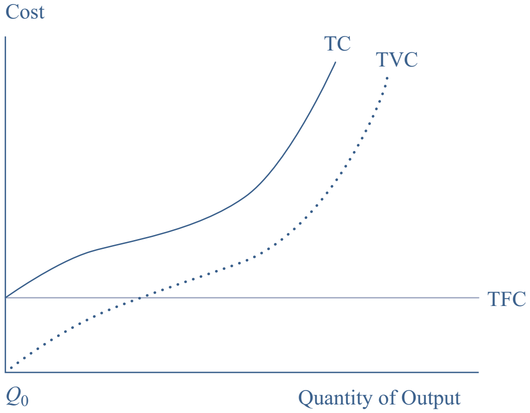
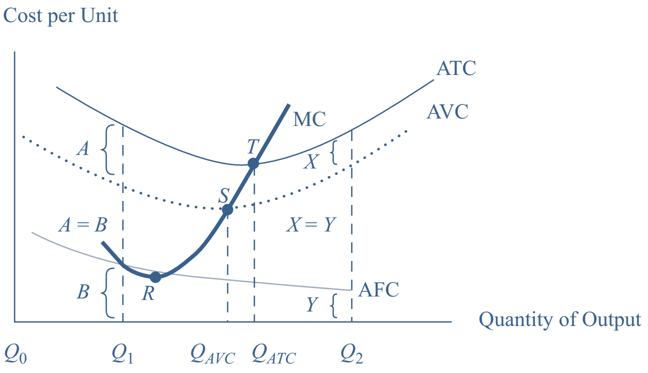


Exhibit 12 shows the relationships between the **average total cost** (ATC), average variable cost (AVC), **average fixed cost** (AFC), and marginal cost (MC) curves in the short run. As output quantity increases, AFC declines because TFCs are spread over a larger number of units. Both ATC and AVC take on a bowl-shaped pattern in which each curve initially declines, reaches a minimum average cost output level, and then increases after that point. The MC curve intersects both the ATC and the AVC at their minimum points—points *S* and *T*. When MC is less than AVC, AVC will be decreasing. When MC is greater than AVC, AVC will be increasing.

Exhibit 12 Average Total Cost, Average Variable Cost, Average Fixed Cost, and Marginal Cost



S, the lowest point on the AVC curve, is where MC equals AVC. Beyond quantity Q_{AVC}, MC is greater than AVC; thus, the AVC curve begins to rise. Note that it occurs at a quantity lower than the minimum point on the ATC curve.

T, the lowest point on the ATC curve, is where MC equals ATC. Beyond quantity Q_{ATC}, MC is greater than ATC; thus, the ATC curve is rising.

A, the difference between ATC and AVC at output quantity Q_1, is the amount of AFC.

R indicates the lowest point on the MC curve. Beyond this point of production, fixed input constraints reduce the productivity of labor.

X indicates the difference between ATC and AVC at quantity Q_2. It is less than *A* because AFC (*Y*) falls with output.

Exhibit 13 shows an example of how total, average, and marginal costs are derived. TC is calculated by summing TFC and TVC. MC is derived by observing the change in TC as the quantity variable changes. There is a relationship that always holds for average and marginal costs: If MC is less than average cost, average cost must fall, and if MC is greater than average cost, average cost must rise. For example, in Exhibit 13, AVC begins to increase as output rises from 2 to 3 units because MC (50) is greater than AVC (41.7). Also from Exhibit 13, ATC declines up to 3 units because MC is less than ATC. After 3 units, ATC increases because the MC of Unit 4 (85) exceeds the ATC of all prior units (75). Initially, the MC curve declines because of increasing marginal returns to labor, but at some point, it begins to increase because of the law of diminishing marginal returns.

Exhibit 13 Total, Average, Marginal, Fixed, and Variable Costs

Quantity (*Q*)	TFC[a]	AFC	TVC	AVC	TC	ATC	MC
0	100	—	0	—	100	—	—
1	100	100.0	50	50.0	150	150.0	50
2	100	50.0	75	37.5	175	87.5	25
3	100	33.3	125	41.7	225	75.0	50
4	100	25.0	210	52.5	310	77.5	85
5	100	20.0	300	60.0	400	80.0	90
6	100	16.7	450	75.0	550	91.7	150
7	100	14.3	650	92.9	750	107.1	200
8	100	12.5	900	112.5	1,000	125.0	250

Exhibit 13 (Continued)

Quantity (*Q*)	TFC[a]	AFC	TVC	AVC	TC	ATC	MC
9	100	11.1	1,200	133.3	1,300	144.4	300
10	100	10.0	1,550	155.0	1,650	165.0	350

[a] Includes all opportunity costs.

As stated earlier, TC increases as the firm expands output and decreases when production is cut. TC increases at a decreasing rate up to a certain output level. Thereafter, the rate of increase accelerates as the firm gets closer to full utilization of capacity. The rate of change in TC mirrors the rate of change in TVC. In Exhibit 13, TC at 5 units is 400—of which 300 is variable cost and 100 is fixed cost. At 10 units, TC is 1,650—of which 1,550 is variable cost and 100 is fixed cost.

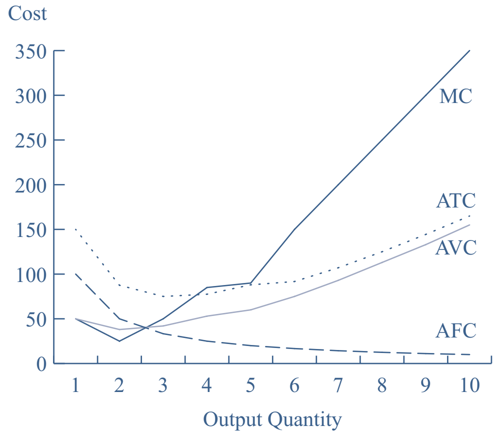
Fixed costs typically are incurred whether the firm produces anything or not. Fixed costs may stay the same over a given range of production but can change to another constant level when production moves outside of that range. The latter is referred to as a **quasi-fixed cost**, although it remains categorized as part of TFC. Examples of fixed costs are debt service, real estate lease agreements, and rental contracts. **Normal profit** is also considered to be a fixed cost because it is a return required by investors on their equity capital regardless of output level. Quasi-fixed cost examples would be certain utilities and administrative salaries that could be lower or avoided altogether when output is zero but would rise to higher constant levels over different production ranges.

Other fixed costs evolve primarily from investments in such fixed assets as real estate, production facilities, and equipment. These fixed costs cannot be arbitrarily cut when production declines. When a firm downsizes, the last expense to be cut is usually fixed cost.

TVC has a direct relationship with quantity. When quantity increases, TVC increases; when quantity decreases, TVC declines. At zero production, TVC is always zero. Variable cost examples are payments for labor, raw materials, and supplies. The change in TVC declines up to a certain output point and then increases as production approaches capacity limits. In Exhibit 13, TVC increases with an increase in quantity. However, the change from 1 to 2 units is 25 (75 – 50), and the change from 9 to 10 units is 350.

Exhibit 14 illustrates the relationships between MC, ATC, AVC, and AFC for the data presented in Exhibit 13.

Exhibit 14 Average Total Cost, Average Variable Cost, Average Fixed Cost, and Marginal Cost for Exhibit 13 Data



Dividing TFC by quantity yields AFC. AFC decreases throughout the production span, reflecting the spreading of a constant cost over more and more production units. At high production volumes, AFC may be so low that it is a small proportion of ATC. In Exhibit 13, AFC declines from 100 at 1 unit to 20 at 5 units, and then to 10 at an output level of 10 units.

In Exhibit 13, AVC at 5 units is 60 (300/5). Over an initial range of production, AVC declines and then reaches a minimum point. Thereafter, AVC increases as the firm uses more of its production capacity. This higher cost results primarily from production constraints imposed by the fixed assets at higher volume levels. The lowest AVC quantity does not correspond to the least-cost quantity for ATC because AFC is still declining. In Exhibit 13, AVC is minimized at 2 units, whereas ATC is minimized at 3 units.

ATC is calculated by dividing TC by quantity (or by summing AFC and AVC). In Exhibit 13, at 3 units, ATC is 75 (TC of 225/3 units of production or AFC of 33.3 + AVC of 41.7). This is the least average cost point of production and the minimum point on the ATC curve. Although cost-minimizing behavior on the part of the firm would dictate operating at the minimum point on its ATC curve, the profit-maximizing quantity may not correspond to this minimum ATC point. Profit per unit, but not necessarily total profit, is maximized at this point.

EXAMPLE 4

Calculation and Interpretation of Total, Average, Marginal, Fixed, and Variable Costs

The first three columns of Exhibit 15 display data on quantity, TFC, and TVC, which are used to calculate TC, AFC, AVC, ATC, and MC. Examine the results for total, average, marginal, fixed, and variable costs. Identify the quantity levels at which the ATC, AVC, and MC values reach their minimum points. Explain the relationship between TFC and TC at a quantity of zero output.

Exhibit 15

Q	TFC[a]	TVC	AFC	AVC	TC	ATC	MC
0	5,000	0	—	—	5,000	—	—
1	5,000	2,000	5,000.0	2,000	7,000	7,000.0	2,000
2	5,000	3,800	2,500.0	1,900	8,800	4,400.0	1,800
3	5,000	5,400	1,666.7	1,800	10,400	3,466.7	1,600
4	5,000	8,000	1,250.0	2,000	13,000	3,250.0	2,600
5	5,000	11,000	1,000.0	2,200	16,000	3,200.0	3,000
6	5,000	15,000	833.3	2,500	20,000	3,333.3	4,000
7	5,000	21,000	714.3	3,000	26,000	3,714.3	6,000
8	5,000	28,800	625.0	3,600	33,800	4,225.0	7,800
9	5,000	38,700	555.6	4,300	43,700	4,855.6	9,900
10	5,000	51,000	500.0	5,100	56,000	5,600.0	12,300

[a] Includes all opportunity costs

Solution:

TFC remains unchanged at 5,000 throughout the entire production range, whereas AFC continuously declines from 5,000 at 1 unit to 500 at 10 units. Both AVC and MC initially decline and then reach their lowest level at 3 units, with costs of 1,800 and 1,600, respectively. Beyond 3 units, both AVC and MC increase, indicating that the cost of production rises with greater output. The least-cost point for ATC is 3,200 at 5 units. At zero output, TC is 5,000, which equals the amount of TFC (at zero output, the firm will need no variable inputs, but it is committed to its fixed plant and equipment in the short run).

PERFECT AND IMPERFECT COMPETITION, PROFIT MAXIMIZATION

10

e determine and interpret breakeven and shutdown points of production;

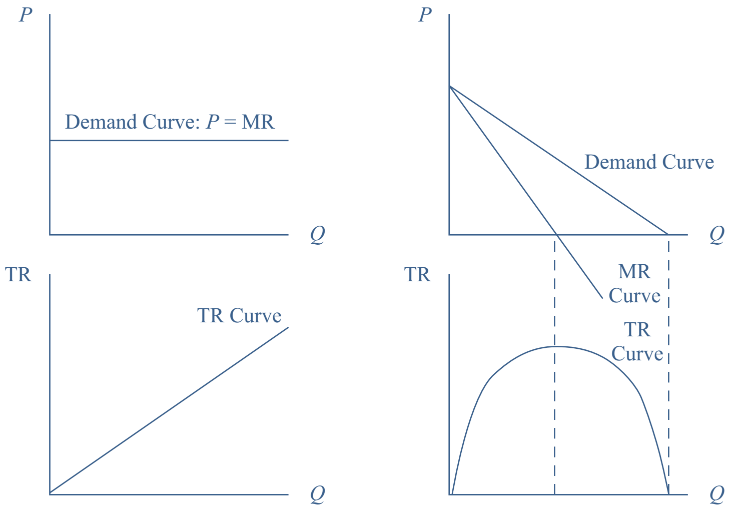
Recall from our earlier discussion of profit-maximizing conditions that a firm can generally be classified as operating in either a perfectly competitive or an imperfectly competitive environment. The difference between the two manifests itself in the slope of the demand curve facing the firm. If the environment of the firm is perfectly competitive, it must take the market price of its output as given, so it faces a perfectly elastic, horizontal demand curve. In this case, as we saw previously, the firm's MR and the price of its product are identical. Additionally, the firm's **average revenue** (AR), or revenue per unit, is also equal to price per unit. However, a firm that faces a negatively sloped demand curve must lower its price to sell an additional unit, so its MR is less than price (*P*).

These characteristics of MR are also applicable to the TR functions. Under conditions of perfect competition, TR (as always) is equal to price times quantity: TR = (*P*)(*Q*). But under conditions of perfect competition, price is dictated by the market; the firm has no control over price. As the firm sells one more unit, its TR rises by the exact amount of price per unit.

Under conditions of imperfect competition, price is a variable under the firm's control, and therefore price is a function of quantity: P = f(Q), and TR = f(Q) × Q. For simplicity, suppose the firm is monopolistic and faces the market demand curve, which we will assume is linear and negatively sloped. Because the monopolist is the only seller, its TR is identical to the total expenditure of all buyers in the market. Earlier, we noted what happens as price is reduced and quantity sold increases in this environment: At first, a decrease in price increases total expenditure by buyers and TR to the firm because the decrease in price is outweighed by the increase in units sold. But as price continues to fall, the decrease in price overshadows the increase in quantity, and total expenditure (revenue) falls. We can now depict the demand and TR functions for firms under conditions of perfect and imperfect competition, as shown in Exhibit 16.

Exhibit 16 Demand and Total Revenue Functions for Firms under Conditions of Perfect and Imperfect Competition

A. Perfectly Competitive Firm *B. Imperfectly Competitive Firm*

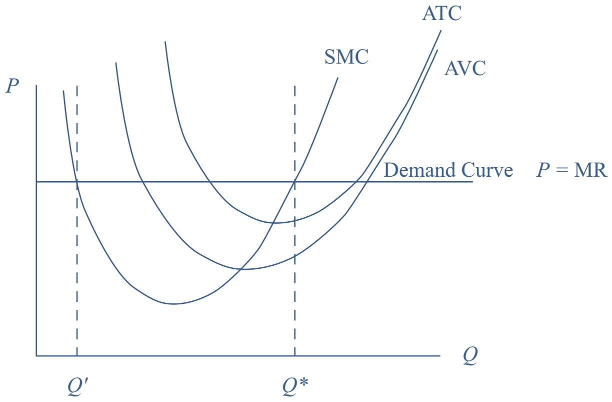


Panel A of Exhibit 16 depicts the demand curve (upper graph) and total revenue curve (lower graph) for the firm under conditions of perfect competition. Notice that the vertical axis in the upper graph is price per unit (e.g., £/bushel), whereas TR is measured on the vertical axis in the lower graph (e.g., £/week.) The same is true for the respective axes in Panel B, which depicts the demand and total revenue curves for the monopolist. The TR curve for the firm under conditions of perfect competition is linear, with a slope equal to price per unit. The TR curve for the monopolist first rises (in the range where MR is positive and demand is elastic) and then falls (in the range where MR is negative and demand is inelastic) with output.

10.1 Profit-Maximization, Breakeven, and Shutdown Points of Production

We can now combine the firm's short-run TC curves with its TR curves to represent profit maximization in the cases of perfect competition and imperfect competition. Exhibit 17 shows both the AR and average cost curves in one graph for the firm under conditions of perfect competition.

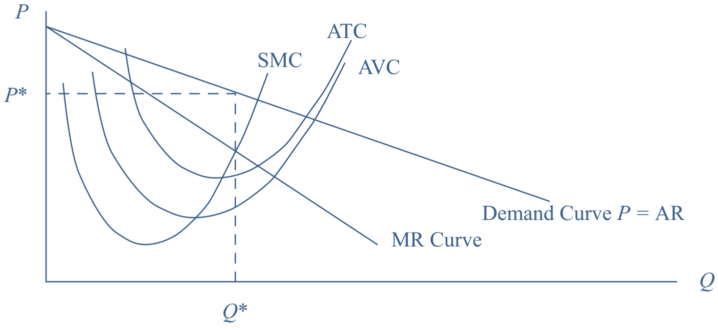
Exhibit 17 Demand and Average and Marginal Cost Curves for the Firm under Conditions of Perfect Competition



The firm is maximizing profit by producing Q^*, where price is equal to SMC and SMC is rising. (Note that there is another output level, Q', where P = SMC, but at that point, SMC is still falling, so this cannot be a profit-maximizing solution.) If market price were to rise, the firm's demand and MR curve would simply shift upward, and the firm would reach a new profit-maximizing output level to the right of Q^*. If, on the other hand, market price were to fall, the firm's demand and MR curve would shift downward, resulting in a new and lower level of profit-maximizing output. As depicted, this firm is currently earning a positive economic profit because market price exceeds ATC at output level Q^*. This profit is possible in the short run, but in the long run, competitors would enter the market to capture some of those profits and would drive the market price down to a level equal to each firm's ATC.

Exhibit 18 depicts the cost and revenue curves for the monopolist that is facing a negatively sloped market demand curve. The MR and demand curves are not identical for this firm. But the profit-maximizing rule is still the same: Find the level of Q that equates SMC to MR—in this case, Q^*. Once that level of output is determined, the optimal price to charge is given by the firm's demand curve at P^*. This monopolist is earning positive economic profit because its price exceeds its ATC. The barriers to entry that give this firm its monopolistic power mean that outside competitors would be unable to compete away this firm's profits.

Exhibit 18 Demand and Average and Marginal Cost Curves for the Monopolistic Firm



11

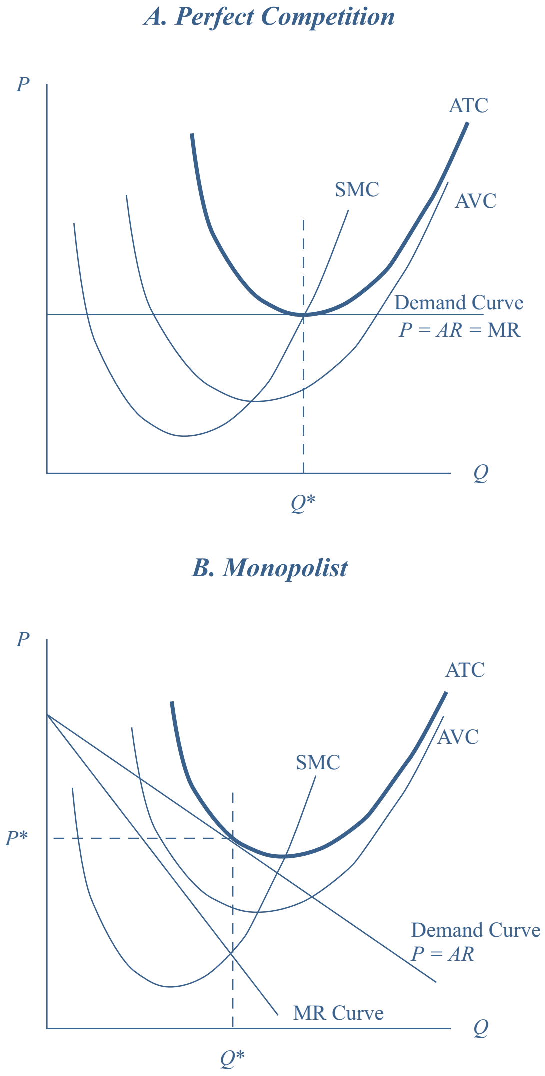
BREAKEVEN ANALYSIS AND SHUTDOWN DECISION

e determine and interpret breakeven and shutdown points of production;

A firm is said to break even if its TR is equal to its TC. It can also be said that a firm breaks even if its price (AR) is exactly equal to its ATC, which is true under conditions of perfect and imperfect competition. Of course, the goal of management is not just to break even but to maximize profit. However, perhaps the best the firm can do is cover all of its economic costs. Economic costs are the sum of total accounting costs and implicit opportunity costs. A firm whose revenue is equal to its economic costs is covering the opportunity cost of all of its factors of production, including capital. Economists would say that such a firm is earning normal profit, but not positive economic profit. It is earning a rate of return on capital just equal to the rate of return that an investor could expect to earn in an equivalently risky alternative investment (opportunity cost). Firms that are operating in a very competitive environment with no barriers to entry from other competitors can expect, in the long run, to be unable to earn a positive economic profit; the excess rate of return would attract entrants who would produce more output and ultimately drive the market price down to the level at which each firm is at best just earning a normal profit. This situation, of course, does not imply that the firm is earning zero accounting profit.

Exhibit 19 depicts the condition for both a firm under conditions of perfect competition (Panel A) and a monopolist (Panel B) in which the best each firm can do is to break even. Note that at the level of output at which SMC is equal to MR, price is just equal to ATC. Hence, economic profit is zero, and the firms are breaking even.

Exhibit 19 Examples of Firms under Perfect Competition and Monopolistic Firms That Can, at Best, Break Even



11.1 The Shutdown Decision

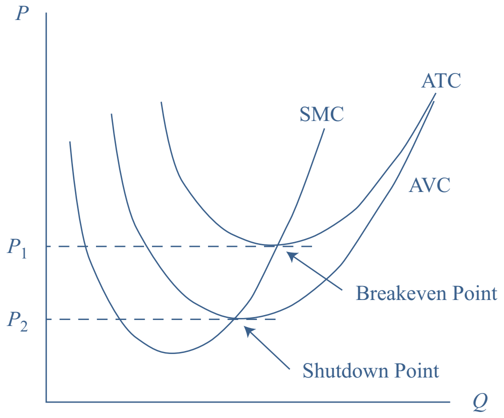
In the long run, if a firm cannot earn at least a zero economic profit, it will not operate because it is not covering the opportunity cost of all of its factors of production, labor, and capital. In the short run, however, a firm might find it advantageous to continue to operate even if it is not earning at least a zero economic profit. The discussion that follows addresses the decision to continue to operate and earn negative profit or shut down operations.

Recall that typically some or all of a firm's fixed costs are incurred regardless of whether the firm operates. The firm might have a lease on its building that it cannot avoid paying until the lease expires. In that case, the lease payment is a sunk cost: It cannot be avoided, no matter what the firm does. Sunk costs must be ignored in the decision to continue to operate in the short run. As long as the firm's revenues cover at least its variable cost, the firm is better off continuing to operate. If price is greater than AVC, the firm is not only covering all of its variable cost but also a portion of fixed cost.

For example, suppose a firm is producing 100 widgets and selling them at a price of €4 each. Obviously, its TR is €400 per time period. Suppose, also, that at that level of output, its ATC is €7, made up of AVC of €3.75 plus AFC of €3.25 per period. This firm is said to be earning negative economic profit (also referred to as **economic loss**, a condition in which revenues fall short of total opportunity cost) of €300 because its TC is €700 and its TR is only €400. Should this firm shut down immediately? If the fixed cost is unavoidable, then the firm is obligated to pay it whether it operates or not. The TFC is €325 (€3.25 per unit on 100 units). If it shuts down and earns zero revenue, then its variable cost would be zero but its losses would still equal the €325 of unavoidable fixed cost. If, however, it continued to operate, it could earn revenue of €400 that would cover its variable cost of €375 and contribute €25 toward the fixed costs. In other words, this firm would lose less by continuing to operate (€300) than by shutting down (€325).

In the long run, unless market price increases, this firm would exit the industry. But in the short run, it will continue to operate at a loss. Exhibit 20 depicts a firm under conditions of perfect competition facing three alternative market price ranges for its output. At any price above P_1, the firm can earn a positive profit and clearly should continue to operate. At a price below P_2, the minimum AVC, the firm could not even cover its variable cost and should shut down. At prices between P_2 and P_1, the firm should continue to operate in the short run because it is able to cover all of its variable cost and contribute something toward unavoidable fixed costs. Economists refer to the minimum AVC point as the **shutdown point** and the minimum ATC point as the **breakeven point**.

Exhibit 20 **A Firm under Conditions of Perfect Competition Will Choose to Shut Down If Market Price Is Less Than Minimum AVC**



EXAMPLE 5

Breakeven Analysis and Profit Maximization When the Firm Faces a Negatively Sloped Demand Curve under Imperfect Competition

Revenue and cost information for a future period is presented in Exhibit 21 for WR International, a newly formed corporation that engages in the manufacturing of low-cost, pre-fabricated dwelling units for urban housing markets in emerging economies. (Note that quantity increments are in blocks of 10 for a 250 change in price.) The firm has few competitors in a market setting of imperfect competition.

1. How many units must WR International sell to initially break even?
2. Where is the region of profitability?
3. At what point will the firm maximize profit? At what points are there economic losses?

Exhibit 21

Quantity (Q)	Price (P)	Total Revenue (TR)	Total Cost (TC)[a]	Profit
0	10,000	0	100,000	–100,000
10	9,750	97,500	170,000	–72,500
20	9,500	190,000	240,000	–50,000
30	9,250	277,500	300,000	–22,500
40	9,000	360,000	360,000	0
50	8,750	437,500	420,000	17,500
60	8,500	510,000	480,000	30,000
70	8,250	577,500	550,000	27,500
80	8,000	640,000	640,000	0
90	7,750	697,500	710,000	–12,500
100	7,500	750,000	800,000	–50,000

[a] Includes all opportunity costs

Solution to 1:

WR International will initially break even at 40 units of production, where TR and TC equal 360,000.

Solution to 2:

The region of profitability will range from greater than 40 units to less than 80 units. Any production quantity of less than 40 units and any quantity greater than 80 units will result in an economic loss.

Solution to 3:

Maximum profit of 30,000 will occur at 60 units. Lower profit will occur at any output level that is higher or lower than 60 units. From 0 units to less than 40 units and for quantities greater than 80 units, economic losses occur.

Given the relationships between TR, TVC, and TFC, Exhibit 22 summarizes the decisions to operate, shut down production, or exit the market in both the short run and the long run. The firm must cover its variable cost to remain in business in the short run; if TR cannot cover TVC, the firm shuts down production to minimize loss. The loss would be equal to the amount of fixed cost. If TVC exceeds TR in the long run, the firm will exit the market to avoid the loss associated with fixed cost at zero production. By exiting the market, the firm's investors do not suffer the erosion of their equity capital from economic losses. When TR is enough to cover TVC but not all of TFC, the firm can continue to produce in the short run but will be unable to maintain financial solvency in the long run.

Exhibit 22 Short Run and Long Run Decisions to Operate or Not

Revenue–Cost Relationship	Short-Run Decision	Long-Term Decision
TR = TC	Stay in market	Stay in market
TR = TVC but < TC	Stay in market	Exit market
TR < TVC	Shut down production	Exit market

EXAMPLE 6

Shutdown Analysis

For the most recent financial reporting period, a business domiciled in Ecuador (which recognizes the US dollar as an official currency) has revenue of $2 million and TC of $2.5 million, which are or can be broken down into TFC of $1 million and TVC of $1.5 million. The net loss on the firm's income statement is reported as $500,000 (ignoring tax implications). In prior periods, the firm had reported profits on its operations.

1 What decision should the firm make regarding operations over the short term?

2 What decision should the firm make regarding operations over the long term?

3 Assume the same business scenario except that revenue is now $1.3 million, which creates a net loss of $1.2 million. What decision should the firm make regarding operations in this case?

Solution to 1:

In the short run, the firm is able to cover all of its TVC but only half of its $1 million in TFC. If the business ceases to operate, its loss would be $1 million, the amount of TFC, whereas the net loss by operating would be minimized at $500,000. The firm should attempt to operate by negotiating special arrangements with creditors to buy time to return operations back to profitability.

Solution to 2:

If the revenue shortfall is expected to persist over time, the firm should cease operations, liquidate assets, and pay debts to the extent possible. Any residual for shareholders would decrease the longer the firm is allowed to operate unprofitably.

Solution to 3:

The firm would minimize loss at $1 million of TFC by shutting down. If the firm decided to continue to do business, the loss would increase to $1.2 million. Shareholders would save $200,000 in equity value by pursuing this option. Unquestionably, the business would have a rather short life expectancy if this loss situation were to continue.

When evaluating profitability, particularly of start-up firms and businesses using turnaround strategies, analysts should consider highlighting breakeven and shutdown points in their financial research. Identifying the unit sales levels at which the firm enters or leaves the production range for profitability and at which the firm can no longer function as a viable business entity provides invaluable insight for investment decisions.

ECONOMIES AND DISECONOMIES OF SCALE WITH SHORT-RUN AND LONG-RUN COST ANALYSIS

12

f describe how economies of scale and diseconomies of scale affect costs.

Rational behavior dictates that the firm select an operating size or scale that maximizes profit over any time frame. The time frame that defines the short run and long run for any firm is based on the ability of the firm to adjust the quantities of the fixed resources it uses. The short run is the time period during which at least one of the factors of production, such as technology, physical capital, and plant size, is fixed. The long run is defined as the time period during which all factors of production are variable. Additionally, in the long run, firms can enter or exit the market based on decisions regarding profitability. The long run is often referred to as the "planning horizon" in which the firm can choose the short-run position or optimal operating size that maximizes profit over time. The firm is always operating in the short run but planning in the long run.

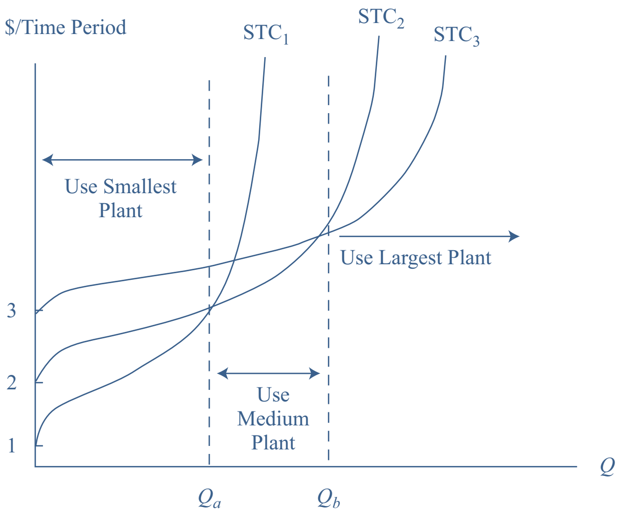
The time required for long-run adjustments varies by industry. For example, the long run for a small business using very little technology and physical capital may be less than a year, whereas for a capital-intensive firm, the long run may be more than a decade. Given enough time, however, all production factors are variable, which allows the firm to choose an operating size or plant capacity based on different technologies and physical capital. In this regard, costs and profits will differ between the short run and the long run.

12.1 Short- and Long-Run Cost Curves

Recall that when we addressed the short-run cost curves of the firm, we assumed that the capital input was held constant. That meant that the only way to vary output in the short run is to change the level of the variable input—in our case, labor. If the capital input—namely, plant and equipment—were to change, however, we would have an entirely new set of short-run cost curves, one for each level of capital input.

The short-run total cost includes all the inputs—labor and capital—the firm is using to produce output. For reasons discussed earlier, the typical short-run total cost (STC) curve might rise with output, first at a decreasing rate because of specialization economies and then at an increasing rate, reflecting the law of diminishing marginal returns to labor. Total fixed cost (the quantity of capital input multiplied by the rental rate on capital) determines the vertical intercept of the STC curve. At higher levels of fixed input, TFC is greater but the production capacity of the firm is also greater. Exhibit 23 shows three different STC curves for the same technology but using three distinct levels of capital input—Points 1, 2, and 3 on the vertical axis.

Exhibit 23 Short-Run Total Cost Curves for Various Plant Sizes



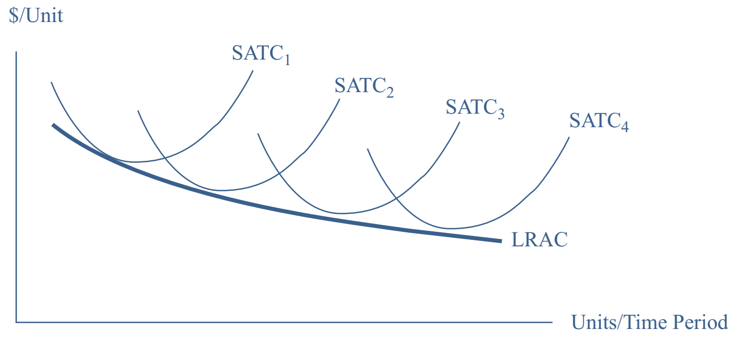
Plant Size 1 is the smallest and, of course, has the lowest fixed cost; hence, its STC_1 curve has the lowest vertical intercept. But note that STC_1 begins to rise more steeply with output, reflecting the lower plant capacity. Plant Size 3 is the largest of the three and reflects that size with both a higher fixed cost and a lower slope at any level of output. If a firm decided to produce an output between zero and Q_a, it would plan on building Plant Size 1 because for any output level in that range, its cost is less than it would be for Plant Size 2 or 3. Accordingly, if the firm were planning to produce output greater than Q_b, it would choose Plant Size 3 because its cost for any of those levels of output would be lower than for Plant Size 1 or 2. And of course, Plant Size 2 would be chosen for output levels between Q_a and Q_b. The long-run total cost curve is derived from the lowest level of STC for each level of output because in the long run, the firm is free to choose which plant size it will operate. This curve is called an "envelope curve." In essence, this curve envelopes—encompasses—all possible combinations of technology, plant size, and physical capital.

For each STC curve, there is also a corresponding **short-run average total cost** (SATC) curve and a corresponding **long-run average total cost** (LRAC) curve, the envelope curve of all possible short-run average total cost curves. The shape of the LRAC curve reflects an important concept called **economies of scale** and **diseconomies of scale**.

12.2 Defining Economies of Scale and Diseconomies of Scale

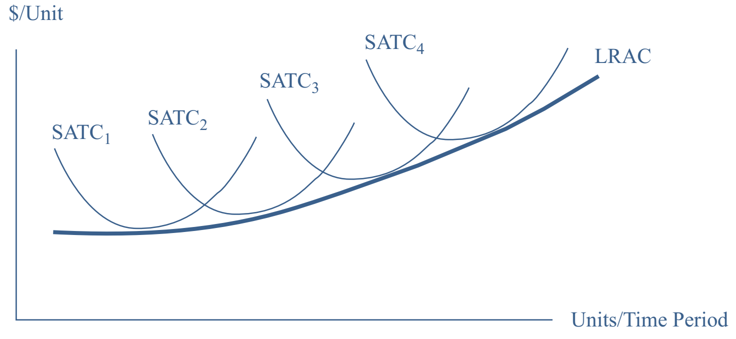
When a firm increases all of its inputs in order to increase its level of output (obviously, a long-run concept), it is said to *scale up* its production. *Scaling down* is the reverse—decreasing all of its inputs in order to produce less in the long run. Economies of scale occur if, as the firm increases its output, cost per unit of production falls. Graphically, this definition translates into a LRAC curve with a negative slope. Exhibit 24 depicts several SATC curves, one for each plant size, and the LRAC curve representing economies of scale.

Exhibit 24 Short-run Average Total Cost Curves for Various Plant Sizes and Their Envelope Curve, LRAC: Economies of Scale



Diseconomies of scale occur if cost per unit rises as output increases. Graphically, diseconomies of scale translate into an LRAC curve with a positive slope. Exhibit 25 depicts several SATC curves, one for each plant size, and their envelope curve, the LRAC curve, representing diseconomies of scale.

Exhibit 25 Short-run Average Total Cost Curves for Various Plant Sizes and Their Envelope Curve, LRAC: Diseconomies of Scale



As the firm grows in size, economies of scale and a lower ATC can result from the following factors:

- **Increasing returns to scale**, which is when a production process allows for increases in output that are proportionately larger than the increase in inputs.
- Having a division of labor and management in a large firm with numerous workers, which allows each worker to specialize in one task rather than perform many duties, as in the case of a small business (as such, workers in a large firm become more proficient at their jobs).
- Being able to afford more expensive, yet more efficient equipment and to adapt the latest in technology that increases productivity.
- Effectively reducing waste and lowering costs through marketable byproducts, less energy consumption, and enhanced quality control.
- Making better use of market information and knowledge for more effective managerial decision making.
- Obtaining discounted prices on resources when buying in larger quantities.

A classic example of a business that realizes economies of scale through greater physical capital investment is an electric utility. By expanding output capacity to accommodate a larger customer base, the utility company's per-unit cost will decline. Economies of scale help explain why electric utilities have naturally evolved from localized entities to regional and multi-region enterprises. Wal-Mart is an example of a business that uses bulk purchasing power to obtain deep discounts from suppliers to keep costs and prices low. Wal-Mart also uses the latest technology to monitor point-of-sale transactions to gather timely market information to respond to changes in customer buying behavior, which leads to economies of scale through lower distribution and inventory costs.

The factors that can lead to diseconomies of scale, inefficiencies, and rising costs when a firm increases in size include the following:

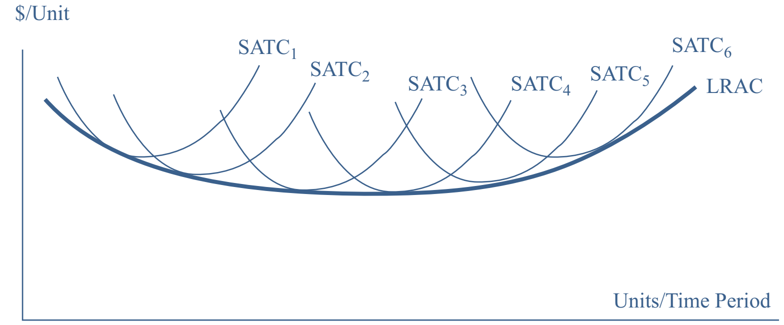
- **Decreasing returns to scale**, which is when a production process leads to increases in output that are proportionately smaller than the increase in inputs.
- Being so large that it cannot be properly managed.
- Overlapping and duplication of business functions and product lines.
- Higher resource prices because of supply constraints when buying inputs in large quantities.

Before its restructuring, General Motors (GM) was an example of a business that had realized diseconomies of scale by becoming too large. Scale diseconomies occurred through product overlap and duplication (i.e., similar or identical automobile models), and the fixed cost for these models was not spread over a large volume of output. In 2009, GM decided to discontinue three brands (Saturn, Pontiac, and Hummer), and in 2018 it was considering dropping various low-volume product models that overlapped with other models by 2020. GM had numerous manufacturing plants throughout the world and sold vehicles in more than a hundred countries. Given this geographical dispersion in production and sales, the company had communication and management coordination problems, which resulted in higher costs. In 2017, GM sold its European arm, Opel, to Groupe PSA, the maker of Peugeot and Citroën. GM also had significantly higher labor costs than its competitors. As the largest producer in the market, it had been a target of labor unions for higher compensation and benefits packages relative to other firms.

Economies and diseconomies of scale can occur at the same time; the impact on long-run average total cost (LRAC) depends on which dominates. If economies of scale dominate, LRAC decreases with increases in output. The reverse holds true when diseconomies of scale prevail. There may be a range of output over which LRAC falls (economies of scale) and then a range over which LRAC might be constant, followed by a range over which diseconomies of scale prevail, as depicted in Exhibit 26.

The minimum point on the LRAC curve is referred to as the **minimum efficient scale**. The minimum efficient scale is the optimal firm size under perfect competition over the long run. Theoretically, perfect competition forces the firm to operate at the minimum point on the LRAC curve because the market price will be established at this level over the long run. If the firm is not operating at this least-cost point, its long-term viability will be threatened.

Exhibit 26 LRAC Can Exhibit Economies and Diseconomies of Scale



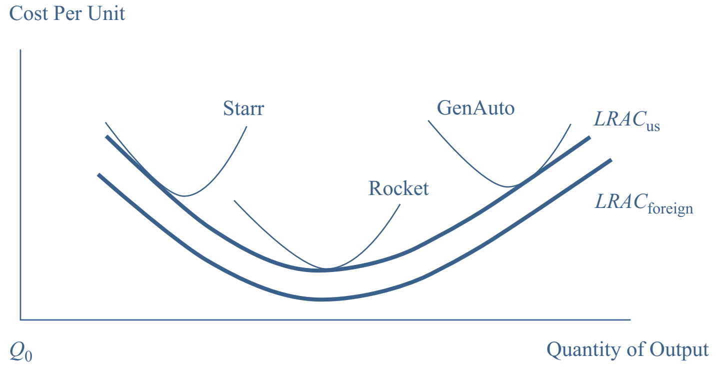
EXAMPLE 7

Long-Run Average Total Cost Curve

Exhibit 27 displays the long-run average total cost curve (LRAC$_{US}$) and the short-run average total cost curves for three hypothetical US-based automobile manufacturers—Starr Vehicles (Starr), Rocket Sports Cars (Rocket), and General Auto (GenAuto). The LRAC curve for foreign-owned automobile companies that compete in the US auto market (LRAC$_{foreign}$) is also indicated in the graph. (The market structure implicit in the exhibit is imperfect competition.)

To what extent are the cost relationships depicted in Exhibit 27 useful for an economic and financial analysis of the three US-based auto firms?

Exhibit 27



Solution:

First, it is observable that the foreign auto companies have a lower LRAC compared with that of the US automobile manufacturers. This competitive position places the US firms at a cost—and possibly, pricing—disadvantage in the market, with the potential to lose market share to the lower-cost foreign competitors. Second, only Rocket operates at the minimum point of the LRAC$_{US}$, whereas GenAuto is situated in the region of diseconomies of scale and Starr is positioned in the economies of scale portion of the curve. To become more efficient and competitive, GenAuto needs to downsize and restructure, which means moving down the LRAC_{US} curve to a smaller, yet lower-cost production volume. In Contrast, Starr has to grow in size to become more efficient and competitive by lowering per-unit costs.

From a long-term investment prospective and given its cost advantage, Rocket has the potential to create more investment value relative to GenAuto and Starr. Over the long run, if GenAuto and Starr can lower their ATC, they will become more attractive to investors. But if any of the three US auto companies cannot match the cost competitiveness of the foreign firms, they may be driven from the market. In the long run, the lower-cost foreign automakers pose a severe competitive challenge to the survival of the US manufacturers and their ability to maintain and grow shareholders' wealth.

SUMMARY

This reading addressed several important concepts that extend the basic market model of demand and supply to assist the analyst in assessing a firm's breakeven and shutdown points of production. Demand concepts covered include own-price elasticity of demand, cross-price elasticity of demand, and income elasticity of demand. Supply concepts covered include total, average, and marginal product of labor; total, variable, and marginal cost of labor; and total and marginal revenue. These concepts are used to calculate the breakeven and shutdown points of production.

- Elasticity of demand is a measure of how sensitive quantity demanded is to changes in various variables.
- Own-price elasticity of demand is the ratio of percentage change in quantity demanded to percentage change in a good or service's own price.
- If own-price elasticity of demand is greater than one in absolute terms, demand is elastic and a decline in price will result in higher total expenditure on that good.
- If own-price elasticity of demand is less than one in absolute terms, demand is inelastic and a decline in price will result in a lower total expenditure on that good.
- If own-price elasticity of demand is equal to negative one, demand is unit, or unitary, elastic and total expenditure on that good is independent of price.
- Own-price elasticity of demand will almost always be negative.
- Income elasticity of demand is the ratio of the percentage change in quantity demanded to the percentage change in consumer income.
- Demand is negatively sloped because of either the substitution effect or the income effect.
- The substitution effect is the phenomenon in which, as a good's price falls, more of this good is substituted for other, more expensive goods.
- The income effect is the phenomenon in which, as a good's price falls, real income rises and, if this good is normal, more of it will be purchased.
- If the good is inferior, the income effect will partially or fully offset the substitution effect.
- There are two exceptions to the law of demand: Giffen goods and Veblen goods.

- Giffen goods are highly inferior and make up a large portion of the consumer budget. As price falls, the substitution effect tends to cause more of the good to be consumed, but the highly negative income effect overwhelms the substitution effect. Demand curves for Giffen goods are positively sloped.
- Veblen goods are highly valued high-priced "status" goods; consumers may tend to buy more of a good if its price rises.
- If income elasticity of demand is positive, the good is a normal good. If income elasticity of demand is negative, the good is an inferior good.
- Cross-price elasticity of demand is the ratio of the percentage change in quantity demanded of one good to the percentage change in the price of a related good.
- If cross-price elasticity between two goods is positive, they are substitutes, and if cross-price elasticity between two goods is negative, they are complements.
- The law of demand states that a decrease in price will cause an increase in quantity demanded.
- Total product of labor is a short-run concept that is the total quantity that is able to be produced for each level of labor input, holding all other inputs constant.
- Average product of labor (APL) is the total product of labor divided by number of labor hours.
- Marginal product of labor (MP_L) is the change in total product divided by the change in labor hours. MP_L might rise as more labor is added to a fixed amount of capital.
- The law of diminishing returns dictates that additional output must fall as more and more labor is added to a fixed amount of capital.
- Production costs increase as input prices rise and fall as inputs become more productive.
- Short-run total cost (STC) is the total expenditure on fixed capital plus the total expenditure on labor.
- Short-run marginal cost (SMC) equals the ratio of wage to marginal product of labor (MP_L).
- Average variable cost (AVC) is the ratio of wage to average product of labor (APL).
- Average total cost (ATC) is total cost (TC) divided by the number of units produced.
- Revenue is price times quantity sold.
- Marginal revenue (MR) is the ratio of change in revenue to change in output.
- Firms under conditions of perfect competition have no pricing power and, therefore, face a perfectly horizontal demand curve at the market price. For firms under conditions of perfect competition, price is identical to marginal revenue (MR).
- Firms under conditions of imperfect competition face a negatively sloped demand curve and have pricing power. For firms under conditions of imperfect competition, marginal revenue (MR) is less than price.
- Economic profit equals total revenue (TR) minus total economic cost, whereas accounting profit equals TR minus total accounting cost.
- Economic cost takes into account the total opportunity cost of all factors of production.
- Opportunity cost is the next best alternative forgone in making a decision.

- Maximum economic profit requires that (1) marginal revenue (MR) equals marginal cost (MC) and (2) MC not be falling with output.
- The breakeven point occurs when total revenue (TR) equals total cost (TC), otherwise stated as the output quantity at which average total cost (ATC) equals price.
- Shutdown occurs when a firm is better off not operating than continuing to operate.
- If all fixed costs are sunk costs, then shutdown occurs when the market price falls below minimum average variable cost. After shutdown, the firm incurs only fixed costs and loses less money than it would operating at a price that does not cover variable costs.
- In the short run, it may be rational for a firm to continue to operate while earning negative economic profit if some unavoidable fixed costs are covered.
- Economies of scale is defined as decreasing long-run cost per unit as output increases. Diseconomies of scale is defined as increasing long-run cost per unit as output increases.
- Long-run average total cost is the cost of production per unit of output under conditions in which all inputs are variable.
- Specialization efficiencies and bargaining power in input price can lead to economies of scale.
- Bureaucratic and communication breakdowns and bottlenecks that raise input prices can lead to diseconomies of scale.
- The minimum point on the long-run average total cost curve defines the minimum efficient scale for the firm.

PRACTICE PROBLEMS

1 If the price elasticity coefficient of the demand curve for paper clips is equal to −1, demand is:

A elastic.

B inelastic.

C unit elastic.

2 The demand for membership at a local health club is determined by the following equation:

$$Q_{hm}^{d} = 400 - 5P_{hm}$$

where Q_{hm}^{d} is the number of health club members and P_{hm} is the price of membership. If the price of health club membership is $35, the price elasticity of demand is *closest* to:

A −0.778.

B −0.500.

C −0.438.

3 Price elasticity of demand for a good will *most likely* be greater if:

A there are no substitutes for the good.

B consumers consider the good as discretionary.

C consumers spend a small portion of their budget on the good.

4 If the income elasticity of demand for a product is −0.6, a:

A 1% increase in income will result in a 0.6% increase in demand.

B 1% increase in income will result in a 0.6% decrease in demand.

C 0.6% increase in income will result in a 1% decrease in demand.

5 An individual's demand for onions is given by the following equation:

$$Q_{o}^{d} = 3 - 0.05P_{o} + 0.009I - 0.16P_{t}$$

where Q_{o}^{d} is the number of onions demanded, P_{o} is the price per pound of onions, I is the household income, and P_{t} is the price per pound of tomatoes. If the price of onions is $1.25, household income is $2,500, and the price of tomatoes is $3.75, the cross-price elasticity of demand for onions with respect to the price of tomatoes is *closest* to:

A −1.0597.

B −0.0242.

C −0.0081.

6 Movement along the demand curve for good X occurs due to a change in:

A income.

B the price of good X.

C the price of a substitute for good X.

© 2017 CFA Institute. All rights reserved.

7 A wireless phone manufacturer introduced a next-generation phone that received a high level of positive publicity. Despite running several high-speed production assembly lines, the manufacturer is still falling short in meeting demand for the phone nine months after introduction. Which of the following statements is the *most* plausible explanation for the demand/supply imbalance?

A The phone price is low relative to the equilibrium price.

B Competitors introduced next-generation phones at a similar price.

C Consumer incomes grew faster than the manufacturer anticipated.

The following information relates to Questions 8–11

The market demand function for four-year private universities is given by the equation

$$Q_{pr}^{d} = 84 - 3.1P_{pr} + 0.8I + 0.9P_{pu}$$

where Q_{pr}^{d} is the number of applicants to private universities per year in thousands, P_{pr} is the average price of private universities (in thousands of USD), I is the household monthly income (in thousands of USD), and P_{pu} is the average price of public (government-supported) universities (in thousands of USD). Assume that P_{pr} is equal to 38, I is equal to 100, and P_{pu} is equal to 18.

8 The price elasticity of demand for private universities is *closest* to:

A –3.1.

B –1.9.

C 0.6.

9 The income elasticity of demand for private universities is *closest* to:

A 0.5.

B 0.8.

C 1.3.

10 The cross-price elasticity of demand for private universities with respect to the price of public universities is *closest* to:

A 0.3.

B 3.1.

C 3.9.

11 If the cross-price elasticity between two goods is negative, the two goods are classified as:

A normal.

B substitutes.

C complements.

---

12 In the case of a normal good with a decrease in own price, which of the following statements is *most likely* true?

A Both the substitution and income effects lead to an increase in the quantity purchased.

B The substitution effect leads to an increase in the quantity purchased, while the income effect has no impact.

C The substitution effect leads to an increase in the quantity purchased, while the income effect leads to a decrease.

13 For a Giffen good, the:

A demand curve is positively sloped.

B substitution effect overwhelms the income effect.

C income and substitution effects are in the same direction.

14 Normal profit is best described as:

A zero economic profit.

B total revenue minus all explicit costs.

C the sum of accounting profit plus economic profit.

15 A company plans to hire additional factory employees. In the short run, marginal returns are most likely to decrease if:

A the factory is operating at full capacity.

B the factory is experiencing a labor shortage.

C workers are required to multitask and share duties.

16 The production relationship between the number of machine hours and total product for a company is presented below.

Machine Hours	Total Product	Average Product
1	3	3.00
2	8	4.00
3	14	4.67
4	19	4.75
5	21	4.20

Diminishing marginal returns first occur beyond machine hour:

A 3.

B 4.

C 5.

17 The marketing director for a Swiss specialty equipment manufacturer estimates the firm can sell 200 units and earn total revenue of CHF500,000. However, if 250 units are sold, revenue will total CHF600,000. The marginal revenue per unit associated with marketing 250 units instead of 200 units is *closest* to:

A CHF 2,000.

B CHF 2,400.

C CHF 2,500.

18 An agricultural firm operating in a perfectly competitive market supplies wheat to manufacturers of consumer food products and animal feeds. If the firm were able to expand its production and unit sales by 10% the *most likely* result would be:

A a 10% increase in total revenue.

B a 10% increase in average revenue.

C an increase in total revenue of less than 10%.

19 An operator of a ski resort is considering offering price reductions on weekday ski passes. At the normal price of €50 per day, 300 customers are expected to buy passes each weekday. At a discounted price of €40 per day 450 customers are expected to buy passes each weekday. The marginal revenue per customer earned from offering the discounted price is *closest* to:

A €20.

B €40.

C €50.

20 The marginal revenue per unit sold for a firm doing business under conditions of perfect competition will *most likely* be:

A equal to average revenue.

B less than average revenue.

C greater than average revenue.

The following information relates to Questions 21–23

A firm's director of operations gathers the following information about the firm's cost structure at different levels of output:

Exhibit 1

Quantity (*Q*)	Total Fixed Cost (*TFC*)	Total Variable Cost (*TVC*)
0	200	0
1	200	100
2	200	150
3	200	200
4	200	240
5	200	320

21 Refer to the data in Exhibit 1. When quantity produced is equal to 4 units, the average fixed cost (*AFC*) is *closest* to:

A 50.

B 60.

C 110.

22 Refer to the data in Exhibit 1. When the firm increases production from 4 to 5 units, the marginal cost (*MC*) is *closest* to:

A 40.

B 64.

C 80.

23 Refer to the data in Exhibit 1. The level of unit production resulting in the lowest average total cost (*ATC*) is *closest* to:

A 3.

B 4.

C 5.

---

24 The short-term breakeven point of production for a firm operating under perfect competition will *most likely* occur when:

A price is equal to average total cost.

B marginal revenue is equal to marginal cost.

C marginal revenue is equal to average variable costs.

25 The short-term shutdown point of production for a firm operating under perfect competition will *most likely* occur when:

A price is equal to average total cost.

B marginal revenue is equal to marginal cost.

C marginal revenue is equal to average variable costs.

26 Under conditions of perfect competition, a company will break even when market price is equal to the minimum point of the:

A average total cost curve.

B average variable cost curve.

C short-run marginal cost curve.

27 A company will shut down production in the short run if total revenue is less than total:

A fixed costs.

B variable costs.

C opportunity costs.

28 A company has total variable costs of $4 million and fixed costs of $3 million. Based on this information, the company will stay in the market in the long term if total revenue is at least:

A $3.0 million.

B $4.5 million.

C $7.0 million.

29 When total revenue is greater than total variable costs but less than total costs, in the short term a firm will *most likely*:

A exit the market.

B stay in the market.

C shut down production.

30 A profit maximum is *least likely* to occur when:

A average total cost is minimized.

B marginal revenue equals marginal cost.

C the difference between total revenue and total cost is maximized.

31 A firm that increases its quantity produced without any change in per-unit cost is experiencing:

A economies of scale.

B diseconomies of scale.

C constant returns to scale.

32 A company is experiencing economies of scale when:

- **A** cost per unit increases as output increases.
- **B** it is operating at a point on the LRAC curve where the slope is negative.
- **C** It is operating beyond the minimum point on the long-run average total cost curve.

33 Diseconomies of scale *most likely* result from:

- **A** specialization in the labor force.
- **B** overlap of business functions and product lines.
- **C** discounted prices on resources when buying in larger quantities.

34 A firm is operating beyond minimum efficient scale in a perfectly competitive industry. To maintain long-term viability the *most likely* course of action for the firm is to:

- **A** operate at the current level of production.
- **B** increase its level of production to gain economies of scale.
- **C** decrease its level of production to the minimum point on the long-run average total cost curve.

35 Under conditions of perfect competition, in the long run firms will *most likely* earn:

- **A** normal profits.
- **B** positive economic profits.
- **C** negative economic profits.

The following information relates to Questions 36 and 37

The manager of a small manufacturing firm gathers the following information about the firm's labor utilization and production:

Exhibit 2

Labor (*L*)	Total Product (*TP*)
0	0
1	150
2	320
3	510
4	660
5	800

36 Refer to the data in Exhibit 2. The number of workers resulting in the highest level of average product of labor is *closest* to:

- **A** 3.
- **B** 4.
- **C** 5.

37 Refer to the data in Exhibit 2. The marginal product of labor demonstrates increasing returns for the firm if the number of workers is *closest* to but not more than:

A 2.

B 3.

C 4.

SOLUTIONS

1. C is correct. When the price elasticity of demand coefficient is –1, demand is said to be unit elastic, or unitary elastic.
2. A is correct. Inserting the price of \$35 into the demand function, quantity demanded is calculated as

   $$Q_{hm}^{d} = 400 - 5(35) = 225$$

   At a price of \$35 per health club membership, the elasticity of demand is

   $$\text{Price elasticity of demand} = \left(\Delta Q_{hm}^{d} / \Delta P_{hm}\right) \times \left(P_{hm} / Q_{hm}^{d}\right)$$

   $$\text{Price elasticity of demand} = -5 \times (35/225) = -0.778$$

3. B is correct. Price elasticity of demand is likely to be greater for items that are seen as optional or discretionary.
4. B is correct. Income elasticity is a measure of how sensitive quantity demanded is to a change in income. If the income elasticity of demand for the product is –0.6, whenever income increases by 1%, the quantity demanded of the product at each price decreases by 0.6%. Consequently, as income rises, consumers will purchase less of the product.
5. B is correct. The cross-price elasticity of demand measures the responsiveness of the demand for onions in response to a change in the price of tomatoes. From the demand function equation:

   $$Q_{o}^{d} = 3 - 0.05P_{o} + 0.009I - 0.16P_{t}$$

   $$Q_{o}^{d} = 3 - 0.05(1.25) + 0.009(2{,}500) - 0.16(3.75) = 24.8375$$

   At a price of onions of \$1.25 and a price of tomatoes of \$3.75, the cross-price elasticity of demand is calculated as follows:

   $$\text{Cross-price elasticity of demand} = \left(\Delta Q_{o}^{d} / \Delta P_{t}\right) \times \left(P_{t} / Q_{o}^{d}\right)$$

   $$\text{Cross-price elasticity of demand} = -0.16 \times (3.75/24.8375) = -0.0242$$

6. B is correct. The demand curve shows quantity demanded as a function of own price only.
7. A is correct. The situation described is one of excess demand because, in order for markets to clear at the given level of quantity supplied, the company would need to raise prices.
8. B is correct. From the demand function:

   Solve for Q_{pr}^{d}:

   $$\Delta Q_{pr}^{d} / \Delta P_{pr} = -3.1 \text{ (the coefficient in front of own price)}$$

   $$\begin{aligned} Q_{pr}^{d} &= 84 - 3.1P_{pr} + 0.8I + 0.9P_{pu} \\ &= 84 - 3.1(38) + 0.8(100) + 0.9(18) \\ &= 62.4 \end{aligned}$$

At $P_{pr} = 38$,

$$\text{price elasticity of demand} = \left(\Delta Q^d_{pr} / \Delta P_{pr}\right)\left(P_{pr} / Q^d_{pr}\right)$$
$$= (-3.1)(38/62.4)$$
$$= -1.9$$

9 C is correct. From the demand function:

Solve for Q^d_{pr}:

$$\Delta Q^d_{pr} / \Delta I = 0.8 \text{ (coefficient in front of the income variable)}$$

$$Q^d_{pr} = 84 - 3.1P_{pr} + 0.8I + 0.9P_{pu}$$
$$= 84 - 3.1(38) + 0.8(100) + 0.9(18)$$
$$= 62.4$$

At $I = 100$,

$$\text{the income elasticity of demand} = \left(\Delta Q^d_{pr} / \Delta I\right)\left(I / Q^d_{pr}\right)$$
$$= (0.8)(100/62.4)$$
$$= 1.3$$

10 A is correct. From the demand function:

Solve for Q^d_{pr}:

$$\Delta Q^d_{pr} / \Delta P_{pu} = 0.9 \text{ (the coefficient in front of } P_{pu}\text{)}$$

$$Q^d_{pr} = 84 - 3.1P_{pr} + 0.8I + 0.9P_{pu}$$
$$= 84 - 3.1(38) + 0.8(100) + 0.9(18)$$
$$= 62.4$$

At $P = 38$, and $P_{pu} = 18$,

$$\text{the cross-price elasticity of demand} = \left(\Delta Q^d_{pr} / \Delta P_{pu}\right)\left(P_{pu} / Q^d_{pr}\right)$$
$$= (0.9)(18/62.4)$$
$$= 0.3$$

11 C is correct. With complements, consumption goes up or down together. With a negative cross-price elasticity, as the price of one good goes up, the demand for both falls.

12 A is correct. In the case of normal goods, the income and substitution effects are reinforcing, leading to an increase in the amount purchased after a drop in price.

13 A is correct. The income effect overwhelms the substitution effect such that an increase in the price of the good results in greater demand for the good, resulting in a positively sloped demand curve.

14 A is correct. Normal profit is the level of accounting profit such that implicit opportunity costs are just covered; thus, it is equal to a level of accounting profit such that economic profit is zero.

15 A is correct. The law of diminishing returns occurs in the short run when additional output falls as more and more labor is added to a fixed amount of capital. When a factory is operating at full capacity, adding additional employees will not increase production because the physical plant is already 100% employed. More labor hours will add to costs without adding to output, thus resulting in diminishing marginal returns.

16 A is correct. Diminishing marginal returns occur when the marginal product of a resource decreases as additional units of that input are employed. Marginal product, which is the additional output resulting from using one more unit of input, is presented below.

Machine Hours	Total Product	Average Product	Marginal Product
1	3	3.00	3
2	8	4.00	5
3	14	4.67	6
4	19	4.75	5
5	21	4.20	2

The marginal product of the third machine hour is 6 and declines thereafter. Consequently, diminishing marginal returns are first evident beyond three machine hours.

17 A is correct. Marginal revenue per unit is defined as the change in total revenue divided by the change in quantity sold. MR = ΔTR ÷ ΔQ. In this case, change in total revenue equals CHF100,000, and change in total units sold equals 50. CHF100,000 ÷ 50 = CHF2,000.

18 A is correct. In a perfectly competitive market, an increase in supply by a single firm will not affect price. Therefore, an increase in units sold by the firm will be matched proportionately by an increase in revenue.

19 A is correct. Marginal revenue per unit is defined as the change in total revenues divided by the change in quantity sold. MR = ΔTR ÷ ΔQ. In this case, change in total revenue per day equals €3,000 [(450 × €40) – (300 × €50)], and change in units sold equals 150 (450 – 300). €3,000 ÷ 150 = €20.

20 A is correct. Under perfect competition, a firm is a price taker at any quantity supplied to the market, and AR = MR = Price.

21 A is correct. Average fixed cost is equal to total fixed cost divided by quantity produced: AFC = TFC/Q = 200/4 = 50.

22 C is correct. Marginal cost is equal to the change in total cost divided by the change in quantity produced. MC = ΔTC/ΔQ = 80/1 = 80.

23 C is correct. Average total cost is equal to total cost divided by quantity produced. At 5 units produced the average total cost is 104. ATC = TC/Q = 520/5 = 104.

24 A is correct. Under perfect competition, price equals marginal revenue. A firm breaks even when marginal revenue equals average total cost.

25 C is correct. The firm should shut down production when marginal revenue is less than or equal to average variable cost.

26 A is correct. A company is said to break even if its total revenue is equal to its total cost. Under conditions of perfect competition, a company will break even when market price is equal to the minimum point of the average total cost curve.

27 B is correct. A company will shut down production in the short run when total revenue is below total variable costs.

28 C is correct. A company will stay in the market in the long term if total revenue is equal to or greater than total cost. Because total costs are \$7 million (\$4 million variable costs and \$3 million fixed costs), the company will stay in the market in the long term if total revenue equals at least \$7 million.

29 B is correct. When total revenue is enough to cover variable costs but not total fixed costs in full, the firm can survive in the short run but would be unable to maintain financial solvency in the long run.

30 A is correct. The quantity at which average total cost is minimized does not necessarily correspond to a profit maximum.

31 C is correct. Output increases in the same proportion as input increases occur at constant returns to scale.

32 B is correct. Economies of scale occur if, as the firm increases output, cost per unit of production falls. Graphically, this definition translates into a long-run average cost curve (LRAC) with a negative slope.

33 B is correct. As the firm increases output, diseconomies of scale and higher average total costs can result when there is overlap and duplication of business functions and product lines.

34 C is correct. The firm operating at greater than long-run efficient scale is subject to diseconomies of scale. It should plan to decrease its level of production.

35 A is correct. Competition should drive prices down to long-run marginal cost, resulting in only normal profits being earned.

36 A is correct. Three workers produce the highest average product equal to 170. AP = 510/3 = 170.

37 B is correct. Marginal product is equal to the change in total product divided by the change in labor. The increase in MP from 2 to 3 workers is 190: $\text{MP} = \Delta\text{TP}/\Delta L = (510 - 320)/(3 - 2) = 190/1 = 190$.

READING

9

The Firm and Market Structures

by Richard Fritz, PhD, and Michele Gambera, PhD, CFA

Richard Fritz, PhD, is at the School of Economics at Georgia Institute of Technology (USA). Michele Gambera, PhD, CFA, is at UBS Asset Management and the University of Illinois at Urbana-Champaign (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;
☐	**b.** explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;
☐	**c.** describe a firm's supply function under each market structure;
☐	**d.** describe and determine the optimal price and output for firms under each market structure;
☐	**e.** explain factors affecting long-run equilibrium under each market structure;
☐	**f.** describe pricing strategy under each market structure;
☐	**g.** describe the use and limitations of concentration measures in identifying market structure;
☐	**h.** identify the type of market structure within which a firm operates.

1

INTRODUCTION & ANALYSIS OF MARKET STRUCTURES

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

The purpose of this reading is to build an understanding of the importance of market structure. As different market structures result in different sets of choices facing a firm's decision makers, an understanding of market structure is a powerful tool in

© 2019 CFA Institute. All rights reserved.

analyzing issues such as a firm's pricing of its products and, more broadly, its potential to increase profitability. In the long run, a firm's profitability will be determined by the forces associated with the market structure within which it operates. In a highly competitive market, long-run profits will be driven down by the forces of competition. In less competitive markets, large profits are possible even in the long run; in the short run, any outcome is possible. Therefore, understanding the forces behind the market structure will aid the financial analyst in determining firms' short- and long-term prospects.

Section 1.1 introduces the analysis of market structures. The section addresses questions such as: What determines the degree of competition associated with each market structure? Given the degree of competition associated with each market structure, what decisions are left to the management team developing corporate strategy? How does a chosen pricing and output strategy evolve into specific decisions that affect the profitability of the firm? The answers to these questions are related to the forces of the market structure within which the firm operates.

Sections 2–16 analyze demand, supply, optimal price and output, and factors affecting long-run equilibrium for perfect competition, monopolistic competition, oligopoly, and pure monopoly, respectively.

Section 17 reviews techniques for identifying the various forms of market structure. For example, there are accepted measures of market concentration that are used by regulators of financial institutions to judge whether or not a planned merger or acquisition will harm the competitive nature of regional banking markets. Financial analysts should be able to identify the type of market structure a firm is operating within. Each different structure implies a different long-run sustainability of profits. A summary and practice problems conclude the reading.

1.1 Analysis of Market Structures

Traditionally, economists classify a market into one of four structures: perfect competition, monopolistic competition, oligopoly, and monopoly. Section 1.1.1 explains that four-way classification in more detail. Section 1.1.2 completes the introduction by providing and explaining the major points to evaluate in determining the structure to which a market belongs.

1.1.1 *Economists' Four Types of Structure*

Economists define a market as a group of buyers and sellers that are aware of each other and can agree on a price for the exchange of goods and services. While the internet has extended a number of markets worldwide, certain markets are limited by geographic boundaries. For example, the internet search engine Google operates in a worldwide market. In contrast, the market for premixed cement is limited to the area within which a truck can deliver the mushy mix from the plant to a construction site before the compound becomes useless. Thomas L. Friedman's international best seller *The World Is Flat*[1] challenges the concept of the geographic limitations of the market. If the service being provided by the seller can be digitized, its market expands worldwide. For example, a technician can scan your injury in a clinic in Switzerland. That radiographic image can be digitized and sent to a radiologist in India to be read. As a customer (i.e., patient), you may never know that part of the medical service provided to you was the result of a worldwide market.

Some markets are highly concentrated, with the majority of total sales coming from a small number of firms. For example, in the market for internet search, three firms controlled 98.9 percent of the US market (Google 63.5 percent, Microsoft 24 percent, and Oath (formerly Yahoo) 11.4 percent) as of January 2018.[2] Other markets are very fragmented, such as automobile repairs, where small independent shops often dominate and large chains may or may not exist. New products can lead to market concentration: It is estimated that the Apple iPod had a world market share of over 70 percent among MP3 players in 2009.

---

[1] Friedman (2006).

THE IMPORTANCE OF MARKET STRUCTURE

Consider the evolution of television broadcasting. As the market environment for television broadcasting evolved, the market structure changed, resulting in a new set of challenges and choices. In the early days, there was only one choice: the "free" analog channels that were broadcast over the airwaves. In most countries, there was only one channel, owned and run by the government. In the United States, some of the more populated markets were able to receive more channels because local channels were set up to cover a market with more potential viewers. By the 1970s, new technologies made it possible to broadcast by way of cable connectivity and the choices offered to consumers began to expand rapidly. Cable television challenged the "free" broadcast channels by offering more choice and a better-quality picture. The innovation was expensive for consumers and profitable for the cable companies. By the 1990s, a new alternative began to challenge the existing broadcast and cable systems: satellite television. Satellite providers offered a further expanded set of choices, albeit at a higher price than the free broadcast and cable alternatives. In the early 2000s, satellite television providers lowered their pricing to compete directly with the cable providers.

Today, cable program providers, satellite television providers, and terrestrial digital broadcasters that offer premium and pay-per-view channels compete for customers who are increasingly finding content on the internet and on their mobile devices. Companies like Netflix, Apple, and Amazon offered alternative ways for consumers to access content. By 2018, these companies had moved beyond the repackaging of existing shows to developing their own content, mirroring the evolution of cable channels such as HBO and ESPN a decade earlier.

This is a simple illustration of the importance of market structure. As the market for television broadcasting became increasingly competitive, managers have had to make decisions regarding product packaging, pricing, advertising, and marketing in order to survive in the changing environment. In addition, mergers and acquisitions as a response to these competitive pressures have changed the essential structure of the industry.

Market structure can be broken down into four distinct categories: perfect competition, monopolistic competition, oligopoly, and monopoly.

We start with the most competitive environment, **perfect competition**. Unlike some economic concepts, perfect competition is not merely an ideal based on assumptions. Perfect competition is a reality—for example, in several commodities markets, where sellers and buyers have a strictly homogeneous product and no single producer is large enough to influence market prices. Perfect competition's characteristics are well recognized and its long-run outcome unavoidable. Profits under the conditions of perfect competition are driven to the required rate of return paid by the entrepreneur to borrow capital from investors (so-called normal profit or rental cost of capital). This does not mean that all perfectly competitive industries are doomed to extinction by a lack of profits. On the contrary, millions of businesses that do very well are living under the pressures of perfect competition.

Monopolistic competition is also highly competitive; however, it is considered a form of imperfect competition. Two economists, Edward H. Chamberlin (US) and Joan Robinson (UK), identified this hybrid market and came up with the term because there are not only strong elements of competition in this market structure but also some monopoly-like conditions. The competitive characteristic is a notably large number of firms, while the monopoly aspect is the result of product differentiation. That is, if the seller can convince consumers that its product is uniquely different from other, similar products, then the seller can exercise some degree of pricing power over the market. A good example is the brand loyalty associated with soft drinks such as Coca-Cola. Many of Coca-Cola's customers believe that their beverages are truly different from and better than all other soft drinks. The same is true for fashion creations and cosmetics.

---

[2] Source: www.statista.com/statistics/267161/market-share-of-search-engines-in-the-united-states/.

The **oligopoly** market structure is based on a relatively small number of firms supplying the market. The small number of firms in the market means that each firm must consider what retaliatory strategies the other firms will pursue when prices and production levels change. Consider the pricing behavior of commercial airline companies. Pricing strategies and route scheduling are based on the expected reaction of the other carriers in similar markets. For any given route—say, from Paris, France, to Chennai, India—only a few carriers are in competition. If one of the carriers changes its pricing package, others will likely retaliate. Understanding the market structure of oligopoly markets can help in identifying a logical pattern of strategic price changes for the competing firms.

Finally, the least competitive market structure is **monopoly**. In pure monopoly markets, there are no other good substitutes for the given product or service. There is a single seller, which, if allowed to operate without constraint, exercises considerable power over pricing and output decisions. In most market-based economies around the globe, pure monopolies are regulated by a governmental authority. The most common example of a regulated monopoly is the local electrical power provider. In most cases, the monopoly power provider is allowed to earn a normal return on its investment and prices are set by the regulatory authority to allow that return.

1.1.2 *Factors That Determine Market Structure*

Five factors determine market structure:

1. The number and relative size of firms supplying the product;
2. The degree of product differentiation;
3. The power of the seller over pricing decisions;
4. The relative strength of the barriers to market entry and exit; and
5. The degree of non-price competition.

The number and relative size of firms in a market influence market structure. If there are many firms, the degree of competition increases. With fewer firms supplying a good or service, consumers are limited in their market choices. One extreme case is the monopoly market structure, with only one firm supplying a unique good or service. Another extreme is perfect competition, with many firms supplying a similar product. Finally, an example of relative size is the automobile industry, in which a small number of large international producers (e.g., Volkswagen and Toyota) are the leaders in the global market, and a number of small companies either have market power because they are niche players (e.g., Ferrari or McLaren) or have little market power because of their narrow range of models or limited geographical presence (e.g., Mazda or Fiat-Chrysler).

In the case of monopolistic competition, there are many firms providing products to the market, as with perfect competition. However, one firm's product is differentiated in some way that makes it appear better than similar products from other firms. If a firm is successful in differentiating its product, the differentiation will provide pricing leverage. The more dissimilar the product appears, the more the market will resemble the monopoly market structure. A firm can differentiate its product through aggressive advertising campaigns; frequent styling changes; the linking of its product with other, complementary products; or a host of other methods.

When the market dictates the price based on aggregate supply and demand conditions, the individual firm has no control over pricing. The typical hog farmer in Nebraska and the milk producer in Bavaria are **price takers**. That is, they must accept whatever price the market dictates. This is the case under the market structure of perfect competition. In the case of monopolistic competition, the success of product differentiation determines the degree with which the firm can influence price. In the case of oligopoly, there are so few firms in the market that price control becomes possible. However, the small number of firms in an oligopoly market invites complex pricing strategies. Collusion, price leadership by dominant firms, and other pricing strategies can result.

The degree to which one market structure can evolve into another and the difference between potential short-run outcomes and long-run equilibrium conditions depend on the strength of the barriers to entry and the possibility that firms fail to recoup their original costs or lose money for an extended period of time and are therefore forced to exit the market. Barriers to entry can result from very large capital investment requirements, as in the case of petroleum refining. Barriers may also result from patents, as in the case of some electronic products and drug formulas. Another entry consideration is the possibility of high exit costs. For example, plants that are specific to a special line of products, such as aluminum smelting plants, are non-redeployable, and exit costs would be high without a liquid market for the firm's assets. High exit costs deter entry and are therefore also considered barriers to entry. In the case of farming, the barriers to entry are low. Production of corn, soybeans, wheat, tomatoes, and other produce is an easy process to replicate; therefore, those are highly competitive markets.

Non-price competition dominates those market structures where product differentiation is critical. Therefore, monopolistic competition relies on competitive strategies that may not include pricing changes. An example of non-price competition is product differentiation through marketing. In other circumstances, non-price competition may occur because the few firms in the market feel dependent on each other. Each firm fears retaliatory price changes that would reduce total revenue for all of the firms in the market. Because oligopoly industries have so few firms, each firm feels dependent on the pricing strategies of the others. Therefore, non-price competition becomes a dominant strategy.

Exhibit 1 Characteristics of Market Structure

Market Structure	Number of Sellers	Degree of Product Differentiation	Barriers to Entry	Pricing Power of Firm	Non-price Competition
Perfect competition	Many	Homogeneous/ Standardized	Very Low	None	None
Monopolistic competition	Many	Differentiated	Low	Some	Advertising and Product Differentiation
Oligopoly	Few	Homogeneous/ Standardized	High	Some or Considerable	Advertising and Product Differentiation
Monopoly	One	Unique Product	Very High	Considerable	Advertising

From the perspective of the owners of the firm, the most desirable market structure is that with the most control over price, because this control can lead to large profits. Monopoly and oligopoly markets offer the greatest potential control over price; monopolistic competition offers less control. Firms operating under perfectly competitive market conditions have no control over price. From the consumers' perspective, the most desirable market structure is that with the greatest degree of competition because prices are generally lower. Thus, consumers would prefer as many goods and services as possible to be offered in competitive markets.

As often happens in economics, there is a trade-off. While perfect competition gives the largest quantity of a good at the lowest price, other market forms may spur more innovation. Specifically, there may be high costs in researching a new product, and firms will incur such costs only if they expect to earn an attractive return on their research investment. This is the case often made for medical innovations, for example—the cost of clinical trials and experiments to create new medicines would bankrupt perfectly competitive firms but may be acceptable in an oligopoly market structure. Therefore, consumers can benefit from less-than-perfectly-competitive markets.

PORTER'S FIVE FORCES AND MARKET STRUCTURE

A financial analyst aiming to establish market conditions and consequent profitability of incumbent firms should start with the questions framed by Exhibit 1: How many sellers are there? Is the product differentiated? and so on. Moreover, in the case of monopolies and quasi monopolies, the analyst should evaluate the legislative and regulatory framework: Can the company set prices freely, or are there governmental controls? Finally, the analyst should consider the threat of competition from potential entrants.

This analysis is often summarized by students of corporate strategy as "Porter's five forces," named after Harvard Business School professor Michael E. Porter. His book, *Competitive Strategy*, presented a systematic analysis of the practice of market strategy. Porter (2008) identified the five forces as:

- Threat of entry;
- Power of suppliers;
- Power of buyers (customers);
- Threat of substitutes; and
- Rivalry among existing competitors.

It is easy to note the parallels between four of these five forces and the columns in Exhibit 1. The only "orphan" is the power of suppliers, which is not at the core of the theoretical economic analysis of competition, but which has substantial weight in the practical analysis of competition and profitability.

Some stock analysts (e.g., Dorsey 2004) use the term "economic moat" to suggest that there are factors protecting the profitability of a firm that are similar to the moats (ditches full of water) that were used to protect some medieval castles. A deep moat means that there is little or no threat of entry by invaders, i.e. competitors. It also means that customers are locked in because of high switching costs.

PERFECT COMPETITION & DEMAND ANALYSIS IN PERFECTLY COMPETITIVE MARKETS

2

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

Perfect competition is characterized by the five conditions presented in Exhibit 1, above:

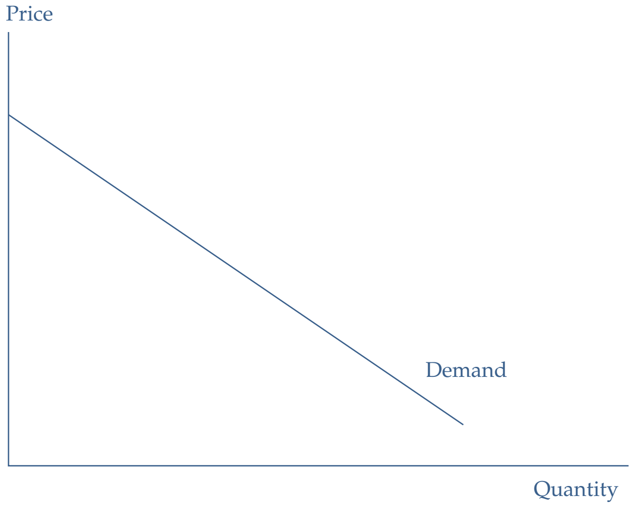
1. There are a large number of potential buyers and sellers.
2. The products offered by the sellers are virtually identical.
3. There are few or easily surmountable barriers to entry and exit.
4. Sellers have no market-pricing power.
5. Non-price competition is absent.

While few markets achieve the distinction of being perfectly competitive, it is useful to establish the outcome associated with this market structure as a benchmark against which other market structures can be compared. The most typical example of perfect competition is found in certain aspects of the agriculture industry, such as the large number of farmers growing corn for animal feed. Corn is a primary source of food for pork, beef, and poultry production. A bushel of corn from Farmer Brown is virtually identical to a bushel of corn from Farmer Lopez. If a hog farmer needs corn to feed his hogs, it does not matter whether the corn comes from Farmer Brown or Farmer Lopez. Furthermore, the aggregate corn market is well defined, with active futures and spot markets. Information about the corn market is easy and inexpensive to access, and there is no way to differentiate the product, such as by advertising. Agribusiness is capital intensive, but where arable land is relatively abundant and water is available, the barriers to entry (e.g., capital and expertise) for corn production are relatively low.

2.1 Demand Analysis in Perfectly Competitive Markets

The price of a homogeneous product sold in a competitive market is determined by the demand and supply in that market. Economists usually represent demand and supply in a market through demand and supply curves in a two-axis plane, where quantity and price are shown on the x-axis and y-axis, respectively. Economists believe that demand functions have negative slopes, as shown in Exhibit 2. That is, at high prices, less is demanded. For normal goods and services, as the price declines, the quantity demanded increases. This concept is based on two effects: the income effect and the substitution effect. The income effect results from the increased purchasing power the consumer has when prices fall. With lower prices, the consumer can afford to purchase more of the product. The substitution effect comes from the increasing attractiveness of the lower-priced product. If soybean prices are unchanged and corn prices decrease, hog farmers will substitute corn for soybeans as feed for their animals.

Exhibit 2 Market Demand in Perfect Competition



Assume the demand for this product can be specified as

$$Q_D = 50 - 2P$$

where Q_D is the quantity of demand and P is the product's price. This demand function can be rearranged in terms of price:

$$P = 25 - 0.5Q_D$$

In this form, total revenue (TR) is equal to price times quantity, or $P \times Q_D$. Thus,

$$\text{TR} = PQ_D = 25Q_D - 0.5Q_D^2$$

Average revenue (AR) can be found by dividing TR by Q_D. Therefore,

$$\text{AR} = \text{TR}/Q_D = \left(25Q_D - 0.5Q_D^2\right)\Big/Q_D = 25 - 0.5Q_D$$

Note that the AR function is identical to the market demand function. The assumption here is that the relationship between price and quantity demanded is linear. Clearly, that may not be the case in the real market. Another simplifying assumption made is that the price of the product is the only determinant of demand. Again, that is not likely in the real market. For example, economic theory suggests that consumer income is another important factor in determining demand. The prices of related goods and services, such as substitutes and complements, are also considered factors affecting demand for a specific product.

Marginal revenue (MR) is the change in total revenue per extra increment sold when the quantity sold changes by a small increment, ΔQ_D. Substituting $(Q_D + \Delta Q_D)$ into the total revenue (TR) equation, marginal revenue can be expressed as:

$$\text{MR} = \frac{\Delta\text{TR}}{\Delta Q_D} = \frac{\left[25\left(Q_D + \Delta Q_D\right) - 0.5\left(Q_D^2 + 2Q_D\Delta Q_D + \Delta Q_D^2\right)\right] - \left[25Q_D - 0.5Q_D^2\right]}{\Delta Q_D}$$

$$= \frac{25\Delta Q_D - Q_D\Delta Q_D - 0.5\Delta Q_D^2}{\Delta Q_D} = 25 - Q_D - 0.5\Delta Q_D$$

For example, suppose $Q_D = 5$ and $\Delta Q_D = 1$, then total revenue increases from 112.50 [= 25(5) – 0.5 (5^2)] to 132 [= 25(6) – 0.5(6^2)], and marginal revenue is 19.5 = (132 – 112.5)/1. Note that marginal revenue is equal to ($25 - Q_D - 0.5\Delta Q_D$). Now suppose that ΔQ_D is much smaller, for example $\Delta Q_D = 0.1$. In this case, total revenue increases to 114.495 [= 25(5.1) – 0.5(5.1^2)], and marginal revenue is 1.995/0.1 = 19.95. It is straightforward to confirm that as ΔQ_D gets smaller marginal revenue gets closer to $20 = 25 - Q_D$. So, for very small changes in the quantity sold we can write marginal revenue as[3]

$$\mathrm{MR} = 25 - Q_D$$

Although we have introduced the concept of marginal revenue in the context of the demand curve for the market as a whole, its usefulness derives from its role in the output and pricing decisions of individual firms. As we will see, marginal revenue and an analogous concept, marginal cost, are critical in determining firms' profit-maximizing strategies.

ELASTICITY OF DEMAND

3

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

Consumers respond differently to changes in the price of different kinds of products and services. The quantity demanded for some products is very price sensitive, while for other products, price changes result in little change in the quantity demanded. Economists refer to the relationship between changes in price and changes in the quantity demanded as the price elasticity of demand. Therefore, the demand for the former group of products—those that are very price sensitive—is said to have high price elasticity, whereas the demand for the latter group is said to have low price elasticity. Understanding the sensitivity of demand changes to changes in price is critical to understanding market structures.

Price elasticity of demand measures the percentage change in the quantity demanded given a percentage change in the price of a given product. Because the relationship of demand to price is negative, the price elasticity of demand would be negative. *Many economists, however, present the price elasticity as an absolute value, so that price elasticity has a positive sign. We will follow that convention.* Higher price elasticity indicates that consumers are very responsive to changes in price. Lower values for price elasticity imply that consumers are not very responsive to price changes. Price elasticity can be measured with the following relationship:

$$\varepsilon_P = -(\%\ \text{change in}\ Q_D) \div (\%\ \text{change in}\ P)$$

where ε_P is price elasticity of demand, Q_D is the quantity demanded, and P is the product's price.

Price elasticity of demand falls into three categories. When demand is very responsive to price change, it is identified as *elastic*. When demand is not responsive to price change, it is identified as *inelastic*. When the percentage change in quantity demanded is exactly the same as the percentage change in price, the demand is called *unitary elastic*.

$\varepsilon_P > 1$ Demand is elastic

$\varepsilon_P = 1$ Demand is unitary elastic

$\varepsilon_P < 1$ Demand is inelastic

---

[3] Readers who are familiar with calculus will recognize this as the derivative of total revenue with respect to the quantity sold.

Price elasticity of demand depends on several factors. *Price elasticity will be higher if there are many close substitutes for the product.* If a product has many good alternatives, consumers will be more sensitive to price changes. For example, carbonated beverages (“soft drinks”) have many close substitutes. It takes strong brand loyalty to keep customer demand high in the soft drink market when one brand’s price is strategically lowered; the price elasticity of demand for Coca-Cola has been estimated to be 3.8. For products with numerous close substitutes, demand is highly elastic. For products with few close substitutes, demand is lower in price elasticity and would be considered price inelastic. The demand for first-class airline tickets is often seen as inelastic because only very wealthy people are expected to buy them; the demand for economy-class tickets is elastic because the typical consumer for this product is more budget-conscious. Consumers do not consider economy-class airline tickets a close substitute for first-class accommodations, particularly on long flights.

The airline ticket example introduces another determinant of price elasticity of demand. *The greater the share of the consumer’s budget spent on the item, the higher the price elasticity of demand*. Expensive items, such as durable goods (e.g., refrigerators and televisions), tend to have higher elasticity measures, while less expensive items, such as potatoes and salt, have lower elasticity values. Consumers will not change their normal salt consumption if the price of salt decreases by 10 percent. Instead, they will buy their next package of salt when they run out, with very little regard to the price change.

The airline ticket also makes a good example for the final factor determining price elasticity. *Price elasticity of demand also depends on the length of time within which the demand schedule is being considered*. Holiday airline travel is highly price elastic. Consumers shop vigorously for vacation flights because they have time to plan their holiday. Business airline travelers typically have less flexibility in determining their schedules. If your business requires a face-to-face meeting with a client, then the price of the ticket is somewhat irrelevant. If gasoline prices increase, there is very little you can do in the short run but pay the higher price. However, evidence of commuter choices indicates that many use alternative transportation methods after the gasoline price spikes. In the long run, higher gasoline prices will lead consumers to change their modes of transportation, trading in less efficient vehicles for automobiles with higher gas mileage or public transit options where available.

There are two extreme cases of price elasticity of demand. One extreme is the **horizontal demand schedule**. This term implies that at a given price, the response in the quantity demanded is infinite. *This is the demand schedule faced by a perfectly competitive firm because it is a price taker*, as in the case of a corn farmer. If the corn farmer tried to charge a higher price than the market price, nobody would buy her product. On the other hand, the farmer has no incentive to sell at a lower price because she can sell all she can produce at the market price. In a perfectly competitive market the quantity supplied by an individual firm has a negligible effect on the market price. In the case of *perfect price elasticity*, the measure is $\varepsilon_P = \infty$.

The other extreme is the **vertical demand schedule**. The vertical demand schedule implies that some fixed quantity is demanded, regardless of price. An example of such demand is the diabetic consumer with the need for a certain amount of insulin. If the price of insulin goes up, the patient will not consume less of it. The amount desired is set by the patient’s medical condition. The measure for *perfect price inelasticity* is $\varepsilon_P = 0$.

The nature of the elasticity calculation and consumer behavior in the marketplace imply that for virtually any product (excluding cases of perfect elasticity and perfect inelasticity), demand is more elastic at higher prices and less elastic (more inelastic) at lower prices. For example, at current low prices, the demand for table salt is very inelastic. However, if table salt increased in price to hundreds of dollars per ounce, consumers would become more responsive to its price changes. Exhibit 3 reports several empirical estimates of price elasticity of demand.

Exhibit 3 Empirical Price Elasticities[4]

Commodity (Good/Service)	Price Elasticity of Market Demand
Alcoholic beverages consumed at home	
Beer	0.84
Wine	0.55
Liquor	0.50
Coffee	
Regular	0.16
Instant	0.36
Credit charges on bank cards	2.44
Furniture	3.04
Glassware/china	1.20
International air transportation United States/Europe	1.20
Shoes	0.73
Soybean meal	1.65
Tomatoes	2.22

OTHER FACTORS AFFECTING DEMAND

4

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

There are two other important forces that influence shifts in consumer demand. One influential factor is consumer income and the other is the price of a related product. For normal goods, as consumer income increases, the demand increases. The degree to which consumers respond to higher incomes by increasing their demand for goods and services is referred to as income elasticity of demand. **Income elasticity of demand** measures the responsiveness of demand to changes in income. The calculation is similar to that of price elasticity, with the percentage change in income replacing the percentage change in price. Note the new calculation below:

$$\varepsilon_Y = (\%\text{ change in } Q_D) \div (\%\text{ change in } Y)$$

where ε_Y is income elasticity of demand, Q_D is the quantity demanded, and Y is consumer income. For normal goods, the measure ε_Y will be a positive value. That is, as consumers' income rises, more of the product is demanded. For products that are considered luxury items, the measure of income elasticity will be greater than one. There are other goods and services that are considered inferior products. For inferior products, as consumer income rises, less of the product is demanded. Inferior products will have negative values for income elasticity. For example, a person on a small income may watch television shows, but if this person had more income, she would prefer going to live concerts and theater performances; in this example, television shows would be the inferior good.

---

[4] Various sources, as noted in McGuigan, Moyer, and Harris (2008), p. 95. These are the elasticities with respect to the product's own price; by convention, they are shown here as positive numbers.

As a technical issue, the difference between price elasticity of demand and income elasticity of demand is that the demand adjustment for price elasticity represents a movement *along the demand schedule* because the demand schedule represents combinations of price and quantity. The demand adjustment for income elasticity represents a *shift in the demand curve* because with a higher income one can afford to purchase more of the good at any price. For a normal good, an increase in income would shift the demand schedule out to the right, away from the origin of the graph, and a decrease in income would shift the demand curve to the left, toward the origin.

The final factor influencing demand for a product is the change in price of a related product, such as a strong substitute or a complementary product. If a close competitor in the beverage market lowers its price, consumers will substitute that product for your product. Thus, your product's demand curve will shift to the left, toward the origin of the graph. **Cross-price elasticity of demand** is the responsiveness of the demand for product A that is associated with the change in price of product B:

$$\varepsilon_X = (\%\text{ change in } Q_{DA}) \div (\%\text{ change in } P_B)$$

where ε_X is cross-price elasticity of demand, Q_{DA} is the quantity demanded of product A, and P_B is the price of product B.

When the cross-price elasticity of demand between two products is *positive*, the two products are considered to be **substitutes**. For example, you may expect to have positive cross-price elasticity between honey and sugar. If the measure of cross-price elasticity is *negative*, the two products are referred to as **complements** of each other. For example, if the price of DVDs goes up, you would expect consumers to buy fewer DVD players. In this case, the cross-price elasticity of demand would have a negative value.

Reviewing cross-price elasticity values provides a simple test for the degree of competition in the market. The more numerous and the closer the substitutes for a product, the lower the pricing power of firms selling in that market; the fewer the substitutes for a product, the greater the pricing power. One interesting application was a US Supreme Court case involving the production and sale of cellophane by DuPont.[5] The court noted that the relevant product market for DuPont's cellophane was the broader flexible packaging materials market. The Supreme Court found the cross-price elasticity of demand between cellophane and other flexible packaging materials to be sufficiently high and exonerated DuPont from a charge of monopolizing the market.

Because price elasticity of demand relates changes in price to changes in the quantity demanded, there must be a logical relationship between marginal revenue and price elasticity. Recall that marginal revenue equals the change in total revenue given a change in output or sales. An increase in total revenue results from a decrease in price that results in an increase in sales. In order for the increase in the quantity demanded to be sufficient to offset the decline in price, the percentage change in quantity demanded must be greater than the percentage decrease in price. The relationship between TR and price elasticity is as follows:

$\varepsilon_P > 1$ Demand is elastic	$\uparrow P \rightarrow \text{TR} \downarrow$ and $\downarrow P \rightarrow \text{TR} \uparrow$
$\varepsilon_P = 1$ Demand is unitary elastic	$\updownarrow P \rightarrow$ no change in TR
$0 < \varepsilon_P < 1$ Demand is inelastic	$\uparrow P \rightarrow \text{TR} \uparrow$ and $\downarrow P \rightarrow \text{TR} \downarrow$

---

[5] *US v. DuPont*, 351 US 377 (1956), as noted in McGuigan, Moyer, and Harris (2008).

Total revenue is maximized when marginal revenue is zero. The logic is that as long as marginal revenue is positive (i.e., each additional unit sold contributes to additional total revenue), total revenue will continue to increase. Only when marginal revenue becomes negative will total revenue begin to decline. Therefore, the percentage decrease in price is greater than the percentage increase in quantity demanded. The relationship between marginal revenue (MR) and price elasticity can be expressed as

$$\text{MR} = P[1 - (1/\varepsilon_P)]$$

An understanding of price elasticity of demand is an important strategic tool. It would be very useful to know in advance what would happen to your firm's total revenue if you increased the product's price. If you are operating in the inelastic portion of the demand curve, increasing the price of the product will increase total revenue. On the other hand, if you are operating in the elastic portion of the product's demand curve, increasing the price will decrease total revenue.

Decision makers can also use the relationship between marginal revenue and price elasticity of demand in other ways. For example, suppose you are a farmer considering planting soybeans or some other feed crop, such as corn. From Exhibit 3, we know that soybean meal's price elasticity of demand has been estimated to be 1.65. We also know that the current (May 2018) soybean meal price is $465.00 per metric ton.[6] Therefore, by solving the equation above, we find that the expected marginal revenue per metric ton of soybean meal is $183.16. Soybeans may prove to be a profitable crop for the farmer. Just a few years earlier, in May of 2014, the price of a metric ton of soybean meal was $578.75. Given the crop's price elasticity of demand, the estimated marginal revenue per metric ton was then $227.97. The higher price translates into higher marginal revenue and might have induced the farmer to plant even more soybeans rather than another feed crop instead.

How do business decision makers decide what level of output to bring to the market? To answer that question, the firm must understand its cost of resources, its production relations, and its supply function. Once the supply function is well defined and understood, it is combined with the demand analysis to determine the profit-maximizing levels of output.

CONSUMER SURPLUS: VALUE MINUS EXPENDITURE

5

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

To this point, we have discussed the fundamentals of supply and demand curves and explained a simple model of how a market can be expected to arrive at an equilibrium combination of price and quantity. While it is certainly necessary for the analyst to understand the basic workings of the market model, it is also crucial to have a sense of why we might care about the nature of the equilibrium. In this section we review the concept of **consumer surplus**, which is helpful in understanding and evaluating business pricing strategies. Consumer surplus is defined as the difference between the value that a consumer places on the units purchased and the amount of money that was required to pay for them. It is a measure of the value gained by the buyer from the transaction.

---

[6] Source: World-Bank Commodity Market Report 2018.

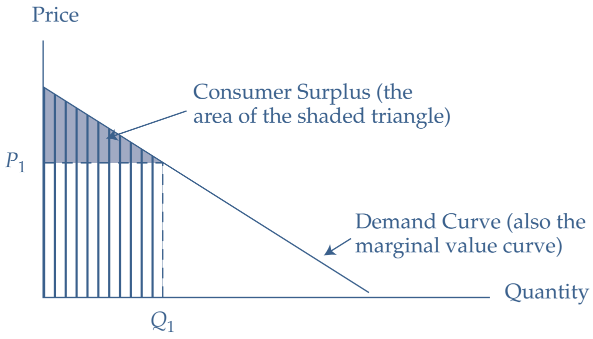
To get an intuitive feel for the concept of consumer surplus, consider the last thing you purchased. Whatever it was, think of how much you paid for it. Now contrast that price with the maximum amount you *would have been willing to pay* rather than go without the item altogether. If those two numbers are different, we say you received some consumer surplus from your purchase. You got a "bargain" because you would have been willing to pay more than you had to pay.

Earlier, we referred to the law of demand, which says that as price falls, consumers are willing to buy more of the good. This observation translates into a negatively sloped demand curve. Alternatively, we could say that the highest price that consumers are willing to pay for an additional unit declines as they consume more and more of a good. In this way, we can interpret their *willingness to pay* as a measure of how much they *value* each additional unit of the good. This is a very important point: To purchase a unit of some good, consumers must give up something else they value. So, the price they are willing to pay for an additional unit of a good is a measure of how much they value that unit, in terms of the other goods they must sacrifice to consume it.

If demand curves are negatively sloped, it must be because the value of each additional unit of the good falls as more of the good is consumed. We shall explore this concept further below, but for now, it is enough to recognize that the demand curve can therefore be considered a **marginal value curve**, because it shows the highest price consumers would be willing to pay for each additional unit. In effect, the demand curve is the willingness of consumers to pay for each additional unit.

This interpretation of the demand curve allows us to measure the total value of consuming any given quantity of a good: It is the sum of all the marginal values of each unit consumed, up to and including the last unit. Graphically, this measure translates into the area under the consumer's demand curve, up to and including the last unit consumed, as shown in Exhibit 4, where the consumer is choosing to buy Q_1 units of the good at a price of P_1. The marginal value of the $Q_1{}^{th}$ unit is clearly P_1 because that is the highest price the consumer is willing to pay for that unit. Importantly, however, the marginal value of each unit *up to* the $Q_1{}^{th}$ is greater than P_1.

Exhibit 4 Consumer Surplus



Note: Consumer surplus is the area beneath the demand curve and above the price paid.

Because the consumer would have been willing to pay more for each of those units than she paid (P_1), we can say she received more value than the cost to her of buying them. This extra value is the buyer's consumer surplus. The *total value* of quantity Q_1 to the buyer is the area of the vertically crosshatched trapezoid in Exhibit 4. The *total expenditure* is only the area of the rectangle with height P_1 and base Q_1 (bottom section). The total consumer surplus received from buying Q_1 units at a level price of

P_1 per unit is the difference between the area under the demand curve and the area of the rectangle $P_1 \times Q_1$. The resulting area is shown as the lightly shaded triangle (upper section).

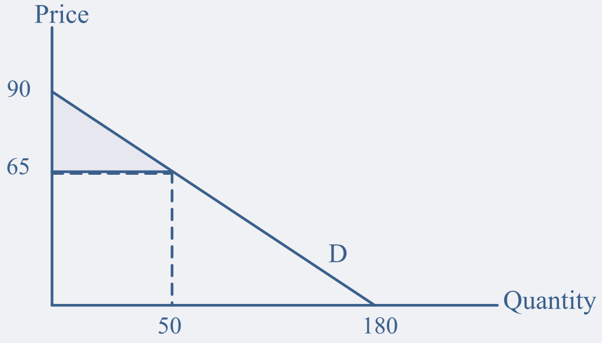
EXAMPLE 1

Consumer Surplus

A market demand function is given by the equation $Q_D = 180 - 2P$. Find the value of consumer surplus if price is equal to 65.

Solution:

First, input 65 into the demand function to find the quantity demanded at that price: $Q_D = 180 - 2(65) = 50$. Then, to make drawing the demand curve easier, invert the demand function by solving for P in terms of Q_D: $P = 90 - 0.5Q_D$. Note that the price intercept is 90 and the quantity intercept is 180. Draw the demand curve:



Find the area of the triangle above the price of 65 and below the demand curve, up to quantity 50: Area = ½ (Base)(Height) = ½ (50)(25) = 625.

SUPPLY ANALYSIS & OPTIMAL PRICE AND OPTIMAL OUTPUT IN PERFECTLY COMPETITIVE MARKETS

6

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

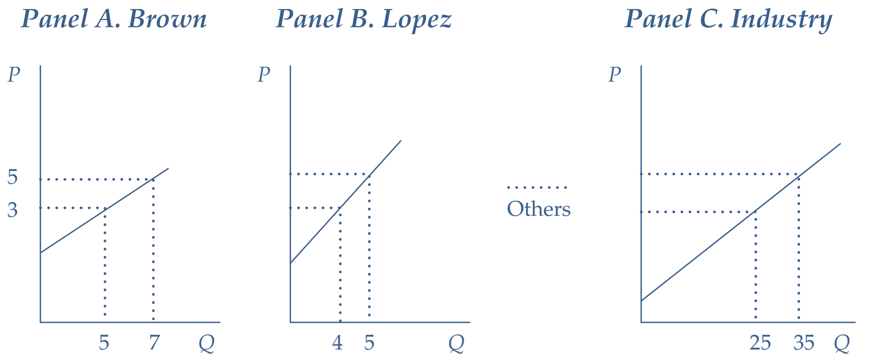
c describe a firm's supply function under each market structure;

d describe and determine the optimal price and output for firms under each market structure;

f describe pricing strategy under each market structure;

Consider two corn farmers, Mr. Brown and Ms. Lopez. They both have land available to them to grow corn and can sell at one price, say 3 currency units per kilogram. They will try to produce as much corn as is profitable at that price. If the price is driven up to 5 currency units per kilogram by new consumers entering the market—say, ethanol producers—Mr. Brown and Ms. Lopez will try to produce more corn. To increase their output levels, they may have to use less productive land, increase irrigation, use more fertilizer, or all three. Their production costs will likely increase. They will both still try to produce as much corn as possible to profit at the new, higher price of 5 currency units per kilogram. Exhibit 5 illustrates this example. Note that the supply functions for the individual firms have positive slopes. Thus, as prices increase, the firms supply greater quantities of the product.

Exhibit 5 Firm and Market Supply in Perfect Competition



Notice that the market supply curve is the sum of the supply curves of the individual firms—Brown, Lopez, and others—that make up the market. Assume that the supply function for the market can be expressed as a linear relationship, as follows:

$$Q_S = 10 + 5P, \text{ or } P = -2 + 0.2Q_S,$$

where Q_S is the quantity supplied and P is the price of the product.

Before we analyze the optimal supply level for the firm, we need to point out that economic costs and profits differ from accounting costs and profits in a significant way. **Economic costs** include all the remuneration needed to keep the productive resource in its current employment or to acquire the resource for productive use.

To evaluate the remuneration needed to keep the resource in its current use and attract new resources for productive use, economists refer to the resource's **opportunity cost**. Opportunity cost is measured by determining the resource's next best opportunity. If a corn farmer could be employed in an alternative position in the labor market with an income of 50,000, then the opportunity cost of the farmer's labor is 50,000. Similarly, the farmer's land and capital could be leased to another farmer or sold and reinvested in another type of business. The return foregone by not doing so is an opportunity cost. In economic terms, total cost includes the full normal market return on all the resources utilized in the business. **Economic profit** is the difference between TR and total cost (TC). Economic profit differs from accounting profit because accounting profit does not include opportunity cost. Accounting profit includes only explicit payments to outside providers of resources (e.g. workers, vendors, lenders) and depreciation based on the historic cost of physical capital.

6.1 Optimal Price and Output in Perfectly Competitive Markets

Carrying forward our examples from earlier sections, we can now combine the market supply and demand functions to solve for the equilibrium price and quantity, where Q^* represents the equilibrium level of both supply and demand.

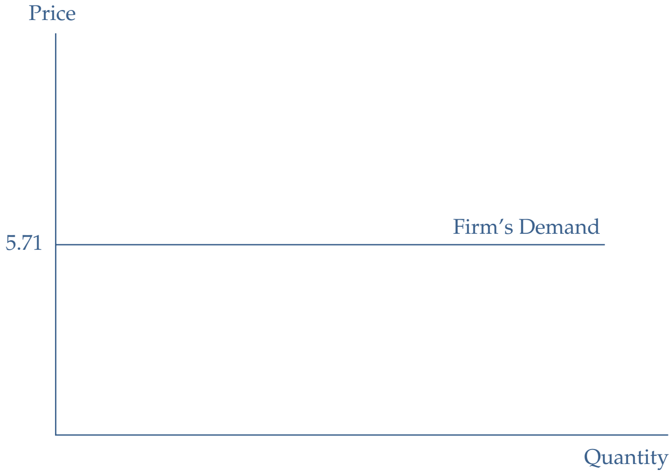
$$\begin{aligned} P = 25 - 0.5Q_D &= -2 + 0.2Q_S = P \\ 25 - 0.5Q_D &= -2 + 0.2Q_S \\ 27 &= 0.7Q^* \\ Q^* &= 38.57 \end{aligned}$$

According to the market demand curve, the equilibrium price is

$$P = 25 - 0.5Q^* = 25 - 0.5(38.57) = 25 - 19.29 = 5.71.$$

With many firms in the market and total output in the market of almost 39 units of the product, the effective market price would be 5.71. This result becomes the demand function for each perfectly competitive firm. Even if a few individual producers could expand production, there would not be a noticeable change in the market equilibrium price. In fact, if any one firm could change the equilibrium market price, the market would not be in perfect competition. Therefore, the demand curve that each perfectly competitive firm faces is a horizontal line at the equilibrium price, as shown in Exhibit 6, even though the demand curve for the whole market is downward sloping.

Exhibit 6 Individual Firm's Demand in Perfect Competition



EXAMPLE 2

Demand Curves in Perfect Competition

Is it possible that the demand schedule faced by Firm A is horizontal while the demand schedule faced by the market as a whole is downward sloping?

A No, because Firm A can change its output based on demand changes.

B No, because a horizontal demand curve means that elasticity is infinite.

C Yes, because consumers can go to another firm if Firm A charges a higher price, and Firm A can sell all it produces at the market price.

Solution:

C is correct. Firm A cannot charge a higher price and has no incentive to sell at a price below the market price.

To analyze the firm's revenue position, recall that average revenue is equivalent to the firm's demand function. Therefore, the horizontal line that represents the firm's demand curve is the firm's AR schedule.

Marginal revenue is the incremental increase in total revenue associated with each additional unit sold. For every extra unit the firm sells, it receives 5.71. Thus, the firm's MR schedule is also the horizontal line at 5.71. TR is calculated by multiplying AR by the quantity of products sold. Total revenue is the area under the AR line at the point where the firm produces the output. In the case of perfect competition, the following conditions hold for the individual firm:

Price = Average revenue = Marginal revenue

The next step is to develop the firm's cost functions. The firm knows that it can sell the entire product it produces at the market's equilibrium price. How much should it produce? That decision is determined by analysis of the firm's costs and revenues. A corn farmer uses three primary resources: land, labor, and capital. In economics, capital is any man-made aid to production. For the corn farmer, his or her capital includes the irrigation system, tractors, harvesters, trucks, grain bins, fertilizer, and so forth. The labor includes the farmer, perhaps members of the farmer's family, and hired labor. In the initial stages of production, only the farmer and the farmer's family are cultivating the land, with a significant investment in capital. They have a tractor, fertilizer, irrigation equipment, grain bins, seed, and a harvester. The investment in land and capital is relatively high compared with the labor input. In this production phase, the average cost of producing a bushel of corn is high. As they begin to expand by adding labor to the collection of expensive land and capital, the average cost of producing corn begins to decline—for example, because one tractor can be used more intensively to plow a larger amount of land. When the combination of land, labor, and capital approaches an efficient range, the average cost of producing a bushel of corn declines.

Given a certain level of technology, there is a limit to the increase in productivity. Eventually something begins to cause declining marginal productivity. That is, each additional unit of input produces a progressively smaller increase in output. This force is called the **law of diminishing returns**. This "law" helps define the shape of the firm's cost functions. Average cost and marginal cost will be U-shaped. Over the initial stages of output, average and marginal costs will decline. At some level of output, the law of diminishing returns will overtake the efficiencies in production and average and marginal costs will increase.

Average cost (AC) is Total cost (TC) divided by Output (Q). Therefore,

$$\text{AC} = \text{TC}/Q$$

Note that we have defined average cost (AC) in terms of total costs. Many authors refer to this as "average total cost" to distinguish it from a related concept, "average variable cost," which omits fixed costs. In the remainder of this reading, *average cost should be understood to mean average total cost*.

Marginal cost (MC) is the change in TC associated with an incremental change in output:

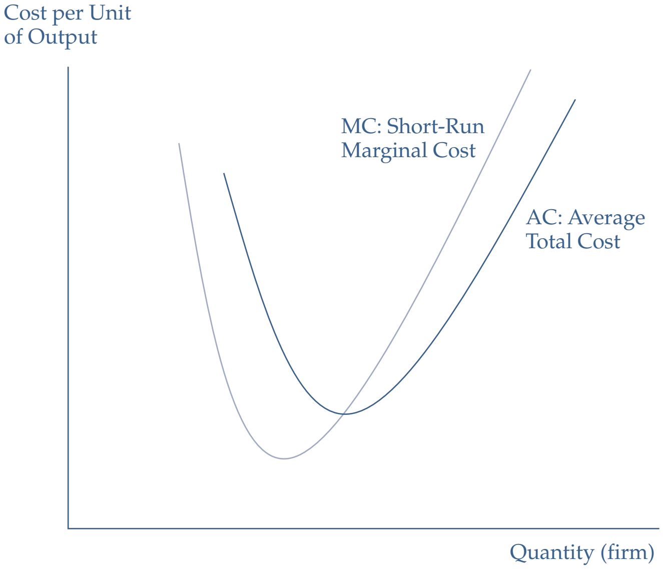
$$\text{MC} = \Delta\text{TC}/\Delta Q$$

By definition, fixed costs do not vary with output, so marginal cost reflects only changes in variable costs.[7] MC declines initially because processes can be made more efficient and specialization makes workers more proficient at their tasks. However, at some higher level of output, MC begins to increase (e.g., must pay workers a higher wage to have them work overtime and, in agriculture, less fertile land must be brought into production). MC and AC will be equal at the level of output where AC is minimized. This is a mathematical necessity and intuitive. If you employ the least expensive labor in the initial phase of production, average and marginal cost will decline. Eventually, additional labor will be more costly. For example, if the labor market is at or near full employment, in order to attract additional workers, you must pay higher wages than they are currently earning elsewhere. Thus, the additional (marginal) labor is more costly, and the higher cost increases the overall average as soon as MC exceeds AC. Exhibit 7 illustrates the relationship between AC and MC.

---

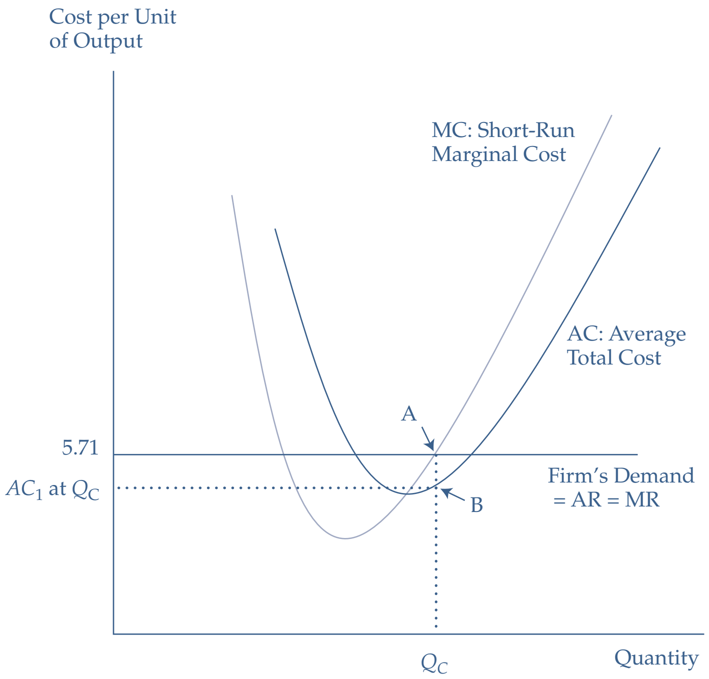
[7] Readers who are familiar with calculus will recognize that MC is simply the derivative of total cost with respect to quantity produced.

Exhibit 7 Individual Firm's Short-Run Cost Schedules



Now combine the revenue and cost functions from Exhibits 6 and 7. In short-run equilibrium, the perfectly competitive firm can earn an economic profit (or an economic loss). In this example, the equilibrium price, 5.71, is higher than the minimum AC. The firm will always maximize profit at an output level where MR = MC. Recall that in perfect competition, the horizontal demand curve is the marginal revenue and average revenue schedules. By setting output at point A in Exhibit 8, where MR = MC, the firm will maximize profits. Total revenue is equal to $P \times Q$—in this case, 5.71 times Q_C. Total cost is equal to Q_C times the average cost of producing Q_C, at point B in Exhibit 8. The difference between the two areas is economic profit.

Exhibit 8 Perfectly Competitive Firm's Short-Run Equilibrium



7

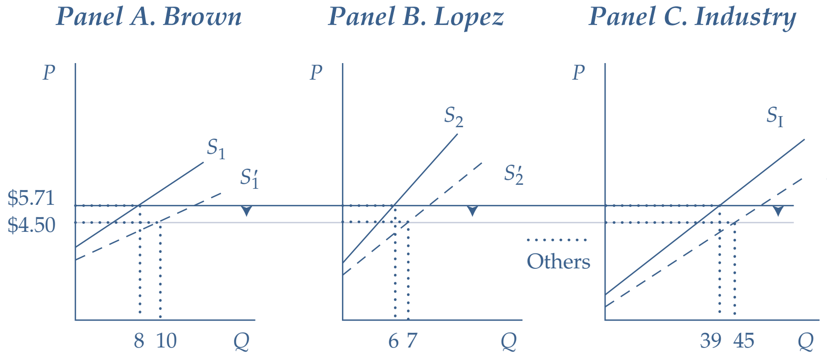
FACTORS AFFECTING LONG-RUN EQUILIBRIUM IN PERFECTLY COMPETITIVE MARKETS

e explain factors affecting long-run equilibrium under each market structure;

f describe pricing strategy under each market structure;

In the long run, economic profit will attract other entrepreneurs to the market, resulting in the production of more output. The aggregate supply will increase, shifting the industry supply (S_1) curve to the right, away from the origin of the graph. For a given demand curve, this increase in supply at each price level will lower the equilibrium price, as shown in Exhibit 9.

Exhibit 9 Perfectly Competitive Market with Increased Supply

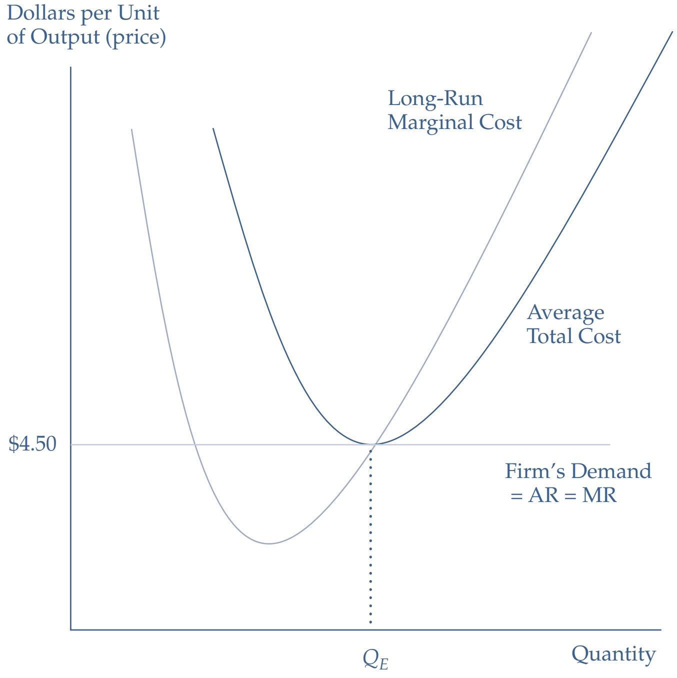

Panel A. Brown *Panel B. Lopez* *Panel C. Industry*

P
S_1
S_1'
$5.71
$4.50
8 10
Q
P
S_2
S_2'
6 7
Q
Others
P
S_I
S_I'
39 45
Q

In the long run, the perfectly competitive firm will operate at the point where marginal cost equals the minimum of average cost, because at that point, entry is no longer profitable: In equilibrium, price equals not only marginal cost (firm equilibrium) but also minimum average cost, so that total revenues equal total costs. This result implies that the perfectly competitive firm operates with zero economic profit. That is, the firm receives its normal profit (rental cost of capital), which is included in its economic costs. Recall that economic profits occur when total revenue exceeds total cost (and therefore differ from accounting profits). With low entry cost and homogeneous products to sell, the perfectly competitive firm earns zero economic profit in the long run.

Exhibit 10 illustrates the long-run equilibrium position of the perfectly competitive firm. Note that total revenue equals price ($4.50) times quantity ($Q_E$) and total cost equals average cost ($4.50) times quantity ($Q_E$).

Exhibit 10 Perfectly Competitive Firm's Long-Run Equilibrium


Dollars per Unit
of Output (price)
Long-Run
Marginal Cost
Average
Total Cost
$4.50
Firm's Demand
= AR = MR
Q_E
Quantity

The long-run marginal cost schedule is the perfectly competitive firm's supply curve. The firm's demand curve is dictated by the aggregate market's equilibrium price. The basic rule of profit maximization is that MR = MC, as is the case in long-run equilibrium. The firm's demand schedule is the same as the firm's marginal revenue and average revenue. Given its cost of operation, the only decision the perfectly competitive firm faces is how much to produce. The answer is the level of output that maximizes its return, and that level is where MR = MC. The demand curve is perfectly elastic. Of course, the firm constantly tries to find ways to lower its cost in the long run.

SCHUMPETER ON INNOVATION AND PERFECT COMPETITION

The Austrian-American economist Joseph A. Schumpeter[8] pointed out that technical change in economics can happen in two main ways:

1. Innovation of process: a new, more efficient way to produce an existing good or service.
2. Innovation of product: a new product altogether or an innovation upon an existing product.

Innovation of process is related to production methods. For example, instead of mixing cement by hand, since the invention of the electric engine it has been possible to use electric mixers. A more recent innovation has been to use the internet to provide technical support to personal computer users: A technician can remotely log on to the customer's PC and fix problems instead of providing instructions over the phone. The result is likely the same, but the process is more efficient.

Innovation of product is related to the product itself. MP3 players, smart phones, robot surgery, and GPS vehicle monitoring have existed only for a few years. They are new products and services. While portable music players existed before the MP3 player, no similar service existed before GPS monitoring of personal vehicles and freight trucks was invented.

How does the reality of continuous innovation of product and process, which is a characteristic of modern economies, fit into the ideal model of perfect competition, where the product is made by a huge number of tiny, anonymous suppliers? This seems a contradiction because the tiny suppliers cannot all be able to invent new products—and indeed, the markets for portable music players and smart phones do not look like perfect competition.

Schumpeter suggested that perfect competition is more of a long-run type of market. In the short run, a company develops a new process or product and is the only one to take advantage of the innovation. This company likely will have high profits and will outpace any competitors. A second stage is what Schumpeter called the swarming (as when a group of bees leaves a hive to follow a queen): In this case, some entrepreneurs notice the innovation and follow the innovator through imitation. Some of them will fail, while others will succeed and possibly be more successful than the initial innovator. The third stage occurs when the new technology is no longer new because everyone has imitated it. At this point, no economic profits are realized, because the new process or product is no longer a competitive advantage, in the sense that everyone has it—which is when perfect competition prevails and we have long-run equilibrium until a new innovation of process or product is introduced.

---

[8] See part 2 of Schumpeter (1942) for the famous "creative destruction" process.

SUPPLY, DEMAND, OPTIMAL PRICING, AND OPTIMAL OUPUT UNDER MONOPOLISTIC COMPETITION

8

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

c describe a firm's supply function under each market structure;

d describe and determine the optimal price and output for firms under each market structure;

f describe pricing strategy under each market structure;

Early in the 20th century, economists began to realize that most markets did not operate under the conditions of perfect competition.[9] Many market structures exhibited characteristics of strong competitive forces; however, other distinct non-competitive factors played important roles in the market. As the name implies, monopolistic competition is a hybrid market. *The most distinctive factor in monopolistic competition is product differentiation.* Recall the characteristics from Exhibit 1:

1. There are a large number of potential buyers and sellers.
2. The products offered by each seller are close substitutes for the products offered by other firms, and each firm tries to make its product look different.
3. Entry into and exit from the market are possible with fairly low costs.
4. Firms have some pricing power.
5. Suppliers differentiate their products through advertising and other non-price strategies.

While the market is made up of many firms that compose the product group, each producer attempts to distinguish its product from that of the others. Product differentiation is accomplished in a variety of ways. For example, consider the wide variety of communication devices available today. Decades ago, when each communication market was controlled by a regulated single seller (the telephone company), all telephones were alike. In the deregulated market of today, the variety of physical styles and colors is extensive. All versions accomplish many of the same tasks.

The communication device manufacturers and providers differentiate their products with different colors, styles, networks, bundled applications, conditional contracts, functionality, and more. Advertising is usually the avenue pursued to convince consumers there is a difference between the goods in the product group. Successful advertising and trademark branding result in customer loyalty. A good example is the brand loyalty associated with Harley-Davidson motorcycles. Harley-Davidson's customers believe that their motorcycles are truly different from and better than all other motorcycles. The same kind of brand loyalty exists for many fashion creations and cosmetics.

The extent to which the producer is successful in product differentiation determines pricing power in the market. Very successful differentiation results in a market structure that resembles the single-seller market (monopoly). However, because there are relatively low entry and exit costs, competition will, in the long run, drive prices and revenues down toward an equilibrium similar to perfect competition. Thus, the hybrid market displays characteristics found in both perfectly competitive and monopoly markets.

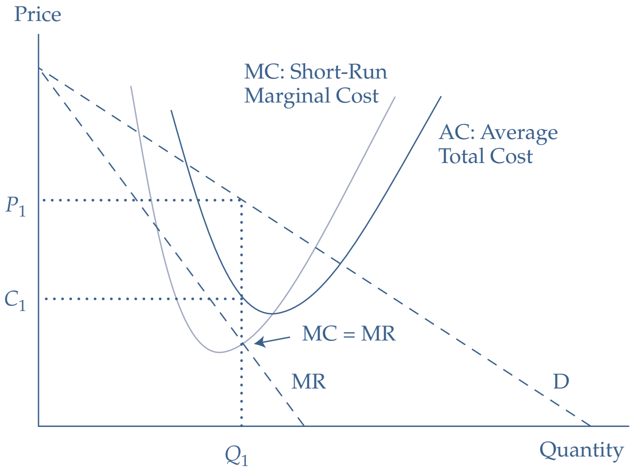
---

[9] Chamberlin (1933).

8.1 Demand Analysis in Monopolistically Competitive Markets

Because each good sold in the product group is somewhat different from the others, the demand curve for each firm in the monopolistic competition market structure is downward sloping to the right. Price and the quantity demanded are negatively related. Lowering the price will increase the quantity demanded, and raising the price will decrease the quantity demanded. There will be ranges of prices within which demand is elastic and (lower) prices at which demand is inelastic. Exhibit 11 illustrates the demand, marginal revenue, and cost structures facing a monopolistically competitive firm in the short run.

Exhibit 11 Short-Run Equilibrium in Monopolistic Competition



In the short run, the profit-maximizing choice is the level of output where MR = MC. Because the product is somewhat different from that of the competitors, the firm can charge the price determined by the demand curve. Therefore, in Exhibit 11, Q_1 is the ideal level of output and P_1 is the price consumers are willing to pay to acquire that quantity. Total revenue is the area of the rectangle $P_1 \times Q_1$.

8.2 Supply Analysis in Monopolistically Competitive Markets

In perfect competition, the firm's supply schedule is represented by the marginal cost schedule. In monopolistic competition, there is no well-defined supply function. The information used to determine the appropriate level of output is based on the intersection of MC and MR. However, the price that will be charged is based on the market demand schedule. The firm's supply curve should measure the quantity the firm is willing to supply at various prices. That information is not represented by either marginal cost or average cost.

8.3 Optimal Price and Output in Monopolistically Competitive Markets

As seen in Section 8.1, in the short run, the profit-maximizing choice is the level of output where MR = MC and total revenue is the area of the rectangle $P_1 \times Q_1$ in Exhibit 11.

The average cost of producing Q_1 units of the product is C_1, and the total cost is the area of the rectangle $C_1 \times Q_1$. The difference between TR and TC is economic profit. The profit relationship is described as

$$\pi = \text{TR} - \text{TC}$$

where π is total profit, TR is total revenue, and TC is total cost.

THE BENEFITS OF IMPERFECT COMPETITION

Is monopolistic competition indeed imperfect—that is, is it a bad thing? At first, one would say that it is an inefficient market structure because prices are higher and the quantity supplied is less than in perfect competition. At the same time, in the real world, we see more markets characterized by monopolistic competition than markets meeting the strict conditions of perfect competition. If monopolistic competition were that inefficient, one wonders, why would it be so common?

A part of the explanation goes back to Schumpeter. Firms try to differentiate their products to meet the needs of customers. Differentiation provides a profit incentive to innovate, experiment with new products and services, and potentially improve the standard of living.

Moreover, because each customer has differing tastes and preferences, slight variations of each good or service are likely to capture the niche of the market that prefers them. An example is the market for candy, where one can find chocolate, licorice, mint, fruit, and many other flavors.

A further reason why monopolistic competition may be good is that people like variety. Traditional economic theories of international trade suggested that countries should buy products from other countries that they cannot produce domestically. Therefore, Norway should buy bananas from a tropical country and sell crude oil in exchange. But this is not the only kind of exchange that happens in reality: For example, Germany imports Honda, Subaru, and Toyota cars from Japan and sells Volkswagen, Porsche, Mercedes, and BMW cars to Japan. In theory, this should not occur because each of the countries produces good cars domestically and does not need to import them. The truth, however (see, for example, Krugman 1989), is that consumers in both countries enjoy variety. Some Japanese drivers prefer to be at the steering wheel of a BMW; others like Hondas, and the same happens in Germany. Variety and product differentiation, therefore, are not necessarily bad things.

LONG-RUN EQUILIBRIUM FOR MONOPOLISTICALLY COMPETITIVE FIRM

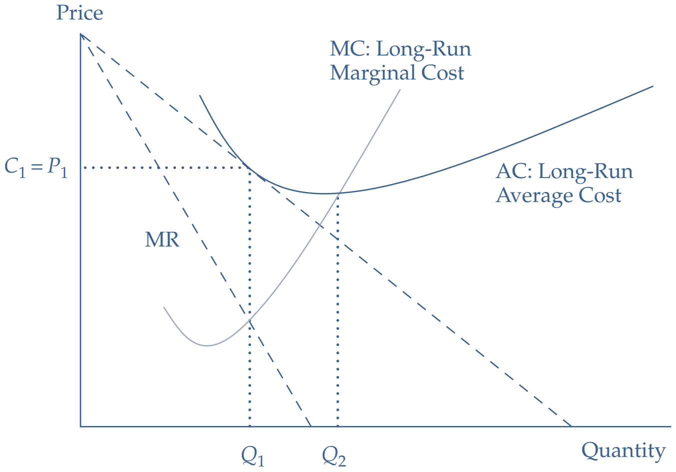
9

e explain factors affecting long-run equilibrium under each market structure;

f describe pricing strategy under each market structure;

Because TC includes all costs associated with production, including opportunity cost, economic profit is a signal to the market, and that signal will attract more competition. Just as with the perfectly competitive market structure, with relatively low entry costs, more firms will enter the market and lure some customers away from the firm making an economic profit. The loss of customers to new entrant firms will drive down the demand for all firms producing similar products. In the long run for the monopolistically competitive firm, economic profit will fall to zero. Exhibit 12 illustrates the condition of long-run equilibrium for monopolistic competition.

Exhibit 12 Long-Run Equilibrium in Monopolistic Competition



In long-run equilibrium, output is still optimal at the level where MR = MC, which is Q_1 in Exhibit 12. Again, the price consumers are willing to pay for any amount of the product is determined from the demand curve. That price is P_1 for the quantity Q_1 in Exhibit 12, and total revenue is the area of the rectangle $P_1 \times Q_1$. Notice that unlike long-run equilibrium in perfect competition, in the market of monopolistic competition, the equilibrium position is at a higher level of average cost than the level of output that minimizes average cost. Average cost does not reach its minimum until output level Q_2 is achieved. Total cost in this long-run equilibrium position is the area of the rectangle $C_1 \times Q_1$. Economic profit is total revenue minus total cost. In Exhibit 12, economic profit is zero because total revenue equals total cost: $P_1 \times Q_1 = C_1 \times Q_1$.

In the hybrid market of monopolistic competition, zero economic profit in long-run equilibrium resembles perfect competition. However, the long-run level of output, Q_1, is less than Q_2, which corresponds to the minimum average cost of production and would be the long run level of output in a perfectly competitive market. In addition, the economic cost in monopolistic competition includes some cost associated with product differentiation, such as advertising. In perfect competition, there are no costs associated with advertising or marketing because all products are homogeneous. Prices are lower, but consumers may have little variety.

10

OLIGOPOLY & DEMAND ANALYSIS AND PRICING STRATEGIES IN OLIGOPOLY MARKETS: PRICING INDEPENDENCE PORTION

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

f describe pricing strategy under each market structure;

An oligopoly market structure is characterized by only a few firms doing business in a relevant market. The products must all be similar and generally be substitutes for one another. In some oligopoly markets, the goods or services may be differentiated by marketing and strong brand recognition, as in the markets for breakfast cereals and for bottled or canned beverages. Other examples of oligopoly markets are made up of homogeneous products with little or no attempt at product differentiation, such as petroleum and cement. *The most distinctive characteristic of oligopoly markets is the small number of firms that dominate the market. There are so few firms in the relevant market that their pricing decisions are interdependent.* That is, each firm's pricing decision is based on the expected retaliation by the other firms. Recall from Exhibit 1 the characteristics of oligopoly markets:

1. There are a small number of potential sellers.
2. The products offered by each seller are close substitutes for the products offered by other firms and may be differentiated by brand or homogeneous and unbranded.
3. Entry into the market is difficult, with fairly high costs and significant barriers to competition.
4. Firms typically have substantial pricing power.
5. Products are often highly differentiated through marketing, features, and other non-price strategies.

Because there are so few firms, each firm can have some degree of pricing power, which can result in substantial profits. Another by-product of the oligopoly market structure is the attractiveness of price collusion. Even without price collusion, a dominant firm may easily become the price maker in the market. Oligopoly markets without collusion typically have the most sophisticated pricing strategies. Examples of non-colluding oligopolies include the US tobacco market and the Thai beer market. In 2004, four firms controlled 99 percent of the US tobacco industry.[10] Brands owned by Singha Co. and by ThaiBev controlled over 90 percent of the Thai beer market in 2009. (This situation is expected to change soon, as the Association of Southeast Asian Nations trade agreement will open the doors to competition from other ASEAN producers.) Perhaps the most well-known oligopoly market with collusion is the OPEC cartel, which seeks to control prices in the petroleum market by fostering agreements among oil-producing countries.

10.1 Demand Analysis and Pricing Strategies in Oligopoly Markets

Oligopoly markets' demand curves depend on the degree of pricing interdependence. In a market where collusion is present, the aggregate market demand curve is divided up by the individual production participants. Under non-colluding market conditions, each firm faces an individual demand curve. Furthermore, non-colluding oligopoly market demand characteristics depend on the pricing strategies adopted by the participating firms. There are three basic pricing strategies: pricing interdependence, the Cournot assumption, and the Nash equilibrium.

---

[10] These examples are based on "Industry Surveys," Net Advantage Database, Standard & Poor's; and Market Share Reports, Gale Research, annual issues, as noted in McGuigan, Moyer, and Harris (2016).

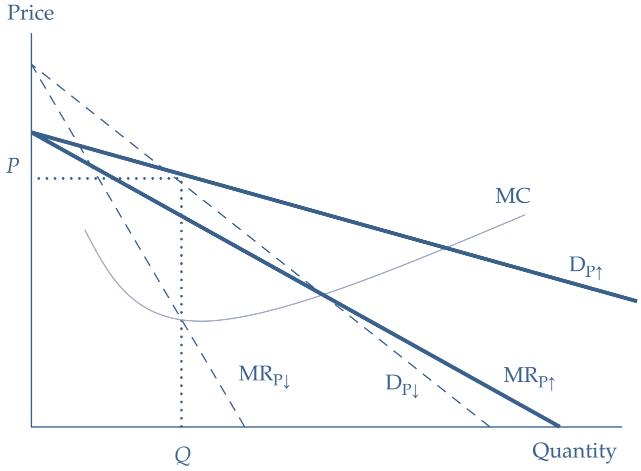
The first pricing strategy is to assume pricing interdependence among the firms in the oligopoly. A good example of this situation is any market where there are "price wars," such as the commercial airline industry. For example, flying out of their hubs in Atlanta, both Delta Air Lines and AirTran Airways jointly serve several cities. AirTran is a low-cost carrier and typically offers lower fares to destinations out of Atlanta. Delta tends to match the lower fares for those cities also served by AirTran when the departure and arrival times are similar to its own. However, when Delta offers service to the same cities at different time slots, Delta's ticket prices are higher.

The most common pricing strategy assumption in these price war markets is that competitors will match a price reduction and ignore a price increase. The logic is that by lowering its price to match a competitor's price reduction, the firm will not experience a reduction in customer demand. Conversely, by not matching the price increase, the firm stands to attract customers away from the firm that raised its prices. The oligopolist's demand relationship must represent the potential increase in market share when rivals' price increases are not matched and no significant change in market share when rivals' price decreases are matched.

Given a prevailing price, the price elasticity of demand will be much greater if the price is increased and less if the price is decreased. The firm's customers are more responsive to price increases because its rivals have lower prices. Alternatively, the firm's customers are less responsive to price decreases because its rivals will match its price change.

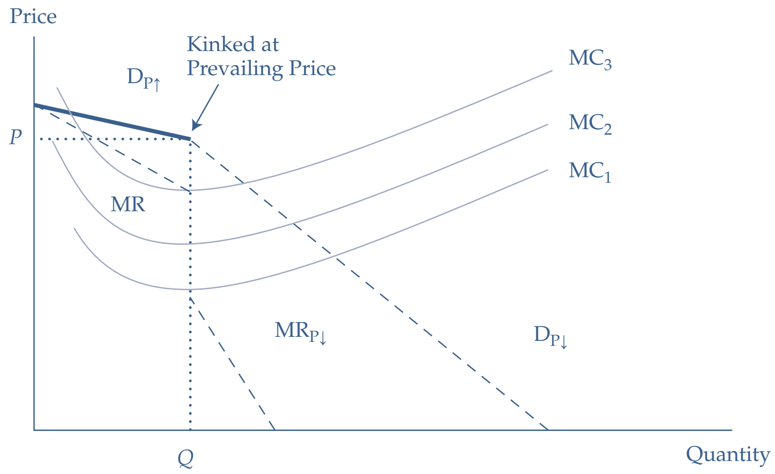
This implies that the oligopolistic firm faces two different demand structures, one associated with price increases and another relating to price reductions. Each demand function will have its own marginal revenue structure as well. Consider the demand and marginal revenue functions in Exhibit 13(A). The functions $D_{P\uparrow}$ and $MR_{P\uparrow}$ represent the demand and marginal revenue schedules associated with higher prices, while the functions $D_{P\downarrow}$ and $MR_{P\downarrow}$ represent the lower prices' demand and marginal revenue schedules. The two demand schedules intersect at the prevailing price (i.e., the price where price increase and price decrease are both equal to zero).

Exhibit 13(A) Kinked Demand Curve in Oligopoly Market



This oligopolistic pricing strategy results in a kinked demand curve, with the two segments representing the different competitor reactions to price changes. The kink in the demand curve also yields a discontinuous marginal revenue structure, with one part associated with the price increase segment of demand and the other relating to the price decrease segment. Therefore, the firm's overall demand equals the relevant portion of $D_{P\uparrow}$ and the relevant portion of $D_{P\downarrow}$. Exhibit 13(B) represents the firm's new demand and marginal revenue schedules. The firm's demand schedule in Exhibit 13(B) is segment $D_{P\uparrow}$ and $D_{P\downarrow}$, where overall demand $D = D_{P\uparrow} + D_{P\downarrow}$.

Exhibit 13(B) Kinked Demand Curve in Oligopoly Market



Notice in Exhibit 13(B) that a wide variety of cost structures are consistent with the prevailing price. If the firm has relatively low marginal costs, MC_1, the profit-maximizing pricing rule established earlier, MR = MC, still holds for the oligopoly firm. Marginal cost can rise to MC_2 and MC_3 before the firm's profitability is challenged. If the marginal cost curve MC_2 passes through the gap in marginal revenue, the most profitable price and output combination remains unchanged at the prevailing price and original level of output.

Criticism of the kinked demand curve analysis focuses on its inability to determine what the prevailing price is from the outset. The kinked demand curve analysis does help explain why stable prices have been observed in oligopoly markets and is therefore a useful tool for analyzing such markets. However, because it cannot determine the original prevailing price, it is considered an incomplete pricing analysis.

OLIGOPOLY & DEMAND ANALYSIS AND PRICING STRATEGIES IN OLIGOPOLY MARKETS: THE COURNOT ASSUMPTION PART

11

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

f describe pricing strategy under each market structure;

The second pricing strategy was first developed by French economist Augustin Cournot in 1838. In the **Cournot assumption**, each firm determines its profit-maximizing production level by assuming that the other firms' output will not change. This assumption simplifies pricing strategy because there is no need to guess what the other firm will do to retaliate. It also provides a useful approach to analyzing real-world behavior in oligopoly markets. Take the most basic oligopoly market situation, a two-firm duopoly market.[11] In equilibrium, neither firm has an incentive to change output, given the other firm's production level. Each firm attempts to maximize its own profits under the assumption that the other firm will continue producing the same level of output in the future. The Cournot strategy assumes that this pattern continues until each firm reaches its long-run equilibrium position. In long-run equilibrium, output and price are stable: There is no change in price or output that will increase profits for either firm.

Consider this example of a duopoly market. Assume that the aggregate market demand has been estimated to be

$$Q_D = 450 - P$$

The supply function is represented by constant marginal cost MC = 30.

The Cournot strategy's solution can be found by setting $Q_D = q_1 + q_2$, where q_1 and q_2 represent the output levels of the two firms. Each firm seeks to maximize profit, and each firm believes the other firm will not change output as it changes its own output (Cournot's assumption). The firm will maximize profit where MR = MC. Rearranging the aggregate demand function in terms of price, we get:

$$P = 450 - Q_D = 450 - q_1 - q_2, \text{ and MC} = 30$$

Total revenue for each of the two firms is found by multiplying price and quantity:

$$\text{TR}_1 = Pq_1 = \left(450 - q_1 - q_2\right)q_1 = 450q_1 - q_1^2 - q_1q_2, \text{ and}$$

$$\text{TR}_2 = Pq_2 = \left(450 - q_1 - q_2\right)q_2 = 450q_2 - q_2q_1 - q_2^2$$

Marginal revenue is defined as the change in total revenue, given a change in sales (q_1 or q_2).[12] For the profit-maximizing output, set MR = MC, or

$$450 - 2q_1 - q_2 = 30$$

and

$$450 - q_1 - 2q_2 = 30$$

Find the simultaneous equilibrium for the two firms by solving the two equations with two unknowns:

$$450 - 2q_1 - q_2 = 450 - q_1 - 2q_2$$

Because $q_2 = q_1$ under Cournot's assumption, insert this solution into the demand function and solve as

$$450 - 2q_1 - q_1 = 450 - 3q_1 = 30$$

Therefore, $q_1 = 140$, $q_2 = 140$, and $Q = 280$.
The price is $P = 450 - 280 = 170$.

In the Cournot strategic pricing solution, the market equilibrium price will be 170 and the aggregate output will be 280 units. This result, known as the Cournot equilibrium, differs from the perfectly competitive market equilibrium because the perfectly competitive price will be lower and the perfectly competitive output will be higher. In general, non-competitive markets have higher prices and lower levels of output in equilibrium when compared with perfect competition. In competition, the equilibrium is reached where price equals marginal cost.

$$P_C = \text{MR}_C = \text{MC, so } 450 - Q = 30$$

---

[11] The smallest possible oligopoly market is a duopoly, which is made up of only two sellers.

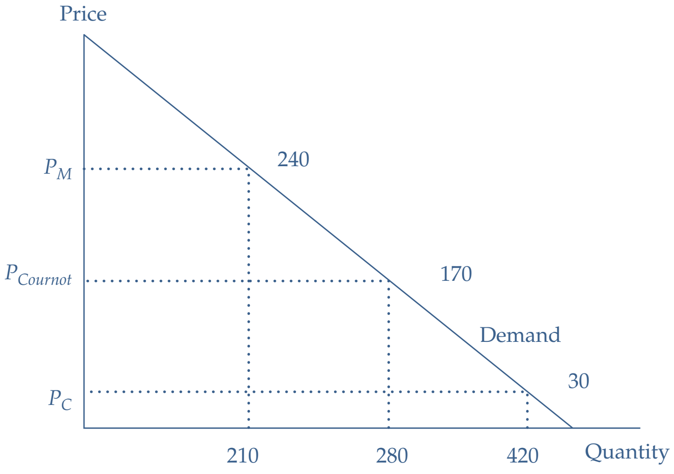
[12] The marginal revenue formulas can be obtained using the technique introduced in Section 2.1. For the market demand function, total revenue is $P \times Q = 450Q - Q^2$ and our technique yields $\text{MR} = \Delta\text{TR}/\Delta Q = 450 - 2Q$. For the individual firms in the Cournot duopoly, $\text{MR}_1 = \Delta\text{TR}_1/\Delta q_1 = 450 - 2q_1 - q_2$, and $\text{MR}_2 = \Delta\text{TR}_2/\Delta q_2 = 450 - q_1 - 2q_2$. Each of these marginal revenue formulas is, of course, the derivative of the relevant total revenue formula with respect to the relevant quantity.

where P_C is the competitive firm's equilibrium price.

$Q = 420$, and $P_C = 30$.

Exhibit 14 describes the oligopoly, competitive, and monopoly market equilibrium positions, where P_M is the monopoly optimum price, P_C is the competitive price, and $P_{Cournot}$ is the oligopoly price under the Cournot assumption.

Exhibit 14 Cournot Equilibrium in Duopoly Market



In the later discussion regarding monopoly market structure, equilibrium will be established where MR = MC. That solution is also shown in Exhibit 14. The monopoly firm's demand schedule is the aggregate market demand schedule. Therefore, the solution is

$$\text{MR} = \text{MC}$$

From Footnote 10, MR = $450 - 2Q$; therefore,

$$450 - 2Q = 30 \quad \text{and} \quad Q = 210$$

From the aggregate demand function, solve for price:

$$P_M = 450 - 210 = 240$$

Note that the Cournot solution falls between the competitive equilibrium and the monopoly solution.

It can be shown that as the number of firms increases from two to three, from three to four, and so on, the output and price equilibrium positions move toward the competitive equilibrium solution. This result has historically been the theoretical basis for the antitrust policies established in the United States.

OLIGOPOLY & DEMAND ANALYSIS AND PRICING STRATEGIES IN OLIGOPOLY MARKETS: THE NASH EQUILIBRIUM PART

12

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

f describe pricing strategy under each market structure;

The third pricing strategy is attributed to one of the 1994 Nobel Prize winners, John Nash, who first developed the general concepts. In the previous analysis, the concept of market equilibrium occurs when firms are achieving their optimum remuneration under the circumstances they face. In this optimum environment, the firm has no motive to change price or output level. Existing firms are earning a normal return (zero economic profit), leaving no motive for entry to or exit from the market. All firms in the market are producing at the output level where price equals the average cost of production.

In **game theory** (the set of tools that decision makers use to consider responses by rival decision makers), the **Nash equilibrium** is present when two or more participants in a non-cooperative game have no incentive to deviate from their respective equilibrium strategies after they have considered and anticipated their opponent's rational choices or strategies. In the context of oligopoly markets, the Nash equilibrium is an equilibrium defined by the characteristic that none of the oligopolists can increase its profits by unilaterally changing its pricing strategy. The assumption is made that each participating firm does the best it can, given the reactions of its rivals. Each firm anticipates that the other firms will react to any change made by competitors by doing the best they can under the altered circumstances. The firms in the oligopoly market have interdependent actions. The actions are non-cooperative, with each firm making decisions that maximize its own profits. The firms do not collude in an effort to maximize joint profits. The equilibrium is reached when all firms are doing the best they can, given the actions of their rivals.

Exhibit 15 illustrates the duopoly result from the Nash equilibrium. Assume there are two firms in the market, ArcCo and BatCo. ArcCo and BatCo can charge high prices or low prices for the product. The market outcomes are shown in Exhibit 15.

Exhibit 15 Nash Equilibrium in Duopoly Market

ArcCo – Low Price 50 70 BatCo – Low Price	ArcCo – Low Price 80 0 BatCo – High Price
ArcCo – High Price 300 350 BatCo – Low Price	ArcCo – High Price 500 300 BatCo – High Price

For example, the top left solution indicates that when both ArcCo and BatCo offer the product at low prices, ArcCo earns a profit of 50 and BatCo earns 70. The top right solution shows that if ArcCo offers the product at a low price, BatCo earns zero profits. The solution with the maximum joint profits is the lower right equilibrium, where both firms charge high prices for the product. Joint profits are 800 in this solution.

However, the Nash equilibrium requires that each firm behaves in its own best interest. BatCo can improve its position by offering the product at low prices when ArcCo is charging high prices. In the lower left solution, BatCo maximizes its profits at 350. While ArcCo can earn 500 in its best solution, it can do so only if BatCo also agrees to charge high prices. This option is clearly not in BatCo's best interest because it can increase its return from 300 to 350 by charging lower prices.

This scenario brings up the possibility of collusion. If ArcCo agrees to share at least 51 of its 500 when both companies are charging high prices, BatCo should also be willing to charge high prices. While, in general, such collusion is unlawful in most countries, it remains a tempting alternative. Clearly, conditions in oligopolistic industries encourage collusion, with a small number of competitors and interdependent pricing behavior. Collusion is motivated by several factors: increased profits, reduced cash flow uncertainty, and improved opportunities to construct barriers to entry.

When collusive agreements are made openly and formally, the firms involved are called a **cartel**. In some cases, collusion is successful; other times, the forces of competition overpower collusive behavior. There are six major factors that affect the chances of successful collusion.[13]

1. *The number and size distribution of sellers.* Successful collusion is more likely if the number of firms is small or if one firm is dominant. Collusion becomes more difficult as the number of firms increases or if the few firms have similar market shares. When the firms have similar market shares, the competitive forces tend to overshadow the benefits of collusion.
2. *The similarity of the products.* When the products are homogeneous, collusion is more successful. The more differentiated the products, the less likely it is that collusion will succeed.
3. *Cost structure.* The more similar the firms' cost structures, the more likely it is that collusion will succeed.
4. *Order size and frequency.* Successful collusion is more likely when orders are frequent, received on a regular basis, and relatively small. Frequent small orders, received regularly, diminish the opportunities and rewards for cheating on the collusive agreement.
5. *The strength and severity of retaliation.* Oligopolists will be less likely to break the collusive agreement if the threat of retaliation by the other firms in the market is severe.
6. *The degree of external competition.* The main reason to enter into the formal collusion is to increase overall profitability of the market, and rising profits attract competition. For example, in 2016 the average extraction cost of a barrel of crude oil from Saudi Arabia was approximately $9, while the average cost from United States shale oil fields was roughly $23.50. The cost of extracting oil from the Canadian tar sands in 2016 was roughly $27 per barrel. It is more likely that crude oil producers in the gulf countries will successfully collude because of the similarity in their cost structures (roughly $9–$10 per barrel). If OPEC had held crude oil prices down below $30 per barrel, there would have been a viable economic argument to develop US shale oil fields through fracking or expand extraction from Canada's tar sands. OPEC's successful cartel raised crude oil prices to the point where outside sources became economically possible and in doing so increased the competition the cartel faces.[14]

---

[13] McGuigan, Moyer, and Harris (2016).

[14] "Barrel Breakdown," *Wall Street Journal*, April 15, 2016.

There are other possible oligopoly strategies that are associated with decision making based on game theory. The Cournot equilibrium and the Nash equilibrium are examples of specific strategic games. A strategic game is any interdependent behavioral choice employed by individuals or groups that share a common goal (e.g., military units, sports teams, or business decision makers). Another prominent decision-making strategy in oligopolistic markets is the first-mover advantage in the **Stackelberg model**, named after the economist who first conceptualized the strategy.[15] The important difference between the Cournot model and the Stackelberg model is that Cournot assumes that in a duopoly market, decision making is simultaneous, while Stackelberg assumes that decisions are made sequentially. In the Stackelberg model, the leader firm chooses its output first and then the follower firm chooses after observing the leader's output. It can be shown that the leader firm has a distinct advantage, being a first mover.[16] In the Stackelberg game, the leader can aggressively overproduce to force the follower to scale back its production or even punish or eliminate the weaker opponent. This approach is sometimes referred to as a "top dog" strategy.[17] The leader earns more than in Cournot's simultaneous game, while the follower earns less. Many other strategic games are possible in oligopoly markets. The important conclusion is that the optimal strategy of the firm depends on what its adversary does. The price and marginal revenue the firm receives for its product depend on both its decisions and its adversary's decisions.

13 SUPPLY ANALYSIS & OPTIMAL PRICE AND OUTPUT & LONG-RUN EQUILIBRIUM IN OLIGOPOLY MARKETS

As in monopolistic competition, the oligopolist does not have a well-defined supply function. That is, there is no way to determine the oligopolist's optimal levels of output and price independent of demand conditions and competitor's strategies. However, the oligopolist still has a cost function that determines the optimal level of supply. Therefore, the profit-maximizing rule established earlier is still valid: The level of output that maximizes profit is where MR = MC. The price to charge is determined by what price consumers are willing to pay for that quantity of the product. Therefore, the equilibrium price comes from the demand curve, while the output level comes from the relationship between marginal revenue and marginal cost.

Consider an oligopoly market in which one of the firms is dominant and thus able to be the price leader. Dominant firms generally have 40 percent or greater market share. When one firm dominates an oligopoly market, it does so because it has greater capacity, has a lower cost structure, was first to market, or has greater customer loyalty than other firms in the market.

Assuming there is no collusion, the dominant firm becomes the price maker, and therefore its actions are similar to monopoly behavior in its segment of the market. The other firms in the market follow the pricing patterns of the dominant firm. Why wouldn't the price followers attempt to gain market share by undercutting the dominant firm's price? The most common explanation is that the dominant firm's supremacy often stems from a lower cost of production. Usually, the price followers would rather charge a price that is even higher than the dominant firm's price choice. If they attempt to undercut the dominant firm, the followers risk a price war with a lower-cost producer that can threaten their survival. Some believe that one explanation for the price leadership position of the dominant firm is simply convenience. Only one firm has to make the pricing decisions, and the others can simply follow its lead.

---

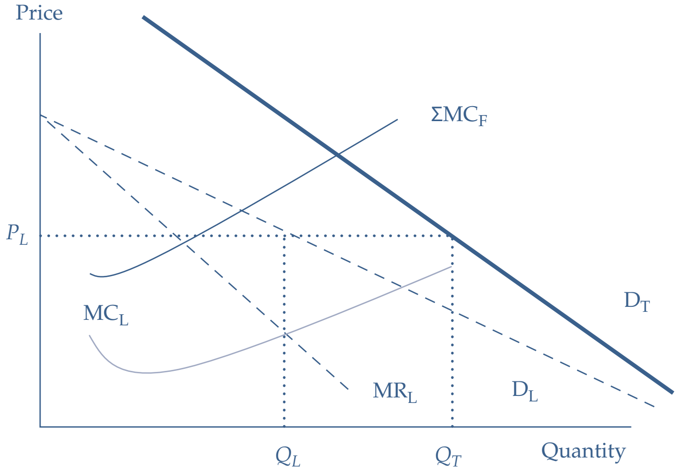
[15] Von Stackelberg (1952). See also Kelly (2011) for a comparison between the Cournot and Stackelberg equilibriums.

[16] Nicholson and Snyder (2016).

[17] Fudenberg and Tirole (1984).

Exhibit 16 establishes the dominant firm's pricing decision. The dominant firm's demand schedule, D_L, is a substantial share of the total market demand, D_T. The low-cost position of the dominant firm is represented by its marginal cost, MC_L. The sum of the marginal costs of the price followers is established as ΣMC_F and represents a higher cost of production than that of the price leader.

Exhibit 16 Dominant Oligopolist's Price Leadership



There is an important reason why the total demand curve and the leader demand curve are not parallel in Exhibit 16: Remember that the leader is the low-cost producer. Therefore, as price decreases, fewer of the smaller suppliers will be able to profitably remain in the market, and several will exit because they do not want to sell below cost. Therefore, the leader will have a larger market share as P decreases, which implies that Q_L increases at a low price, exactly as shown by a steeper D_T in the diagram.

The price leader identifies its profit-maximizing output where $MR_L = MC_L$, at output Q_L. This is the quantity it wants to supply; however, the price it will charge is determined by its segment of the total demand function, D_L. At price P_L, the dominant firm will supply quantity Q_L of total demand, D_T. The price followers will supply the difference to the market, $(Q_T - Q_L) = Q_F$. Therefore, neither the dominant firm nor the follower firms have a single functional relationship that determines the quantity supplied at various prices.

13.1 Optimal Price and Output in Oligopoly Markets

From the discussion above, clearly there is no single optimum price and output analysis that fits all oligopoly market situations. The interdependence among the few firms that make up the oligopoly market provides a complex set of pricing alternatives, depending on the circumstances in each market. In the case of the kinked demand curve, the optimum price is the prevailing price at the kink in the demand function. However, as previously noted, the kinked demand curve analysis does not provide insight into what established the prevailing price in the first place.

Perhaps the case of the dominant firm, with the other firms following the price leader, is the most obvious. In that case, the optimal price is determined at the output level where MR = MC. The profit-maximizing price is then determined by the output position of the segment of the demand function faced by the dominant firm. The price followers have little incentive to change the leader's price. In the case of the Cournot assumption, each firm assumes that the other firms will not alter their output following the dominant firm's selection of its price and output level.

Therefore, again, the optimum price is determined by the output level where MR = MC. In the case of the Nash equilibrium, each firm will react to the circumstances it faces, maximizing its own profit. These adjustments continue until there are stable prices and levels of output. Because of the interdependence, there is no certainty as to the individual firm's price and output level.

13.2 Factors Affecting Long-Run Equilibrium in Oligopoly Markets

Long-run economic profits are possible for firms operating in oligopoly markets. However, history has shown that, over time, the market share of the dominant firm declines. Profits attract entry by other firms into the oligopoly market. Over time, the marginal costs of the entrant firms decrease because they adopt more efficient production techniques, the dominant firm's demand and marginal revenue shrink, and the profitability of the dominant firm declines. In the early 1900s, J.P. Morgan, Elbert Gary, Andrew Carnegie, and Charles M. Schwab created the United States Steel Corporation (US Steel). When it was first formed in 1901, US Steel controlled 66 percent of the market. By 1920, US Steel's market share had declined to 46 percent, and by 1925 its market share was 42 percent.

In the long run, optimal pricing strategy must include the reactions of rival firms. History has proven that pricing wars should be avoided because any gains in market share are temporary. Decreasing prices to drive competitors away lowers total revenue to all participants in the oligopoly market. Innovation may be a way—though sometimes an uneconomical one—to maintain market leadership.

OLIGOPOLIES: APPEARANCE VERSUS BEHAVIOR

When is an oligopoly not an oligopoly? There are two extreme cases of this situation. A normal oligopoly has a few firms producing a differentiated good, and this differentiation gives them pricing power.

At one end of the spectrum, we have the oligopoly with a credible threat of entry. In practice, if the oligopolists are producing a good or service that can be easily replicated, has limited economies of scale, and is not protected by brand recognition or patents, they will not be able to charge high prices. The easier it is for a new supplier to enter the market, the lower the margins. In practice, this oligopoly will behave very much like a perfectly competitive market.

At the opposite end of the spectrum, we have the case of the cartel. Here, the oligopolists collude and act as if they were a single firm. In practice, a very effective cartel enacts a cooperative strategy. As shown in Sections 11 and 12, instead of going to a Nash equilibrium, the cartel participants go to the more lucrative (for them) cooperative equilibrium.

A cartel may be explicit (that is, based on a contract) or implicit (based on signals). An example of signals in a duopoly would be that one of the firms reduces its prices and the other does not. Because the firm not cutting prices refuses to start a price war, the firm that cut prices may interpret this signal as a "suggestion" to raise prices to a higher level than before, so that profits may increase for both.

MONOPOLY & DEMAND & SUPPLY & OPTIMAL PRICE AND OUTPUT IN MONOPOLY MARKETS

14

a describe characteristics of perfect competition, monopolistic competition, oligopoly, and pure monopoly;

b explain relationships between price, marginal revenue, marginal cost, economic profit, and the elasticity of demand under each market structure;

c describe a firm's supply function under each market structure;

d describe and determine the optimal price and output for firms under each market structure;

e explain factors affecting long-run equilibrium under each market structure;

f describe pricing strategy under each market structure;

Monopoly market structure is at the opposite end of the spectrum from perfect competition. For various reasons, there are significant barriers to entry such that a single firm produces a highly specialized product and faces no threat of competition. There are no good substitutes for the product in the relevant market, and the market demand function is the same as the individual firm's demand schedule. *The distinguishing characteristics of monopoly are that a single firm represents the market and significant barriers to entry exist.* Exhibit 1 identified the characteristics of monopoly markets:

1. There is a single seller of a highly differentiated product.
2. The product offered by the seller has no close substitute.
3. Entry into the market is very difficult, with high costs and significant barriers to competition.
4. The firm has considerable pricing power.
5. The product is differentiated through non-price strategies such as advertising.

Monopoly markets are unusual. With a single seller dominating the market, power over price decisions is significant. For a single seller to achieve this power, there must be factors that allow the monopoly to exist. One obvious source of monopoly power would be a patent or copyright that prevents other firms from entering the market. Patent and copyright laws exist to reward intellectual capital and investment in research and development. In so doing, they provide significant barriers to entry.

Another possible source of market power is control over critical resources used for production. One example is De Beers Consolidated Mines Limited. De Beers owned or controlled all diamond mining operations in South Africa and established pricing agreements with other important diamond producers. In doing so, De Beers was able to control the prices for cut diamonds for decades. Technically, De Beers was a near-monopoly dominant firm rather than a pure monopoly, although its pricing procedure for cut diamonds resembled monopoly behavior.

Perhaps the most common form of monopolistic market power occurs as the result of government-controlled authorization. In most urban areas, a single source of water and sewer services is offered. In some cases, these services are offered by a government-controlled entity. In other cases, private companies provide the services under government regulation. Such "natural" monopolies require a large initial investment that benefits from economies of scale; therefore, government may authorize a single seller to provide a certain service because having multiple sellers would be too costly. For example, electricity in most markets is provided by a single seller. Economies of scale result when significant capital investment benefits from declining long-run average costs. In the case of electricity, a large gas-fueled power plant producing electricity for a large area is substantially more efficient than having a small diesel generator for every building. That is, the average cost of generating and delivering a kilowatt of electricity will be substantially lower with the single power station, but the initial fixed cost of building the power station and the lines delivering electricity to each home, factory, and office will be very high.

In the case of natural monopolies, limiting the market to a single seller is considered beneficial to society. One water and sewer system is deemed better for the community than dozens of competitors because building multiple infrastructures for running water and sewer service would be particularly expensive and complicated. One electrical power grid supplying electricity for a community can make large capital investments in generating plants and lower the long-run average cost, while multiple power grids would lead to a potentially dangerous maze of wires. Clearly, not all monopolies are in a position to make significant economic profits. Regulators, such as public utility commissions in the United States, attempt to determine what a normal return for the monopoly owners' investment should be, and they set prices accordingly. Nevertheless, monopolists attempt to maximize profits.

Not all monopolies originate from "natural" barriers. For some monopolists, barriers to entry do not derive from increasing returns to scale. We mentioned that marketing and brand loyalty are sources of product differentiation in monopolistic competition. In some highly successful cases, strong brand loyalty can become a formidable barrier to entry. For example, if the Swiss watchmaker Rolex is unusually successful in establishing brand loyalty, so that its customers think there is no close substitute for its product, then the company will have monopoly-like pricing power over its market.

The final potential source of market power is the increasing returns associated with network effects. Network effects result from synergies related to increasing market penetration. By achieving a critical level of adoption, Microsoft was able to extend its market power through the network effect—for example, because most computer users know how to use Microsoft Word. Therefore, for firms, Word is cheaper to adopt than other programs because almost every new hire will be proficient in using the software and will need no further training. At some level of market share, a network-based product or service (think of Facebook or eBay) reaches a point where each additional share point increases the probability that another user will adopt.[18] These network effects increase the value to other potential adopters. In Microsoft's case, the network effects crowded out other potential competitors, including Netscape's internet browser, that might have led to applications bypassing Windows. Eventually, Microsoft's operating system's market share reached 92 percent of the global market. Similar situations occur in financial markets: If a publicly listed share or a derivative contract is more frequently traded on a certain exchange, market participants wishing to sell or buy the security will go to the more liquid exchange because they expect to find a better price and faster execution there.

14.1 Demand Analysis in Monopoly Markets

The monopolist's demand schedule is the aggregate demand for the product in the relevant market. Because of the income effect and the substitution effect, demand is negatively related to price, as usual. The slope of the demand curve is negative and therefore downward sloping. The general form of the demand relationship is

$$Q_D = a - bP \qquad \text{or, rewritten,} \qquad P = a/b - (1/b)Q_D$$

$$\text{Therefore, total revenue} = \text{TR} = P \times Q = (a/b)Q_D - (1/b)Q_D^2$$

---

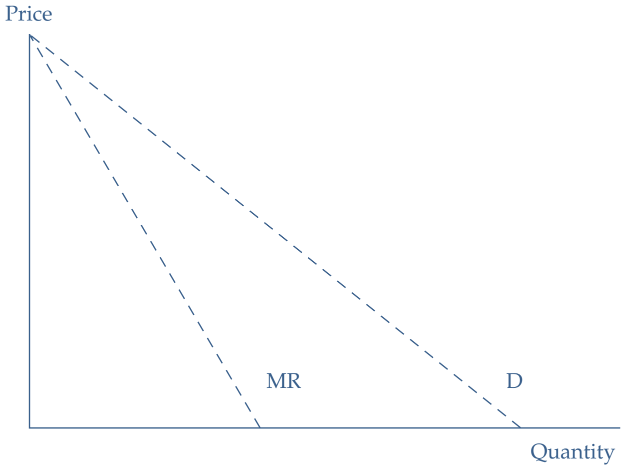
[18] When a network-based device reaches a 30 percent share, the next 50 percentage points are cheaper to promote, according to McGuigan, Moyer, and Harris (2016).

Marginal revenue is the change in revenue given a change in the quantity demanded. Because an increase in quantity requires a lower price, the marginal revenue schedule is steeper than the demand schedule. If the demand schedule is linear, then the marginal revenue curve is twice as steep as the demand schedule.[19]

$$MR = \Delta TR/\Delta Q = (a/b) - (2/b)Q_D$$

The demand and marginal revenue relationship is expressed in Exhibit 17.

Exhibit 17 Monopolist's Demand and Marginal Revenue



Suppose a company operating on a remote island is the single seller of natural gas. Demand for its product can be expressed as

$Q_D = 400 - 0.5P$, which can be rearranged as
$P = 800 - 2Q_D$

Total revenue is $P \times Q = TR = 800Q_D - 2Q_D^2$, and marginal revenue is $MR = 800 - 4Q_D$.[20]

In Exhibit 17, the demand curve's intercept is 800 and the slope is –2. The marginal revenue curve in Exhibit 17 has an intercept of 800 and a slope of –4.

Average revenue is TR/Q_D; therefore, $AR = 800 - 2Q_D$, which is the same as the demand function. In the monopoly market model, average revenue is the same as the market demand schedule.

14.2 Supply Analysis in Monopoly Markets

A monopolist's supply analysis is based on the firm's cost structure. As in the market structures of monopolistic competition and oligopoly, the monopolist does not have a well-defined supply function that determines the optimal output level and the price to charge. The optimal output is the profit-maximizing output level. The profit-maximizing level of output occurs where marginal revenue equals marginal cost, MR = MC.

---

[19] Marginal revenue can be found using the technique shown in Section 2.1 or, for readers who are familiar with calculus, by taking the derivative of the total revenue function: $MR = \Delta TR/\Delta Q = (a/b) - (2/b)Q_D$.

[20] $MR = \Delta TR/\Delta Q = 800 - 4Q$.

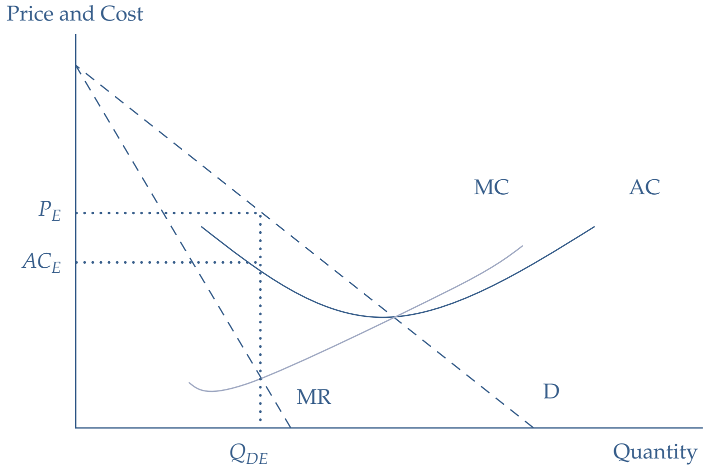
Assume the natural gas company has determined that its total cost can be expressed as

$$\text{TC} = 20{,}000 + 50Q + 3Q^2$$

Marginal cost is $\Delta\text{TC}/\Delta Q = \text{MC} = 50 + 6Q$.[21]

Supply and demand can be combined to determine the profit-maximizing level of output. Exhibit 18 combines the monopolist's demand and cost functions.

Exhibit 18 Monopolist's Demand, Marginal Revenue, and Cost Structures



In Exhibit 18, the demand and marginal revenue functions are clearly defined by the aggregate market. However, the monopolist does not have a supply curve. The quantity that maximizes profit is determined by the intersection of MC and MR, Q_{DE}.

The price consumers are willing to pay for this level of output is P_E, as determined by the demand curve, P_E.

The profit-maximizing level of output is MR = MC: $800 - 4Q_D = 50 + 6Q_D$; therefore, $Q_D = 75$ when profit is maximized.

Total profit equals total revenue minus total cost:

$$\pi = 800Q - 2Q_D^2 - \left(20{,}000 + 50Q_D + 3Q_D^2\right) = -20{,}000 + 750Q_D - 5Q_D^2$$

Profit is represented by the difference between the area of the rectangle $Q_{DE} \times P_E$, representing total revenue, and the area of the rectangle $Q_{DE} \times AC_E$, representing total cost.

MONOPOLISTS AND THEIR INCENTIVES

In theoretical models, which usually take product quality and technology as given, monopolists can choose to vary either price or quantity. In real life, they also can vary their product.

A monopolist can choose to limit quality if producing a higher-quality product is costly and higher quality does not increase profits accordingly. For example, the quality of domestically produced cars in most developed countries improved dramatically once imports became more available. Before the opening of borders to imports, the single incumbent that dominated the market (for example, Fiat in Italy) or the small group of incumbents acting as a collusive oligopoly (such as the Detroit “Big Three” in the United States) were the effective monopolists of their domestic automobile markets. Rust corrosion, limited reliability, and poor gas mileage were common.[22]

---

[21] The marginal cost equation can be found in this case by applying the technique used to find the marginal revenue equation in Section 2.1, or by taking the derivative of the total cost function.

Similarly, regulated utilities may have limited incentives to innovate. Several studies, including Gómez-Ibáñez (2003), have found that state-owned and other monopoly telephone utilities tended to provide very poor service before competition was introduced. Poor service may not be limited to poor connection quality but may also include extensive delays in adding new users and limited introduction of new services, such as caller ID or automatic answering services.

Intuitively, a monopolist will not spend resources on quality control, research and development, or even customer relations unless there is a threat of entry of a new competitor or unless there is a clear link between such expenses and a profit increase. In contrast, in competitive markets, including oligopoly, innovation and quality are often ways to differentiate the product and increase profits.

14.3 Optimal Price and Output in Monopoly Markets

Continuing the natural gas example from above, the total profit function can be solved using the quadratic formula.[23] Another method to solve the profit function is to evaluate $\Delta\pi/\Delta Q_D$ and set it equal to zero. This identifies the point at which profit is unaffected by changes in output.[24] Of course, this will give the same result as we found by equating marginal revenue with marginal cost. The monopoly will maximize profits when $Q^* = 75$ units of output and the price is set from the demand curve at 650.

$$P^* = 800 - 2(75) = 650 \text{ per unit}$$

To find total maximum profits, substitute these values into the profit function above:

$$\pi = -20{,}000 + 750Q_D - 5Q_D^2 = -20{,}000 + 750(75) - 5\left(75^2\right) = 8{,}125$$

Note that the price and output combination that maximizes profit occurs in the elastic portion of the demand curve in Exhibit 18. This must be so because marginal revenue and marginal cost will always intersect where marginal revenue is positive. This fact implies that quantity demanded responds more than proportionately to prices changes, i.e. demand is elastic, at the point at which MC = MR. As noted earlier, the relationship between marginal revenue and price elasticity, E_P, is:

$$\text{MR} = P[1 - 1/E_P]$$

In monopoly, MR = MC; therefore,

$$P[1 - 1/E_P] = \text{MC}$$

The firm can use this relationship to determine the profit-maximizing price if the firm knows its cost structure and the price elasticity of demand, E_P. For example, assume the firm knows that its marginal cost is constant at 75 and recent market analysis indicates that price elasticity is estimated to be 1.5. The optimal price is solved as

$$P[1 - 1/1.5] = 75 \text{ and}$$

$$P = 225$$

---

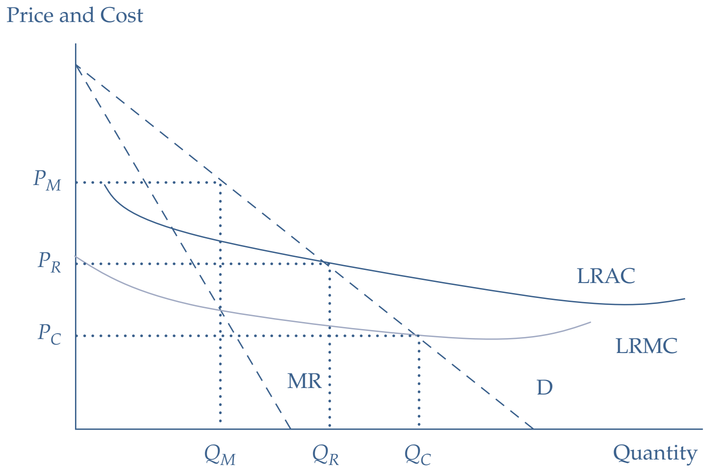
[22] For more on this topic, see Banker, Khosla, and Sinha (1998).

[23] The quadratic formula, where $aQ^2 + bQ + c = 0$, is $Q = \left\{-b \pm \sqrt{\left(b^2 - 4ac\right)}\right\} \Big/ 2a$.

[24] Maximum profit occurs where $\Delta\pi/\Delta Q_D = 0 = 750 - 10Q_D$. Therefore, profits are maximized at $Q_D = 75$.

Exhibit 18 indicated that the monopolist wants to produce at Q_E and charge the price of P_E. Suppose this is a natural monopoly that is operating as a government franchise under regulation. Natural monopolies are usually found where production is based on significant economies of scale and declining cost structure in the market. Examples include electric power generation, natural gas distribution, and the water and sewer industries. These are often called public utilities. Exhibit 19 illustrates such a market in long-run equilibrium.

Exhibit 19 Natural Monopoly in a Regulated Pricing Environment



In Exhibit 19, three possible pricing and output solutions are presented. The first is what the monopolist would do without regulation: The monopolist would seek to maximize profits by producing Q_M units of the product, where long-run marginal cost equals marginal revenue, LRMC = MR. To maximize profits, the monopolist would raise the price to the level the demand curve will accept, P_M.

In perfect competition, the price and output equilibrium occurs where price is equal to the marginal cost of producing the incremental unit of the product. In a competitive market, the quantity produced would be higher, Q_C, and the price lower, P_C. For this regulated monopoly, the competitive solution would be unfair because at output Q_C, the price P_C would not cover the average cost of production. One possibility is to subsidize the difference between the long-run average cost, LRAC, and the competitive price, P_C, for each unit sold.

Another solution is for the regulator to set the price at the point where long-run average cost equals average revenue. Recall that the demand curve represents the average revenue the firm receives at each output level. The government regulator will attempt to determine the monopolistic firm's long-run average cost and set the output and price so that the firm receives a fair return on the owners' invested capital. The regulatory solution is output level Q_R, with the price set at P_R. Therefore, the regulatory solution is found between the unregulated monopoly equilibrium and the competitive equilibrium.

15 PRICE DISCRIMINATION AND CONSUMER SURPLUS

- f describe pricing strategy under each market structure;

Monopolists can either be more or less effective in taking advantage of their market structure. At one extreme, we have a monopolist that charges prices and supplies quantities that are the same as they would be in perfect competition; this scenario may be a result of regulation or threat of entry (if the monopolist charged more, another company could come in and price the former monopolist out of the market). At the opposite extreme, hated by all consumers and economists, is the monopolist that extracts the entire consumer surplus. This scenario is called **first-degree price discrimination**, where a monopolist can charge each customer the highest price the customer is willing to pay. This is called price discrimination because the monopolist charges a different price to each client. How can this be? For example, if the monopolist knows the exact demand schedule of the customer, then the monopolist can capture the entire consumer surplus. In practice, the monopolist can measure how often the product is used and charges the customer the highest price the consumer is willing to pay for that unit of good. Another possibility is that public price disclosure is non-existent, so that no customer knows what the other customers are paying. Interestingly, not every consumer is worse off in this case, because some consumers may be charged a price that is below that of perfect competition, as long as the marginal revenue exceeds the marginal cost.

In **second-degree price discrimination** the monopolist offers a menu of quantity-based pricing options designed to induce customers to self-select based on how highly they value the product. Such mechanisms include volume discounts, volume surcharges, coupons, product bundling, and restrictions on use. In practice, producers can use not just the quantity but also the quality (e.g., “professional grade”) to charge more to customers that value the product highly.

Third-degree price discrimination happens when customers are segregated by demographic or other traits. For example, some econometric software is licensed this way: A student version can handle only small datasets and is sold for a low price; a professional version can handle very large datasets and is sold at a much higher price because corporations need to compute the estimates for their business and are therefore willing to pay more for a license. Another example is that airlines know that passengers who want to fly somewhere and come back the same day are most likely business people; therefore, one-day roundtrip tickets are generally more expensive than tickets with a return flight at a later date or over a weekend.

Price discrimination has many practical applications when the seller has pricing power. The best way to understand how this concept works is to think of consumer surplus: As seen in this reading, a consumer may be willing to pay more for the first unit of a good, but to buy a second unit she will want to pay a lower price, thus getting a better deal on the first unit. In practice, sellers can sometimes use income and substitution effects to their advantage. Think of something you often buy, perhaps lunch at your favorite café. How much would you be willing to pay for a “lunch club membership card” that would allow you to purchase lunches at, say, half price? If the café could extract from you the maximum amount each month that you would be willing to pay for the half-price option, then it would successfully have removed the income effect from you in the form of a monthly fixed fee. Notice that a downward-sloping demand curve implies that you would end up buying more lunches each month than before you purchased the discount card, even though you would be no better or worse off than before. This is a way that sellers are sometimes able to extract consumer surplus by means of creative pricing schemes. It’s a common practice among big-box retailers, sports clubs, and other users of what is called “two-part tariff pricing,” as in the example below.

EXAMPLE 3

Price Discrimination

Nicole's monthly demand for visits to her health club is given by the following equation: $Q_D = 20 - 4P$, where Q_D is visits per month and P is euros per visit. The health club's marginal cost is fixed at €2 per visit.

1. Draw Nicole's demand curve for health club visits per month.
2. If the club charged a price per visit equal to its marginal cost, how many visits would Nicole make per month?
3. How much consumer surplus would Nicole enjoy at that price?
4. How much could the club charge Nicole each month for a membership fee?

Solution to 1:

$Q_D = 20 - 4P$, so when $P = 0$, $Q_D = 20$. Inverting, $P = 5 - 0.25Q_D$, so when $Q = 0$, $P = 5$.

Solution to 2:

$Q_D = 20 - 4(2) = 12$. Nicole would make 12 visits per month at a price of €2 per visit.

Solution to 3:

Nicole's consumer surplus can be measured as the area under her demand curve and above the price she pays for a total of 12 visits, or $(0.5)(12)(3) = 18$. Nicole would enjoy a consumer surplus of €18 per month.

Solution to 4:

The club could extract all of Nicole's consumer surplus by charging her a monthly membership fee of €18 plus a per-visit price of €2. This pricing method is called a two-part tariff because it assesses one price per unit of the item purchased plus a per-month fee (sometimes called an "entry fee") equal to the buyer's consumer surplus evaluated at the per-unit price.

16 FACTORS AFFECTING LONG-RUN EQUILIBRIUM IN MONOPOLY MARKETS

e explain factors affecting long-run equilibrium under each market structure;

The unregulated monopoly market structure can produce economic profits in the long run. In the long run, all factors of production are variable, while in the short run, some factors of production are fixed. Generally, the short-run factor that is fixed is the capital investment, such as the factory, machinery, production technology, available arable land, and so forth. The long-run solution allows for all inputs, including technology, to change. In order to maintain a monopoly market position in the long run, the firm must be protected by substantial and ongoing barriers to entry. If the monopoly position is the result of a patent, then new patents must be continuously added to prevent the entry of other firms into the market.

For regulated monopolies, such as natural monopolies, there are a variety of long-run solutions. One solution is to set the price equal to marginal cost, $P = \text{MC}$. However, that price will not likely be high enough to cover the average cost of production, as Exhibit 19 illustrated. The answer is to provide a subsidy sufficient to compensate the firm. The national rail system in the United States, Amtrak, is an example of a regulated monopoly operating with a government subsidy.

National ownership of the monopoly is another solution. Nationalization of the natural monopoly has been a popular solution in Europe and other parts of the world. The United States has generally avoided this potential solution. One problem with this arrangement is that once a price is established, consumers are unwilling to accept price increases, even as factor costs increase. Politically, raising prices on products from government-owned enterprises is highly unpopular.

Establishing a governmental entity that regulates an authorized monopoly is another popular solution. Exhibit 19 illustrated the appropriate decision rule. The regulator sets price equal to long-run average cost, $P_R = \text{LRAC}$. This solution assures that investors will receive a normal return for the risk they are taking in the market. Given that no other competitors are allowed, the risk is lower than in a highly competitive market environment. The challenge facing the regulator is determining the authentic risk-related return and the monopolist's realistic long-run average cost.

The final solution is to franchise the monopolistic firm through a bidding war. Again, the public goal is to select the winning firm based on price equaling long-run average cost. Retail outlets at rail stations and airports and concession outlets at stadiums are examples of government franchises. The long-run success of the monopoly franchise depends on its ability to meet the goal of pricing its products at the level of its long-run average cost.

EXAMPLE 4

Monopolies and Efficiency

Are monopolies *always* inefficient?

A No, because if they charge more than average cost they are nationalized.

B Yes, because they charge all consumers more than perfectly competitive markets would.

C No, because economies of scale and regulation (or threat of entry) may give a better outcome for buyers than perfect competition.

Solution:

C is correct. Economies of scale and regulation may make monopolies more efficient than perfect competition.

IDENTIFICATION OF MARKET STRUCTURE

17

g describe the use and limitations of concentration measures in identifying market structure;

h identify the type of market structure within which a firm operates.

Monopoly markets and other situations where companies have pricing power can be inefficient because producers constrain output to cause an increase in prices. Therefore, there will be less of the good being consumed and it will be sold at a higher price, which is generally inefficient for the overall market. As a result, many countries have introduced competition law to regulate the degree of competition in many industries.

Market power in the real world is not always as clear as it is in textbook examples. Governments and regulators often have the difficult task of measuring market power and establishing whether a firm has a dominant position that may resemble a monopoly. A few historical examples of this are as follows:

1. In the 1990s, US regulators prosecuted agricultural corporation Archer Daniels Midland for conspiring with Japanese competitors to fix the price of lysine, an amino acid used as an animal feed additive. The antitrust action resulted in a settlement that involved over US$100 million in fines paid by the cartel members.
2. In the 1970s, US antitrust authorities broke up the local telephone monopoly, leaving AT&T the long-distance business (and opening that business to competitors), and required AT&T to divest itself of the local telephone companies it owned. This antitrust decision brought competition, innovation, and lower prices to the US telephone market.
3. European regulators (specifically, the European Commission) have affected the mergers and monopoly positions of European corporations (as in the case of the companies Roche, Rhone-Poulenc, and BASF, which were at the center of a vitamin price-fixing case) as well as non-European companies (such as Intel) that do business in Europe. Moreover, the merger between the US company General Electric and the European company Honeywell was denied by the European Commission on grounds of excessive market concentration.

Quantifying excessive market concentration is difficult. Sometimes, regulators need to measure whether something that has not yet occurred might generate excessive market power. For example, a merger between two companies might allow the combined company to be a monopolist or quasi monopolist in a certain market.

A financial analyst hearing news about a possible merger should always consider the impact of competition law (sometimes called antitrust law)—that is, whether a proposed merger may be blocked by regulators in the interest of preserving a competitive market.

17.1 Econometric Approaches

How should one measure market power? The theoretical answer is to estimate the elasticity of demand and supply in a market. If demand is very elastic, the market must be very close to perfect competition. If demand is rigid (inelastic), companies *may* have market power. This is the approach taken in the cellophane case mentioned in Section 4.

From the econometric point of view, this estimation requires some attention. The problem is that observed price and quantity are the equilibrium values of price and quantity and do not represent the value of either supply or demand. Technically, this is called the problem of endogeneity, in the sense that the equilibrium price and quantity are jointly determined by the interaction of demand and supply. Therefore, to have an appropriate estimation of demand and supply, we will need to use a model with two equations, namely, an equation of demanded quantity (as a function of price, income of the buyers, and other variables) and an equation of supplied quantity (as a function of price, production costs, and other variables). The estimated parameters will then allow us to compute elasticity.

Regression analysis is useful in computing elasticity but requires a large number of observations. Therefore, one may use a time-series approach and, for example, look at 20 years of quarterly sales data for a market. However, the market structure may have changed radically over those 20 years, and the estimated elasticity may not apply to the current situation. Moreover, the supply curve may change due to a merger among large competitors, and the estimation based on past data may not be informative regarding the future state of the market post merger.

An alternative approach is a cross-sectional regression analysis. Instead of looking at total sales and average prices in a market over time (the time-series approach mentioned above), we can look at sales from different companies in the market during the same year, or even at single transactions from many buyers and companies. Clearly, this approach requires substantial data-gathering effort, and therefore, this estimation method can be complicated. Moreover, different specifications of the explanatory variables (for example, using total GDP rather than median household income or per-capita GDP to represent income) may sometimes lead to dramatically different estimates.

17.2 Simpler Measures

Trying to avoid the above drawbacks, analysts often use simpler measures to estimate elasticity. The simplest measure is the concentration ratio, which is the sum of the market shares of the largest *N* firms. To compute this ratio, one would, for example, add the sales values of the largest 10 firms and divide this figure by total market sales. This number is always between zero (perfect competition) and 100 percent (monopoly).

The main advantage of the concentration ratio is that it is simple to compute, as shown above. The disadvantage is that it does not directly quantify market power. In other words, is a high concentration ratio a clear signal of monopoly power? The analysis of entry in Section 1 explains clearly that this is not the case: A company may be the only incumbent in a market, but if the barriers to entry are low, the simple presence of a *potential* entrant may be sufficient to convince the incumbent to behave like a firm in perfect competition. For example, a sugar wholesaler may be the only one in a country, but the knowledge that other large wholesalers in the food industry might easily add imported sugar to their range of products should convince the sugar wholesaler to price its product as if it were in perfect competition.

Another disadvantage of the concentration ratio is that it tends to be unaffected by mergers among the top market incumbents. For example, if the largest and second-largest incumbents merge, the pricing power of the combined entity is likely to be larger than that of the two pre-existing companies. But the concentration ratio may not change much.

CALCULATING THE CONCENTRATION RATIO

Suppose there are eight producers of a certain good in a market. The largest producer has 35 percent of the market, the second largest has 25 percent, the third has 20 percent, the fourth has 10 percent, and the remaining four have 2.5 percent each. If we computed the concentration ratio of the top three producers, it would be 35 + 25 + 20 = 80 percent, while the concentration ratio of the top four producers would be 35 + 25 + 20 + 10 = 90 percent.

If the two largest companies merged, the new concentration ratio for the top three producers would be 60 (the sum of the market shares of the merged companies) + 20 + 10 = 90 percent, and the concentration ratio for the four top producers would be 92.5 percent. Therefore, this merger affects the concentration ratio very mildly, even though it creates a substantial entity that controls 60 percent of the market.

For example, the effect of consolidation in the US retail gasoline market has resulted in increasing degrees of concentration. In 1992, the top four companies in the US retail gasoline market shared 33 percent of the market. By 2001, the top four companies controlled 78 percent of the market (Exxon Mobil 24 percent, Shell 20 percent, BP/Amoco/Arco 18 percent, and Chevron/Texaco 16 percent).

To avoid the known issues with concentration ratios, economists O.C. Herfindahl and A.O. Hirschman suggested an index where the market shares of the top N companies are first squared and then added. If one firm controls the whole market (a monopoly), the Herfindahl–Hirschman index (HHI) equals 1. If there are M firms in the industry with equal market shares, then the HHI equals $(1/M)$. This provides a useful gauge for interpreting an HHI. For example, an HHI of 0.20 would be analogous to having the market shared equally by 5 firms.

The HHI for the top three companies in the example in the box above would be $0.35^2 + 0.25^2 + 0.20^2 = 0.225$ before the merger, while after the merger, it would be $0.60^2 + 0.20^2 + 0.10^2 = 0.410$, which is substantially higher than the initial 0.225. The HHI is widely used by competition regulators; however, just like the concentration ratio, the HHI does not take the possibility of entry into account, nor does it consider the elasticity of demand. Therefore, the HHI has limited use for a financial analyst trying to estimate the potential profitability of a company or group of companies.

EXAMPLE 5

The Herfindahl–Hirschman Index

Suppose a market has 10 suppliers, each of them with 10 percent of the market. What are the concentration ratio and the HHI of the top four firms?

A Concentration ratio 4 percent and HHI 40

B Concentration ratio 40 percent and HHI 0.4

C Concentration ratio 40 percent and HHI 0.04

Solution:

C is correct. The concentration ratio for the top four firms is 10 + 10 + 10 + 10 = 40 percent, and the HHI is $0.10^2 \times 4 = 0.01 \times 4 = 0.04$.

SUMMARY

In this reading, we have surveyed how economists classify market structures. We have analyzed the distinctions between the different structures that are important for understanding demand and supply relations, optimal price and output, and the factors affecting long-run profitability. We also provided guidelines for identifying market structure in practice. Among our conclusions are the following:

- Economic market structures can be grouped into four categories: perfect competition, monopolistic competition, oligopoly, and monopoly.
- The categories differ because of the following characteristics: The number of producers is many in perfect and monopolistic competition, few in oligopoly, and one in monopoly. The degree of product differentiation, the pricing power

of the producer, the barriers to entry of new producers, and the level of non-price competition (e.g., advertising) are all low in perfect competition, moderate in monopolistic competition, high in oligopoly, and generally highest in monopoly.

- A financial analyst must understand the characteristics of market structures in order to better forecast a firm's future profit stream.
- The optimal marginal revenue equals marginal cost. However, only in perfect competition does the marginal revenue equal price. In the remaining structures, price generally exceeds marginal revenue because a firm can sell more units only by reducing the per unit price.
- The quantity sold is highest in perfect competition. The price in perfect competition is usually lowest, but this depends on factors such as demand elasticity and increasing returns to scale (which may reduce the producer's marginal cost). Monopolists, oligopolists, and producers in monopolistic competition attempt to differentiate their products so that they can charge higher prices.
- Typically, monopolists sell a smaller quantity at a higher price. Investors may benefit from being shareholders of monopolistic firms that have large margins and substantial positive cash flows.
- Competitive firms do not earn economic profit. There will be a market compensation for the rental of capital and of management services, but the lack of pricing power implies that there will be no extra margins.
- While in the short run firms in any market structure can have economic profits, the more competitive a market is and the lower the barriers to entry, the faster the extra profits will fade. In the long run, new entrants shrink margins and push the least efficient firms out of the market.
- Oligopoly is characterized by the importance of strategic behavior. Firms can change the price, quantity, quality, and advertisement of the product to gain an advantage over their competitors. Several types of equilibrium (e.g., Nash, Cournot, kinked demand curve) may occur that affect the likelihood of each of the incumbents (and potential entrants in the long run) having economic profits. Price wars may be started to force weaker competitors to abandon the market.
- Measuring market power is complicated. Ideally, econometric estimates of the elasticity of demand and supply should be computed. However, because of the lack of reliable data and the fact that elasticity changes over time (so that past data may not apply to the current situation), regulators and economists often use simpler measures. The concentration ratio is simple, but the HHI, with little more computation required, often produces a better figure for decision making.

REFERENCES

Banker, R.D., I. Khosla, and K.K. Sinha. 1998. "Quality and Competition." *Management Science*, vol. 44, no. 9:1179–1192.

Chamberlin, Edward H. 1933. *The Theory of Monopolistic Competition*. Cambridge, MA: Harvard University Press.

Dorsey, Pat. 2004. *The Five Rules for Successful Stock Investing: Morningstar's Guide to Building Wealth and Winning in the Market*. Hoboken, NJ: John Wiley & Sons.

Friedman, Thomas L. 2006. *The World Is Flat: A Brief History of the Twenty-first Century*. New York: Farrar, Straus and Giroux.

Fudenberg, Drew, and Jean Tirole. 1984. "The Fat Cat Effect, the Puppy Dog Ploy and the Lean and Hungry Look." *American Economic Review*, vol. 74, no. 2:361–366.

Gómez-Ibáñez, José A. 2006. *Regulating Infrastructure: Monopoly, Contracts, and Discretion*. Cambridge, MA: Harvard University Press.

Kelly, Anthony. 2011. *Decision Making Using Game Theory: An Introduction for Managers*. Cambridge, UK: Cambridge University Press.

Krugman, Paul R. 1989. "Industrial Organization and International Trade." In *Handbook of Industrial Organization*, vol. 2. Edited by Richard Schmalensee and Robert Willig. Amsterdam: Elsevier B.V.

McCloskey, Donald. 1985. *The Applied Theory of Price*. 2nd ed. New York: Macmillan.

McGuigan, James R., R. Charles Moyer, and Frederick H. Harris. 2016. *Managerial Economics: Applications, Strategy and Tactics*. 14th ed. Mason, OH: Thomson South-Western.

Porter, Michael E. 2008. "The Five Competitive Forces that Shape Strategy." *Harvard Business Review*, vol. 86, no. 1:78–93.

Nicholson, Walter, and Christopher M. Snyder. 2016. *Microeconomic Theory: Basic Principles and Extensions*. 12th ed. Mason, OH: Thomson South-Western.

Schumpeter, Joseph A. 1942. *Capitalism, Socialism and Democracy*. New York: HarperCollins.

von Stackelberg, Heinrich. 1952. *The Theory of the Market Economy*. New York: Oxford University Press.

PRACTICE PROBLEMS

1 A market structure characterized by many sellers with each having some pricing power and product differentiation is *best* described as:

A oligopoly.

B perfect competition.

C monopolistic competition.

2 A market structure with relatively few sellers of a homogeneous or standardized product is *best* described as:

A oligopoly.

B monopoly.

C perfect competition.

3 Market competitors are *least likely* to use advertising as a tool of differentiation in an industry structure identified as:

A monopoly.

B perfect competition.

C monopolistic competition.

4 Upsilon Natural Gas, Inc. is a monopoly enjoying very high barriers to entry. Its marginal cost is \$40 and its average cost is \$70. A recent market study has determined the price elasticity of demand is 1.5. The company will *most likely* set its price at:

A \$40.

B \$70.

C \$120.

5 The demand schedule in a perfectly competitive market is given by P = 93 – 1.5Q (for Q ≤ 62) and the long-run cost structure of each company is:

Total cost: $256 + 2Q + 4Q^2$

Average cost: $256/Q + 2 + 4Q$

Marginal cost: $2 + 8Q$

New companies will enter the market at any price greater than:

A 8.

B 66.

C 81.

6 Companies *most likely* have a well-defined supply function when the market structure is:

A oligopoly.

B perfect competition.

C monopolistic competition.

7 Aquarius, Inc. is the dominant company and the price leader in its market. One of the other companies in the market attempts to gain market share by undercutting the price set by Aquarius. The market share of Aquarius will *most likely*:

A increase.

© 2011 CFA Institute. All rights reserved.

B decrease.

C stay the same.

8 SigmaSoft and ThetaTech are the dominant makers of computer system software. The market has two components: a large mass-market component in which demand is price sensitive, and a smaller performance-oriented component in which demand is much less price sensitive. SigmaSoft's product is considered to be technically superior. Each company can choose one of two strategies:

- *Open architecture (Open):* Mass market focus allowing other software venders to develop products for its platform.
- *Proprietary (Prop):* Allow only its own software applications to run on its platform.

Depending upon the strategy each company selects, their profits would be:

SigmaSoft – Open **400** **600** **ThetaTech – Open**	**SigmaSoft – Prop** **650** **700** **ThetaTech – Open**
SigmaSoft – Open **800** **300** **ThetaTech – Prop**	**SigmaSoft – Prop** **600** **400** **ThetaTech – Prop**

The Nash equilibrium for these companies is:

A proprietary for SigmaSoft and proprietary for ThetaTech.

B open architecture for SigmaSoft and proprietary for ThetaTech.

C proprietary for SigmaSoft and open architecture for ThetaTech.

9 A company doing business in a monopolistically competitive market will *most likely* maximize profits when its output quantity is set such that:

A average cost is minimized.

B marginal revenue equals average cost.

C marginal revenue equals marginal cost.

10 Oligopolistic pricing strategy *most likely* results in a demand curve that is:

A kinked.

B vertical.

C horizontal.

11 Collusion is *less likely* in a market when:

A the product is homogeneous.

B companies have similar market shares.

C the cost structures of companies are similar.

12 If companies earn economic profits in a perfectly competitive market, over the long run the supply curve will *most likely*:

A shift to the left.

B shift to the right.

C remain unchanged.

13 Over time, the market share of the dominant company in an oligopolistic market will *most likely*:

A increase.

B decrease.

C remain the same.

14 A government entity that regulates an authorized monopoly will *most likely* base regulated prices on:

A marginal cost.

B long run average cost.

C first degree price discrimination.

15 An analyst gathers the following market share data for an industry:

Company	Sales (in millions of €)
ABC	300
Brown	250
Coral	200
Delta	150
Erie	100
All others	50

The industry's four-company concentration ratio is *closest* to:

A 71%.

B 86%.

C 95%.

16 An analyst gathered the following market share data for an industry comprised of five companies:

Company	Market Share (%)
Zeta	35
Yusef	25
Xenon	20
Waters	10
Vlastos	10

The industry's three-firm Herfindahl–Hirschmann Index is *closest* to:

A 0.185.

B 0.225.

C 0.235.

17 One disadvantage of the Herfindahl–Hirschmann Index is that the index:

A is difficult to compute.

B fails to reflect low barriers to entry.

C fails to reflect the effect of mergers in the industry.

18 In an industry comprised of three companies, which are small-scale manufacturers of an easily replicable product unprotected by brand recognition or patents, the *most* representative model of company behavior is:

A oligopoly.

B perfect competition.

C monopolistic competition.

19 Deep River Manufacturing is one of many companies in an industry that make a food product. Deep River units are identical up to the point they are labeled. Deep River produces its labeled brand, which sells for \$2.20 per unit, and "house brands" for seven different grocery chains which sell for \$2.00 per unit. Each grocery chain sells both the Deep River brand and its house brand. The *best* characterization of Deep River's market is:

A oligopoly.

B perfect competition.

C monopolistic competition.

SOLUTIONS

1 C is correct. Monopolistic competition is characterized by many sellers, differentiated products, and some pricing power.

2 A is correct. Few sellers of a homogeneous or standardized product characterizes an oligopoly.

3 B is correct. The product produced in a perfectly competitive market cannot be differentiated by advertising or any other means.

4 C is correct. Profits are maximized when MR = MC. For a monopoly, MR = $P[1 - 1/E_p]$. Setting this equal to MC and solving for P:

$$\$40 = P[1 - (1/1.5)] = P \times 0.333$$
$$P = \$120$$

5 B is correct. The long-run competitive equilibrium occurs where MC = AC = P for each company. Equating MC and AC implies $2 + 8Q = 256/Q + 2 + 4Q$.

Solving for Q gives $Q = 8$. Equating MC with price gives $P = 2 + 8Q = 66$. Any price above 66 yields an economic profit because P = MC > AC, so new companies will enter the market.

6 B is correct. A company in a perfectly competitive market must accept whatever price the market dictates. The marginal cost schedule of a company in a perfectly competitive market determines its supply function.

7 A is correct. As prices decrease, smaller companies will leave the market rather than sell below cost. The market share of Aquarius, the price leader, will increase.

8 C is correct. In the Nash model, each company considers the other's reaction in selecting its strategy. In equilibrium, neither company has an incentive to change its strategy. ThetaTech is better off with open architecture regardless of what SigmaSoft decides. Given this choice, SigmaSoft is better off with a proprietary platform. Neither company will change its decision unilaterally.

9 C is correct. The profit maximizing choice is the level of output where marginal revenue equals marginal cost.

10 A is correct. The oligopolist faces two different demand structures, one for price increases and another for price decreases. Competitors will lower prices to match a price reduction, but will not match a price increase. The result is a kinked demand curve.

11 B is correct. When companies have similar market shares, competitive forces tend to outweigh the benefits of collusion.

12 B is correct. The economic profit will attract new entrants to the market and encourage existing companies to expand capacity.

13 B is correct. The dominant company's market share tends to decrease as profits attract entry by other companies.

14 B is correct. This allows the investors to receive a normal return for the risk they are taking in the market.

15 B is correct. The top four companies in the industry comprise 86 percent of industry sales: (300 + 250 + 200 + 150)/(300 + 250 + 200 + 150 + 100 + 50) = 900/1050 = 86%.

16 B is correct. The three-firm Herfindahl–Hirschmann Index is $0.35^2 + 0.25^2 + 0.20^2 = 0.225$.

17 B is correct. The Herfindahl–Hirschmann Index does not reflect low barriers to entry that may restrict the market power of companies currently in the market.

18 B is correct. The credible threat of entry holds down prices and multiple incumbents are offering undifferentiated products.

19 C is correct. There are many competitors in the market, but some product differentiation exists, as the price differential between Deep River's brand and the house brands indicates.

READING

10

Aggregate Output, Prices, and Economic Growth

by Paul R. Kutasovic, PhD, CFA, and Richard Fritz, PhD

Paul R. Kutasovic, PhD, CFA, is at New York Institute of Technology (USA). Richard Fritz, PhD, is at the School of Economics at Georgia Institute of Technology (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** calculate and explain gross domestic product (GDP) using expenditure and income approaches;
☐	**b.** compare the sum-of-value-added and value-of-final-output methods of calculating GDP;
☐	**c.** compare nominal and real GDP and calculate and interpret the GDP deflator;
☐	**d.** compare GDP, national income, personal income, and personal disposable income;
☐	**e.** explain the fundamental relationship among saving, investment, the fiscal balance, and the trade balance;
☐	**f.** explain how the aggregate demand curve is generated;
☐	**g.** explain the aggregate supply curve in the short run and long run;
☐	**h.** explain causes of movements along and shifts in aggregate demand and supply curves;
☐	**i.** describe how fluctuations in aggregate demand and aggregate supply cause short-run changes in the economy and the business cycle;
☐	**j.** distinguish among the following types of macroeconomic equilibria: long-run full employment, short-run recessionary gap, short-run inflationary gap, and short-run stagflation;
☐	**k.** explain how a short-run macroeconomic equilibrium may occur at a level above or below full employment;
☐	**l.** analyze the effect of combined changes in aggregate supply and demand on the economy;

(continued)

© 2021 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**m.** describe sources, measurement, and sustainability of economic growth;
☐	**n.** describe the production function approach to analyzing the sources of economic growth;
☐	**o.** define and contrast input growth with growth of total factor productivity as components of economic growth.

1

INTRODUCTION

In the field of economics, *microeconomics* is the study of the economic activity and behavior of individual economic units, such as a household, a company, or a market for a particular good or service, and *macroeconomics* is the study of the aggregate activities of households, companies, and markets. Macroeconomics focuses on national aggregates, such as total *investment*, the amount spent by all businesses on plant and equipment; total *consumption*, the amount spent by all households on goods and services; the rate of change in the general level of prices; and the overall level of interest rates.

Macroeconomic analysis examines a nation's aggregate output and income, its competitive and comparative advantages, the productivity of its labor force, its price level and inflation rate, and the actions of its national government and central bank. The objective of macroeconomic analysis is to address such fundamental questions as the following:

- What is an economy's aggregate output, and how is aggregate income measured?
- What factors determine the level of aggregate output/income for an economy?
- What are the levels of aggregate demand and aggregate supply of goods and services within the country?
- Is the level of output increasing or decreasing, and at what rate?
- Is the general price level stable, rising, or falling?
- Is unemployment rising or falling?
- Are households spending or saving more?
- Are workers able to produce more output for a given level of inputs?
- Are businesses investing in and expanding their productive capacity?
- Are exports (imports) rising or falling?

From an investment perspective, investors must be able to evaluate a country's current economic environment and to forecast its future economic environment in order to identify asset classes and securities that will benefit from economic trends occurring within that country. Macroeconomic variables—such as the level of inflation, unemployment, consumption, government spending, and investment—have different effects on the growth and profitability of industries within a country, the companies within those industries, and the returns of the securities issued by those companies.

The following section describes gross domestic product and related measures of domestic output and income. Short-run and long-run aggregate demand and supply curves are discussed next, followed by the causes of shifts and movements along those curves, as well as factors that affect equilibrium levels of output, prices, and interest rates. The last section discusses sources, sustainability, and measures of economic growth.

AGGREGATE OUTPUT AND INCOME

2

a calculate and explain gross domestic product (GDP) using expenditure and income approaches

b compare the sum-of-value-added and value-of-final-output methods of calculating GDP

The **aggregate output** of an economy is the value of all the goods and services produced during a specified period. The **aggregate income** of an economy is the value of all the payments earned by the suppliers of factors used in the production of goods and services. Because the value of the output produced must accrue to the factors of production, aggregate output and aggregate income within an economy must be equal.

There are four broad forms of payments (i.e., income): compensation of employees, rent, interest, and profits. Compensation of employees includes wages and benefits (primarily employer contributions to private pension plans and health insurance) that individuals receive in exchange for providing labor. **Rent** is payment for the use of property. **Interest** is payment for lending funds. **Profit** is the return that owners of a company receive for the use of their capital and the assumption of financial risk when making their investments. We can think of the sum of rent, interest, and profit as the *operating surplus* of a company. It represents the return on all capital used by the business.

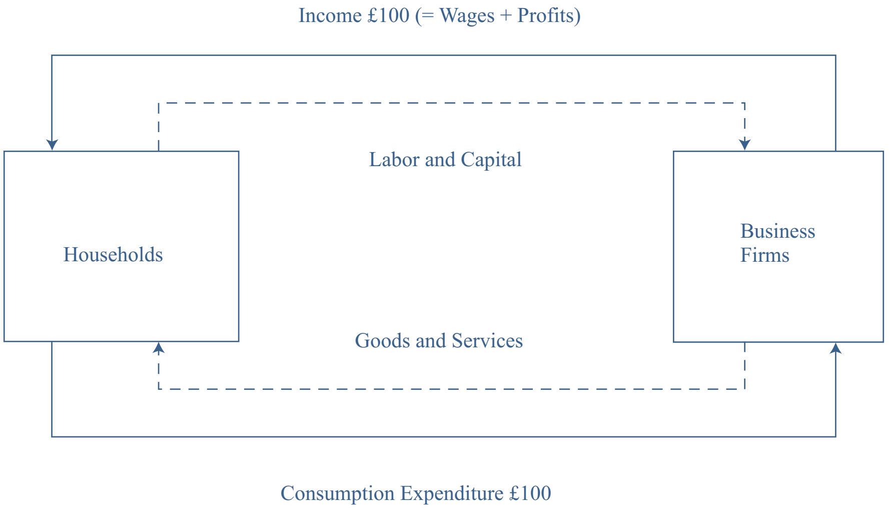
Although businesses are the direct owners of much of the property and physical capital in the economy, by virtue of owning the businesses, households are the ultimate owners of these assets and hence the ultimate recipients of the profits. In reality, of course, a portion of profits are usually retained within businesses to help finance maintenance and expansion of capacity. Similarly, because the government is viewed as operating on a non-profit basis, any revenue it receives from ownership of companies and/or property may be viewed as being passed back to households in the form of lower taxes. Therefore, for simplicity, it is standard in macroeconomics to attribute all income to the household sector unless the analysis depends on a more precise accounting.

Aggregate *expenditure*, the total amount spent on the goods and services produced in the (domestic) economy during the period, must also be equal to aggregate output and aggregate income. Some of this expenditure, however, may come from other countries in the form of net exports. Thus, aggregate output, aggregate income, and aggregate expenditure all refer to different ways of decomposing the same quantity.

Exhibit 1 illustrates the flow of inputs, output, income, and expenditures in a very simple economy. Households supply the factors of production (labor and capital) to businesses in exchange for wages and profit (aggregate income) totaling £100. These flows are shown by the top two arrows. Companies use the inputs to produce goods and services (aggregate output) that they sell to households (aggregate expenditure) for £100. The output and expenditure flows are shown by the bottom two arrows. Aggregate output, income, and expenditure are all equal to £100.

In this simplified example, households spend all of their income on domestically produced goods and services. They do not buy non-domestic goods, save for the future, or pay taxes. Similarly, businesses do not sell internationally or to the government, and they do not invest to increase their productive capacity. These important components of the economy will be added in a later section. But first we need to discuss how output and income are measured.

Exhibit 1 Output, Income, and Expenditure in a Simple Economy: The Circular Flow



2.1 Gross Domestic Product

Gross domestic product (GDP) measures

- the market value of all final goods and services produced within the economy during a given period (output definition) or, equivalently,
- the aggregate income earned by all households, all companies, and the government within the economy during a given period (income definition).

Intuitively, GDP measures the flow of output and income in the economy. GDP represents the broadest measure of the value of economic activity occurring within a country during a given period.

Therefore, GDP can be determined in two different ways. In the income approach, GDP is calculated as the total amount earned by households and companies in the economy. In the expenditure approach, GDP is calculated as the total amount spent on the goods and services produced within the economy during a given period. For the economy as a whole, total income must equal total expenditures, so the two approaches yield the same result.

Many developed countries use a standardized methodology for measuring GDP. This methodology is described in the official handbook of the Organisation for Economic Co-Operation and Development (Paris: OECD Publishing). The OECD reports the national accounts for many developed nations. In the United States, the National Income and Product Accounts (also called NIPA, or national accounts, for short) is the official US government accounting of all the income and expenditure flows in the US economy. The national accounts are the responsibility of the US Department of Commerce and are published in its *Survey of Current Business*. In Canada, similar data are available from Statistics Canada, whereas in China, the National Bureau of Statistics of China provides GDP data.

To ensure that GDP is measured consistently over time and across countries, the following three broad criteria are used:

- All goods and services included in the calculation of GDP *must be produced during the measurement period*. Therefore, items produced in previous periods—such as houses, cars, machinery, or equipment—are excluded. In addition, transfer payments from the government sector to individuals, such as unemployment compensation or welfare benefits, are excluded. Capital gains that accrue to individuals when their assets appreciate in value are also excluded.
- The only goods and services included in the calculation of GDP are those whose value *can be determined by being sold in the market*. This criterion enables the price of goods or services to be objectively determined. For example, a liter of extra virgin olive oil is more valuable than a liter of spring water because extra virgin olive oil has a higher market price than spring water does. The value of labor used in activities that are not sold on the market, such as commuting or gardening, is also excluded from GDP. Byproducts of production processes are also excluded if they have no explicit market value, such as air pollution, water pollution, and acid rain.
- Only the market value of final goods and services is included in GDP. Final goods and services are those that are not resold. *Intermediate goods* are goods that are resold or used to produce another good. Note that "intermediate goods" should not be confused with inventories and "final goods" should not be confused with so-called final sales. The value of intermediate goods is excluded from GDP because additional value is added during the production process, and all the value added during the entire production process is reflected in the final sale price of the finished good. An alternative approach to measuring GDP is summing all the value added during the production and distribution processes. The most direct approach, however, is to sum the market value of all the final goods and services produced within the economy during a given period.

Two distinct, but closely related, measurement methods can be used to calculate GDP based on expenditures: value of final output and sum of value added. Exhibit 2 illustrates these two methods. In this example, a farmer sells wheat to a miller. The miller grinds the wheat into flour and sells it to a baker, who makes bread and sells it to a retailer. Finally, the bread is sold to retail customers. The wheat and flour are both intermediate goods in this example because they are used as inputs to produce another good. Thus, they are not counted (directly) in GDP. For the purposes of GDP, the value of the final product is €1.00, which includes the value added by the bread retailer as a distributor of the bread. If, in contrast, the baker sold directly to the public, the value counted in GDP would be the price at which the baker sold the bread, €0.78. The left column of the exhibit shows the total revenue received at each stage of the process, whereas the right column shows the value added at each stage. Note that the market value of the final product (€1.00) is equal to the sum of the value added at each of the stages. Thus, the contribution to GDP can be measured as either the final sale price or the sum of the value added at each stage.

Exhibit 2 Value of Final Product Equals Income Created

	Receipts at Each Stage (€)	Value Added (= Income Created) at Each Stage (€)	
Receipts of farmer from miller	0.15	0.15	Value added by farmer
Receipts of miller from baker	0.46	0.31	Value added by miller

(continued)

Exhibit 2 (Continued)

	Receipts at Each Stage (€)	Value Added (= Income Created) at Each Stage (€)	
Receipts of baker from retailer	0.78	0.32	Value added by baker
Receipts of retailer from final customer	1.00	0.22	Value added by retailer
	1.00	1.00	
	Value of final output	Total value added = Total income created	

EXAMPLE 1

Contribution of Automobile Production to GDP

Exhibit 3 provides simplified information on the cost of producing an automobile in the United States at various stages of the production process. The example assumes the automobile is produced and sold domestically and assumes no imported material is used. Calculate the contribution of automobile production to GDP using the value-added method, and show that it is equivalent to the expenditure method. What effect would the use of imported steel or plastics have on GDP?

Exhibit 3 Cost of Producing Automobiles

Stage of Production	Sales Value ($)
1 Production of basic materials	
Steel	1,000
Plastics	3,000
Semiconductors	1,000
2 Assembly of automobile (manufacturer price)	15,000
3 Wholesale price for automobile dealer	16,000
4 Retail price	18,000

Solution:

GDP includes only the value of final goods and ignores intermediate goods in order to avoid double counting. Thus, the final sale price of $18,000 and not the total sales value of $54,000 (summing sales at all the levels of production) would be included in GDP. Alternatively, we can avoid double counting by calculating and summing the value added at each stage. At each stage of production, the difference between what a company pays for its inputs and what it receives for the product is its contribution to GDP. The value added for each stage of production is computed as follows:

Stage of Production	Sales Value ($)	Value Added ($)	
1 Production of basic materials			
Steel	1,000	1,000	
Plastics	3,000	3,000	
Semiconductors	1,000	1,000	
Total inputs		5,000	(sum of three inputs)
2 Assembly of car (manufacturer price)	15,000	10,000	= (15,000 – 5,000)
3 Wholesale price for car dealer	16,000	1,000	= (16,000 – 15,000)
4 Retail price	18,000	2,000	= (18,000 – 16,000)
Total expenditures	18,000		
Total value added		18,000	

Thus, the sum of the value added by each stage of production is equal to $18,000, which is equal to the final selling price of the automobile. If some of the inputs (steel, plastics, or semiconductors) are imported, the value added would be reduced by the amount paid for the imports.

2.1.1 *Goods and Services Included at Imputed Values*

As a general rule, only the value of goods and services whose *value can be determined by being sold in the market* are included in the measurement of GDP. Owner-occupied housing and government services, however, are two examples of services that are not sold in the marketplace but are still included in the measurement of GDP.

When a household (individual) rents a place to live, she is buying housing services. The household pays the owner of the property rent in exchange for shelter. The income that a property owner receives is included in the calculation of GDP. When a household purchases a home, however, it is implicitly paying itself in exchange for the shelter. As a result, the government must estimate (impute) a value for this owner-occupied rent, which is then added to GDP.

The value of government services provided by police officers, firemen, judges, and other government officials is a key factor that affects the level of economic activity. Valuing these services is difficult, however, because they are not sold in a market like other services; individual customers cannot decide how much to consume or how much they are willing to pay. Therefore, these services are simply included in GDP at their cost (e.g., wages paid) with no value added attributed to the production process.

For simplicity and global comparability, the number of goods and services with imputed values that are included in the measurement of GDP are limited. In general, non-market activity is excluded from GDP. Thus, activities performed for one's own benefit, such as cooking, cleaning, and home repair, are excluded. Activities in the so-called underground economy are also excluded. The underground economy reflects economic activity that people hide from the government either because it is illegal or because they are attempting to evade taxation. Undocumented laborers who are paid "off the books" are one example. The illegal drug trade is another. Similarly, barter transactions, such as neighbors exchanging services with each other (e.g., helping your neighbor repair his fence in exchange for him plowing your garden), are excluded from GDP.

Exhibit 4 shows a historical study on the estimated size of the underground economy in various countries as a percentage of nominal GDP. The estimates range from 8% in the United States to 60% in Peru. Based on these estimates, the US national income accounts fail to account for roughly 7.4% (= 8/108) of economic activity, whereas in Peru, the national accounts miss roughly 37.5% (= 60/160) of the economy. For most of the countries shown, the national accounts miss 12%–20% of the economy.

Exhibit 4 Underground Economy As a Percentage of Nominal GDP (2006)

Country	Underground Economy As a Percentage of Nominal GDP (%)
Peru	60.0
Mexico	32.1
South Korea	27.5
Costa Rica	26.8
Greece	26.0
India	24.4
Italy	23.1
Spain	20.2
Sweden	16.3
Germany	15.4
Canada	14.1
China	14.0
France	13.2
Japan	8.9
United States	8.0

Source: Friedrich Schneider and Andreas Buehm, Linz University, 2009.

It should be clear from these estimates of the underground economy that the reliability of official GDP data varies considerably across countries. Failure to capture a significant portion of activity is one problem. Poor data collection practices and unreliable statistical methods within the official accounts are also potential problems.

2.1.2 *Nominal and Real GDP*

c compare nominal and real GDP and calculate and interpret the GDP deflator

To evaluate an economy's health, it is often useful to remove the effect of changes in the general price level on GDP because higher (lower) income driven solely by changes in the price level is not indicative of a higher (lower) level of economic activity. To account for these effects, economists use **real GDP**, which indicates what would have been the total expenditures on the output of goods and services if prices were unchanged. **Per capita real GDP** (real GDP divided by the size of the population) has often been used as a measure of the average standard of living in a country.

Suppose we are interested in measuring the GDP of an economy. For the sake of simplicity, suppose that the economy consists of a single automobile maker and that in 2019, 300,000 vehicles are produced with an average market price of €18,750. GDP in 2019 would be €5,625,000,000. Economists define the value of goods and services measured at current prices as **nominal GDP**. Suppose that in 2020, 300,000 vehicles are again produced but that the average market price for a vehicle increases by 7% to €20,062.50. GDP in 2020 would be €6,018,750,000. Even though no more cars were produced in 2020 than in 2019, it appears that the economy grew by (€6,018,750,000/€5,625,000,000) – 1 = 7% between 2019 and 2020, although it actually did not grow at all.

Nominal and real GDP can be expressed as

$$\text{Nominal GDP}_t = P_t \times Q_t$$

where

P_t = prices in year t
Q_t = quantity produced in year t

$$\text{Real GDP}_t = P_B \times Q_t$$

where

P_B = prices in the base year

Taking the base year to be 2019 and putting in the 2019 and 2020 numbers gives the following:

$$\text{Nominal GDP}_{2019} = (€18{,}750 \times 300{,}000) = €5{,}625{,}000{,}000$$

$$\text{Real GDP}_{2019} = (€18{,}750 \times 300{,}000) = €5{,}625{,}000{,}000$$

$$\text{Nominal GDP}_{2020} = (€20{,}062.50 \times 300{,}000) = €6{,}018{,}750{,}000$$

$$\text{Real GDP}_{2020} = (€18{,}750 \times 300{,}000) = €5{,}625{,}000{,}000$$

In this example, real GDP did not change between 2019 and 2020 because the total output remained the same: 300,000 vehicles. The difference between nominal GDP in 2020 and real GDP in 2020 was the 7% inflation rate.

Now suppose that the auto manufacturer produced 3% more vehicles in 2020 than in 2019 (i.e., production in 2020 was 309,000 vehicles). Real GDP would increase by 3% from 2019 to 2020. With a 7% increase in prices, nominal GDP for 2020 would now be as follows:

$$\begin{aligned}\text{Nominal GDP}_{2020} &= (1.03 \times 300{,}000) \times (1.07 \times €18{,}750)\\ &= (309{,}000 \times €20{,}062.50)\\ &= €6{,}199{,}312{,}500\end{aligned}$$

The **implicit price deflator for GDP**, or simply the **GDP deflator**, is defined as follows:

$$\text{GDP deflator} = \frac{\text{Value of current year output at current year prices}}{\text{Value of current year output at base year prices}} \times 100$$

Thus, in the example, the GDP deflator for 2020 is [(309,000 × €20,062.50)/ (309,000 × €18,750)](100) = (1.07)(100) = 107. The GDP deflator broadly measures the aggregate changes in prices across the overall economy, and hence changes in the deflator provide a useful gauge of inflation within the economy.

Real GDP is equal to nominal GDP divided by the GDP deflator scaled by 100:

$$\text{Real GDP} = \text{Nominal GDP}/(\text{GDP deflator}/100)$$

This relation gives the GDP deflator its name. That is, the measure of GDP in terms of current prices, nominal GDP, is adjusted for inflation by dividing it by the deflator. The expression also shows that the GDP deflator is the ratio of nominal GDP to real GDP scaled by 100:

$$\text{GDP deflator} = (\text{Nominal GDP}/\text{Real GDP}) \times 100$$

Thus, real GDP for 2020 would be

$$\begin{aligned}\text{Real GDP}_{2020} &= \text{Nominal GDP/(GDP deflator/100)}\\ &= \text{€}6{,}199{,}312{,}500/(107/100)\\ &= \text{€}5{,}793{,}750{,}000\end{aligned}$$

Note that €5,793,750,000 represents 3% real growth over 2019 GDP and 3% higher real GDP for 2020 than under the assumption of no growth in unit car sales in 2020.

What would be the increase in *nominal* GDP for 2020 compared with 2019 with the 3% greater automobile production and 7% inflation?

$$(\text{Nominal GDP}_{2020}/\text{Nominal GDP}_{2019}) - 1$$

$$= (\text{€}6{,}199{,}312{,}500/\text{€}5{,}625{,}000{,}000) - 1$$

$$= 0.102$$

So, nominal GDP would increase by 10.2%, which equals (1.07 × 1.03) − 1, or approximately 7% + 3% = 10%. Which number is more informative about growth in economic activity: 3% real growth or 10.2% nominal growth? The real growth rate is more informative because it exactly captures increases in output. Nominal growth, by blending price changes with output changes, is less directly informative about output changes.In summary, real economic growth is measured by the percentage change in real GDP. When measuring real economic activity or when comparing one nation's economy with another, real GDP and real GDP growth should be used because they more closely reflect the quantity of output available for consumption and investment.

EXAMPLE 2

Calculating the GDP Deflator

John Lambert is an equity analyst with Equitytrust, an investment management firm that primarily invests in domestic stocks and bonds in a hypothetical country. Equitytrust's investment policy committee is concerned about the possibility of inflation. The implicit GDP deflator is an important measure of the overall price level in the economy, and changes in the deflator provide an important gauge of inflation within the economy. Released GDP data are shown in Exhibit 5. The committee asks Lambert to use the GDP data to calculate the implicit GDP price deflator from 2016 to 2020 and the inflation rate for 2020.

Exhibit 5 Real and Nominal GDP

	Seasonally adjusted at annual rates (SAAR)				
	2016	**2017**	**2018**	**2019**	**2020**
GDP at market prices	1,822,808	1,897,531	1,990,183	1,994,911	2,035,506
Real GDP	1,659,195	1,698,153	1,747,478	1,762,561	1,786,677

Solution:

The implicit GDP price deflator measures inflation across all sectors of the economy, including the consumer, business, government, exports, and imports. It is calculated as the ratio of nominal to real GDP and reported as an index number with the base year deflator equal to 100. The implicit GDP price deflator for the economy for 2020 is calculated as (2,035,506/1,786,677) × 100 = 113.9. The results for the other years are shown in the following table:

	2016	2017	2018	2019	2020
GDP at market prices ($ millions)	1,822,808	1,897,531	1,990,183	1,994,911	2,035,506
Real GDP (2007 $ millions)	1,659,195	1,698,153	1,747,478	1,762,561	1,786,677
Implicit GDP price deflator	113.9	111.7	113.9	113.2	113.9

For 2016, the annual inflation rate is equal to [(113.9/113.2) – 1] or 0.62 percent. This shows that Canada experienced a very low rate of inflation in 2016.

2.2 The Components of GDP

d compare GDP, national income, personal income, and personal disposable income

Having defined GDP and discussed how it is measured, we can now consider the major components of GDP, the flows among the four major sectors of the economy—the household sector, the business sector, the government sector, and the foreign or external sector (consisting of transactions with the "rest of the world")—and the markets through which they interact. An expression for GDP, based on the expenditure approach, is

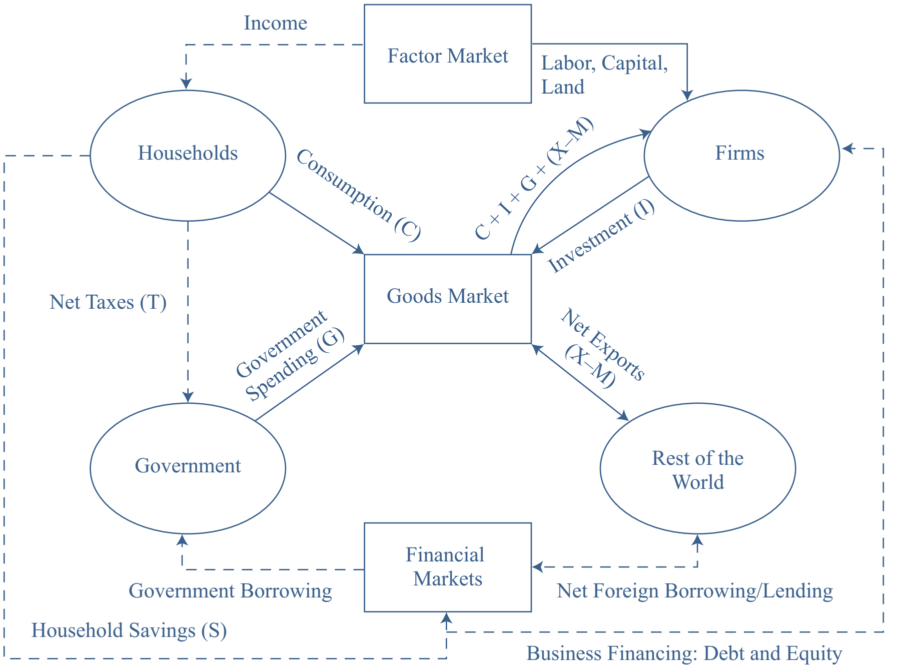
$$\text{GDP} = C + I + G + (X - M) = (C + G^C) + (I + G^I) + (X - M) \tag{1}$$

where

C = consumer spending on final goods and services
I = gross private domestic investment, which includes business investment in capital goods (e.g., fixed capital such as plant and equipment) and changes in inventory (**inventory investment**)
G = government spending on final goods and services for both current consumption and investment in capital goods = $G^C + G^I$
X = exports
M = imports

Exhibit 6 shows the flow of expenditures, income, and financing among the four sectors of the economy and the three principal markets. In the exhibit, solid arrows point in the direction of expenditure on final goods and services. For simplicity, corresponding flows of output are not shown separately. The flow of factors of production is also shown with a solid arrow. Financial flows, including income and net taxes, are shown with dashed arrows pointing to the recipient of funds.

Exhibit 6 Output, Income, and Expenditure Flows


Income
Factor Market
Labor, Capital, Land
Households
Firms
Consumption (C)
C + I + G + (X–M)
Investment (I)
Goods Market
Net Taxes (T)
Government Spending (G)
Net Exports (X–M)
Government
Rest of the World
Financial Markets
Government Borrowing
Net Foreign Borrowing/Lending
Household Savings (S)
Business Financing: Debt and Equity

2.2.1 *The Household and Business Sectors*

The very top portion of Exhibit 6 shows the services of labor, land, and capital flowing through the *factor market* to business firms and the flow of income back from firms to households. Households spend part of their income on consumption (C) and save (S) part of their income for future consumption. Current consumption expenditure flows through the *goods market* to the business sector. Household saving flows into the *financial markets*, where it provides funding for businesses that need to borrow or raise equity capital. Firms borrow or raise equity primarily to finance investment (I) in inventory, property, plant, and equipment. Investment (I) is shown flowing from firms through the goods market and back to firms because the business sector both demands and produces the goods needed to build productive capacity (*capital goods*).

In most developed economies, such as Italy and the United States, expenditures on capital goods represent a significant portion of GDP. Investments (expenditures) on capital goods accounted for approximately 12.3% of Italy's GDP in 2015, while in the United States investments accounted for approximately 16.2% of GDP during the same period. In some developing countries, notably China (45.4%) and India (32.9%), investment spending accounts for a substantially larger share of the economy (according to OECD.Stat). As we will examine in greater detail later, investment spending is an important determinant of an economy's long-term growth rate. At the same time, investment spending is the most volatile component of the economy, and changes in capital spending—especially spending on inventories—are one of the main factors causing short-run economic fluctuations.

2.2.2 *The Government Sector*

The government sector collects taxes from households and businesses. For simplicity, only the taxes collected from the household sector are shown in Exhibit 6. In turn, the government sector purchases goods and services (G) for both consumption and investment from the business sector. For example, the government sector hires construction companies to build roads, schools, and other infrastructure goods. Government expenditure (*G*) also reflects spending on the military, police and fire protection, the postal service, and other government services. To keep Exhibit 6 simple, however, we combine consumption and investment expenditures into government expenditure, *G*.

Governments also make transfer payments to households. In general, such payments are designed to address social objectives such as maintaining minimum living standards, providing health care, and assisting the unemployed with retraining and temporary support. In Exhibit 6, transfer payments are subtracted from taxes and reflected in net taxes (*T*).

Transfer payments are not included in government expenditures on goods and services (*G*) because they represent a monetary transfer by the government of tax revenue back to individuals, with no corresponding receipt of goods or services. The household spending facilitated by the transfer payments is, of course, included in consumption (*C*) and, hence, GDP. It is worth noting that transfers do not always take the form of direct payments to beneficiaries. Instead, the government may pay for or even directly provide goods or services to individuals. For example, universal health care programs often work in this way.

If, as is usually the case, government expenditure (*G*) exceeds net taxes (*T*), then the government has a *fiscal deficit* and must borrow in the financial markets. Thus, the government may compete with businesses in the financial markets for the funds generated by household saving. The only other potential source of funds in an economy is capital flows from the rest of the world, discussed in the next section.

In 2015, the ratio of general government spending (which includes central government as well as state, provincial, and local government) to GDP in Italy was 50.4%, while in the United States it was 37.7%. In countries where the government provides more services, such as universal health care in Italy, the government's contribution to GDP is greater. France's government sector represents 57.0% of GDP. In other countries, the public sector makes up a smaller share. In Mexico, for example, government spending is 24.5% of GDP. Exhibit 7 shows data on tax revenues, general government spending, and transfer payments as a share of nominal GDP.

Exhibit 7 General Government Spending and Taxes as a Percentage of GDP (2015)

		General Government Spending as a Percentage of GDP		
Country	**General Government Tax Revenues as a Percentage of GDP**	**Total**	**Goods and Services and Debt Service**	**Transfer Payments**
Canada	39.8%	41.1%	n.a.	n.a.
Mexico	23.6	24.5	n.a.	n.a.
United States	33.5	37.7	29.9	7.8
Japan	35.8	39.4	23.3	16.1
South Korea	33.8	32.4	n.a.	n.a.
France	53.4	57.0	32.5	24.5
Germany	44.7	44.0	25.0	19.0
Greece	48.3	54.2	33.7	20.5
Italy	47.7	50.4	29.0	21.4
Sweden	50.5	50.2	29.3	20.9
Australia	34.3	37.2	27.0	10.2
OECD – Average	42.2	43.8	27.3	16.5

Note: n.a. = not available
Source: Government at a Glance, 2017 edition (stats.oecd.org).

2.2.3 *The External Sector*

Trade and capital flows involving the rest of the world are shown in the bottom right quadrant of Exhibit 6. Net exports ($X - M$) reflects the difference between the value of goods and services sold internationally—exports (X)—and the portion of domestic consumption (C), investment (I), and government expenditure (G) that represents purchases of goods and services from the rest of the world—imports (M).

A **balance of trade deficit** means that the domestic economy is spending more on international goods and services than other nations' economies are spending on domestic goods and services. It also means that the country is spending more than it produces because domestic saving is insufficient to finance domestic investment plus the government's fiscal balance. A trade deficit must be funded by borrowing from the rest of the world through the financial markets. The rest of the world is able to provide this financing because, by definition, it must be running a corresponding trade surplus and is spending less than it produces.

It bears emphasizing that trade and capital flows between an economy and the rest of the world must balance. One area's deficit is another's surplus, and vice versa. This dynamic is an accounting identity that must hold. In effect, having allowed a country to run a trade deficit, other nations must, in aggregate, finance it. The financing terms, however, may or may not be attractive.

Exhibit 8 reports trade balances for the United States with selected countries. Note that in 2016, China was the United States' largest trading partner, in terms of sending goods to the United States (US imports). Mexico and Canada, members of NAFTA along with the United States, were second and third in sending goods to the United Sates but were also important consumers of US goods (US exports). Over all its trading partners, the US balance of trade deficit in 2016 was US$734,316.3 million.

Exhibit 8 US International Trade in Goods—Selected Countries, 2016 (US$ millions)

	Exports	Imports	Balance
Total, all countries	1,454,624.2	2,188,940.5	−734,316.3
Europe	318,447.1	483,454.8	−165,007.7
Euro area	200,166.6	325,880.1	−125,713.4
France	30,941.2	46,764.6	−15,823.4
Germany	49,362.0	114,227.4	−64,865.4
Italy	16,753.6	45,210.1	−28,456.5
Canada	266,826.7	278,066.8	−11,240.1
Mexico	230,969.1	294,151.0	−63,191.9
China	115,775.1	462,813.0	−347,037.9
India	21,689.0	45,998.4	−24,309.5
Japan	63,264.3	132,201.8	−68,937.6

Source: FT-900 Supplement, US International Trade in Goods and Services, December 2016.

2.3 GDP, National Income, Personal Income, and Personal Disposable Income

This section examines the calculation of GDP and other income measures in detail by means of an analysis of data from Statistics Canada.

Exhibit 9a provides data on the level of Canadian GDP and its components measured at market prices (nominal GDP) as part of Example 3.

Exhibit 9a GDP Measures for the Canadian Economy
(C$ millions at market prices, seasonally adjusted at annual rates)

	2014	2015	2016	2017	2018
Expenditure-Based:					
Final consumption expenditure	**1,517,376**	**1,566,215**	**1,610,954**	**1,684,563**	**1,754,326**
Household final consumption expenditure	1,086,253	1,122,079	1,153,668	1,208,438	1,255,327
Non-profit institutions serving households' final consumption expenditure	26,826	28,575	30,951	31,976	33,207
General governments final consumption expenditure	404,297	415,561	426,335	444,149	465,792
Gross fixed capital formation	**486,542**	**474,732**	**461,262**	**486,760**	**501,247**
Business gross fixed capital formation	410,591	395,432	382,196	401,686	409,835
Non-profit institutions serving households' gross fixed capital formation	3,145	2,956	2,212	2,290	2,325
General governments' gross fixed capital formation	72,806	76,344	76,854	82,784	89,087
Investment in inventories	**9,603**	**–577**	**–228**	**17,088**	**11,904**
Exports to other countries	**633,112**	**633,955**	**638,095**	**672,533**	**713,351**
Less: Imports from other countries	**651,176**	**683,019**	**685,868**	**719,909**	**756,913**
Statistical discrepancy	**–559**	**–865**	**1,320**	**72**	**–59**
GDP at market prices	**1,994,898**	**1,990,441**	**2,025,535**	**2,141,107**	**2,223,856**
Income-Based:					
Compensation of employees	**998,463**	**1,026,846**	**1,026,483**	**1,069,956**	**1,120,520**
Gross operating surplus	**557,244**	**505,167**	**526,024**	**573,257**	**590,510**
Net operating surplus: corporations	289,123	220,212	233,763	280,267	287,280
Consumption of fixed capital: corporations	203,080	217,285	222,954	221,894	228,666
Consumption of fixed capital: general governments and non-profit institutions serving households	65,041	67,670	69,307	71,096	74,564
Gross mixed income	**227,166**	**235,365**	**244,403**	**256,757**	**265,147**
Net mixed income	172,079	177,997	183,577	194,463	199,212
Consumption of fixed capital: unincorporated businesses	55,087	57,368	60,826	62,294	65,935
Taxes less subsidies on production	**84,321**	**88,570**	**90,516**	**95,563**	**98,698**
Taxes less subsidies on products and imports	**127,144**	**133,627**	**139,429**	**145,647**	**148,921**

(continued)

Exhibit 9a (Continued)

	2014	2015	2016	2017	2018
Statistical discrepancy	560	866	−1,320	−73	60
Gross domestic income at market prices	1,994,898	1,990,441	2,025,535	2,141,107	2,223,856

Source: Statistics Canada. Table 36-10-0222-01 and Table: 36-10-0221-01 GDP, expenditure-based & income-based, annual (x C$1,000,000).

Exhibit 9a shows the two approaches to measuring GDP: 1) expenditures on final output measured as the sum of sales to the final users and 2) the sum of the factor incomes generated in the production of final output. In theory, the two approaches should provide the same estimate of GDP. As shown in the exhibit, however, they differ in practice because of the use of different data sources. The difference is accounted for by a *statistical discrepancy*. Market analysts more closely follow the expenditure approach because the expenditure data are more timely and reliable than data for the income components.

Using the expenditure approach, Statistics Canada measures Canadian GDP as follows:

GDP = Consumer spending on goods and services
+ Business gross fixed investment
+ Change in inventories
+ Government spending on goods and services
+ Government gross fixed investment
+ Exports – Imports
+ Statistical discrepancy

Canadian national income accounts explicitly allocate government expenditures between consumption expenditure and gross fixed capital formation. Not all countries make this distinction. The United States, for example, does not. Also note that the investment in business inventories must be included in expenditures. Otherwise, goods produced but not yet sold would be left out of GDP.

The income-based approach calculates gross domestic income (GDI) as the sum of factor incomes and essentially measures the cost of producing final output. Two of the costs entering into the gross value of output, however, are not really earned by a factor of production. These items, depreciation and indirect taxes, are discussed next. GDP is estimated in the income approach as follows (note that construction of the national income accounts varies across countries):

GDP = Gross domestic income (GDI)
= Net domestic income + Consumption of fixed capital (CFC) + Statistical discrepancy

where gross domestic income is the income received by all factors of production used in the generation of final output:

Gross domestic income = Compensation of employees
+ Gross operating surplus
+ Gross mixed income
+ Taxes less subsidies on production
+ Taxes less subsidies on products and imports

Compensation of employees consists of 1) wages and salaries including direct compensation in cash or in kind plus 2) "employers' social contributions" that supplement wages, which are primarily payments for government-sponsored social security schemes including pensions and health insurance.

Gross operating surplus is related to corporate profits and includes private corporations, non-profit corporations, and government corporations. Gross operating surplus is the surplus arising from operations and does not take out charges for rent, interest, or similar charges on financial assets or natural resources used by the business. As such, gross operating surplus essentially measures the return on capital used by the business as a whole, rather than the return to the owners of the business (profit).

Gross mixed income is the same concept applied to unincorporated business in the economy. It is measured in the same way as gross operating surplus and has three major components: 1) farm income, 2) non-farm income excluding rent, and 3) rental income.

"Indirect business taxes less subsidies" reflects taxes and subsidies included in the final price of the good or service. It is the (net) portion of **national income** that is directly paid to the government. In the Canadian accounts, these are measured in two ways: 1) "taxes less subsidies on products and imports," which include sales taxes, fuel taxes, and import duties, and 2) "taxes less subsidies on factors of production," which are mainly property taxes and payroll taxes.

The consumption of fixed capital (CFC) is a measure of the wear and tear (depreciation) of the capital stock that occurs in the production of goods and services. This measure acknowledges the fact that some income/output must be allocated to the replacement of the existing capital stock as it wears out. Loosely speaking, one may think of Profit + CFC as the gross surplus earned by capital, with the CFC being the amount that must be earned and reinvested to maintain the existing productivity of the capital.

Along with the GDP report, Statistics Canada and other government statistical agencies provide information on personal income (called Primary Household Income in the Canadian accounts) and personal saving (called Household Net Saving), as shown in Exhibit 9b. **Personal income** is a broad measure of household income and measures the ability of consumers to make purchases. As such, it is one of the key determinants of consumption spending. Primary Household Income includes all income received by households, whether earned or unearned. It includes compensation of employees plus net mixed income from unincorporated businesses plus net property income.

Exhibit 9b Household Income and Saving for the Canadian Economy (C$ millions at market prices, seasonally adjusted at annual rates)

	2014	2015	2016	2017	2018
Compensation of employees	**996,247**	**1,024,743**	**1,024,142**	**1,066,600**	**1,117,091**
Plus: net mixed income	172,079	177,997	183,577	194,463	**fill in**
Plus: net property income	126,019	142,185	133,098	140,942	140,370
Primary household income	**1,294,345**	**1,344,925**	**1,340,817**	**1,402,005**	**calc**
Plus: current transfers received	266,153	288,200	307,077	320,861	333,602
Less: current transfers paid	478,843	503,245	515,136	533,256	555,884
Household disposable income	**1,081,655**	**1,129,880**	**1,132,758**	**1,189,610**	**calc**
Less: household final consumption expenditure	1,086,253	1,122,079	1,153,668	1,208,438	fill in
Plus: change in pension entitlements	46,439	43,260	43,139	43,391	43,087
Household net saving	**41,841**	**51,061**	**22,229**	**24,563**	**calc**

(continued)

Exhibit 9b (Continued)

	2014	2015	2016	2017	2018
Household saving rate (%)	3.9	4.5	2.0	2.1	calc

Source: Statistics Canada. Table 36-10-0224-01 Household sector, current accounts, provincial and territorial, annual.

Household disposable income (HDI) is equal to personal income less personal taxes. In reality, households both pay taxes to the government and receive some income from governments (transfer payments such as social insurance payments, unemployment compensation, and disability payments) that is not earned. Thus HDI is calculated as household primary income less net current transfers paid. This calculation measures the amount of after-tax income that households have to spend on goods and services or to save. Thus, it is the most relevant, and most closely watched, measure of income for household spending and saving decisions.

Finally, household net saving is equal to HDI adjusted for two items: subtracting household final consumption expenditures and adding the net change in pension entitlements.

EXAMPLE 3

Canadian GDP Release and Other Measures of Production and Income

The investment policy committee at Equitytrust asks John Lambert to review the Canadian GDP data shown in Exhibits 9a and 9b.

1. Calculate 2018 GDP using the expenditure approach, and indicate how the expenditures are represented in Exhibit 6.
2. Calculate 2018 GDP using the income approach.
3. Calculate personal income for 2018.
4. Using the Canadian data for 2018, calculate the level of household saving (S), the saving rate, and net payments (T) by the household sector to government.
5. Calculate the impact of international trade on the Canadian economy in 2018 and Canada's net foreign borrowing/lending in 2018.
6. Calculate the amount of net fixed capital formation across all sectors in 2018.

Solutions:

(all numbers in millions)

Solution to 1:

In the expenditure approach, nominal GDP is calculated as the sum of spending by the major sectors in the economy:

GDP = Final consumption expenditures + Gross fixed capital formation + Investment in Inventories + Exports − Imports + Statistical discrepancySubstituting the numbers from Exhibit 9a,

GDP = 1,754,326 + 501,247 + 11,904 + 713,351 − 756,913 − 59
= C$2,223,856.

In Exhibit 6, these expenditures are represented by the arrows pointing to the goods market and by the arrow pointing back to firms labeled as $C + I + G + (X - M)$. Note that G in Exhibit 6 is captured here partly in final consumption expenditures and partly in gross fixed capital formation for the Canadian national income accounts.

Solution to 2:

On the income side, nominal GDP is equal to gross domestic income—the sum of income received by the factors of production—given by the following equation:

Gross domestic income = Compensation of employees + Gross operating surplus + Gross mixed income+ Indirect business taxes − Subsidies + Statistical discrepancy

Substituting in the numbers from Exhibit 9a, we get indirect business taxes less subsidies as equal to the sum of taxes less subsidies on production plus taxes less subsidies on products and imports = 98,698 + 148,921 = C$247,619. Using this result,

GDP = GDI = 1,120,520 + 590,510 + 265,147 + 247,619 + 60
= C$2,223,856.

Solution to 3:

Personal income (household primary income) is calculated as

Household Primary Income = Compensation of employees + Net mixed income + Net property incomeSubstituting in the numbers from Exhibit 9b,

Household primary income = 1,117,091 + 199,212 + 140,370
= C$1,456,673.

Solution to 4:

Household saving is equal to personal disposable income less household final consumption expenditures plus the change in pension entitlements. Household final consumption (C) is given in Exhibit 9a as C$1,255,327. Substituting in the appropriate numbers, Household net saving = Household primary income − Net current transfers paid − Household final consumption expenditure + Net change in pension entitlements = 1,456,673 − (555,884 − 333,602) − 1,255,327 + 43,087 = C$22,151.

The Canadian saving rate for 2018 = (22,151/1,117,091) = 2%.

Net "taxes" paid by the household sector consists of two components: 1) transfers paid by households to the government minus 2) government transfer payments received by households. From Exhibit 9b, government transfer payments to households for 2018 were C$333,602. The transfer payments by households to government in 2018 was C$555,884. Thus the net payments *to* government by households in 2018 was:

555,884 – 333,602 = C$222,282

Solution to 5:

It is clear from Exhibit 9a that the international sector generally has a large effect each year on the Canadian economy. In 2018, exports were C$713,351, or roughly 32% of Canadian GDP. Imports were higher than exports at C$756,913 in 2018, indicating that Canada had a trade deficit of 713,351 – 756,913 = C$43,562. This figure is roughly the same size as the trade deficit in 2017.

Canada funds its trade deficit by borrowing from the rest of the world through the financial markets. In 2018, this involved borrowing C$43,562 from the rest of the world. Trade and capital flows between an economy and the rest of the world must balance. A trade deficit must be funded by a capital inflow.

Solution to 6:

Net fixed capital formation is equal to gross fixed capital formation found in the expenditure accounts less consumption of fixed capital (CFC) by corporations, government and non-profits, and unincorporated businesses found in the income accounts. In 2018, gross fixed capital formation was C$501,247, and CFC was (228,666 + 74,564 + 65,935) for corporations, government and non-profits, and unincorporated businesses, respectively.

Thus net fixed capital formation in the Canadian economy for 2018 was 501,247 – 369,165 = C$132,082.

2.4 Relationship among Saving, Investment, the Fiscal Balance and the Trade Balance

e explain the fundamental relationship among saving, investment, the fiscal balance, and the trade balance

Total expenditure on domestically produced output comes from four sources: household consumption (C), investments (I), government spending (G), and net exports ($X - M$). This can be expressed as follows:

$$\text{Expenditure} = C + I + G + (X - M)$$

Personal disposable income is equal to GDP (Y) plus transfer payments (F) minus retained earnings and depreciation (= business saving, S_B) minus direct and indirect taxes (R). Households allocate disposable income between consumption of goods and services (C) and household saving (S_H). Therefore,

$$Y + F - S_B - R = C + S_H$$

Rearranging this equation, we get

$$Y = C + S + T$$

where $T = (R - F)$ denotes net taxes and $S = (S_B + S_H)$ denotes total private sector saving.

Because total expenditures must be identical to aggregate income (Y), we have the following relationship:

$$C + S + T = C + I + G + (X - M)$$

By rearranging this equation, we get the following fundamental relationship among domestic saving, investment, the fiscal balance, and the trade balance:

$$S = I + (G - T) + (X - M) \quad (2)$$

This equation shows that domestic private saving is used or absorbed in one of three ways: investment spending (I), financing government deficits ($G - T$), and building up financial claims against overseas economies [positive trade balance, $(X - M) > 0$]. If there is a trade deficit [$(X - M) < 0$], then domestic private saving is being supplemented by inflows of foreign saving and other countries' economies are building up financial claims against the domestic economy.

By rearranging the identity, we can examine the implications of government deficits and surpluses:

$$G - T = (S - I) - (X - M)$$

A fiscal deficit [$(G - T) > 0$] implies that the private sector must save more than it invests [$(S - I) > 0$] or the country must run a trade deficit [$(X - M) < 0$] with corresponding inflow of foreign saving—or both.

EXAMPLE 4

Foreign Capital Inflows Help Finance Government Deficits

The budgetary situation changed dramatically in Canada during 2009, the first year of the financial crisis. The Department of Finance Canada reported that in 2009, the combined federal–provincial governments had a deficit of C$84,249 (all amounts in C$ millions). Thus, the government sector operated at a deficit that needed to be financed. How was this deficit financed?

Solution:

Using the formula $G - T = (S - I) - (X - M)$ shows that a budget deficit is financed through either higher domestic saving (S), lower business investment (I), or international borrowing ($X - M$).

In 2009, private sector saving exceeded investment spending by C$58,588 (319,802 – 277,574). Thus, domestic private saving financed more than 69.5% of the 2009 government deficit (58,588/84,249).

To finance the rest of the government deficit, foreign imports (M) would have to exceed exports (X) by C$25,661. The actual trade deficit (amount of foreign borrowing) in 2009 was C$26,169, slightly greater than the amount required. This difference largely results from the statistical discrepancy caused by different data sources being used for expenditure-based and income-based estimates of GDP.

Equation 2 is the key relationship that must hold in order for aggregate income and aggregate expenditure to be equal. Up to this point, we have treated it as simply an accounting identity. We now need to think of it as the outcome of explicit decisions on the part of households, businesses, government, and people in other countries. When we do so, we face the question of what underlies these decisions and how the requisite balance is established.

Economists have found that the dominant determinant of consumption spending is disposable income ($Y - S_B - T$). This relationship can be expressed formally by indicating that consumption is a function $\mathbf{C}(\cdot)$ of disposable income,

$$C = \mathbf{C}(Y - S_B - T)$$

or, dropping the technically correct but practically insignificant adjustment for retained earnings and depreciation (S_B), a function of GDP minus net taxes:

$$C = \mathbf{C}(Y - T)$$

When households receive an additional unit of income, some proportion of this additional income is spent and the remainder is saved. The **marginal propensity to consume** (MPC) represents the proportion of an additional unit of disposable income that is consumed or spent. Because the amount that is not spent is saved, the **marginal propensity to save** (MPS) is MPS = 1 − MPC.

According to the consumption function, either an increase in real income or a decrease in taxes will increase aggregate consumption. Somewhat more sophisticated models of consumption recognize that consumption depends not only on current disposable income but also on wealth. Except for the very rich, individuals tend to spend a higher fraction of their current income as their wealth increases because with higher current wealth, there is less need to save to provide for future consumption.

Exhibit 10 shows household consumption expenditures as a percentage of GDP for selected countries.

Exhibit 10 Household Final Consumption Expenditures as a Percentage of GDP, 2018

United States	68.0%
Mexico	64.8
Italy	60.3
Japan	55.6
Canada	57.9
France	53.9
Germany	52.1

Source: The World Bank, https://data.worldbank.org/indicator/NE.CON.PRVT.ZS

These figures reflect the *average propensity to consume* (APC)—that is, the ratio C/Y—rather than a measure of how the next unit of income would be divided between spending and saving, the MPC. However, they are reasonable proxies for the MPC in each country. Comparing Germany's 55.0% APC with Mexico's 67.8%, the implication is that the Mexican economy is more sensitive to changes in disposable household income than is the German economy. All else being equal, macroeconomic policies that increase disposable household income, such as lowering government taxes, would have a larger effect on the economies of Mexico (67.8%) and the United States (68.3%) than similar policies would have in Germany (55.0%) or France (55.4%).

Companies are the primary source of investment spending (I). They make investment decisions in order to expand their stock of physical capital, such as building new factories or adding new equipment to existing facilities. A definition of physical capital is *any manmade aid to production*. Companies also buy investment goods, such as manufacturing plants and equipment to replace existing facilities and equipment that wear out. Total investment, including replacement of worn-out capital, is called *gross investment*, as opposed to *net investment*, which reflects only the addition of new capacity. GDP includes gross investment; hence the name *gross* domestic product. Total investment spending in such developed countries as Italy, Germany, the United Kingdom, and the United States ranged between 12% and 16% of GDP in 2015 (OECD. Stat Extracts: Country Statistical Profiles 2017, stats.oecd.org).

Investment decisions depend primarily on two factors: the level of interest rates and aggregate output/income. The level of interest rates reflects the cost of financing investment. The level of aggregate output serves as a proxy for the expected profitability of new investments. When an economy is underutilizing its resources, interest rates are typically very low, and yet investment spending often remains dormant because the expected return on new investments is also low. Conversely, when output is high and companies have little spare capacity, the expected return on new investments is high. Thus, investment decisions may be modeled as a decreasing function $\boldsymbol{I}(\cdot,\cdot)$ of the **real interest rate** (nominal interest rate minus the expected rate of inflation) and an increasing function of the level of aggregate output. Formally,

$$I = \boldsymbol{I}(r, Y)$$

where I is investment spending, r is the real interest rate, and Y is, as usual, aggregate income. This investment function leaves out some important drivers of investment decisions, such as the availability of new and better **technology**. Nonetheless, it reflects the two most important considerations: the cost of funding (represented by the real interest rate) and the expected profitability of the new capital (proxied by the level of aggregate output).

Many government spending decisions are insensitive to the current level of economic activity, the level of interest rates, the currency exchange rate, and other economic factors. Thus, economists often treat the level of government spending on goods and services (G) as an *exogenous* policy variable determined outside the macroeconomic model. In essence, this means that the adjustments required to maintain the balance among aggregate spending, income, and output must occur primarily within the private sector.

Tax policy may also be viewed as an exogenous policy tool. The actual amount of net taxes (T) collected, however, is closely tied to the level of economic activity. Most countries impose income taxes or value-added taxes (VAT) or both that increase with the level of income or expenditure. Similarly, at least some transfer payments to the household sector are usually based on economic need and are hence inversely related to aggregate income. Each of these factors makes net taxes (T) rise and fall with aggregate income, Y. The government's fiscal balance can be represented as

$$G - T = \bar{\boldsymbol{G}} - t(Y)$$

where $\bar{\boldsymbol{G}}$ is the exogenous level of government expenditure and $t(Y)$ indicates that net taxes are an (increasing) function of aggregate income, Y. The fiscal balance decreases (smaller deficit or larger surplus) as aggregate income (Y) increases and increases as income declines. This effect is called an *automatic stabilizer* because it tends to mitigate changes in aggregate output.

Net exports ($X - M$) are primarily a function of income in the domestic country and in the rest of the world and of the relative prices of domestic and non-domestic goods and services. As domestic income rises, some of the additional demand that is induced will be for imported goods. Thus, net exports will decline. An increase in income in the rest of the world will lead to an increase in demand for the domestic country's products and hence an increase net exports. A decrease in the relative price of domestically produced goods and services, perhaps because of a depreciation in the domestic currency, will shift demand toward these products and hence increase net exports.

3 AGGREGATE DEMAND AND AGGREGATE SUPPLY

f explain how the aggregate demand curve is generated

Short-term fluctuations in an economy's output arise for a number of reasons. The fluctuations of output and price level can have a significant effect on issuers of securities and therefore come under close scrutiny by practitioners and policymakers.

We will build a model of aggregate demand and aggregate supply and use it to discuss how aggregate output and the level of prices are determined in the economy. **Aggregate demand** (AD) represents the quantity of goods and services that households, businesses, government, and international customers want to buy at any given level of prices. **Aggregate supply** (AS) represents the quantity of goods and services that producers are willing to supply at any given level of prices. It also reflects the amount of labor and capital that households are willing to offer into the marketplace at given real wage rates and cost of capital.

3.1 Aggregate Demand

f explain how the aggregate demand curve is generated

As we will see, the **aggregate demand curve** looks like the ordinary demand curves that we encounter in microeconomics: quantity demanded increases as the price level declines. But our intuitive understanding of that relationship—lower price allows us to buy more of a good *with a given level of income*—does not apply here because income is not fixed. Instead, aggregate income/expenditure (GDP) is to be determined within the model along with the price level. Thus, we need to explain the relationship between price and quantity demanded somewhat differently.

The aggregate demand curve represents the combinations of aggregate income and the price level at which two conditions are satisfied. First, aggregate expenditure equals aggregate income. As indicated in our discussion of GDP accounting, this relation must always be true after the fact. Second, the available real money supply is willingly held by households and businesses. Recall that GDP is defined as

$$\text{GDP} = C + I + G + (X - M),$$

where

C = consumer spending on final goods and services
I = gross private domestic investment
G = government spending on final goods and services
X = exports
M = imports

We assume that government spending (G) is exogenous, set by government policy. To explain why the aggregate demand curve slopes downward, we analyze how the price level in the economy affects the consumption (C), investment (I), and net export ($X - M$) components of GDP. The downward slope of the aggregate demand curve results from three effects: the **wealth effect**, the **interest rate effect**, and the **real exchange rate effect**. We assume that the nominal money supply is held constant.

3.1.1 *Wealth Effect*

The **wealth effect** is based on the concept of purchasing power of nominal wealth, including nominal value of the money held by consumers, physically or in a bank account. Nominal wealth does not change: One British pound is always worth one British pound. Similarly, one euro is worth one euro, and one yen is worth one yen.

The real value, however—the value of money in terms of goods and services—is *not* fixed. It fluctuates with the prices of goods and services. Assume that our favorite loaf of bread costs one pound (£1 = one loaf of bread). If we have £10 in a pocket, then we have money that is equivalent to 10 loaves of bread. The real money holdings are therefore 10 loaves of bread.

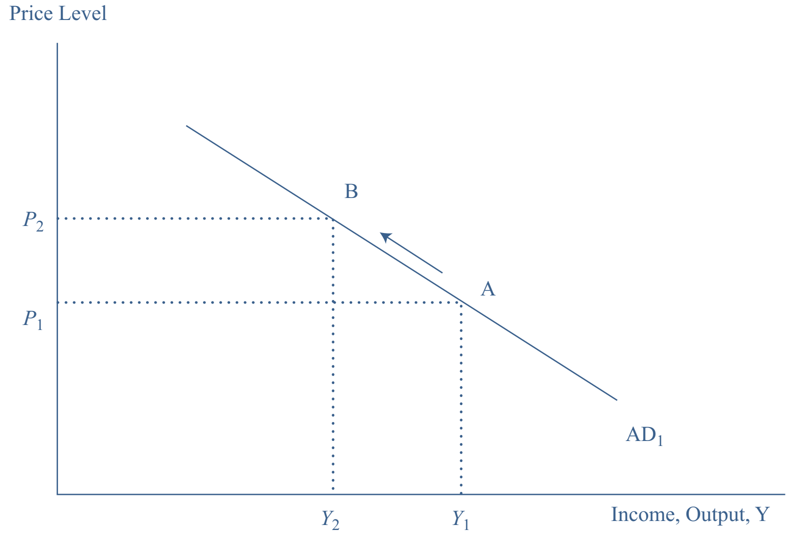
$$\frac{£10}{\left(\frac{£1}{\text{loaf}}\right)} = 10 \text{ loaves}$$

If the price of a loaf of bread increases to £2 (£2 = one loaf of bread), the same £10 is now worth only five loaves of bread. Our real money holdings, and therefore our real wealth, have decreased. Conversely, if the price of a loaf of bread decreases to one-half pound (£0.50 = one loaf of bread), the same £10 is now worth 20 loaves of bread. Our real money holdings, and therefore our real wealth, have increased.

The effect of the change in the price level and the resulting change in the real value of money holdings give the wealth effect its name, because as the real value of our money holdings changes, so does our real wealth.

The wealth effect is one reason that the aggregate demand curve is downward sloping. An increase in the price level decreases the quantity of goods and services that can be purchased with the fixed quantity of nominal wealth—consumers are less wealthy (in real terms) and therefore demand fewer goods and services. Exhibit 11 illustrates this dynamic as the movement from Point A to Point B. An increase in the price level from P_1 to P_2 reduces income from Y_1 to Y_2. Conversely, a decrease in the price level increases the quantity of goods and services that can be purchased with the fixed quantity of money available—consumers are wealthier (in real terms) and therefore demand more goods and services.

Exhibit 11 The Aggregate Demand Curve



3.1.2 *The Interest Rate Effect*

When the price level changes, the demand for money also changes. For example, if the price for your lunch today is 10 euros (€10 = lunch), then we need to have 10 euros in our pocket to eat lunch. Now, assume that the price of lunch doubles to 20 euros (€20 = lunch). If we want to eat lunch, we need to get more money—our demand for money increases. With a fixed supply of money, the price of money increases. Because the price of money is the interest rate, then as the price level increases, the interest rate increases.

The converse is also true. If the price of lunch decreases to five euros (€5 = lunch), we don't need to carry around as much money as before – our demand for money decreases. With a fixed supply of money, the price of money decreases. Therefore, as the price level decreases, the interest rate decreases.

Why not just keep extra "lunches" (money) in our pocket? Because holding extra money in such liquid form means that we are giving up the interest that we could be earning by lending the money out to others (*e.g.* through purchasing a government or corporate bond or just leaving the money in a bank account and letting the bank do the lending for us). The interest we would earn is the opportunity cost of keeping money in our pocket. So, when the price level decreases, we will want to keep more of our money in interest-bearing assets and our demand for money decreases.

Changes in the interest rate affect the quantity of goods and services demanded. If interest rates increase, businesses invest less because their borrowing costs increase. This shift leads to fewer profitable projects to invest in and also negatively affects the demand for commercial real estate. In addition, higher interest rates lead to lower consumption—especially for large purchases such as automobiles or residential real estate, which are usually purchased with loans. If interest rates decrease, then businesses have more profitable projects to invest in, and consumers would be expected to spend more.

This leads us to the **interest rate effect**. A higher price level creates greater demand for money, which raises the interest rate. The higher interest rate decreases demand for investment and consumption expenditures, which leads to less demand for goods and services. Conversely, a lower price level leads to a lower demand for money, which leads to a lower interest rate. The lower interest rate increases demand for investment and consumption expenditures, which leads to more demand for goods and services.

3.1.3 *The Real Exchange Rate Effect*

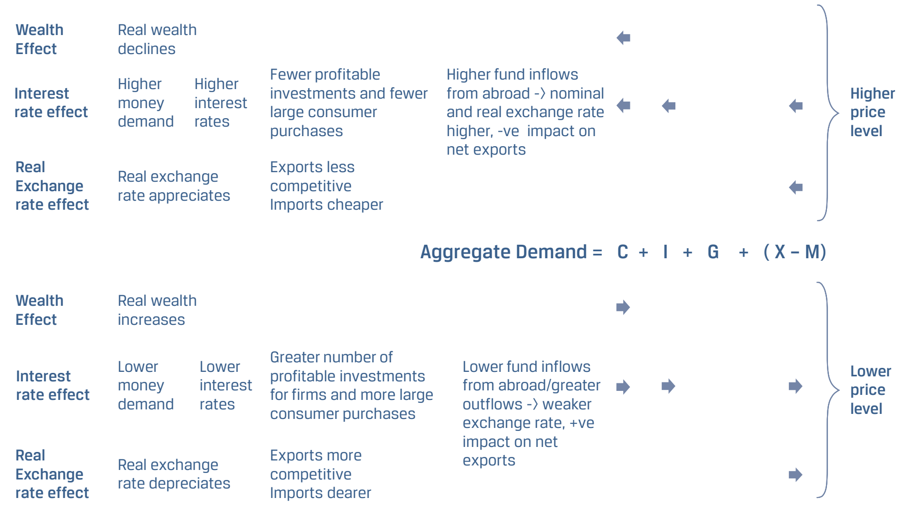
An increase in the domestic price level causes appreciation of the real exchange rate. Real exchange rates will be explained later, but for the purpose of this discussion, it is important to understand that the real exchange rate is the observed, quoted nominal exchange rate adjusted for price level. So a higher price level affects the real exchange rate and makes domestic goods more expensive in other countries, reducing exports. It also makes non-domestic goods less expensive domestically, increasing imports. The result is lower demand for domestic goods and services. Conversely, a decrease in the domestic price level (assuming the price level abroad remains unchanged) leads to a depreciation of the real exchange rate. This decrease in the real exchange rate makes domestic goods less expensive in other countries, increasing exports, and non-domestic goods more expensive domestically, decreasing imports. The result is higher demand for domestic goods and services. This dynamic gives us the **real exchange rate effect**.

An additional channel affecting exchange rates involves the interest rate effect. When interest rates increase (because of a higher price level), non-domestic investors increase their demand for the domestic currency in the foreign exchange market because they want to earn the higher return on their savings. This increased demand causes the domestic currency to appreciate, which in turn increases the real exchange rate.

When interest rates decrease (because of a lower price level), a similar but opposite effect occurs. Domestic savers seek higher returns in non-domestic markets. This increased supply of the domestic currency in the foreign exchange market causes the domestic currency to depreciate, which in turn decreases the real exchange rate.

The interest rate effect just described is an additional channel through which a higher price level leads to a higher real exchange rate. The two effects are complementary, operate in the same direction and magnify the impact of the change in price level on aggregate demand. Exhibit 12 provides a summary.

Exhibit 12 Deriving Aggregate Demand Curve: Summary of Key Effects


Wealth Effect
Real wealth declines
Interest rate effect
Higher money demand
Higher interest rates
Fewer profitable investments and fewer large consumer purchases
Higher fund inflows from abroad -) nominal and real exchange rate higher, -ve impact on net exports
Real Exchange rate effect
Real exchange rate appreciates
Exports less competitive Imports cheaper
Higher price level
Aggregate Demand = C + I + G + (X – M)
Wealth Effect
Real wealth increases
Interest rate effect
Lower money demand
Lower interest rates
Greater number of profitable investments for firms and more large consumer purchases
Lower fund inflows from abroad/greater outflows -) weaker exchange rate, +ve impact on net exports
Real Exchange rate effect
Real exchange rate depreciates
Exports more competitive Imports dearer
Lower price level

EXAMPLE 5

Slope of the Aggregate Demand Curve

Considering the movement along the aggregate demand curve, which of the following statements best explains the interest rate effect when the price level decreases?

A Decrease in money demand, with a fixed money supply, leads to lower interest rates. Lower interest rates induce businesses to invest more because there is now a greater number of profitable projects. Consumers also spend more when interest rates are lower. Lower interest rates also weaken the exchange rate, boosting net exports. These effects cause output to be higher at a lower price level.

B Decrease in money demand, with a fixed money supply, leads to lower interest rates. Lower interest rates induce businesses to invest less because there are now fewer profitable projects. Lower interest rates also weaken the exchange rate, boosting net exports.

C Decrease in money demand, with a fixed money supply, leads to higher interest rates, which induce businesses to invest less and consumers to save more. Higher interest rates cause the exchange rate to appreciate, boosting net exports.

Solution:

A is correct. B is incorrect because the statement suggests that lower interest rates lead to less business investment. C is incorrect because the statement suggests that a lower price level leads to higher interest rates.

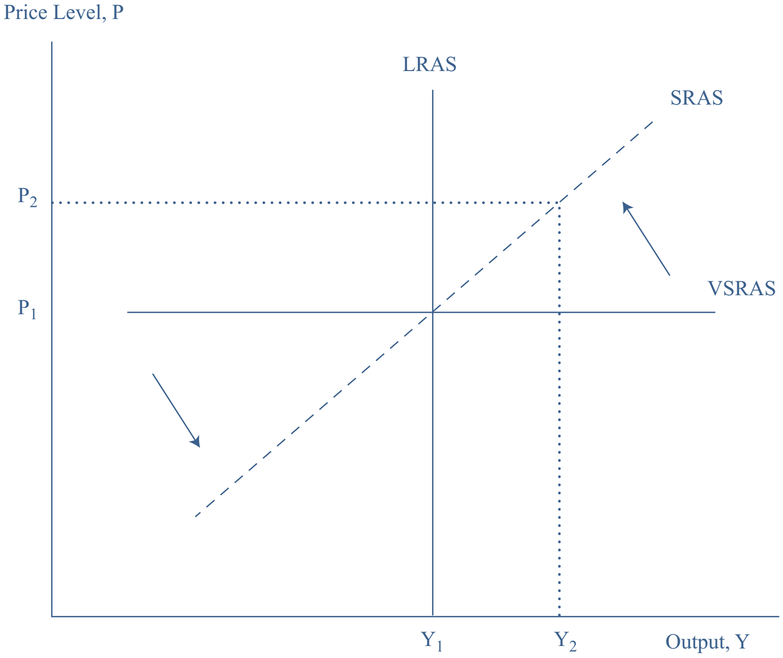
3.2 Aggregate Supply

g explain the aggregate supply curve in the short run and long run

Aggregate demand tells us only the relationship between the price level and the amount of output demanded at those prices. To understand what price and output level will prevail in the economy, we need to add aggregate supply, the amount of output producers are willing to provide at various prices. The **aggregate supply curve** (AS curve) represents the level of domestic output that companies will produce at each price level. Unlike the demand side, we must distinguish between the short- and long-run AS curves, which differ with respect to how wages and other input prices respond to changes in final output prices. "Long run" and "short run" are relative terms and are necessarily imprecise with respect to calendar time. The "long run" is long enough that wages, prices, and expectations can adjust but not long enough that physical capital is a variable input. Capital and the available technology to use that capital remain fixed. This condition implies a period of at least a few years and perhaps a decade. The truly long run, in which even the capital stock is variable, may be thought of as covering multiple decades. Consideration of the very long run is postponed to our subsequent discussion of economic growth.

In the very short run, perhaps a few months or quarters, companies will increase or decrease output to some degree without changing prices. Exhibit 13 illustrates this dynamic with the horizontal line labeled VSRAS. If demand is somewhat stronger than expected, companies earn higher profit by increasing output as long as they can cover their variable costs. So they will run their plant and equipment more intensively, demand more effort from their salaried employees, and increase the hours of employees who are paid on the basis of hours worked. If demand is somewhat weaker than projected, companies can run their plants less intensively, cut labor hours, and use staff to perform maintenance and carry out efficiency-enhancing projects that are often postponed during busier periods.

Exhibit 13 Aggregate Supply Curve



Over somewhat longer periods, the AS curve is upward sloping because more costs become variable. This dynamic is represented by the short-run aggregate supply (SRAS) curve in Exhibit 13. In most businesses, wages are adjusted once a year, but for companies with union contracts, several years may pass before the contracts expire. The prices for raw materials and other inputs may also be established under long-term contracts. Hence, wages and other input costs are relatively inflexible in the short run and do not fully adjust to changes in output prices. As the price level rises, most companies enjoy higher profit margins and hence expand production. In Exhibit 13, when prices move from P_1 to P_2, the quantity of aggregate output supplied increases from Y_1 to Y_2. Conversely, a reduction in the price level squeezes profit margins and causes companies to reduce production.

Over time, however, wages and other input prices tend to "catch up" with the prices of final goods and services. In other words, wages and prices that are inflexible or slow to adjust in the short run adjust to changes in the price level over the long run. Thus, over the long run, when the aggregate price level changes, wages and other input prices change proportionately so that the higher aggregate price level has no effect on aggregate supply. This dynamic is illustrated by the vertical long-run aggregate supply (LRAS) curve in Exhibit 13. As prices move from P_1 to P_2, the quantity of output supplied remains at Y_1 in the long run. The only change that occurs is that prices shift to a higher level (from P_1 to P_2).

The position of the LRAS curve is determined by the potential output of the economy. The amount of output produced depends on the fixed amount of capital and labor and the available technology. This classical model of aggregate supply can be expressed as

$$Y = \boldsymbol{F}\left(\overline{K}, \overline{L}\right) = \overline{Y}$$

where $\overline{K}$ is the fixed amount of capital and $\overline{L}$ is the available labor supply. The stock of capital is assumed to incorporate the existing technological base. Note that investment (I) reflects replacement of worn-out capital plus the change in capital, ΔK. Over short periods, net investment is assumed to have a negligible effect on aggregate supply. The available labor supply is also held constant, and workers are assumed to have a given set of skills. The long-run equilibrium level of output, Y_1 in Exhibit 13, is referred to as the *full employment*, or *natural*, level of output. At this level of output, the economy's resources are deemed to be fully employed, and (labor) *unemployment is at its natural rate*. This concept of a natural rate of unemployment assumes the macroeconomy is currently operating at an efficient and unconstrained level of production. Companies have enough spare capacity to avoid bottlenecks, and there is a modest, stable pool of unemployed workers (job seekers equal job vacancies) looking for and transitioning into new jobs.

3.3 Shifts in the Aggregate Demand Curve

h explain causes of movements along and shifts in aggregate demand and supply curves

i describe how fluctuations in aggregate demand and aggregate supply cause short-run changes in the economy and the business cycle

In the next two sections, the aggregate demand (AD) and aggregate supply (AS) models are used to address three critical macroeconomic questions:

1. What causes an economy to expand or contract?

2 What causes inflation and changes in the level of unemployment?

3 What determines an economy's rate of sustainable growth, and how can it be measured?

Before addressing these questions, we need to distinguish between 1) the long-run growth rate of real GDP and 2) short-run fluctuations in real GDP around this long-run trend.

The business cycle is a direct result of short-term fluctuations of real GDP. It consists of periods of economic expansion and contraction. In an expansion, real GDP is increasing, the unemployment rate is declining, and capacity utilization is rising. In a contraction, real GDP is decreasing, the unemployment rate is rising, and capacity utilization is declining. Shifts in the AD and AS curves determine the short-run changes in the economy associated with the business cycle. In addition, the AD–AS model provides a framework for estimating the sustainable growth rate of an economy, which is addressed in Section 4.

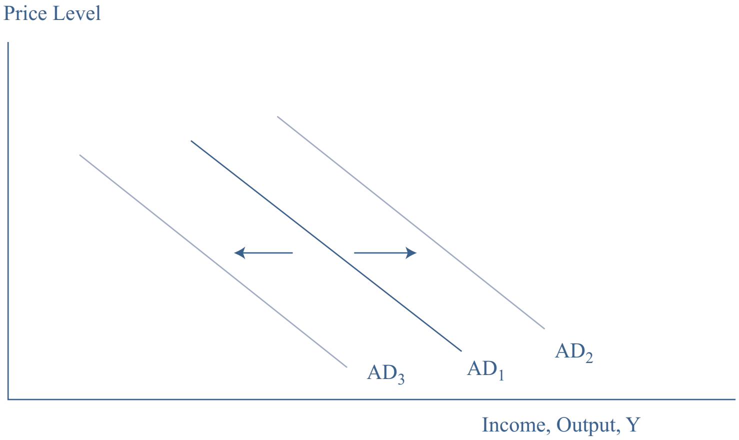
From an asset allocation perspective, it is important to determine the current phase of the business cycle as well as how fast the economy is growing relative to its sustainable growth rate. The expected rate of return on equities and fixed-income securities, for example, depends on estimates of the growth rate of GDP and inflation. For equities, GDP growth is the primary determinant of aggregate corporate profits. For fixed-income securities, the expected rate of inflation determines the spread between real and nominal rates of return. In order to use the AD and AS model to analyze the economy and to make investment decisions, we need to first understand what factors cause the curves to shift.

3.3.1 *Shifts in Aggregate Demand*

In addition to price, factors that influence the level of spending by households, companies, governments, and people in other countries (i.e., the aggregate level of expenditures) will cause the AD curve to shift. A shift to the right represents an increase in aggregate demand at any price level. Exhibit 14 shows this dynamic as a shift from AD_1 to AD_2. A shift to the left represents a decrease in aggregate demand at any price level. This dynamic is indicated by a move from AD_1 to AD_3. Key factors that directly or indirectly influence the level of aggregate expenditures and cause the aggregate demand curve to shift include changes in the following:

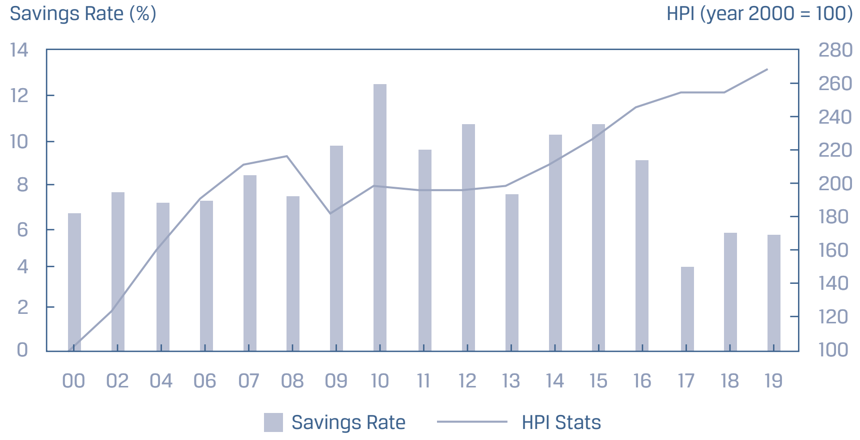
- household wealth,
- consumer and business expectations,
- capacity utilization,
- monetary policy,
- the exchange rate,
- growth in the global economy, and
- fiscal policy (government spending and taxes).

Exhibit 14 Shifts in the Aggregate Demand Curve



Household Wealth Household wealth includes the value of both financial assets (e.g., cash, savings accounts, investment securities, and pensions) and real assets (e.g., real estate). The primary reason households save a portion of their current income is to accumulate wealth for consumption in the future. The proportion of disposable income that households save depends partly on the value of the financial and real assets that they have already accumulated. If these assets increase in value, households will tend to save less and spend a greater proportion of their income because they will still be able to meet their wealth accumulation goals. As a result, an increase in household wealth increases consumer spending and shifts the aggregate demand curve to the right. In contrast, a decline in wealth will reduce consumer spending and shift the AD curve to the left. This dynamic, often referred to as the **wealth effect**, provides one explanation for how changes in equity prices affect economic activity. Higher equity prices increase household wealth, which increases consumer spending and reduces the amount saved out of current income. Economic studies estimate that an increase or decrease in wealth in developed countries increases or decreases annual consumer spending by 3%–7% of the change in wealth (Case, Quigley, and Shiller 2005). A smaller but still statistically significant wealth effect has been found in a number of emerging markets (developing countries) (Funke 2004).

Exhibit 15 Housing Prices and the Saving Rate in the United Kingdom, 2000–2019



Source: Office of National Statistics, United Kingdom.

EXAMPLE 6

The Wealth Effect on Saving and Consumption

The importance of the wealth effect on consumption, and its relationship to housing prices, was evident in the sharp global economic downturn. A major factor associated with the economic downturn was the sharp fall in housing prices, especially in countries that experienced a housing boom earlier in the decade, such as the United States, the United Kingdom, Spain, and Ireland. In each of these countries, consumers reduced spending sharply and raised the level of saving in response to the decline in wealth. Do the data in Exhibit 15 provide support for the wealth effect?

Solution:

Housing prices in the United Kingdom rose by nearly 111% [(211.1 – 100)/100] between 2000 and 2007. During this period, the saving rate averaged 7.2% of income up until 2008. As predicted, when housing prices fell by 11% between 2008 and 2009, the saving rate rose dramatically from 7.2% in 2008 to 12.4% in 2010. Of course, the decline in housing prices was not the only factor contributing to the increase in the saving rate. Stock prices also declined during this period, further reducing wealth in the United Kingdom, and the slowdown raised uncertainty over future jobs and income. Then, as the economy recovered, house prices rose again from 2013, which led to a (lagged) decline in the saving rate.

Consumer and Business Expectations Psychology has an important impact on consumer and business spending. When consumers are confident about their future income and the stability/safety of their jobs, they tend to spend a higher portion of their disposable income. This dynamic shifts the AD curve to the right. Consumer spending declines and the AD curve shifts to the left when consumers become less confident. Similarly, when businesses are optimistic about their future growth and profitability, they spend (invest) more on capital projects, which also shifts the AD curve to the right.

Capacity Utilization Capacity utilization is a measure of how fully an economy's production capacity is being used. Companies with excess capacity have little incentive to invest in new property, plant, and equipment. In contrast, when companies are operating at or near full capacity, they will need to increase investment spending in order to expand production. Data from the OECD and the US Federal Reserve indicate that when aggregate capacity utilization reaches 82%–85%, production blockages arise, prompting companies to increase their level of investment spending. This dynamic shifts the AD curve to the right.

Fiscal Policy **Fiscal policy** is the use of taxes and government spending to affect the level of aggregate expenditures. An increase in government spending, one of the direct components of AD, shifts the AD curve to the right, whereas a decrease in government spending shifts the AD curve to the left. Taxes affect GDP indirectly through their effect on consumer spending and business investment. Lower taxes will increase the proportion of personal income and corporate pre-tax profits that consumers and businesses have available to spend and will shift the AD curve to the right. In contrast, higher taxes will shift the AD curve to the left.

Monetary Policy *Money* is generally defined as currency in circulation plus deposits at commercial banks. **Monetary policy** refers to action taken by a nation's central bank to affect aggregate output and prices through changes in bank reserves, reserve requirements, or its target interest rate.

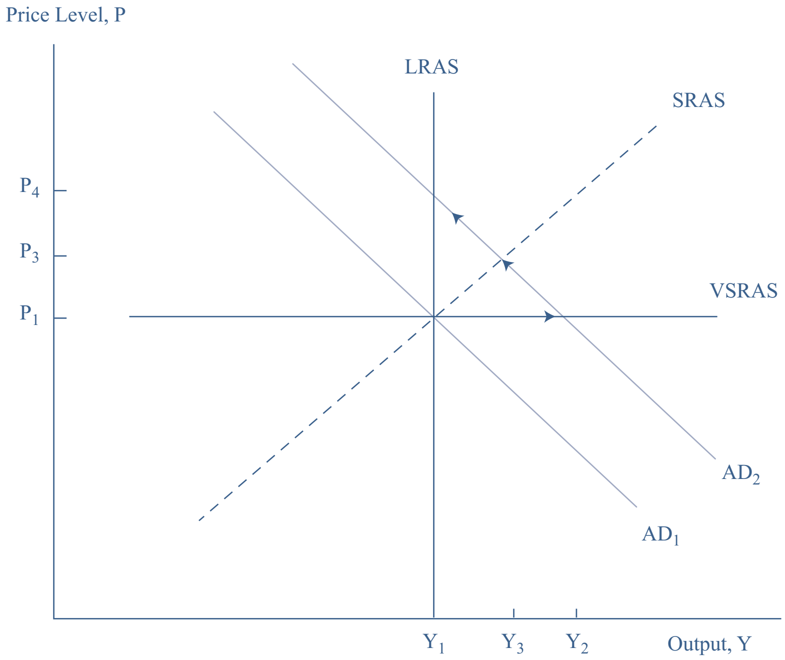
Most countries have fractional reserve banking systems in which each bank must hold reserves (vault cash plus deposits at the central bank) at least equal to the required reserve ratio times its customer deposits. Banks with excess reserves can lend them to banks that need reserves to meet their reserve requirements. The central bank can increase the money supply by 1) buying securities from banks, 2) lowering the required reserve ratio if such a tool is available, and/or 3) reducing its target for the interest rate at which banks borrow and lend reserves among themselves. In each case, the opposite action would decrease the money supply.

When the central bank buys securities from banks in an open-market operation, it pays for them with a corresponding increase in bank reserves. This dynamic increases the amount of deposits banks can accept from their customers—that is, the money supply. Similarly, cutting the required reserve ratio increases the level of deposits (i.e., money) consistent with a given level of reserves in the system. If the central bank chooses to target an interbank lending rate, as the Federal Reserve targets the federal funds rate in the United States, then it must add or drain reserves via open-market operations to maintain the target interest rate. If it raises (lowers) its target interest rate, it will have to drain (add) reserves in order to make reserves more (less) expensive in the interbank market. Thus, open-market operations and interest rate targeting are very closely related. The main distinction is whether the central bank chooses to target a level of reserves and let the market determine the interest rate or chooses to target the interest rate and let the market (i.e., banks) determine the level of reserves they desire to hold at that rate.

An increase in the money supply shifts the AD curve to the right so that each price level corresponds to a higher level of income and expenditure. There are various channels through which the additional expenditures may be induced. For example, the interest rate reduction required to induce investors to hold the additional money balances will encourage companies to invest more and households to borrow to purchase durable goods, such as cars. In addition, banks may facilitate greater expenditure by raising credit limits and loosening credit standards. Conversely, a reduction in the money supply shifts the AD curve to the left.

Exhibit 16 illustrates the short-run and long-run effect of expansionary monetary policy. Suppose the central bank expands the money supply in an attempt to stimulate demand when the economy is already in long-run equilibrium. The expansionary policy will shift the AD curve to the right, from AD_1 to AD_2. In the very short run, output will expand from Y_1 to Y_2 without an increase in the price level. After operating at higher-than-normal production rates for a few months or quarters, companies will begin to push for price increases and input prices will begin to rise as well. The aggregate supply curve will steepen, and prices will increase to P_3 while output declines to Y_3. As input prices become more flexible, the AS curve will steepen until, in the long run, it is vertical and output has returned to the long-run natural level, Y_1, with prices rising to P_4. Thus, expanding the money supply increases output in the short run, but in the long run it affects only the price level.

Exhibit 16 Short-Run and Long-Run Effects of Monetary Expansion



Exchange Rate Changes in the exchange rate affect the price of exports and imports and thus aggregate demand. For example, a lower euro relative to other currencies makes European exports cheaper in world markets and non-European products sold in Europe (European imports) more expensive. Therefore, a lower euro should cause European exports to increase and imports to decline, leading the AD curve to shift to the right. Conversely, a stronger euro reduces exports and raises imports, and the AD curve shifts to the left.

Growth in the Global Economy International trade is what links countries together and creates a global economy. Faster economic growth in non-domestic markets encourages people in other countries to buy more products from domestic producers and increases exports. For example, rapid GDP growth in ASEAN member countries has increased their demand for products from other nations. Japan has benefited from this rapid growth because it has exported more products to ASEAN countries. In terms of the AD and AS model, the AD curve for Japan has shifted to the right because of increased demand for Japanese products in ASEAN countries, resulting in higher exports. A decline in the growth rates of ASEAN members' economies would have a negative effect on the Japanese economy because exports would be lower. This change would cause the Japanese AD curve to shift to the left.

What happens to interest rates when the AD curve shifts? In the case of an increase in the money supply, the interest rate declines at each price level because the increase in income (Y) increases saving, and rates must decline to induce a corresponding increase in investment spending (I). In each of the other cases considered above, a rightward shift in the AD curve will increase the interest rate at each price level. With the real money supply held constant, the interest rate must rise as income increases. The increase in the interest rate reduces the demand for money at each level of expenditure/income and, therefore, allows expenditure/income to increase without an increase in the money supply. In terms of the quantity theory of money equation, this shift corresponds to a higher velocity of money, V.

Exhibit 17 summarizes the main factors that shift the AD curve. In each case, the impact of the factor is considered in isolation. In practice, however, various factors may be at work simultaneously and there may be interaction among them. This is especially true with regard to expectational factors—consumer and business confidence—which are likely to be influenced by other developments.

Exhibit 17 Impact of Factors Shifting Aggregate Demand

An Increase in the Following Factors	Shifts the AD Curve	Reason
Stock prices	Rightward: Increase in AD	Higher consumption
Housing prices	Rightward: Increase in AD	Higher consumption
Consumer confidence	Rightward: Increase in AD	Higher consumption
Business confidence	Rightward: Increase in AD	Higher investment
Capacity utilization	Rightward: Increase in AD	Higher investment
Government spending	Rightward: Increase in AD	Government spending a component of AD
Taxes	Leftward: Decrease in AD	Lower consumption and investment
Bank reserves	Rightward: Increase in AD	Lower interest rate, higher investment and possibly higher consumption
Exchange rate (foreign currency per unit domestic currency)	Leftward: Decrease in AD	Lower exports and higher imports
Global growth	Rightward: Increase in AD	Higher exports

EXAMPLE 7

Shifts in Aggregate Demand

Francois Ubert is a portfolio manager with EuroWorld, a French investment management firm. Ubert is considering increasing his clients' portfolio exposure to Brazilian equities. Before doing so, he asks you to prepare a report on the following recent economic events in Brazil and to summarize the effect of each event on the Brazilian economy and on Brazilian equity and fixed-income securities.

1 The Brazilian central bank reduced bank reserves, resulting in a lower money supply.

2 The capacity utilization rate in Brazil is currently estimated to be 86.4%, a 2.7% increase from the previous year.

3 Corporate profits reported by Brazilian companies increased by 30% over last year's levels, and corporations have revised their forecasts of future profitability upward.

4 The government recently announced that it plans to start construction on a number of hydroelectric projects to increase renewable energy sources.

5 Forecasts by private sector economists project that the European economy will enter a sharp slowdown and decline in the next year.

Solution to 1:

This monetary policy action is designed to reduce consumption and business investment spending. The reduction in real money balances will increase interest rates and discourage lending within the banking system. Higher interest rates and tighter credit will reduce both investment and consumption expenditures and shift the AD curve to the left. The prices of fixed-income securities will fall because of the rise in interest rates. The reduction in aggregate output should lower corporate profits, and it is likely that equity prices will also fall.

Solution to 2:

Capacity utilization is a key factor determining the level of investment spending. A current utilization rate of more than 86% and an increase from the previous year indicate a growing lack of spare capacity in the Brazilian economy. As a result, businesses will probably increase their level of capital spending. This change will increase AD and shift the AD curve to the right. Higher economic activity (income/output) will cause upward pressure on interest rates and may have a negative effect on fixed-income securities. Higher income/output should increase corporate profits and is likely to have a positive effect on equity securities.

Solution to 3:

Expected corporate profits are an important determinant of the level of investment spending. The large increase in expected profits will raise the level of investment spending and increase aggregate demand. This change will shift the AD curve to the right. The increase in corporate profits and the resulting increase in economic output should have a positive effect on equities. The increase in output will put upward pressure on interest rates and downward pressure on the prices of fixed-income securities.

Solution to 4:

Fiscal policy uses government spending to influence the level and growth rate of economic activity. The announcement indicates an increase in government spending, which is a direct component of AD. Therefore, higher spending on the projects will increase AD and shift the AD curve to the right. The increase in output and expenditure should be positive for equities. But it will be negative for existing fixed-income investments because higher interest rates will be required to induce investors to buy and hold the government debt issued to fund the new projects.

Solution to 5:

A sharp slowdown in Europe will decrease the demand for Brazilian exports by European households and businesses and shift the AD curve to the left. The resulting decline in income and downward pressure on prices will be positive for fixed-income securities but negative for equities.

3.3.2 *Shifts in the Short-Run Aggregate Supply Curve*

g explain the aggregate supply curve in the short run and long run

h explain causes of movements along and shifts in aggregate demand and supply curves

Factors that change the cost of production or expected profit margins will cause the SRAS curve to shift. These factors include changes in

- nominal wages,
- input prices, including the price of natural resources,
- expectations about future output prices and the overall price level,
- business taxes and subsidies, and
- the exchange rate.

In addition, factors that shift the long-run AS curve (discussed elsewhere) will also shift the SRAS curve by a corresponding amount because the SRAS and LRAS reflect the same underlying resources and technology. As the economy's resources and technology change, the full employment (or natural) level of output changes, and both the LRAS and SRAS shift accordingly.

Change in Nominal Wages Changes in nominal wages shift the short-run AS curve because wages are often the largest component of a company's costs. An increase in nominal wages raises production costs, resulting in a decrease in AS and a leftward shift in the SRAS curve. Lower wages shift the AS curve to the right. It is important to note that changes in nominal wages have no effect on the LRAS curve.

A better way to measure the impact of labor costs on the AS curve is to measure the change in unit labor cost. We define the change in unit labor cost as

$$\begin{aligned}\%\text{ Change in unit labor cost} = {} & \%\text{ Change in nominal wages} \\ & - \%\text{ Change in productivity}\end{aligned}$$

EXAMPLE 8

Unit Labor Cost and Short-Run Aggregate Supply

Suppose Finnish workers are paid €20 an hour and are able to produce 100 cell phones in an hour. The labor cost per cell phone is €0.20 (€20 divided by 100 units). If the wages per hour for Finnish workers rise by 10% from €20 to €22 and they are able to raise their productivity by 10%, what is the impact on unit labor cost and the short-run aggregate supply curve?

Solution:

The workers can now produce 110 cell phones per hour, and unit labor cost will not change (22/110 = 0.20). In this case, the SRAS curve will remain in its original position. If wages had increased by 20% instead of 10%, then unit labor cost would have increased and the SRAS would shift to the left. Conversely, if the wage increase were only 5%, then unit labor cost would have decreased and the SRAS would shift to the right.

Change in Input Prices The price of raw materials is an important component of cost for many businesses. Lower input prices reduce the cost of production, which, in turn, makes companies willing to produce more at any output price. This dynamic is reflected in a rightward shift of the SRAS curve. Conversely, higher input prices increase production costs, which, in turn, causes companies to reduce production at any output price. This dynamic shifts the SRAS curve to the left. During the 1970s, high oil prices caused the SRAS curve in most countries to shift to the left. In contrast, in the mid-1980s, declining oil prices lowered the cost of production and shifted the SRAS curve in most countries to the right. Oil prices currently have a smaller effect on the global economy than in the 1970s and 1980s because most countries have reduced their reliance on oil and improved their energy efficiency so that they now use less energy per unit of GDP.

Change in Expectations about Future Prices The impact of expected future prices on current output decisions is not as straightforward as it might seem. First, each company is primarily concerned about the price of its own output rather than the general price level. The latter may be more reflective of its costs. If it expects its own output price to rise (fall) relative to the general price level, then it may increase (decrease) production in response to the perceived change in its profit margin. As more and more companies become optimistic (pessimistic) about their ability to raise the relative price of their product, the SRAS will shift to the right (left). In the aggregate, of course, companies can neither raise nor lower their prices relative to the general price level. Hence, shifts in the SRAS driven by such price expectations are likely to be modest and temporary. Second, considering future prices introduces a temporal aspect into decision making. If the future price level is expected to be higher, companies may decide to produce more today in order to expand inventory available for future sale. But they will only do so if the cost of carrying inventory (financing, storage, and spoilage) is less than they expect to save on production costs by producing more today and less in the future. Conversely, they may cut current production and sell out of existing inventory if they expect future prices (and costs) to be lower.

The upshot is that expectations of higher (lower) future prices are likely to shift the SRAS curve to the right (left), but the effect may be modest and/or temporary.

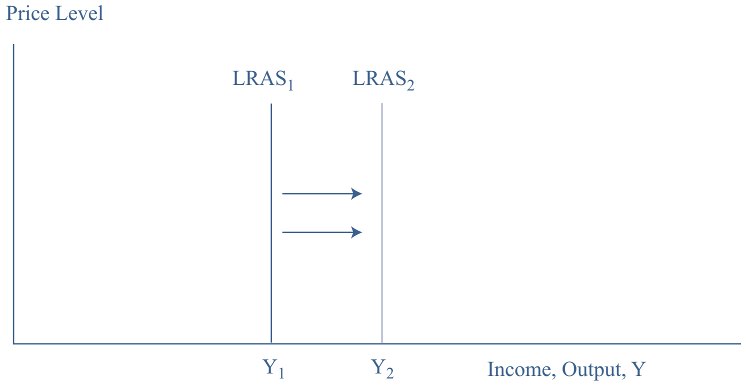
Change in Business Taxes and Subsidies Higher business taxes increase production costs per unit and shift the short-run AS curve to the left. Business subsidies are a payment from the government to the producer. Subsidies for businesses lower their production costs and shift the SRAS curve to the right.

Change in the Exchange Rate Many countries import raw materials, including energy and intermediate goods. As a result, changes in the exchange rate can affect the cost of production and, therefore, aggregate supply. A higher yen relative to the euro, for instance, would lower the cost of raw materials and intermediate goods imported to Japan from Europe. This, in turn, will lower the production costs of Japanese producers and shift the AS curve in Japan to the right. A lower yen would have the opposite effect.

3.3.3 *Shifts in the Long-Run Aggregate Supply Curve*

As discussed, the position of the LRAS curve is determined by the potential output of the economy. **Potential GDP** measures the productive capacity of the economy and is the level of real GDP that can be produced at full employment. Potential GDP is not a static concept but can increase each year at a steady rate as the economy's resource capacity grows. Therefore, any factor increasing the resource base of an economy causes the LRAS curve to shift as shown in Exhibit 18.

Exhibit 18 Shift in the Long-Run Aggregate Supply (LRAS) Curve



These factors include changes in

- supply of labor and quality of labor forces (human capital),
- supply of natural resources,
- supply of physical capital, and
- productivity and technology.

Supply of Labor The larger the supply of labor, the more output the economy can produce. The labor supply depends on growth in the population, the labor force participation rate (the percentage of the population working or looking for work), and net immigration. The determinants of the labor supply are discussed in more detail later in our coverage. Increases in the labor supply shift the LRAS curve to the right. Decreases shift the curve to the left.

Supply of Natural Resources Natural resources are essential inputs to the production process and include everything from available land to oil to water. Increased availability of natural resources shifts the LRAS curve to the right.

Supply of Physical Capital Investment in new property, plant, equipment, and software is an essential ingredient for growth. An increase in the stock of physical capital will increase the capacity of the economy to produce goods and services. Simply put, if workers are provided with more and better equipment to use, they should be able to produce more output than they could with the older equipment. Thus, strong growth in business investment, which increases the supply of physical capital, shifts the LRAS curve to the right.

Supply of Human Capital Another way to raise the productive capacity of a country is to increase human capital—the quality of the labor force—through training, skills development, and education. Improvement in the quality of the labor force shifts the LRAS curve to the right.

Labor Productivity and Technology Another important factor affecting the productive capacity of an economy is how efficient labor is in transforming inputs into final goods and services. **Productivity** measures the efficiency of labor and is the amount of output produced by workers during a given period—for example, output per hour worked. An increase in productivity decreases labor cost, improves profitability, and results in higher output. Two of the main drivers of labor productivity—physical capital per worker and the quality of the workforce—have been discussed already. The third key determinant of productivity is technology. Advances in technology shift the LRAS curve to the right.

EXAMPLE 9

Unit Labor Cost and Long-Run Aggregate Supply

Finnish workers are paid €20 per hour and are able to produce 100 cell phones in an hour. If workers develop a new technique for assembly and are able to produce 200 cell phones per hour, what is the effect on the long-run aggregate supply curve?

Solution:

Labor cost per unit will decline to €0.10 (20/200 = €0.10 per cell phone). As a result, profit per unit will rise and companies will have an incentive to increase production. Thus, the LRAS curve will shift to the right.

Exhibit 19 summarizes the factors shifting the AS curve. Rightward shifts in the SRAS or LRAS curves are defined as an increase in supply, while leftward shifts represent a decrease in supply.

Exhibit 19 Impact of Factors Shifting Aggregate Supply

An Increase in	Shifts SRAS	Shifts LRAS	Reason
Supply of labor	Rightward	Rightward	Increases resource base
Supply of natural resources	Rightward	Rightward	Increases resource base
Supply of human capital	Rightward	Rightward	Increases resource base
Supply of physical capital	Rightward	Rightward	Increases resource base
Productivity and technology	Rightward	Rightward	Improves efficiency of inputs
Nominal wages	Leftward	No effect	Increases labor cost
Input prices (e.g., energy)	Leftward	No effect	Increases cost of production
Expectation of future prices	Rightward	No effect	Anticipation of higher costs and/or perception of improved pricing power
Business taxes	Leftward	No effect	Increases cost of production
Subsidy	Rightward	No effect	Lowers cost of production
Exchange rate	Rightward	No effect	Lowers cost of production

As with our summary of factors that shift the AD curve, Exhibit 19 considers each of the factors affecting aggregate supply in isolation. In practice, various factors will be at work simultaneously, and there may be interaction among them. This is especially important with respect to interaction between factors listed as affecting only SRAS and those that also affect LRAS.

For example, consider an increase in the cost of natural resource inputs (e.g., energy). This change shifts the SRAS curve to the left, but according to Exhibit 19, it has no effect on LRAS. This presumes that the relative prices of the factors of production have not changed permanently. If a permanent change has occurred, companies will be forced to conserve on the now more expensive input and will be unable to produce as efficiently. The LRAS curve would, therefore, shift to the left, just as it would if the available supply of natural resources had declined relative to the supply of other inputs. Indeed, that is the most likely cause of a permanent change in relative input prices.

EXAMPLE 10

Shifts in Aggregate Supply

In this hypothetical scenario, John Donovan is a portfolio manager for a global mutual fund. Currently, his fund has 10% of its assets invested in Chinese equities. He is considering increasing the fund's allocation to the Chinese equity market. His decision will be based on an analysis of the following economic developments and their effect on the Chinese economy and equity market. What is the effect on SRAS and LRAS from the following factors?

1. Global oil prices, currently near their longer-run trend at US$75 a barrel, have increased from US$35 a barrel over the last three years because of strong demand from emerging markets.
2. The number of students studying engineering has dramatically increased at Chinese universities during the last decade.
3. Wages for China's workers are rising, leading some multinational companies to consider shifting their investments to Vietnam or Cambodia.
4. Data show that business investment as a share of GDP is more than 40% in China.
5. The People's Bank of China is likely to permit the yuan to appreciate by 10% during the next year.

Solution to 1:

Higher energy prices cause a decrease in short-run AS and shift the SRAS curve to the left. Because oil prices are back to their longer-run trend, the leftward shift in SRAS essentially reverses a previous shift that occurred when oil prices fell to US$35 a barrel, and it is likely that there will be no effect on the LRAS curve. Lower output and profit are likely to have a negative effect on Chinese equity prices.

Solution to 2:

More students studying engineering indicates an improvement in the quality of the labor force—an increase in human capital. As a result, AS increases and the AS curve shifts to the right. Both short-run and long-run curves are affected. Higher output and profits may be expected to have a positive effect on Chinese equity prices.

Solution to 3:

The increase in wages increases labor costs for businesses, causes short-run aggregate supply to decline, and shifts the SRAS curve to the left. Lower output and profit should have a negative effect on Chinese equity prices.

Solution to 4:

The high level of business investment indicates that the capital stock in China is growing at a fast rate. This growth means that workers have more capital to use, which increases their productivity. Thus, AS increases and the AS curve shifts to the right. Both short-run AS and long-run AS are affected. Higher output should have a positive effect on Chinese equity prices.

Solution to 5:

The probable appreciation of the yuan means that the cost of imported raw materials, such as iron ore, coal, and oil, will be lower for Chinese companies. As a result, short-run AS increases and the SRAS curve shifts to the right. The LRAS curve may also shift to the right if the appreciation of the yuan is permanent and global commodity prices do not fully adjust. Higher output and profit should have a positive effect on Chinese equity prices. Note that the stronger yuan will also reduce export demand and shift the AD curve to the left. The combined effect of the AD and AS shifts on output, profit, and equity prices is ambiguous.

The implications of the aforementioned factors for equity investment in China are ambiguous. If the long-run effects dominate, however, then the net effect should be positive. The positive factors—the high level of investment and the growing pool of engineering students—have a lasting effect on output and profit. The negative factors—higher wages and oil prices—should be temporary, because wages will realign with the price level and the increase in oil prices appears to offset a previous temporary decline. The reduction in raw material prices resulting from the stronger currency is positive for output, profit, and equities in the short run and perhaps in the long run as well.

3.4 Equilibrium GDP and Prices

i. describe how fluctuations in aggregate demand and aggregate supply cause short-run changes in the economy and the business cycle

j. distinguish among the following types of macroeconomic equilibria: long-run full employment, short-run recessionary gap, short-run inflationary gap, and short-run stagflation

k. explain how a short-run macroeconomic equilibrium may occur at a level above or below full employment

l. analyze the effect of combined changes in aggregate supply and demand on the economy

Now that we have discussed the components of the AD and AS model, we can combine them to determine the real level of GDP and the price level. Equilibrium occurs where the AD and AS curves intersect. At this point, the quantity of aggregate output demanded (or the level of aggregate expenditures) is equal to the quantity of aggregate output supplied. In Exhibit 20, equilibrium price and GDP occur at P_1 and Y_1. If the price level is above P_1, then the quantity of output supplied exceeds the amount demanded. This situation would result in unsold inventories and would require a reduction in production and in prices. If the price level is below P_1, then the quantity of aggregate output demanded exceeds the quantity of aggregate output supplied. This situation would result in a shortage of goods that would put upward pressure on prices.

It is important to understand that short-run macroeconomic equilibrium may occur at a level above or below full employment. We consider four possible types of macroeconomic equilibrium:

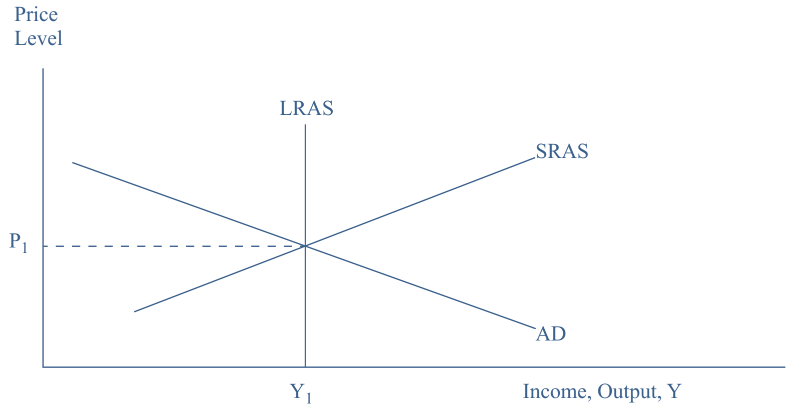
1. Long-run full employment
2. Short-run recessionary gap
3. Short-run inflationary gap
4. Short-run stagflation

From an investment perspective, the performance of asset classes and financial markets will differ in each case as the economy makes the adjustment toward the macroeconomic equilibrium. We look at these differences later in the reading.

3.4.1 *Long-Run Equilibrium*

Exhibit 20 shows the long-run full employment equilibrium for an economy. In this case, equilibrium occurs where the AD curve intersects the SRAS curve at a point on the LRAS curve. Because equilibrium occurs at a point on the LRAS curve, the economy is at potential real GDP. Both labor and capital are fully employed, and everyone who wants a job has one. *In the long run, equilibrium GDP is equal to potential GDP.*

Exhibit 20 Long-Run Macroeconomic Equilibrium

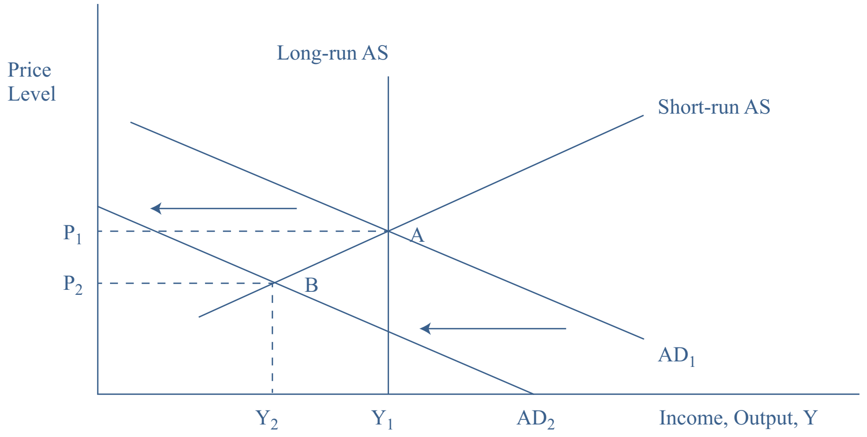


In practice, the level of potential GDP is difficult to measure with precision. Because of fluctuations arising from shifts in the AD and SRAS curves, the economy rarely operates at potential GDP. Thus, potential GDP is not observable from the data on actual GDP. In addition, potential GDP is determined by factors that are themselves difficult to measure. Thus, "bottom-up" estimates of the *level* of potential output are also quite imprecise. As we will cover subsequently, however, economists have confidence that the long-run *growth rate* of potential GDP can be estimated well enough to provide meaningful guidance for analysts and policymakers. Hence, in the short run, economists generally focus on factors that cause actual GDP to grow faster or slower than their estimate of the long-run growth rate of potential output. In addition, they focus on measures that indicate, albeit imprecisely, the extent to which the economy is operating above or below its productive capacity, such as unemployment and capacity utilization.

3.4.2 *Recessionary Gap*

Cyclical fluctuations in real GDP and prices are caused by shifts in both the AD and SRAS curves. A decline in AD or a leftward shift in the AD curve results in lower GDP and lower prices. Such declines in AD lead to economic contractions, and if such declines drive demand below the economy's potential GDP, the economy goes into a **recession** (defined as a period during which real GDP decreases for at least two consecutive quarters). In Exhibit 21, when aggregate demand falls, the equilibrium shifts from Point A to Point B. Real GDP contracts from Y_1 to Y_2, and the aggregate price level falls from P_1 to P_2. Because of the decline in demand, companies reduce their workforce and the unemployment rate rises. The economy is in recession, and the recessionary gap is measured as the difference between Y_2 and Y_1 or the amount by which equilibrium output is below potential GDP. Thus, a recessionary gap occurs when the AD curve intersects the short-run AS curve at a short-run equilibrium level of GDP below potential GDP. *Most importantly, in contrast to full employment, equilibrium GDP is below potential GDP.*

Exhibit 21 Recessionary Gap



Any of the factors discussed earlier could cause the shift in the AD curve. Tightening of monetary policy, higher taxes, more pessimistic consumers and businesses, and lower equity and housing prices all reduce AD and are all possible causes of a recession.

The question is, how does the economy return to full employment? There is considerable debate among economists about the answer to this question. Some economists argue that an automatic, self-correcting mechanism will push the economy back to its potential, without the need for government action. The idea is that because of the decline in prices and higher unemployment, workers will be willing to accept lower nominal wages. Workers will do this because each currency unit of wages now buys more goods and services because of their lower price. As a result, lower wages and input prices will shift the SRAS curve to the right (see Exhibit 19) and push the economy back to full employment and potential GDP.

The problem is that this price mechanism can take several years to work. As an alternative, government can use the tools of fiscal and monetary policy to shift the AD curve to the right (from Point B to Point A in Exhibit 21) and move the economy back to full employment. On the fiscal side, policymakers can reduce taxes or increase government spending. On the monetary side, the central bank can lower interest rates or increase the money supply. The problem, however, is that variable lags in the effectiveness of these policy measures imply that policy adjustments may end up reinforcing rather than counteracting underlying shifts in the economy.

Investment Implications of a Decrease in AD Aggregate demand and aggregate supply are theoretical measures that are very hard to measure directly. Most governments, however, publish statistics that provide an indication of the direction that aggregate demand and supply are moving over time. For example, statistics on consumer sentiment, factory orders for durable and nondurable goods, the value of unfilled orders, the number of new housing starts, the number of hours worked, and changes in inventories provide an indication of the direction of aggregate demand. If these statistics suggest that a recession is caused by a decline in AD, the following conditions are likely to occur:

- Corporate profits will decline.
- Commodity prices will decline.
- Interest rates will decline.
- Demand for credit will decline.

This situation suggests the following investment strategy:

- Reduce investments in **cyclical companies**, companies with sales and profits that regularly expand and contract with the business cycle or state of economy (for example, automobile and chemical companies), because their earnings are likely to decline the most in an economic slowdown.
- Reduce investments in commodities and/or commodity-oriented companies because the decline in commodity prices will slow revenue growth and reduce profit margins.
- Increase investments in **defensive companies**, companies with sales and profits that have little sensitivity to the business cycle or state of the economy (e.g., food and pharmaceutical companies), because they are likely to experience only modest earnings declines in an economic slowdown.
- Increase investments in investment-grade or government-issued fixed-income securities. The prices of these securities should increase as interest rates decline.
- Increase investments in long-maturity fixed-income securities because their prices will be more responsive to the decline in interest rates than the prices of shorter-maturity securities.
- Reduce investments in speculative equity securities and in fixed-income securities with low credit quality ratings.

As with most investment strategies, this strategy will be most successful if it is implemented before other market participants recognize the opportunities and asset prices adjust.

EXAMPLE 11

Using AD and AS: A Historical Example: 2007–2009

Many Asian economies were more adversely affected than the United States by the global recession that began in late 2007. In the first quarter of 2009, real GDP fell at an annualized rate of 16% in Japan and 11% in Singapore, compared with a 6% annualized decline in the United States. Using the data on exports as a share of GDP shown in Exhibit 22, explain how the following economic factors contributed to the recession in the Asian economies:

1. Collapse of house prices and home construction in the United States.
2. Oil prices rising from around US$30 a barrel in 2004 to nearly US$150 a barrel in 2008. (Note that most Asian economies rely on imports for almost all of their oil and energy needs. In contrast, the United States has a large domestic energy industry and imports only about half of its oil.)
3. The dramatic reduction in credit availability following the collapse or near collapse of major financial institutions in 2008.

Exhibit 22 Exports as a Share of GDP, 2007 and 2016

	2007		2016	
Economy	Exports as a Percentage of GDP	Percentage of Exports Going to United States	Exports as a Percentage of GDP	Percentage of Exports Going to United States
Hong Kong SAR	186	11.2	192	8.3
Singapore	166	11.5	175	6.5
Thailand	62	11.6	69	11.2
Germany	53	10.9	47	9
South Korea	47	7.1	44	12
Mexico	37	26.4	37	81.2
Canada	28	80.2	31	76.2
Chinese mainland	27	19	21	19
India	17	20.1	20	16
Japan	14	17	17	20.2
Kenya	12	—	16	6.7
United States	12	—	12	—
Ethiopia	11	6.7	9	9.9

Sources: World Bank: World Development Indicators NE.EXP.GNFS.ZS and atlas.media.mit.edu/en/profile/country.

Solution to 1:

The collapse in housing prices caused housing construction spending, a component of business investment, to decline in the United States. The decline in housing prices also caused a sharp fall in household wealth. As a result, consumption spending in the United States declined because of the wealth effect. The decline in both consumption and housing construction shifted the AD curve for the United States to the left, resulting in a US recession. The link to the Asian economies was through global trade because exports represented such a large share of the Asian economies' GDP (Exhibit 23). In turn, these economies exported a significant amount of goods and services to the United States. Thus, the recession in the United States and especially the decline in US consumption spending caused a sharp fall in exports among Asian economies. This decrease lowered their AD and shifted the AD curve to the left, resulting in a recessionary gap in these economies.

Solution to 2:

The rise in oil prices increased input cost and shifted the short-run AS curve to the left. Because the eastern Asian economies depend heavily on imported oil, their economies were more adversely affected than that of the United States.

Solution to 3:

The decline in housing prices caused financial institutions in the United States to suffer large losses on housing-related loans and securities. Several large lenders collapsed, and the US Treasury and the Federal Reserve had to intervene to prevent a wave of bankruptcies among large financial institutions. As a result of the crisis, it became difficult for households and businesses to obtain credit to finance their spending. This dynamic caused AD to fall and increased the severity of the recession in the United States, resulting in a significant decline in US imports and thus exports from the Asian economies. In addition, the financial crisis made it more difficult to obtain trade finance, further reducing exports from Asia.

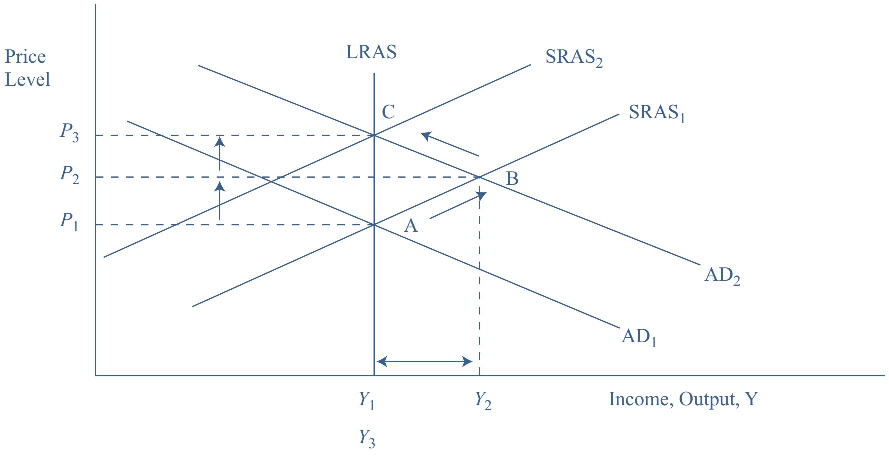
In summary, global investors need to be aware of the growing linkages among economies and the extent that one economy's growth depends on demand from within as well as from outside of that economy. Data on exports as a percentage of an economy's GDP provide an indication of this dependence. Although Japan is often viewed as an export-driven economy, Exhibit 23 shows that in 2018, exports were only 18% of its GDP. Similarly, the economy of India depends largely on domestic spending for growth because in 2018, exports accounted for only 20% of GDP.

3.4.3 *Inflationary Gap*

Increases in AD lead to economic expansions as real GDP and employment increase. If the expansion drives the economy beyond its production capacity, however, **inflation** will occur. The inflation rate is defined as the increase in the general price level from one period to the next. As summarized in Exhibit 17, higher government spending, lower taxes, a more optimistic outlook among consumers and businesses, a weaker domestic currency, rising equity and housing prices, and an increase in the money supply would each stimulate aggregate demand and shift the AD curve to the right. If aggregate supply does not increase to match the increase in AD, a rise in the overall level of prices will result.

As shown in Exhibit 23, an increase in AD will shift the equilibrium level of GDP from Point A to Point B. Real output increases from Y_1 to Y_2, and the aggregate price level rises from P_1 to P_2. As a result of the increase in aggregate demand, companies increase their production and hire more workers. The unemployment rate declines. Once an economy reaches its potential GDP, however, companies must pay higher wages and other input prices to further increase production. The economy now faces an inflationary gap, measured by the difference between Y_2 and Y_1 in Exhibit 23. *An inflationary gap occurs when the economy's short-run level of equilibrium GDP is above potential GDP, resulting in upward pressure on prices.*

Exhibit 23 Inflationary Gap



GDP cannot remain at Y_2 for long because the economy is overutilizing its resources—that is, extra shifts of workers are hired and plant and equipment are operating at their maximum capacity. Eventually, workers become tired and plants and equipment wear out. The increase in the general price level and input prices will set in motion the process of returning the economy back to potential GDP. Higher wages and input prices shift the SRAS curve to the left (from $SRAS_1$ to $SRAS_2$), moving the economy to Point C in Exhibit 23. Again, this self-correcting mechanism may work slowly.

A nation's government and/or its central bank can attempt to use the tools of fiscal and monetary policy to control inflation by shifting the AD curve to the left (AD_2 to AD_1 in Exhibit 23) so that the return to full employment occurs without the price increase. From a fiscal perspective, policymakers can raise taxes or cut government spending. From a monetary perspective, the central bank can reduce bank reserves, resulting in a decrease in the growth of the money supply and higher interest rates.

Investment Implications of an Increase in AD Resulting in an Inflationary Gap If economic statistics (consumer sentiment, factory orders for durable and nondurable goods, etc.) suggest that there is an expansion caused by an increase in AD, the following conditions are likely to occur:

- Corporate profits will rise.
- Commodity prices will increase.
- Interest rates will rise.
- Inflationary pressures will build.

This situation suggests the following investment strategy:

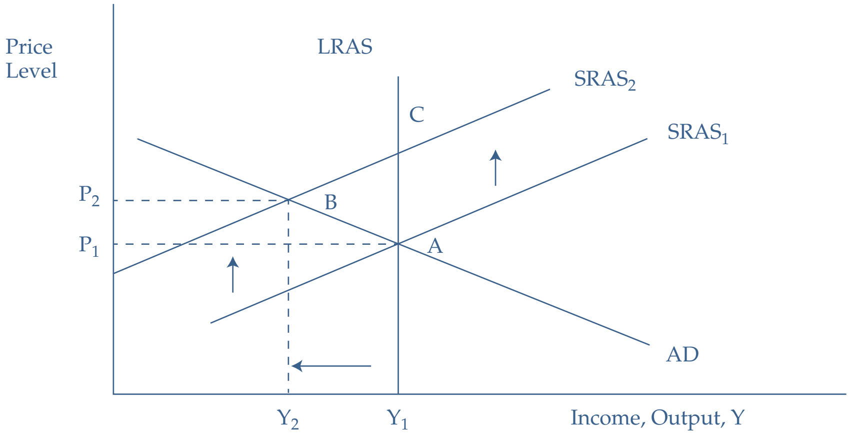
- Increase investment in cyclical companies because they are expected to have the largest increase in earnings.
- Reduce investments in defensive companies because they are expected to have only a modest increase in earnings.
- Increase investments in commodities and commodity-oriented equities because they will benefit from higher production and output.
- Reduce investments in fixed-income securities, especially longer-maturity securities, because they will decline in price as interest rates rise. Raise exposure to speculative fixed-income securities (junk bonds) because default risks decrease in an economic expansion.

3.4.4 *Stagflation: Both High Inflation and High Unemployment*

Structural fluctuations in real GDP are caused by fluctuations in SRAS. Declines in aggregate supply bring about **stagflation**—high unemployment and increased inflation. Increases in aggregate supply conversely give rise to high economic growth and low inflation.

Exhibit 24 shows the case of a decline in aggregate supply, perhaps caused by an unexpected increase in basic material and oil prices. The equilibrium level of GDP shifts from Point A to B. The economy experiences a recession as GDP falls from Y_1 to Y_2, but the price level, instead of falling, rises from P_1 to P_2. Over time, the reduction in output and employment should put downward pressure on wages and input prices and shift the SRAS curve back to the right, re-establishing full employment equilibrium at Point A. However, this mechanism may be painfully slow. Policymakers may use fiscal and monetary policy to shift the AD curve to the right, as previously discussed, but at the cost of a permanently higher price level at Point C.

Exhibit 24 Stagflation



The global economy experienced stagflation in the mid-1970s and early 1980s. Both unemployment and inflation soared. The problem was caused by a sharp decline in aggregate supply fueled by higher input prices, especially the price of oil. In 1973, the price of oil quadrupled. A steep global recession began in late 1973 and lasted through early 1975. The recession was unusual because prices rose rather than declined as would be expected in a typical demand-caused downturn. In 1979–1980, the price of oil doubled. Higher energy prices shifted the SRAS curve to the left, as shown in Exhibit 24, leading to a global recession in 1980–1982. In the United States, the contraction in output was reinforced by the Federal Reserve's decision to tighten monetary policy to fight the supply-induced inflation.

Investment Implications of a Shift in AS Labor and raw material costs, including energy prices, determine the direction of shifts in short-run aggregate supply: Higher costs for labor, raw materials, and energy lead to a decrease in aggregate supply, resulting in lower economic growth and higher prices. Conversely, lower labor costs, raw material prices, and energy prices lead to an increase in aggregate supply, resulting in higher economic growth and a lower aggregate price level. Productivity is also an important factor. Higher rates of productivity growth shift the AS to the right, resulting in higher output and lower unit input prices. Lower rates of productivity growth do the opposite and shift the AS curve to the left.

From an investment perspective, a decline in AS (leftward shift of the SRAS curve) suggests:

- reducing investment in fixed income because rising output prices (i.e., inflation) put upward pressure on nominal interest rates,
- reducing investment in most equity securities because profit margins are squeezed and output declines, and
- increasing investment in commodities or commodity-based companies because prices and profits are likely to rise.

On the other hand, an increase in AS (rightward shift of the SRAS curve) resulting from higher productivity growth or lower labor, raw material, and energy costs is favorable for most asset classes other than commodities.

3.4.5 *Conclusions on AD and AS*

The business cycle and the resulting fluctuations in real GDP are caused by shifts in the AD and AS curves. The effects of these shifts can be summarized as follows:

- An increase in AD raises real GDP, lowers the unemployment rate, and increases the aggregate level of prices.
- A decrease in AD lowers real GDP, increases the unemployment rate, and decreases the aggregate level of prices.
- An increase in AS raises real GDP, lowers the unemployment rate, and lowers the aggregate level of prices.
- A decrease in AS lowers real GDP, raises the unemployment rate, and raises the aggregate level of prices.

If both curves shift, the effect is a combination of the aforementioned individual effects. We look at four possible cases:

1. *Both AD and AS increase.* If both AD and AS increase, real GDP will increase but the effect on inflation is unclear unless we know the magnitude of the changes, because an increase in AD will increase the price level, whereas an increase in AS will decrease the price level. If AD increases more than AS, the price level will increase. If AS increases more than AD, however, the price level will decline.
2. *Both AD and AS decrease.* If both AD and AS decrease, real GDP and employment will decline but the effect on the price level is unclear unless we know the magnitude of the changes, because a decrease in AD decreases the price level, whereas a decrease in AS increases the price level. If AD decreases more than AS, the price level will fall. If AS decreases more than AD, the price level will rise.
3. *AD increases and AS decreases.* If AD increases and AS declines, the price level will rise but the effect on real GDP is unclear unless we know the magnitude of the changes, because an increase in AD increases real GDP, whereas a decrease in AS decreases real GDP. If AD increases more than AS declines, GDP will rise. If AS decreases more than AD increases, real GDP will fall.
4. *AD decreases and AS increases.* If AD decreases and AS increases, the price level will decline but the effect on real GDP is unclear unless we know the magnitudes of the changes, because a decrease in AD decreases real GDP, whereas an increase in AS increases real GDP. If AD decreases more than AS increases, real GDP will fall. If AS increases more than AD declines, real GDP will rise.

   Exhibit 25 summarizes these four cases.

Exhibit 25 Effect of Combined Changes in AS and AD

Change in AS	Change in AD	Effect on Real GDP	Effect on Aggregate Price Level
Increase	Increase	Increase	Indeterminate
Decrease	Decrease	Decrease	Indeterminate
Increase	Decrease	Indeterminate	Decrease
Decrease	Increase	Indeterminate	Increase

Whether the growth of the economy is demand- or supply-driven has an effect on asset prices. Demand-driven expansions are normally associated with rising interest rates and inflation, whereas contractions are associated with lower inflation and interest rates. Supply-driven expansions are associated with lower inflation and interest rates, whereas supply-driven contractions are associated with rising inflation and interest rates.

EXAMPLE 12

Investment Strategy Based on AD and AS Curves

An analyst is evaluating the possibility of investing in China, Italy, Mexico, or Brazil. What are the equity and fixed-income investment opportunities in these countries based on the following events?

1. The Chinese government announced a spending plan of US$1.2 trillion, or 13% of GDP. In addition, the central bank of China eased monetary policy, resulting in a surge of lending.
2. The Italian government announced a decline in labor productivity, and it expects this trend to continue into the future.
3. In response to rising inflationary pressure, the Mexican central bank tightened monetary policy, and the government announced tax increases and spending cuts to balance the budget.
4. A major discovery of oil off the coast of Brazil lowered oil prices, while the Brazilian government announced a major increase in spending on public infrastructure to stimulate the economy.

Solution to 1:

Stimulative fiscal and monetary policies should result in a demand-driven expansion. Investors should reduce investments in fixed-income securities and defensive companies and invest in cyclical companies and commodities. As a result, the prospects for growth-oriented equity investments look favorable in China.

Solution to 2:

A decline in labor productivity will result in a decline in AS; that is, the AS curve will shift to the left. This scenario is typically a poor investment environment. Investors should reduce investments in both fixed-income and equity securities and invest in commodities. Entry into Italian stocks and bonds does not look attractive.

Solution to 3:

The policy measures put in place by the Mexican government and central bank will cause a drop in AD and likely result in a recession. Investors should increase their investments in fixed-income securities because interest rates will most likely decline as the recession deepens. This scenario is a poor environment for equity securities.

Solution to 4:

In this situation, both the AD and AS curves will shift. The increase in spending on public infrastructure will shift the AD curve to the right, resulting in higher aggregate expenditures and prices. Lower oil prices will shift the AS curve to the right, resulting in higher GDP but lower prices. Thus, GDP will clearly increase, but the effect on prices and inflation is indeterminate. As a result, investors should increase their investment in equity securities; however, the effect on fixed-income securities is unclear.

EXAMPLE 13

Historical Example: Using AD and AS to Explain Japan's Economic Problem

Japan experienced sluggish growth in real GDP for nearly two decades following the bursting of an asset and investment bubble in the late 1980s. At the same time, Japan experienced deflation (declining prices) over the two-decade period. The reasons for this protracted period of stagnation continue to be debated among economists. Failure to recognize a change in the Japanese growth rate has hurt many investors, especially those taking a long-term perspective. From their peak in 1989, Japanese equity prices, as measured by the Nikkei index, fell by more than 60% before bottoming out in mid-1992. During the following two decades, the market was essentially flat despite considerable volatility.

The performance of the Japanese economy can be explained using the AD and AS model. The protracted slowdown of growth in Japan beginning in the early 1990s can be linked to the effect of the collapse of the equity and commercial real estate markets in the late 1980s and to excessive investment in capital goods (new factories and equipment) in the 1980s. These problems were compounded by persistent weakness in the banking sector, a profound lack of confidence among businesses and consumers, and negative demographics with slow growth in the working age population.

The sum of these developments caused a decline in both the AD and AS curves. Aggregate demand declined, shifting the AD curve to the left for the following reasons:

- The wealth effect resulting from the decline in equity and real estate prices sharply reduced consumption spending. Asset prices took more than two decades to recover.
- Excessive investment in capital goods caused a sharp decline in business investment during the period.
- Lack of confidence among businesses and consumers during the period.
- Problems in the banking sector made monetary policy ineffective because banks were unable to lend, which negatively affected both consumer and business spending.

AS declined for the following reasons:

- Marked slowing in private investment spending during the period reduced the capital stock. This dynamic also reduced potential GDP.
- Slow population growth limited the growth in the labor supply, which also reduced potential GDP.
- Higher energy prices slowed growth during the period because of Japan's heavy dependence on imported energy.

As would be expected, the declines in both AD and AS resulted in slow GDP growth. The fact that prices fell indicates that the AD curve shifted more than the AS curve.

ECONOMIC GROWTH AND SUSTAINABILITY

4

m describe sources, measurement, and sustainability of economic growth

n describe the production function approach to analyzing the sources of economic growth

o define and contrast input growth with growth of total factor productivity as components of economic growth

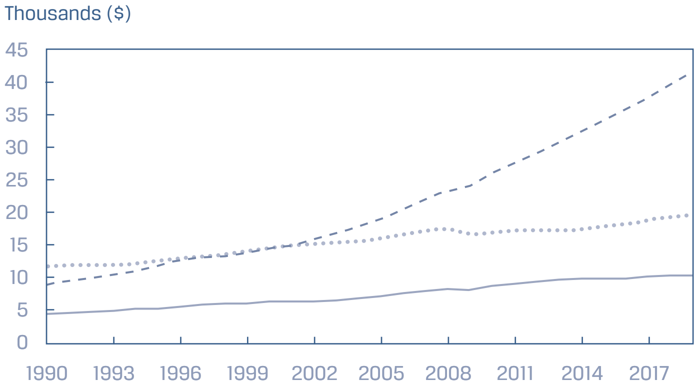
Understanding sources of economic growth is crucial for practitioners involved in various types of investment decision making. Economic growth is calculated as the annual percentage change in real GDP or the annual change in real per capita GDP:

- Growth in real GDP measures how rapidly the total economy is expanding.
- Per capita GDP, defined as real GDP divided by population, measures the standard of living in each country and the ability of the average person to buy goods and services.

Exhibit 26 illustrates growth in GDP in selected regions since 1990. It shows that although total GDP doubled in Latin America during the period, GDP per capita in the region increased by only about 50%, as a result of population growth.

Exhibit 26 Real GDP Growth and Per Capita GDP Growth for Selected Regions

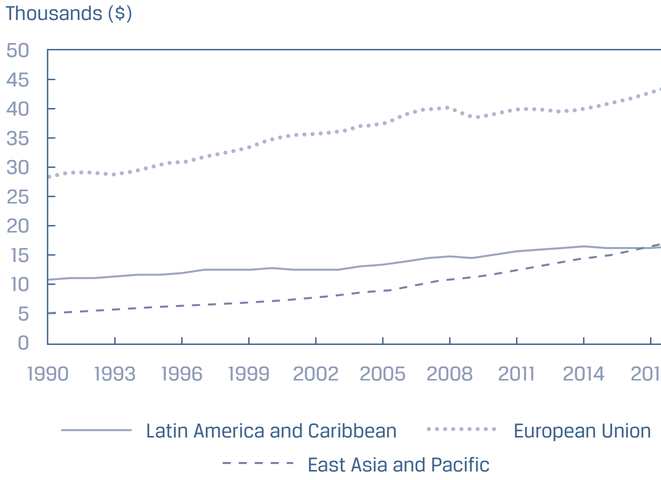
A. In PPP (constant 2017 US$ billion)



(continued)

Exhibit 26 (Continued)

B. GDP per Capita, PPP (constant 2017 US$)



Note: PPP is Purchasing Power Parity.
Source: World Bank data.

Economic growth is important because rapid growth in per capita real GDP can transform a poor nation into a wealthy one. Even small differences in the growth rate of per capita GDP, if sustained over time, have a large effect on an economy's standard of living. Nevertheless, there is a limit to how fast an economy can grow. Faster growth is not always better for an economy because there are costs associated with excess growth, such as higher inflation, potential environmental damage, and the lower consumption and higher savings needed to finance the growth.

These costs raise the issue of sustainable growth, which requires an understanding of the concept of potential GDP. Recall that potential GDP measures the economy's productive capacity and is the level of real GDP that an economy could produce if capital and labor are fully employed. In order to grow over time, an economy must add to its productive capacity. Thus, the **sustainable rate of economic growth** is measured by the rate of increase in the economy's productive capacity or potential GDP. It is important to note that economists cannot directly measure potential output. Instead, they estimate it using a variety of techniques that we will discuss.

For global investors, estimating the sustainable rate of economic growth for an economy is important in valuing a variety of assets. Growth in potential GDP matters for long-run equity market appreciation and for individual industries and company forecasts. Potential GDP can be used to gauge inflationary pressures, determine credit quality, and assess the likelihood of a change in central bank policy, all of which are relevant for fixed income investors. Economic growth in large economies or whole regions is also useful for forecasting demand for commodities.

4.1 The Production Function and Potential GDP

We focus on the simplest form of what is known as the neoclassical or Solow growth model, the framework used to understand and analyze the underlying sources of growth for an economy. The model shows that the economy's productive capacity and potential GDP increase for two reasons:

1. accumulation of such inputs as capital, labor, and raw materials used in production, and
2. discovery and application of new technologies that make the inputs in the production process more productive—that is, able to produce more goods and services for the same amount of input.

4.1.1 *The Production Function*

The model is based on a **production function** that provides the quantitative link between the levels of output that the economy can produce and the inputs used in the production process, given the state of technology. A two-factor production function with labor and capital as the inputs is expressed mathematically as

$$Y = A\boldsymbol{F}(L,K),$$

where Y denotes the level of aggregate output in the economy, L is the quantity of labor or number of workers in the economy, K is the capital stock or the equipment and structures used to produce goods and services, and A represents **total factor productivity** (TFP). TFP is a scale factor that captures the exogenous effect of technological change and other factors that raise output beyond the measured contribution of the capital and labor inputs. Like potential GDP, TFP is not directly observed in the economy but instead is estimated.

The production function shows that output in the economy depends on inputs and the level of technology. The economy's capacity to produce goods grows when these inputs increase and/or technology advances. The more technologically advanced an economy is, the more output it is able to produce from a given amount of inputs.

4.1.2 *The Model Assumptions*

The neoclassical model makes three assumptions about the production function that provide a link to microeconomics. First, it assumes that the production function has constant returns to scale. This means that if all the inputs in the production process are increased by the same percentage, then output will rise by that percentage. Thus, doubling all inputs would double output. Second, the model assumes that the production function exhibits **diminishing marginal productivity** with respect to any individual input. Finally, the model assumes no positive or negative externalities are associated with the use of the inputs. These assumptions play an important role in assessing the contribution of labor and capital to economic growth. Marginal productivity looks at the extra output produced from a one-unit increase in an input if the other inputs are unchanged. It applies to any input as long as the other inputs are held constant. For example, if we have a factory of a fixed size and we add more workers to the factory, the marginal productivity of labor measures how much additional output each additional worker will produce.

Diminishing marginal productivity means that at some point the extra output obtained from each additional unit of the input will decline. In the prior example, if we hire more workers at the existing factory (fixed capital input in this case), output will rise by a smaller and smaller amount with each additional worker.

In the case of capital, if we add more and more capital to a fixed number of workers, the amount of additional output contributed by each additional amount of capital will fall. Thus, if capital grows faster than labor, capital will become less productive, resulting in slower and slower growth. Diminishing marginal productivity of capital has two major implications for potential GDP and long-term growth:

1 Long-term sustainable growth cannot rely solely on **capital deepening investment** that increases the stock of capital relative to labor. More generally, increasing the supply of some input(s) relative to other inputs will lead to diminishing returns and cannot be the basis for sustainable growth.
2 Given that developing countries have relatively less capital, their productivity of capital is high. All else the same, the growth rates of developing countries should exceed those of developed countries. As a result, there should be a **convergence** of incomes between developed and developing countries over time.

Because of the assumption of diminishing returns to capital, the only way to sustain growth in potential GDP per capita in the long run is through growth in TFP. This results in an upward shift in the production function: The economy produces more goods and services using the same level of labor and capital inputs. In terms of the formal production function $Y = A\boldsymbol{F}(L,K)$, this dynamic is reflected by an increase in the parameter A.

4.1.3 *Growth Accounting Equation*

Using the production function, Robert Solow developed a model that explained the contribution of labor, capital, and TFP to economic growth. The growth accounting equation shows that the rate of growth of potential output equals growth in TFP plus the weighted average growth rate of labor and capital.

$$\text{Growth in potential GDP} = \text{Growth in TFP} + W_L\,(\text{Growth in labor}) + W_C\,(\text{Growth in capital})$$

where W_L and W_C are the relative shares of capital and labor in national income. The capital share is the sum of corporate profits, net interest income, net rental income, and depreciation divided by GDP. The labor share is employee compensation divided by GDP. For many developed countries, W_L and W_C are roughly 0.7 and 0.3, respectively.

The growth accounting equation highlights a key point: The contribution of labor and capital to long-term growth depends on their respective shares of national income. For the United States, because labor's share is higher, an increase in the growth rate of labor will have a significantly larger effect (roughly double) on potential GDP growth than will an equivalent increase in the growth rate of capital.

The growth accounting equation can be further modified to explain growth in per capita GDP. Because it measures the standard of living and purchasing power of the average person in an economy, per capita GDP is more relevant than the absolute level of GDP in comparing economic performance among countries. Transforming the growth accounting equation into per capita terms results in the following equation:

$$\text{Growth in per capita potential GDP} = \text{Growth in TFP} + W_C\,(\text{Growth in capital-to-labor ratio})$$

The capital-to-labor ratio measures the amount of capital available per worker and is weighted by the share of capital in national income. Because capital's share in national income in the US economy is 0.3, a 1% increase in the amount of capital available for each worker increases per capita output by only 0.3%.

EXAMPLE 14

Impact of Capital Investment in Laos

Annual GDP growth in Laos has averaged about 7% over the 10 years since 2009. The economy has benefitted from foreign direct investment in hydroelectric dams and mining projects. Using the neoclassical growth model, an emerging market analyst states that these capital inflows are unlikely to boost the long-term potential GDP growth rate of Laos.

Explain using the neoclassical model whether the analyst's statement is correct.

Solution:

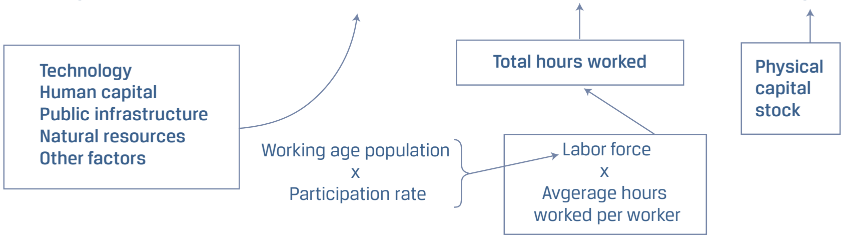
The capital inflows in the form of foreign direct investment are increasing capital spending in Laos and, through capital deepening, will raise the short-term GDP growth rate. Long-term potential GDP, however, depends on growth in TFP, the key driver of which is technical progress. Adding to capital through the foreign direct investment will be subject to diminishing returns. Thus, the analyst is correct because long-term potential GDP growth cannot rely on capital accumulation or capital deepening. To increase potential GDP, Laos must implement policies to boost the growth in TFP.

4.2 Sources of Economic Growth

The foregoing growth accounting equation focuses on the main determinants of growth—capital, labor, and TFP (the key driver being technology)—and omits a number of other sources of growth to simplify the analysis. For many countries, however, human capital, public infrastructure, natural resources, and the economic environment and institutions play an important role in explaining economic growth. In the standard neoclassical model discussed earlier, these inputs are measured as part of TFP growth along with technological changes. Using this model, the sources of growth for an economy that determine its capacity to supply goods and services are as shown in Exhibit 27. Note that the amount of labor, measured as total hours worked, is determined by the size of the average hours worked per worker multiplied by the labor force. Labor force itself is a product of working age population and the participation rate (the proportion of working age population choosing to work or seeking work).

Exhibit 27 Sources of Economic Growth

Growth in potential GDP = Growth in TFP + WL (Growth in labor) + WC (Growth in capital)



4.2.1 *Technology*

The most important factor affecting economic growth is technology, especially in developed countries. Technology refers to the process a company uses to transform inputs into outputs. Technological advances are discoveries that make it possible to produce more or higher-quality goods and services with the same resources or inputs. At the same time, technological progress results in the creation of new goods and services. Finally, technological progress improves how efficiently businesses are organized and managed.

Technological advances allow an economy to overcome the limits imposed by diminishing marginal returns. Thus, an economy will face limits to growth if it relies exclusively on expanding the inputs or factors of production. One of the key drivers of growth in developed countries is the information technology (IT) sector, whose growth has been driven by technological innovation that has caused the price of key technologies, such as semiconductors, to fall dramatically. The steep declines in prices have encouraged investment in IT at the expense of other assets. The sector has grown very fast and has made a significant contribution to economic growth, employment, and exports.

Countries can innovate through expenditures, both public and private, on research and development (R&D). Expenditures on R&D and the number of patents issued, although not directly measuring innovation, provide some useful insight into innovative performance. Countries can also acquire new technology through imitation or copying the technology developed elsewhere. The embodiment of technology in capital goods can also enable relatively poor countries to catch up to the technology leaders.

TFP is often used as a proxy for technological progress and organizational innovation. TFP is the amount by which output would rise because of improvements in the production process. It is not directly observed and is calculated as a residual, the difference between the growth rate of potential output and the weighted average growth rate of capital and labor. Specifically,

$$\text{TFP growth} = \text{Growth in potential GDP} - [W_L\ (\text{Growth in labor}) + W_C\ (\text{Growth in capital})]$$

The interpretation of the TFP as a pure technology measure is incorrect. Given the measurement and estimation difficulties, TFP will measure not only technological change but also the effect of other inputs into the production process such as human capital, public infrastructure, and natural resources. In addition to the other inputs, TFP measures the effect of externalities and spillover effects, omitted inputs, organizational efficiencies, and institutional and governance factors. Nevertheless, it is a useful indicator of the underlying technological forces.

4.2.2 *Labor Supply*

Growth in the number of people available for work (quantity of workforce) is an important source of economic growth and partially accounts for the superior growth performance of the US economy versus the European and Japanese economies. Most developing countries, such as China, India, and Mexico, have a large potential labor supply. We can measure the potential size of the labor input as the total number of hours available for work, which is given by

$$\text{Total hours worked} = \text{Labor force} \times \text{Average hours worked per worker}$$

The **labor force** is defined as the portion of the working age population (over the age of 16) that is employed or available for work but not working (unemployed). The contribution of labor to overall output is also affected by changes in the average hours worked per worker. The average number of hours worked is highly sensitive to the business cycle. The long-term trend, however, has been toward a shorter work week in the advanced countries. This development is the result of legislation, collective bargaining agreements, and the growth of part-time and temporary work.

4.2.3 *Human Capital*

In addition to the quantity of labor (the "*L*" in the earlier mentioned production function), the quality of the labor force is important. Human capital reflects the accumulated knowledge and skill that workers acquire from education, training, and life experience. It measures the quality of the workforce, and we treat it as part of TFP. In general, better-educated and skilled workers will be more productive and more adaptable to changes in technology.

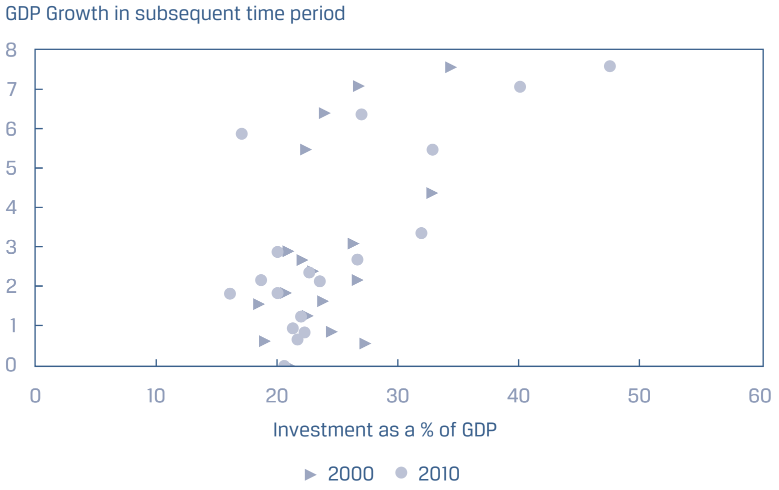
An economy's human capital is increased through investment in education and on-the-job training. Like physical capital, investment in education is costly. Studies show, however, that education offers a significant return—that is, people with more education earn higher wages. Moreover, education may also have a spillover or externality effect: Increasing the educational level of one person not only raises that person's output but also the output of those around her. The spillover effect operates through the link between education and advances in technology. Education not only improves the quality of the labor force but also encourages growth through innovation. Investment in health is also a major contributor to human capital, especially in developing countries.

4.2.4 *Physical Capital Stock*

The physical **capital stock** (accumulated amount of buildings, machinery, and equipment used to produce goods and services) increases from year to year as long as net investment (gross investment less depreciation of capital) is positive. Thus, countries with a higher rate of investment should have a growing physical capital stock and a higher rate of GDP growth.

Exhibit 28 shows the level of business investment as a share of GDP across a selection of 18 developed and developing economies. The exhibit shows significant variation across countries. As is evident in Exhibit 28, the correlation between investment (as a percentage of GDP) and economic growth in the subsequent period (19 years after 2000 and 9 years after 2010) is high. Economies that devote a large share of GDP to investment, such as China, India, and South Korea (numerical data not shown), have high growth rates. Ireland, Europe's fastest-growing economy from 2010–2018, has among the highest investment-to-GDP ratio in Europe. Economies that devote a smaller share of GDP to investment, such as Brazil and Mexico, have slower growth rates. The data helps explain why the Chinese economy has expanded so rapidly, achieving an annual GDP growth rate of nearly 10% over the last two decades. Investment spending in China on new factories, equipment, and infrastructure as a percentage of GDP is the highest in the world.

Exhibit 28 Business Investment As a Percentage of GDP and GDP Growth in the Subsequent Period



Notes: The term "investment" is represented by non-residential gross fixed capital formation as a share of GDP used. GDP growth is the average annual real GDP growth in the decade following the year of the investment as a percentage of GDP data.
Sources: GDP: OECD National Accounts Statistics (database), April 2019. Gross fixed capital: World Development Indicator, ne.gdi.totl.zs

4.2.5 *Natural Resources*

Raw materials are an essential input to growth and include everything from available land to oil to water. Historically, consumption of raw materials has increased as economies have grown. There are two categories of natural resources:

1 **Renewable resources** are those that can be replenished, such as a forest. For example, if a tree is cut, a seedling can be planted and a new forest harvested in the future.

2 **Non-renewable resources** are finite resources that are depleted once they are consumed. Oil and coal are examples.

Natural resources account for some of the differences in growth among countries. Today, such countries as Brazil and Australia, as well as those in the Middle East, have relatively high per capita incomes because of their resource base. Countries in the Middle East have large pools of oil. Brazil has an abundance of land suitable for large-scale agricultural production, making it a major exporter of coffee, soybeans, and beef.

Although natural resources are an important factor in growth, they are not necessary for a country to achieve a high level of income provided it can acquire the requisite inputs through trade. Countries in eastern Asia, such as Japan and South Korea, have in the past experienced rapid economic growth but own few natural resources. It is also important to note that having plentiful natural resources may hurt an economy's growth, a topic that will be explored at a later stage.

4.2.6 *Public Infrastructure*

Roads, water systems, mass transportation, airports, and utilities are all examples of public infrastructure or public capital. Even if the infrastructure assets are privately owned or operated, the economic impact will be the same. Infrastructure assets have few substitutes and generate significant economies of scale, so they have the characteristics of a natural monopoly. A key feature of public capital is that they create **externalities**. Externalities are spillover effects of production and consumption activities onto others who are not directly involved in a particular transaction, activity, or decision. In the case of public infrastructure, the externalities complement the production of private sector goods and services. The full benefit of public infrastructure investment may extend beyond its direct expenditure because it boosts the productivity of private capital.

4.2.7 *Other Factors Driving Growth*

Researchers have focused on the positive externalities associated with research and development and public education. Research and development efforts by one firm spill over and affect the stock of knowledge available to all firms, raising potential GDP growth. Thus, external or social benefits arise from increased spending on research and development that go beyond the benefit for the individual firm. The key point for economic growth is that there may be constant or increasing returns to the inputs, and thus the economy is no longer constrained by diminishing marginal productivity.

Externalities also have a negative effect on growth, with pollution being the primary example. Air and water quality in advanced countries have improved significantly over the last few decades, and there is a strong negative correlation between per capita GDP and pollution. But these success stories largely reflect the local economic effects of economic growth. The environmental problems we face today are global in nature and affect the entire planet. The burning of fossil fuels releases carbon dioxide into the atmosphere, causing a greenhouse effect that raises the earth's overall temperature. Climate change is closely linked to economic growth, because carbon dioxide emissions rise with economic growth. These emissions impose a cost on the global economy in terms of economic activity, food production, health, and habitability.

Without government action, individuals and firms may have no incentive to reduce these negative externalities. The more global the pollution problem, the less any single country benefits from reducing its emissions. Moreover, nations face an important trade-off on emissions: Reducing emissions today may result in lower short-term growth, but not reducing emissions today will negatively affect long-term sustainable GDP growth. Policy measures that create incentives to reduce the carbon intensity of output are the potential solution.

In addition to the aforementioned factors, a country's economic environment can play an important role in influencing economic growth. Lack of appropriate institutions and a poor legal and political environment restrains growth, especially in the developing countries. Factors such as internal competitive structure; poorly developed financial markets; a country's openness to trade; and its political stability, property rights, and rule of law can influence innovative activity and economic growth.

EXAMPLE 15

Sources of Growth for Westerland

Westerland is an advanced economy with a high capital-to-labor ratio. Working-age population growth has been close to zero for the last decade. During the same period, annual GDP growth has slowed to about 1%. More than 70% of Westerland's electricity is produced by renewable energy resources. The nation has reduced carbon dioxide emissions by 25% over the last decade. Policymakers hope to raise Westerland's sustainable rate of growth or its potential GDP.

Explain how the following factors are a potential source of growth for Westerland.

1. Tax reductions are proposed on business investment and R&D.

2 Labor market incentives are planned to create flexible work schedules and develop policies for maternity/parental leave and childcare.

3 Westerland plans to further reduce its carbon dioxide emissions, but global carbon dioxide emissions are expected to rise significantly as global GDP growth accelerates.

Solution to 1:

Because Westerland is an advanced economy operating at a high level of capital per worker, diminishing returns to capital are likely to be significant. If this is the case, additions to capital will have little effect on growth and are not a significant source of growth. In contrast, incentives encouraging increased R&D expenditures by increasing the growth rate of TFP and technological change are a potential source of growth.

Solution to 2:

Westerland is facing a demographic challenge because slow population growth will limit its potential growth rate. To offset this issue, the country can encourage immigration and/or increase its workforce participation rate. The objective of creating flexible work schedules and developing policies for parental leave and childcare is to raise the labor force participation rate (note that Working age population × Participation rate × Average hours worked gives L, the amount of labor). In the short run, the labor market incentives will increase the growth rate in the labor supply and thus are a potential source of growth. After the initial response to these incentives, however, the labor force participation rate will level off. As a result, the policies will have no long-term effect on the labor input and will not increase sustainable growth.

Solution to 3:

Westerland has successfully reduced its carbon dioxide emissions and plans further reduction, probably at the cost of slower GDP growth. Nevertheless, it still faces the negative effects of climate change despite its own effort to reduce emissions. This dynamic highlights climate change as a classic example of a negative externality that no single country can solve and requires global cooperation. This international burden sharing presents a stumbling block for success.

4.3 Measures of Sustainable Growth

Measuring how fast an economy can grow is an important exercise. Economists project potential GDP in order to forecast the economy's sustainable growth path. Potential GDP is an unobserved concept that is approximated using a number of alternative methods. Note that estimates of the economy's potential growth can change as new data become available. Being able to understand such changes is critical for financial analysts, because long-term forecasts for cash flows that companies generate are dependent on the sustainable rate of economic growth.

4.3.1 *Labor Productivity*

One approach to estimating potential GDP uses the productivity of the labor force, which is generally reliable data. **Labor productivity** consists of the quantity of goods and services (real GDP) that a worker can produce in one hour of work. Our standard of living improves if we produce more goods and services for each hour of work. Labor productivity is calculated as real GDP for a given year divided by the total number of hours worked in that year, counting all workers. We use total hours, rather than the number of workers, to adjust for the fact that not everyone works the same number of hours.

Labor productivity = Real GDP/Aggregate hours

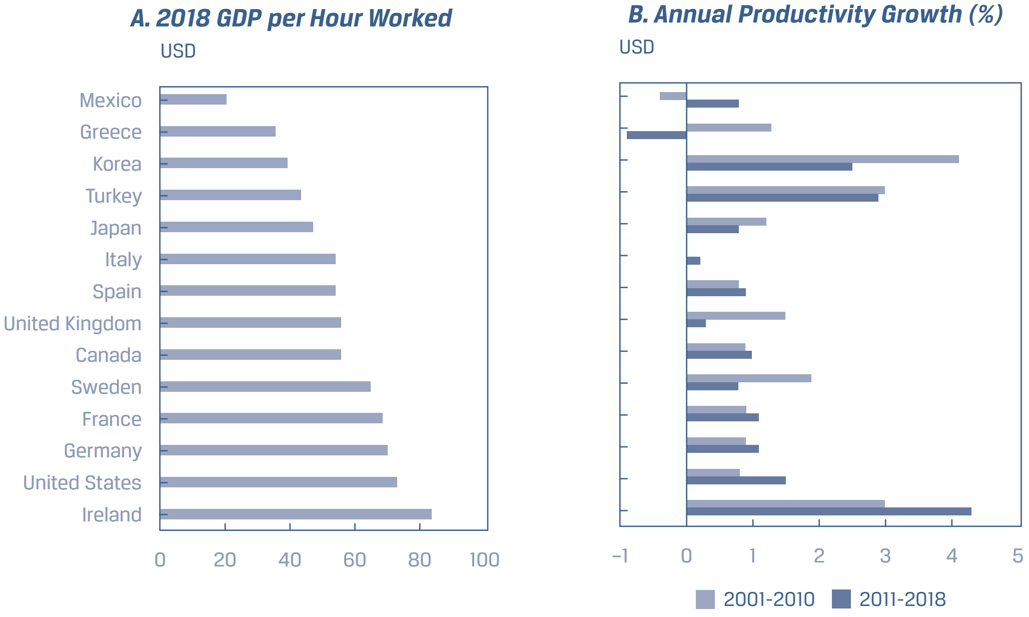
Therefore, we need to understand the forces that make labor more productive. Productivity is determined by the factors that we examined earlier: education and skill of workers (human capital), investments in physical and public capital, natural resources, improvements in technology, and other factors. An increase in any of these factors will increase the productivity of the labor force. The factors determining labor productivity can be derived from the production functions under the assumption of constant returns to scale, where a doubling of inputs causes output to double as well. Dividing the production function by L, we obtain the following:

$$Y/L = A\boldsymbol{F}(1,K/L),$$

where Y/L is output per worker, which is a measure of labor productivity. The equation states that labor productivity depends on physical capital per worker (K/L) and total factor productivity, which largely measures technological change (A). TFP is a scale factor that does not depend on the mix of inputs. Changes in TFP are measured as a residual, capturing growth that cannot be attributed to the labor supply and physical capital inputs. On the other hand, as shown in this equation, output per worker depends on both the general level of productivity (reflected in TFP) and the mix of inputs. Increases in either TFP or the capital-to-labor ratio boost output per worker. Because both output and labor input can be observed, labor productivity can be measured directly.

Labor productivity is a key concept for measuring the health and prosperity of an economy and its sustainable rate of growth. An analyst examining the growth prospects for an economy needs to focus on the labor productivity data for that country. Growth of labor productivity over time largely explains the differences in the living standards among countries. Labor productivity growth also helps analysts estimate the long-term sustainable growth rates among countries. The distinction between the level and growth rate of productivity is important to understand. Exhibit 29 shows that many countries have relatively high levels of productivity, even though productivity growth during the last decade or two has not been significant (e.g., Italy, the United Kingdom, Spain). Ireland, however, entered this century with a lower level of productivity and had a lower starting point compared with the other countries. Its substantial productivity growth in the last two decades has elevated the country to the top of this particular list.

Exhibit 29 Labor Productivity: Level vs. Growth Rate in Select Countries



Source: OECD Productivity Statistics (database), April 2019

4.3.2 *Level of Labor Productivity*

The higher the level of labor productivity, the more goods and services the economy can produce with the same number of workers. The level of labor productivity depends on the accumulated stock of human and physical capital and is much higher in the developed countries. For example, India has a population of about 1.3 billion people, compared with more than 83 million people in Germany (United Nations, World Population Prospects 2019). Because of its much larger population, India has significantly more workers than Germany; however, the German economy, as measured by real GDP, is larger. As shown in Exhibit 29, Germany has among the highest level of productivity in the world, producing more than US$70 of GDP per hour worked. Similarly, workers in France, the United States, and Ireland have high levels of productivity. In comparison, Mexican workers produce only US$20.70 worth of GDP per hour worked. Thus, German workers are more than three times more productive as Mexican workers.

4.3.3 *Growth Rate of Labor Productivity*

The growth rate of labor productivity is the percentage increase in productivity over a year. It is among the economic statistics that economists and financial analysts watch most closely. In contrast to the level of productivity, the growth rate of productivity is typically higher in developing countries where human and physical capital is scarce but growing rapidly.

If productivity growth is rapid, it means the same number of workers can produce more and more goods and services. In this case, companies can afford to pay higher future wages and still make a profit. Thus, high rates of productivity growth will translate into rising profits. Rising profits, if not already anticipated by the market, will support higher stock prices.

In contrast, persistently low productivity growth suggests the economy is in bad shape. Without productivity gains, businesses have to either cut wages or boost prices in order to increase profit margins. Low rates of productivity growth should be associated with slow growth in profits and flat or declining stock prices.

EXAMPLE 16

Prospects for Equity Returns in Mexico

John Todd, CFA, manages a global mutual fund with nearly 30% of its assets invested in Europe. Because of the low population growth rate, he is concerned about the long-term outlook for the European economies. With potentially slower economic growth in Europe, the environment for equities may be less attractive. Therefore, Todd is considering reallocating some of the assets from Europe to Mexico. Data shows that the level of business investment has been around 24% of GDP, while that of China is more than 40% and that of European countries is close to 20%. The level of productivity for Europe as a whole is more than $60 per hour worked (as of 2018). Data also shows that labor productivity since 2001 in Mexico is growing at about 0.2% annual rate, below that of Germany, France, and Spain. Do you think that investment opportunities are favorable in Mexico? According to the OECD, the Mexican population increased by 1.1% in 2018, compared with a 0.2% increase in the European Union population (27 countries).

Solution:

Other than the higher population growth rate, the potential sources of growth for Mexico are not favorable. The level of business investment in Mexico is quite low, especially compared with China, and not much higher than that of many of the advanced economies in Europe. The level of labor productivity in Mexico is well below that in most European countries. This is not surprising given that the amount of capital per worker in Mexico is much lower than that in Europe.

What is surprising and of concern is the low rate of labor productivity growth in Mexico. This means that the rightward shift in the aggregate supply curve is greater for the European countries than for Mexico, despite the more favorable demographic trend in the country. In addition, it implies that there is more potential for expanding profit margins in Europe than in Mexico. Thus, the analysis of potential growth does not suggest a favorable outlook for equity returns in Mexico. In this simplified example, in the absence of more favorable considerations—for example, compelling equity valuations—Todd should decide not to reallocate assets from Europe to Mexico.

4.4 Measuring Sustainable Growth

Labor productivity data can be used to estimate the rate of sustainable growth of the economy. A useful way to describe potential GDP is as a combination of aggregate hours worked and the productivity of those workers:

$$\text{Potential GDP} = \text{Aggregate hours worked} \times \text{Labor productivity}$$

Transforming the above equation into growth rates, we obtain the following:

$$\begin{aligned}\text{Potential growth rate} = {} & \text{Long-term growth rate of aggregate hours worked} + \\ & \text{Long-term labor productivity growth rate}\end{aligned}$$

Thus, potential growth is a combination of the long-term growth rate of the aggregate hours worked and the long-term growth rate of labor productivity. Therefore, if aggregate hours worked are growing at 1% per year and productivity per worker is rising at 2% per year, then potential GDP (adjusted for inflation) is rising at 3% per year.

EXAMPLE 17

Estimating the Rate of Growth in Potential GDP

Exhibit 30 provides data on the growth rate in aggregate hours worked and productivity growth for Canada, Germany, Japan, and the United States. Estimate the growth rate of potential GDP for each country by averaging the growth rates for these variables since 2001. Also, describe the role that the labor input plays in determining potential growth for each country.

Exhibit 30 Aggregate Hours Worked and Productivity: Average Annual Growth Rate

	Aggregate Hours Worked		Productivity	
	2011–2018	**2001–2010**	**2011–2018**	**2001–2010**
Canada	0.9%	1.4%	1.0%	0.9%
Germany	0.5	0.2	1.1	0.9
Japan	0.3	−0.2	0.8	1.2
United States	0.7	0.7	1.5	0.8

Solution:

Potential GDP is calculated as the sum of the trend growth rate in the aggregate hours worked (labor input) and the trend growth rate in labor productivity. The growth in the labor input depends on the population growth rate, changes in the labor force participation rate, and changes in hours worked per person. Estimating based on the average for the period from 2001–2018 gives the following data:

	Projected Growth in Aggregate Hours Worked	Projected Growth in Labor Productivity	Projected Growth in Potential GDP
Canada	1.2%	0.9%	2.1%
Germany	0.3	1.0	1.3
Japan	0.0	1.0	1.0
United States	0.7	1.1	1.8

Note: Rounding used throughout.

Most of the difference between the growth rates in potential GDP among these countries can be explained by the labor input. The most significant result is the difference in the growth rate in aggregate hours in Germany and Japan, in contrast to that in the United States and Canada. The results suggest that Japan's sluggish economic growth is likely to continue because of the lack of growth in the labor input.

EXAMPLE 18

Prospects for Fixed-Income Investments

As a fixed income analyst for a large Canadian bank, you are preparing the 2019–2020 forecast and have just received the latest GDP forecast from the OECD for Canada, Germany, Japan, and the United States. The forecast is given in Exhibit 31, along with the potential GDP growth estimates that the analyst has calculated.

Exhibit 31 Projected vs. Potential GDP Growth

	Projected Average Annual GDP Growth (2019–2020)	Projected Growth in Potential GDP
Canada	4.0%	2.1%
Germany	1.5	1.3
Japan	0.0	1.0
United States	3.5	1.8

To evaluate the future prospects for fixed-income investments, analysts must estimate the future rate of inflation and assess the possibility of changes in monetary policy by the central bank. An important indicator for both of these factors is the degree of slack in the economy, which can be measured by comparing the growth rates of actual GDP and potential GDP.

Based on the information in Exhibit 31, evaluate the prospects for fixed-income investments in each country.

Solution:

In comparing the OECD forecast for GDP growth with the estimated growth rate in potential GDP, there are two cases to consider:

1. If actual GDP is growing at a faster rate than potential GDP, it signals growing inflationary pressures, and an increased likelihood that the central bank will raise interest rates.
2. If actual GDP is growing at a slower rate than potential GDP, it signals growing resource slack, less inflationary pressures, and an increased likelihood that the central bank will reduce rates or leave them unchanged.

Projected and potential GDP for the aforementioned countries can be compared referring to Exhibit 31.

The data suggest that inflationary pressure will grow in the United States and Canada and that both the Federal Reserve and the Bank of Canada will eventually raise interest rates. Thus, the environment for bond investing is not favorable in the United States and Canada, because bond prices are likely to decline.

With Germany growing at close to its potential rate of GDP growth, the rate of inflation should neither rise nor fall. Monetary policy is set by the European Central Bank (ECB), but data on the German economy play a big role in the ECB's decision. Based on the data in Exhibit 31, no change in ECB policy is likely. For bond investors, little change in bond prices is likely in Germany, so investors need to focus on the interest (coupon) income received from the bond.

Finally, growing resource slack in Japan will put downward pressure on inflation and may force the Bank of Japan to keep rates low. Bond prices should rise in this environment.

SUMMARY

- GDP is the market value of all final goods and services produced within a country during a given period.
- GDP can be valued by looking at either the total amount spent on goods and services produced in the economy or the income generated in producing those goods and services.
- GDP counts only final purchases of newly produced goods and services during the current period. Transfer payments and capital gains are excluded from GDP.
- Intermediate goods are excluded from GDP in order to avoid double counting.
- GDP can be measured either from the value of final output or by summing the value added at each stage of the production and distribution process. The sum of the value added by each stage is equal to the final selling price of the good.
- Nominal GDP is the value of production using the prices of the current year. Real GDP measures production using the constant prices of a base year. The GDP deflator equals the ratio of nominal GDP to real GDP.
- Households earn income in exchange for providing—directly or indirectly through ownership of businesses—the factors of production (labor, capital, and natural resources including land). From this income, they consume, save, and pay net taxes.
- Businesses produce most of the economy's output/income and invest to maintain and expand productive capacity. Companies retain some earnings but pay out most of their revenue as income to the household sector and as taxes to the government.
- The government sector collects taxes from households and businesses and purchases goods and services, for both consumption and investment, from the private business sector.
- International trade consists of exports and imports. The difference between the two is net exports. If net exports are positive (negative), then the country spends less (more) than it earns. Net exports are balanced by accumulation of either claims on the rest of the world (net exports > 0) or obligations to the rest of the world (net exports < 0).
- Capital markets provide a link between saving and investment in the economy.
- From the expenditure side, GDP includes personal consumption (C), gross private domestic investment (I), government spending (G), and net exports ($X - M$).

- The major categories of expenditure are often broken down into subcategories. Gross private domestic investment includes both investment in fixed assets (plant and equipment) and the change in inventories. In some countries, government spending on investment is separated from other government spending.
- National income is the income received by all factors of production used in the generation of final output. It equals GDP minus the **capital consumption allowance** (depreciation) and a statistical discrepancy.
- Personal income reflects pre-tax income received by households. It equals national income plus transfers minus undistributed corporate profits, corporate income taxes, and indirect business taxes.
- Personal disposable income equals personal income minus personal taxes.
- Private saving must equal investment plus the fiscal and trade deficits. That is, $S = I + (G - T) + (X - M)$.
- Consumption spending is a function of disposable income. The marginal propensity to consume represents the fraction of an additional unit of disposable income that is spent.
- Investment spending depends on the average interest rate and the level of aggregate income. Government purchases and tax policy are often considered to be exogenous variables determined outside the macroeconomic model. Actual taxes collected depend on income and are, therefore, endogenous—that is, determined within the model.
- Aggregate demand and aggregate supply determine the level of real GDP and the price level.
- The downward slope of the aggregate demand curve arises as the result of three effects: the wealth effect, the interest rate effect, and the real exchange rate effect. The curve is drawn assuming a constant money supply.
- The aggregate demand curve will shift if there is a change in a factor, other than price, that affects aggregate demand. These factors include household wealth, consumer and business expectations, capacity utilization, monetary policy, fiscal policy, exchange rates, and other nations' GDP.
- The aggregate supply curve is the relationship between the quantity of real GDP supplied and the price level, keeping all other factors constant. Movements along the supply curve reflect the effects of price on supply.
- The short-run aggregate supply curve is upward sloping because higher prices result in higher profits and induce businesses to produce more and laborers to work more. In the short run, some prices are sticky, implying that some prices do not adjust to changes in demand.
- In the long run, all prices are assumed to be flexible. The long-run aggregate supply curve is vertical because input costs adjust to changes in output prices, leaving the optimal level of output unchanged. The position of the curve is determined by the economy's level of potential GDP.
- The level of potential output, also called the full employment or natural level of output, is unobservable and difficult to measure precisely. This concept represents an efficient and unconstrained level of production at which companies have enough spare capacity to avoid bottlenecks and there is a balance between the pool of unemployed workers and the pool of job openings.
- The long-run aggregate supply curve will shift because of changes in labor supply, supply of physical and human capital, and productivity/technology.

- The short-run supply curve will shift because of changes in potential GDP, nominal wages, input prices, expectations about future prices, business taxes and subsidies, and the exchange rate.
- Short-term fluctuations in GDP are caused by shifts in aggregate demand and aggregate supply.
- When the level of GDP in the economy is below potential GDP, such a recessionary situation exerts downward pressure on the aggregate price level.
- When the level of GDP is above potential GDP, such an overheated situation puts upward pressure on the aggregate price level.
- Stagflation, a combination of high inflation and weak economic growth, is caused by a decline in short-run aggregate supply.
- Growth in real GDP measures how rapidly the total economy is expanding. Per capita GDP, defined as real GDP divided by population, reflects a country's standard of living. Real GDP growth rates and levels of per capita GDP vary widely among countries.
- The sources of economic growth include the supply of labor, the supply of physical and human capital, raw materials, and technological knowledge.
- Output can be described in terms of a production function. For example, $Y = A\boldsymbol{F}(L,K)$, where L is the quantity of labor, K is the capital stock, and A represents total factor productivity (primarily technology). The function $\boldsymbol{F}(\cdot)$ is assumed to exhibit constant returns to scale but diminishing marginal productivity for each input individually.
- Total factor productivity is a scale factor that reflects the portion of output growth that is not accounted for by changes in the capital and labor inputs. TFP is mainly a reflection of technological change.
- Based on a two-factor production function, Potential GDP growth = Growth in TFP + W_L (Growth in labor) + W_C (Growth in capital), where W_L and W_C $(= 1 - W_L)$ are the shares of labor and capital in GDP.
- Diminishing marginal productivity implies that
  - increasing the supply of some input(s) relative to other inputs will lead to diminishing returns and cannot be the basis for sustainable growth. In particular, long-term sustainable growth cannot rely solely on capital deepening (i.e., on increasing the stock of capital relative to labor).
  - given the relative scarcity and hence high productivity of capital in developing countries, the growth rate of developing countries should exceed that of developed countries.
- The sustainable rate of economic growth is measured by the rate of increase in the economy's productive capacity or potential GDP.
- The labor supply is determined by both population growth (including immigration) and the labor force participation rate. The capital stock in a country increases with investment. Correlation between long-run economic growth and the rate of investment is high.
- In addition to labor, capital, and technology, human capital—essentially, the quality of the labor force—and natural resources are important determinants of output and growth as is the state of public infrastructure, institutional, legal and political environment.
- Technological advances are discoveries that make it possible to produce more and/or higher-quality goods and services with the same resources or inputs. Technology is the main factor affecting economic growth in developed countries.

- Externalities and spillover effects are also recognized to play an important role in growth. Positive externalities associated with research and development and public education. Negative externalities such as pollution have a negative effect on growth.
- Climate change is closely linked to economic growth as carbon dioxide emissions rise with economic growth. These emissions impose a cost on the global economy in terms of economic activity, food production, health and habitability.
- The sustainable rate of growth in an economy is determined by the growth rate of the labor supply plus the growth rate of labor productivity.

REFERENCES

Case, K., J. Quigley, and R. Shiller. 2005. "*Comparing Wealth Effects: The Stock Market versus the Housing Market.*" Advances in Macroeconomics, vol. 5, no. 1.

Funke, N. 2004. "*Is There a Stock Market Wealth Effect in Emerging Markets?*" International Monetary Fund (March).

United Nations, Department of Economic and Social Affairs, Population Division. 2019). World Population Prospects 2019, Online Edition. Rev. 1.

PRACTICE PROBLEMS

1 Which of the following statements is the *most* appropriate description of gross domestic product (GDP)?

A The total income earned by all households, firms, and the government whose value can be verified

B The total amount spent on all final goods and services produced within the economy during a given period

C The total market value of resalable and final goods and services produced within the economy during a given period

2 The component *least likely* to be included in a measurement of gross domestic product (GDP) is:

A the value of owner occupied rent.

B the annual salary of a local police officer.

C environmental damage caused by production.

3 Which of the following conditions is *least likely* to increase a country's GDP?

A An increase in net exports

B Increased investment in capital goods

C Increased government transfer payments

4 Which of the following would be included in Canadian GDP for a given year? The market value of:

A wine grown in Canada by US citizens.

B electronics made in Japan and sold in Canada.

C movies produced outside Canada by Canadian filmmakers.

5 Suppose a painting is produced and sold in 2018 for £5,000. The expenses involved in producing the painting amounted to £2,000. According to the sum-of-value-added method of calculating GDP, the value added by the final step of creating the painting was:

A £2,000.

B £3,000.

C £5,000.

6 A GDP deflator less than 1 indicates that an economy has experienced:

A inflation.

B deflation.

C stagflation.

7 The *most* accurate description of nominal GDP is:

A a measure of total expenditures at current prices.

B the value of goods and services at constant prices.

C a measure to compare one nation's economy to another.

8 From the beginning to the ending years of a decade, the annual value of final goods and services for country *X* increased from €100 billion to €300 billion. During that period, the GDP deflator increased from 111 to 200. Over the decade, real GDP for country *X* increased by approximately:

A 50%.

© 2020 CFA Institute. All rights reserved.

B 67%.

C 200%.

9 If the GDP deflator values for Year 1 and Year 3 were 190 and 212.8, respectively, which of the following *best* describes the annual growth rate of the overall price level?

A 5.8%.

B 6%.

C 12%.

10 The numerator of the GDP price deflator reflects:

A the value of base year output at current prices.

B the value of current year output at current prices.

C the value of current year output at base year prices.

11 Consider the following data for a hypothetical country:

Account name	Amount ($ trillions)
Consumption	15.0
Capital consumption allowance	1.5
Government spending	3.8
Imports	1.7
Gross private domestic investment	4.0
Exports	1.5

Based only on the data given, the gross domestic product and national income are respectively *closest* to:

A 21.1 and 20.6.

B 22.6 and 21.1.

C 22.8 and 20.8.

12 In calculating personal income for a given year, which of the following would *not* be subtracted from national income?

A Indirect business taxes

B Undistributed corporate profits

C Unincorporated business net income

13 Equality between aggregate expenditure and aggregate output implies that the government's fiscal deficit must equal:

A Private saving – Investment – Net exports.

B Private saving – Investment + Net exports.

C Investment – Private saving + Net exports.

14 Because of a sharp decline in real estate values, the household sector has increased the fraction of disposable income that it saves. If output and investment spending remain unchanged, which of the following is *most likely*?

A A decrease in the government deficit

B A decrease in net exports and increased capital inflow

C An increase in net exports and increased capital outflow

15 In explaining the downward slope of the aggregate demand curve, the interest rate effect suggests that a higher price level will:

A lead to lower interest rates, which will lead to higher levels of consumption and investment.

B lead to higher interest rates, which will lead to lower levels of consumption and investment.

C not affect interest rates and therefore will lead to a higher level of aggregate demand.

16 One of the reasons aggregate demand curve is downward sloping is because a lower price level means that the real exchange rate:

A depreciates, making domestic goods cheaper in other countries and imports less competitive, resulting in a higher level of net exports.

B appreciates, making domestic goods more expensive in other countries and imports more competitive, resulting in a higher level of net exports.

C appreciates, making the country's exports and imports less competitive and leading to lower net exports.

17 The full employment, or natural, level of output is *best* described as:

A the maximum level obtainable with existing resources.

B the level at which all available workers have jobs consistent with their skills.

C a level with a modest, stable pool of unemployed workers transitioning to new jobs.

18 Which of the following *best* describes the aggregate supply curve in the short run (one to two years)? The short-run aggregate supply curve is:

A flat, because output is more flexible than prices in the short run.

B vertical, because wages and other input prices fully adjust to the price level.

C upward sloping, because input prices do not fully adjust to the price level in the short run.

19 If wages were automatically adjusted for changes in the price level, the short-run aggregate supply curve would *most likely* be:

A flatter.

B steeper.

C unchanged.

20 The *least likely* cause of a decrease in aggregate demand is:

A higher taxes.

B a weak domestic currency.

C a fall in capacity utilization.

21 Which of the following is *most likely* to cause the long-run aggregate supply curve to shift to the left?

A Higher nominal wages

B A decline in productivity

C An increase in corporate taxes

22 Increased household wealth will *most likely* cause an increase in:

A household saving.

B investment expenditures.

C consumption expenditures.

23 The *most likely* outcome when both aggregate supply and aggregate demand increase is:

A a rise in inflation.

B higher employment.

C an increase in nominal GDP.

24 Which of the following is *least likely* to be caused by a shift in aggregate demand?

A Stagflation

B A recessionary gap

C An inflationary gap

25 Following a sharp increase in the price of energy, the overall price level is *most likely* to rise in the short run:

A and remain elevated indefinitely unless the central bank tightens.

B but remain unchanged in the long run unless the money supply is increased.

C and continue to rise until all prices have increased by the same proportion.

26 Among developed economies, which of the following sources of economic growth is *most likely* to explain superior growth performance?

A Technology

B Capital stock

C Labor supply

27 Which of the following can be measured directly?

A Potential GDP

B Labor productivity

C Total factor productivity

28 The sustainable growth rate is *best* estimated as:

A the weighted average of capital and labor growth rates.

B growth in the labor force plus growth of labor productivity.

C growth in total factor productivity plus growth in the capital-to-labor ratio.

29 In the neoclassical or Solow growth model, an increase in total factor productivity reflects an increase in:

A returns to scale.

B output for given inputs.

C the sustainable growth rate.

The following information relates to Questions 30–31

An economic forecasting firm has estimated the following equation from historical data based on the neoclassical growth model:

$$\text{Potential output growth} = 1.5 + 0.72 \times \text{Growth of labor} + 0.28 \times \text{Growth of capital}$$

30 The intercept (1.5) in this equation is *best* interpreted as:

A the long-run sustainable growth rate.

B the growth rate of total factor productivity.

C above-trend historical growth that is unlikely to be sustained.

31 The coefficient on the growth rate of labor (0.72) in this equation is *best* interpreted as:

A the labor force participation rate.

B the marginal productivity of labor.

C the share of income earned by labor.

---

32 Convergence of incomes over time between emerging market countries and developed countries is *most likely* the result of:

A total factor productivity.

B diminishing marginal productivity of capital.

C the exhaustion of non-renewable resources.

SOLUTIONS

1 B is correct. GDP is the total amount spent on all final goods and services produced within the economy during a specific period.

2 C is correct. Byproducts of production processes that have no explicit market value are not included in GDP.

3 C is correct. Government transfer payments, such as unemployment compensation or welfare benefits, are excluded from GDP.

4 A is correct. Canadian GDP is the total market value of all final goods and services produced during a given period within Canada. The wine was produced in Canada and counts toward Canadian GDP.

5 B is correct. The value added by the artist is £5,000 – £2,000 = £3,000.

6 B is correct. The GDP deflator = Nominal GDP/Real GDP. To obtain a ratio less than 1, real GDP must exceed nominal GDP, which indicates that prices have decreased and, accordingly, deflation has occurred.

7 A is correct. Nominal GDP is defined as the value of goods and services measured at current prices. Expenditure is used synonymously with the value of goods and services because aggregate expenditures must equal aggregate output of an economy.

8 B is correct. Real GDP in the first year was €100 billion/1.11 = €90 billion, and in the last year it was €300 billion/2.00 = €150 billion. Thus, (€150 – €90)/€90 = 0.67, or 67%.

9 A is correct: $(212.8/190)^{1/2} - 1 = 0.0583$, or 5.8%.

10 B is correct.

$$\text{GDP deflator} = \frac{\text{Value of current year output at current year prices}}{\text{Value of current year output at base year prices}} \times 100$$

11 B is correct. GDP = Consumption + Gross private domestic investment + Government spending + (Exports – Imports) = 15 + 4 + 3.8 + (1.5 – 1.7) = 22.6. National income = GDP – CCA = 22.6 – 1.5 = 21.1

12 C is correct. Unincorporated business net income, also known as proprietor's income, is included in personal income.

13 A is correct. The fundamental relationship among saving, investment, the fiscal balance, and the trade balance is $S = I + (G - T) + (X - M)$. This form of the relationship shows that private saving must fund investment expenditures, the government fiscal balance, and net exports (= net capital outflows). Rearranging gives $G - T = (S - I) - (X - M)$. The government's fiscal deficit $(G - T)$ must equal the private sector's saving/investment balance $(S - I)$ minus net exports.

14 C is correct. The fundamental relationship among saving, investment, the fiscal balance, and the trade balance is $S = I + (G - T) + (X - M)$. Given the levels of output and investment spending, an increase in saving (reduction in consumption) must be offset by either an increase in the fiscal deficit or an increase in net exports. Increasing the fiscal deficit is not one of the choices, so an increase in net exports and corresponding increase in net capital outflows (increased lending internationally and/or increased international purchases of assets) is the correct response.

15 B is correct. Assuming a fixed supply of money, a higher price level will lead to a higher price of money. Because the price of money is the interest rate, as the price level increases, the interest rate increases. Higher interest rates will lead to lower consumption and investment.

16 A is correct. A lower price level results in a weaker real exchange rate, making domestic goods cheaper to people in other countries and imports less competitive, resulting in a higher level of net exports.

17 C is correct. At the full employment, or natural, level of output, the economy is operating at an efficient and unconstrained level of production. Companies have enough spare capacity to avoid bottlenecks, and there is a modest, stable pool of unemployed workers (job seekers equal job vacancies) looking for and transitioning into new jobs.

18 C is correct. Because of long-term contracts and other rigidities, wages and other input costs do not fully adjust to changes in the price level in the short run. Given input prices, firms respond to output price changes by expanding or contracting output to maximize profit. Hence, the SRAS is upward sloping.

19 B is correct. The slope of the short-run aggregate supply curve reflects the extent to which wages and other input costs adjust to the overall price level. Automatic adjustment of wages would mitigate the impact of price changes on profitability. Hence, firms would not adjust output as much in response to changing output prices—the SRAS curve would be steeper.

20 B is correct. A weak domestic currency will result in an increase in aggregate demand at each price level—a rightward shift in the AD curve. A weaker currency will cause a country's exports to be cheaper in global markets. Conversely, imports will be more expensive for domestic buyers. Hence, the net exports component of aggregate demand will increase.

21 B is correct. Productivity measures the efficiency of labor and is the amount of output produced by workers during a given period. A decline in productivity implies decreased efficiency. A decline in productivity increases labor costs, decreases profitability, and results in lower output at each output price level—a leftward shift in both the short-run and long-run aggregate supply curves.

22 C is correct. The wealth effect explains the impact of increases or decreases in household wealth on economic activity. Household wealth includes financial and real assets. As asset values increase, consumers save less and spend more out of current income because they will still be able to meet their wealth accumulation goals. Therefore, an increase in household wealth results in a rightward shift in the aggregate demand curve.

23 B is correct. Higher aggregate demand and higher aggregate supply raise real GDP and lower unemployment, meaning employment levels increase.

24 A is correct. Stagflation occurs when output is declining and prices are rising. This dynamic most likely results from a decline in aggregate supply—a leftward shift of the SRAS curve. Depending on the source of the shift, the LRAS may shift too.

25 B is correct. An increase in energy prices will shift the short-run aggregate supply curve (SRAS) to the left, reducing output and increasing prices. If the aggregate demand curve does not change—in particular, if the central bank does not expand the money supply—slack in the economy will put downward pressure on in input prices, shifting the SRAS back to its original position. In the long run, the price level will be unchanged.

26 A is correct. Technology is the most important factor affecting economic growth for developed countries. Technological advances are very important because they allow an economy to overcome the limits imposed by diminishing marginal returns.

27 B is correct. Labor productivity can be directly measured as output/hour.

28 B is correct. Output growth is equal to the growth rate of the labor force plus the growth rate of labor productivity—that is, output per worker. Unlike total factor productivity, output per worker is observable, so this measure is the most practical way to approach estimation of sustainable growth.

29 B is correct. Total factor productivity is a scale factor primarily reflecting technology. An increase in TFP means that output increases for any level of factor inputs.

30 B is correct. The estimated equation is the standard Solow growth accounting equation. The intercept is the growth rate of total factor productivity.

31 C is correct. In the standard Solow growth accounting equation, the coefficient on each factor's growth rate is its share of income.

32 B is correct. Diminishing marginal productivity of capital means that as a country accumulates more capital per worker, the incremental boost to output declines. Thus, all else equal, economies grow more slowly as they become more capital intensive. Given the relative scarcity and hence high marginal productivity of capital in developing countries, these economies tend to grow more rapidly than developed countries. This dynamic leads to convergence in income levels over time.

READING

11

Understanding Business Cycles

by Michele Gambera, PhD, CFA, Milton Ezrati, and Bolong Cao, PhD, CFA

Michele Gambera, PhD, CFA, is with UBS Asset Management and the University of Illinois at Urbana-Champaign (USA). Milton Ezrati (USA). Bolong Cao, PhD, CFA, is at Ohio University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the business cycle and its phases;
☐	**b.** describe credit cycles;
☐	**c.** describe how resource use, consumer and business activity, housing sector activity, and external trade sector activity vary as an economy moves through the business cycle;
☐	**d.** describe theories of the business cycle;
☐	**e.** interpret a set of economic indicators, and describe their uses and limitations;
☐	**f.** describe types of unemployment, and compare measures of unemployment;
☐	**g.** explain inflation, hyperinflation, disinflation, and deflation;
☐	**h.** explain the construction of indexes used to measure inflation;
☐	**i.** compare inflation measures, including their uses and limitations;
☐	**j.** contrast cost-push and demand-pull inflation.

INTRODUCTION

1

A typical economy’s output of goods and services fluctuates around its longer-term path. We now turn our attention to those recurring, cyclical fluctuations in economic output. Some of the factors that influence short-term changes in the economy—such as changes in population, technology, and capital—are the same as those that affect long-term sustainable economic growth. But forces that cause shifts in aggregate

© 2021 CFA Institute. All rights reserved.

demand and aggregate supply curves—such as expectations, political developments, natural disasters, and fiscal and monetary policy decisions—influence economies particularly in the short run.

We first describe a typical business cycle and its phases. While each cycle is different, analysts and investors need to be familiar with the typical cycle phases and what they mean for the expectations and decisions of businesses and households that influence the performance of sectors and companies. These behaviors also impact financial conditions and risk appetite, thus impacting the setting of expectations and choices of portfolio exposures to different investment sectors or styles.

In the sections that follow, we describe credit cycles, introduce several theories of business cycles, and explain how different economic schools of thought interpret the business cycle and their recommendations with respect to it. We also discuss variables that demonstrate predictable relationships with the economy, focusing on those whose movements have value in predicting the future course of the economy. We then proceed to explain measures and features of unemployment and inflation.

2

OVERVIEW OF THE BUSINESS CYCLE

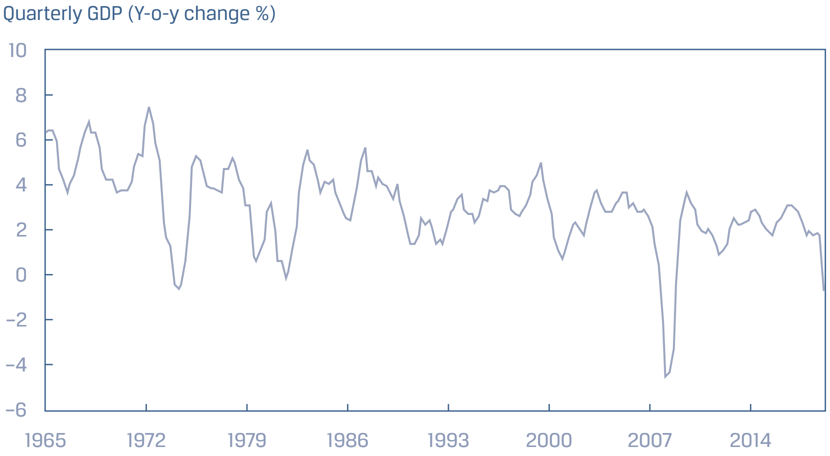
a describe the business cycle and its phases

Business cycles are recurrent expansions and contractions in economic activity affecting broad segments of the economy. Burns and Mitchell (1946) define the business cycle as follows:

> Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; this sequence of events is recurrent but not periodic; in duration, business cycles vary from more than one year to 10 or 12 years.

This definition is rich with important insights. First, business cycles are typical of economies that rely mainly on business enterprises—therefore, not agrarian societies or centrally planned economies. Second, a cycle has an expected sequence of phases, alternating between expansion and contraction, or upswings and downturns. Third, such phases occur at about the same time throughout the economy. Finally, cycles are recurrent; they happen again and again over time but not in a periodic way; they do not all have the exact same intensity and duration. Exhibit 1 provides an illustration of the pattern of economic growth rate in developed markets.

Exhibit 1 Fluctuations of Growth in OECD Countries over Time



Note: The Organisation for Economic Co-Operation and Development (OECD) includes more than 30 large countries.
Source: OECD.Stat (htttps://stats.oecd.org/) YoY change in quarterly GDP in OECD countries.

Burns' and Mitchell's definition remains helpful. History never repeats itself in quite the same way, but it certainly does offer patterns that can be used when analyzing the present and forecasting the future. Business cycle analysis is a wide-ranging topic with conflicting perspectives held by industry participants.

2.1 Phases of the Business Cycle

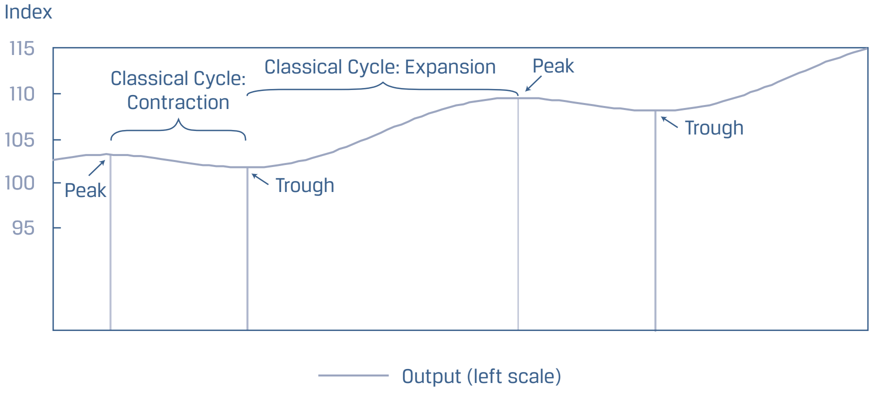
Business cycles are recurring sequences of alternating upswings and downturns. There are various ways to break the business cycle into phases. The most obvious is to divide it into two primary segments: the expansion, or the upswing, and the contraction, or the downturn, with two key turning points, or peaks and troughs (see Exhibits 2 and 3). These two periods are fairly easy to identify in retrospect. Subdividing the cycle more finely, however, becomes ambiguous, even in retrospect, because it requires identifying more nuanced changes, such as acceleration or deceleration of growth without a change in its direction. It is thus useful to divide the cycle into several phases distinguished through both economic and financial market characteristics. Our focus is on economic characteristics of the different phases, but we will highlight their implication for the behavior of different segments of the financial markets along the way.

The timing of these periods will depend on the type of cycle. Before moving on to the description of the four distinct phases to which we will refer in the subsequent sections, we first explain the different cycle concepts that analysts should be aware of given the range of different opinions, interpretations, and descriptions that practitioners use.

2.1.1 *Types of Cycles*

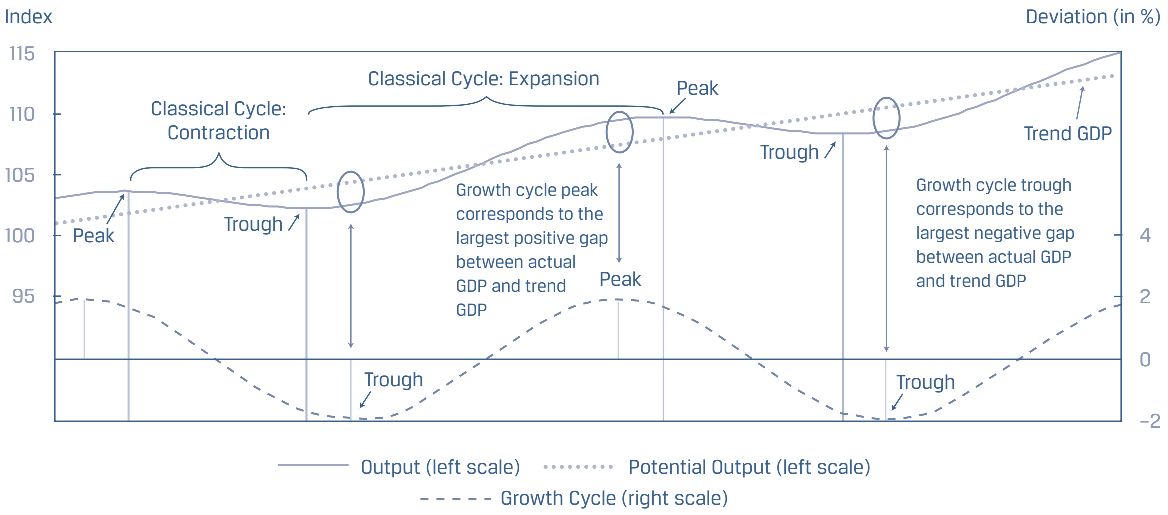
- **Classical cycle** refers to fluctuations in the level of economic activity (e.g., measured by GDP in volume terms). The contraction phases between peaks and troughs are often short, while expansion phases are much longer. Exhibit 2 shows the classical cycle of economic activity. In practice, the classical cycle is not used extensively by academics and practitioners because it does not easily allow the breakdown of movements in GDP between short-term fluctuations and long-run trends. In addition, an absolute decline in activity between peaks and troughs does not occur frequently.

Exhibit 2 Classical Cycle



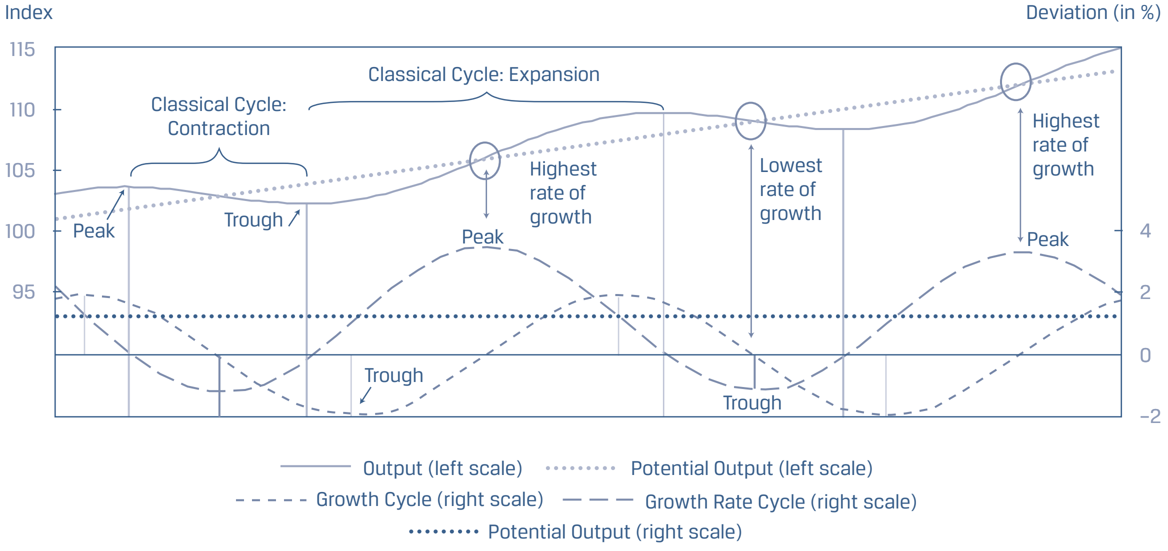
- **Growth cycle** refers to fluctuations in economic activity around the long-term potential or trend growth level. The focus is on how much actual economic activity is below or above trend growth in economic activity. The dashed "wave" in the lower part of Exhibit 3 captures the fluctuation of actual activity from trend growth activity. Exhibit 3 shows "gaps" between actual and trend output. The growth cycle definition comes closest to how mainstream economists think: It dissects overall economic activity into a part driven by long-run trends and a part reflecting short-run fluctuations. Compared to the classical view of business cycles, peaks are generally reached earlier and troughs later in time. The time periods below and above trend growth are of similar length.

Exhibit 3 Classical and Growth Cycles



- **Growth rate cycle** refers to fluctuations in the growth rate of economic activity (e.g., GDP growth rate). Peaks and troughs are mostly recognized earlier than when using the other two definitions (see Exhibit 4). An advantage is that there is not necessarily a need to first estimate a long-run growth path. Nevertheless, economists often refer to economic growth being above or below potential growth rate, reflecting upswings or downturns.

Exhibit 4 Classical, Growth, and Growth Rate Cycles



Notes: The vertical lines indicate troughs and peaks when using either the classical, growth, or growth rate cycle definition of a business cycle. The growth cycle reflects the percentage deviation of output relative to its trend. The growth rates in the growth rate cycle are calculated as annualized month-over-month growth rates.

2.1.2 *Practical Issues*

In practice, the definitions of a business cycle are used interchangeably, which often causes confusion with regard to how one labels the phases and their timing. The classical cycle definition is rarely used. In line with how most economists and practitioners view the cycle, we will generally be using the growth cycle concept in which business cycles can be thought of as fluctuations around potential output (the trend in potential output is shown as the upward sloping dotted line in Exhibit 3).

2.1.3 *Four Phases of the Cycle*

The overall business cycle can be split into four phases:

> **Recovery**: The economy is going through the "trough" of the cycle, where actual output is at its lowest level relative to potential output. Economic activity, including consumer and business spending, is below potential but is starting to increase, closing the negative output gap.
>
> **Expansion**: The recovery gathers pace, output increases, and the rate of growth is above average. Actual output rises above potential output, and the economy enters the so-called "boom" phase. Consumers increase spending, and companies increase production, employment, and investment. Prices and interest rates may start increasing. As the expansion continues, the economy may start to experience shortages in factors of production. Overinvestment in productive capacity may lead companies to reduce further investment spending.
>
> **Slowdown**: Output of the economy reaches its highest level relative to potential output (largest positive output gap). The growth rate begins to slow relative to potential output growth, and the positive output gap begins to narrow. Consumers remain optimistic about the economy, but companies may rely on overtime rather than using new hires to meet demand. Inflation slows at some point, and price levels may decrease.
>
> **Contraction**: Actual economic output falls below potential economic output. Consumer and business confidence declines. Companies reduce costs by eliminating overtime and reducing employment. The economy may experience declines in absolute economic activity, a recession, or if the fall in activity is particularly large, a depression. If the decline is moderate, this phase tends to be shorter than the expansion phase.

Exhibit 5 provides a summary of the key characteristics of each phase and describes how several important economic variables evolve through the course of a business cycle.

Exhibit 5 Business Cycle Phase Characteristics

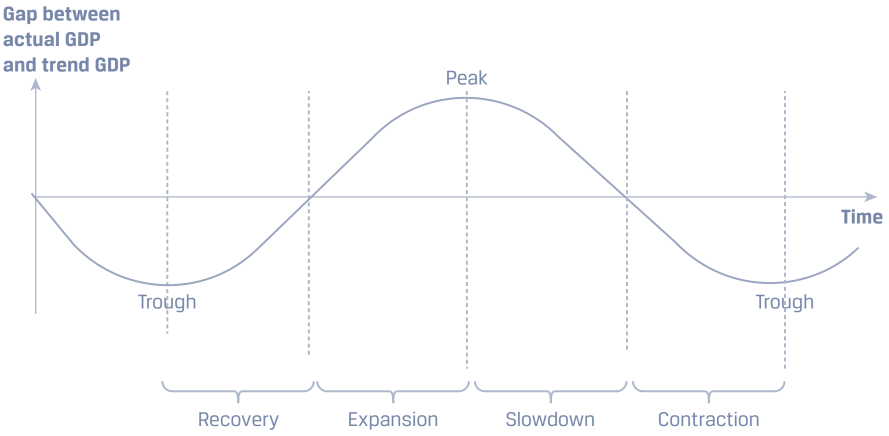


Exhibit 5 (Continued)

Phase	Recovery	Expansion	Slowdown	Contraction
Description	Economy going through a trough. Negative output gap starts to narrow.	Economy enjoying an upswing. Positive output gap opens.	Economy going through a peak. Positive output gap starts to narrow.	Economy weakens and may go into a recession. Negative output gap opens.
Activity levels – consumers and businesses	Activity levels are below potential but start to increase.	Activity measures show above-average growth rates.	Activity measures are above average but decelerating. Moving to below-average rates of growth.	Activity measures are below potential. Growth is lower than normal.
Employment	Layoffs slow. Businesses rely on overtime before moving to hiring. Unemployment remains higher than average.	Businesses move from using overtime and temporary employees to hiring. Unemployment rate stabilizes and starts falling.	Business continue hiring but at a slower pace. Unemployment rate continues to fall but at decreasing rates.	Businesses first cut hours, eliminate overtime, and freeze hiring, followed by outright layoffs. Unemployment rate starts to rise.
Inflation	Inflation remains moderate.	Inflation picks up modestly.	Inflation further accelerates.	Inflation decelerates but with a lag.

2.2 Leads and Lags in Business and Consumer Decision Making

The behavior of businesses and households is key to the cycle and frequently incorporates leads and lags relative to what are established as turning points. For example, at the beginning of an expansion phase, companies may want to fully use their existing workforce and wait to hire new employees until they are sure that the economy is indeed growing. However, gradually all economic variables are going to revert toward their normal range of values (e.g., GDP growth will be close to potential, or average, growth).

2.3 Market Conditions and Investor Behavior

Many economic variables and sectors of the economy have distinctive cyclical patterns. Knowledge of these patterns can offer insight into likely cyclical directions overall, or it can be particularly applicable to an investment strategy that requires more specific rather than general cyclical insights for investment success.

Recovery phase When asset markets expect the end of a recession and the beginning of an expansion phase, risky assets will be repriced upward. When an expansion is expected, the markets will start incorporating higher profit expectations into the prices of corporate bonds and stocks. Typically, equity markets will hit a trough about three to six months before the economy bottoms and well before the economic indicators turn up. Indeed, as we will see later, the equity market is classified as a leading indicator of the economy.

Expansion phase When an economy's expansion is well-established, a later part of an expansion, referred to as a "**boom**," often follows. The boom is an expansionary phase characterized by economic growth "testing the limits" of the economy, strong confidence, profit, and credit growth. For example, companies may expand so much that they have difficulty finding qualified workers and will compete with other prospective employers by raising wages and continuing to expand capacity, relying on strong cash flows and borrowing. The government and/or central bank may step in if it is concerned about the economy overheating.

Slowdown phase During the boom, the riskiest assets will often have substantial price increases. Safe assets, such as government bonds that were more highly prized during recession, may have lower prices and thus higher yields. In addition, investors may fear higher inflation, which also contributes to higher nominal yields.

Contraction phase During contraction, investors place relatively high values on such safer assets as government securities and shares of companies with steady (or growing) positive cash flows, such as utilities and producers of staple goods. Such preferences reflect the fact that the marginal utility of a safe income stream increases in periods when employment is insecure or declining.

EXAMPLE 1

When Do Recessions Begin and End?

A simple and commonly referred to rule is the following: A recession has started when a country or region experiences two consecutive quarters of negative real GDP growth. Real GDP growth is a measure of the "real" or "inflation-adjusted" growth of the overall economy. This rule can be misleading because it does not indicate a recession if real GDP growth is negative in one quarter, slightly positive the next quarter, and again negative in the next quarter. Many analysts question this result. This issue is why some countries have statistical and economic committees that apply the principles stated by Burns and Mitchell to several macroeconomic variables—not just real GDP growth—as a basis to identify business cycle peaks and troughs. The National Bureau of Economic Research (NBER) is an organization that dates business cycles in the United States. Interestingly, the economists and statisticians on NBER's Business Cycle Dating Committee analyze numerous time series of data focusing on employment, industrial production, and sales. Because the data are available with a delay (preliminary data releases can be revised even several years after the period they refer to), it also means that the Committee's determinations may take place well after the business cycle turning points have occurred. As we will see later in the reading, there are practical indicators that may help economists understand in advance if a cyclical turning point is about to happen.

1 Which of the following rules is *most likely* to be used to determine whether the economy is in a recession?
   - A The central bank has run out of foreign reserves.
   - B Real GDP has two consecutive quarters of negative growth.
   - C Economic activity experiences a significant decline in two business sectors.

2 Suppose you are interested in forecasting earnings growth for a company active in a country where no official business cycle dating committee (such as the NBER) exists. The variables you are *most likely* to consider to identify peaks and troughs of a country's business cycle are:
   - A inflation, interest rates, and unemployment.
   - B stock market values and money supply.
   - C unemployment, GDP growth, industrial production, and inflation.

Solution to 1:

B is correct. GDP is a measure of economic activity for the whole economy. Changes in foreign reserves or a limited number of sectors may not have a material impact on the whole economy.

Solution to 2:

C is correct. Unemployment, GDP growth, industrial production, and inflation are measures of economic activity. The discount rate, the monetary base, and stock market indexes are not direct measures of economic activities. The first two are determined by monetary policy, which react to economic activities, whereas the stock market indexes tend to be forward looking or leading indicators of the economy.

CREDIT CYCLES AND THEIR RELATIONSHIP TO BUSINESS CYCLES

3

b describe credit cycles

Whereas business cycles mostly use GDP as a measure of economic activity, a literature has emerged in which cyclical developments of financial variables are analyzed separately. This is most commonly done in terms of credit and property prices. Credit cycles describe the changing availability—and pricing—of credit. They describe growth in private sector credit (availability and usage of loans), which is essential for business investments and household purchases of real estate. They are therefore connected to real economic activity captured by business cycles that describe fluctuations in real GDP.

When the economy is strong or improving, the willingness of lenders to extend credit, and on favorable terms, is high. Conversely, when the economy is weak or weakening, lenders pull back, or "tighten" credit, by making it less available and more expensive. This frequently contributes to the decline of such asset values as real estate, causing further economic weakness and higher defaults. This is because of the importance of credit in the financing of construction and the purchase of property. Credit cycles are a subset of a wider family of so-called financial cycles, a topic that goes beyond the scope of our coverage.

3.1 Applications of Credit Cycles

Financial factors were for a long time not prominent on the radar screens of macroeconomists. Monetary and financial phenomena were largely seen as a veil that could be ignored when trying to understand the economy. But loose private sector credit is considered to have contributed to a number of financial crises, such as the Latam crisis of the 1980s; the Mexican, Brazilian, and Russian crises of the 1990s; the Asian crisis of 1997–1998; and the global financial crisis of 2008–2009. Loose credit conditions often lead to asset price and real estate bubbles that burst when capital market outflows and drawdowns occur mostly due to weaker fundamentals.

It is recognized that in a world with financial frictions, business cycles can be amplified, with deeper recessions and more extensive expansions because of changes in access to external financing. In line with this, it is found that the duration and magnitude of recessions and recoveries are often shaped by linkages between business and credit cycles (Claessens, Kose, and, Terrones 2012). In particular, recessions accompanied by financial disruption episodes (notably, house and equity price busts), tend to be longer and deeper. Recoveries combined with rapid growth in credit, risk-taking, and house prices tend to be stronger.

However, whereas financial variables tend to co-vary closely with each other and can often help explain the size of an economic expansion or contraction, they are not always synchronized with the traditional business cycle. Credit cycles tend to be longer, deeper, and sharper than business cycles. Although the length of a business cycle varies from peak to trough, the average length of a credit cycle is mostly found to be longer than that of the business cycle. Exhibit 6 illustrates how credit cycles can be visualized.

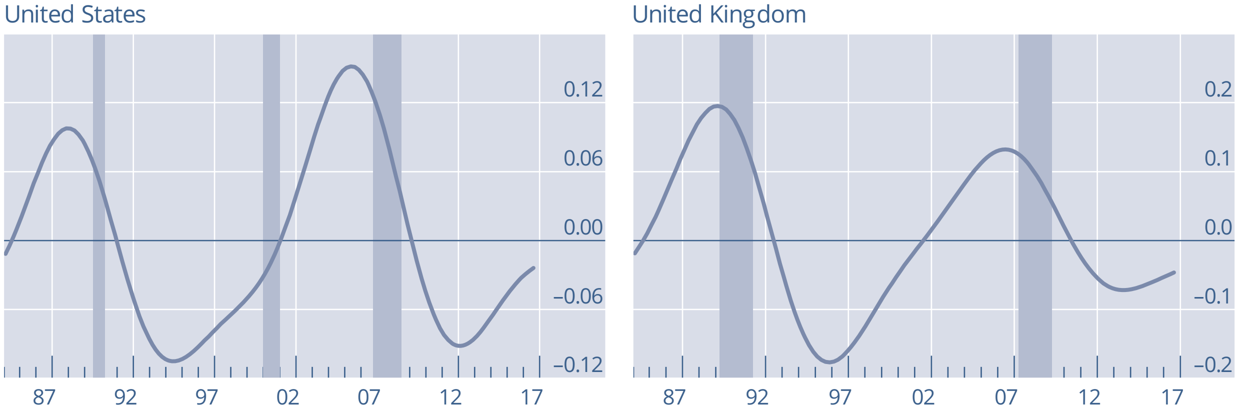
VISUALIZING FINANCIAL CYCLES

In an October 2019 Working Paper titled "Predicting Recessions: Financial Cycle versus Term Spread," the Bank for International Settlements (BIS) provided a visual presentation of credit cycles, which is reproduced in Exhibit 6. It shows that such cycles tend to boom before recessions.

Exhibit 6

Financial cycles tend to boom ahead of recessions[1]

Graph 1



The shaded areas represent recessions.

[1] Financial cycles are measured by the composite financial cycle proxy calculated from frequency-based (bandpass) filters capturing medium-term cycles in real credit, the credit-to-GDP ratio and real house prices.

Source: BIS Material (available on the BIS website free of charge: www.bis.org).

Notes: Credit cycles are measured by a (composite) proxy calculated from variables that include credit-to-GDP ratio and real house prices. The axis on the right shows the year-on-year change in the proxy.

3.2 Consequences for Policy

Investors pay attention to the stage in the credit cycle because (1) it helps them understand developments in the housing and construction markets; (2) it helps them assess the extent of business cycle expansions as well as contractions, particularly the severity of a recession if it coincides with the contraction phase of the credit cycle; and (3) it helps them better anticipate policy makers' actions. Whereas monetary and fiscal policy traditionally concentrate on reducing the volatility of business cycles, macro-prudential stabilization policies that aim to dampen financial booms have gained importance. This is further stressed by findings that strong peaks in credit cycles are closely associated with subsequent systemic banking crises (Schüler, Hiebert, and Peltonen 2017).

EXAMPLE 2

A senior portfolio manager at Carnara Asset Management explains her analysis of business cycles to a junior analyst. She makes two statements:

Statement I	Business cycles measure activity by GDP, whereas credit cycles combine a range of financial variables, such as the amount of and pricing of credit.
Statement II	Credit cycles and business cycles are unrelated and serve different purposes.

A Only Statement I is true.

B Only Statement II is true

C Both statements are true

Solution:

A is correct. Only Statement I is true. Statement II is not true because researchers have found linkages between financial and business cycles that help explain the magnitude of business cycle expansions and contractions depending on the state of the credit cycle.

BUSINESS CYCLE FLUCTUATIONS FROM A FIRM'S PERSPECTIVE

4

c describe how resource use, consumer and business activity, housing sector activity, and external trade sector activity vary as an economy moves through the business cycle

This section provides a broad overview of how the use of resources needed to produce goods and services typically evolves during a business cycle. We start by focusing on circumstances of firms and explore some of the links between fluctuations in inventory, employment, and investment in physical capital with economic fluctuations.

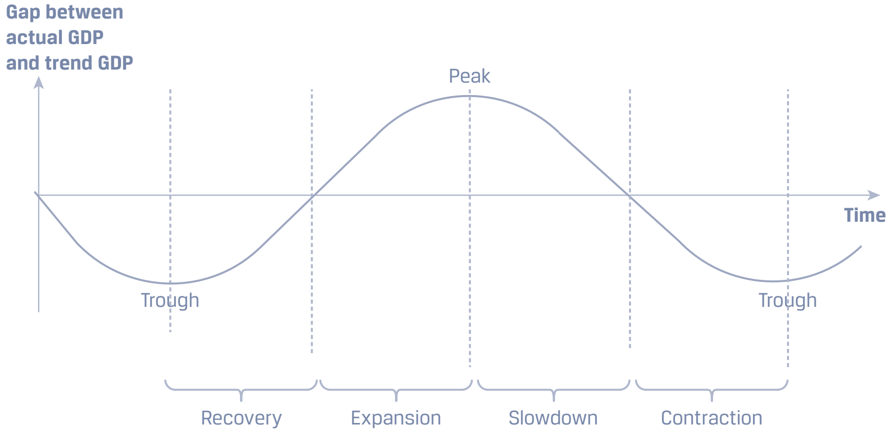
4.1 The Workforce and Company Costs

The pattern of hiring and employment is shown in Exhibit 7. When the economy enters contraction, companies reduce costs and eliminate overtime. They may try to retain workers rather than reducing employment only to increase it later. Finding and training new workers is costly, and it may be more cost efficient to keep workers on the payroll even if they are not fully utilized. Companies may also benefit from an implicit bond of loyalty between a company and its workers, boosting productivity in the process. In prolonged contractions, companies will start reducing costs more aggressively—terminating consultants, advertising campaigns, and workers beyond the strict minimum. Capacity utilization will be low, and few companies will invest in new equipment. Companies will try to liquidate their inventories of unsold products. In addition, banks will be reluctant to lend because bankruptcy risks are perceived to be higher, adding to the weakness in the economy.

Decreases in aggregate demand are likely to depress wages or wage growth as well as prices of inputs and capital goods. After a while, all of these input prices will be relatively low. In addition, interest rates may be cut to try to revive the economy.

As prices and interest rates decrease, consumers and companies may begin to purchase more and aggregate demand may begin to rise. This stage is the turning point of the business cycle: Aggregate demand starts to increase and economic activity increases.

Exhibit 7 Business Cycle Phases—Levels of Employment



Phase	Recovery	Expansion	Slowdown	Contraction
Description of activity levels	Economy starts at trough and output below potential. Activity picks up, and gap starts to close.	Economy enjoying an upswing, with activity measures showing above-average growth rates.	Economy at peak. Activity above average but decelerating. The economy may experience shortages of factors of production as demand may exceed supply.	Economy goes into a contraction, (recession, if severe). Activity measures are below potential. Growth is lower than normal.
Employment	Layoffs slow. Businesses rely on overtime before moving to hiring. Unemployment remains higher than average.	Businesses move from using overtime and temporary employees to hiring. Unemployment rate stabilizes and starts falling.	Businesses continue hiring but at a slower pace. Unemployment rate continues to fall but at slowly decreasing rates.	Businesses first cut hours, eliminate overtime, and freeze hiring, followed by outright layoffs. Unemployment rate starts to rise.

Levels of employment lag the cycle

4.2 Fluctuations in Capital Spending

Capital spending—spending on tangible goods, such as property, plant, and equipment—typically fluctuates with the business cycle. Because business profits and cash flows are sensitive to changes in economic activity, capital spending is also highly sensitive to changes in economic activity. In fact, investment is one of the most procyclical and volatile components of GDP. Companies' spending decisions are driven by business conditions, expectations, and levels of capacity utilization, all of which fluctuate over the cycle. With regard to efficiency, firms will run "lean production" to generate maximum output with the fewest number of workers at the end of contractions. Exhibit 8 provides a description of capital spending over the cycle. Note that new orders statistics include orders that will be delivered over several years. For example, it is common for airlines to order 40 airplanes to be delivered over five years. Where relevant, analysts use "core" orders that exclude defense and aircrafts for a better understanding of the economy's trend.

Exhibit 8 Capital Spending during the Economic Cycle

Phase of the Cycle	Recovery	Expansion	Slowdown	Contraction
Business conditions and expectations	Excess capacity during trough, low utilization, little need for capacity expansion. Interest rates tend to be low—supporting investment.	Companies enjoy favorable conditions. Capacity utilization increases from low levels. Over time, productive capacity may begin to limit ability to respond to demand. Growth in earnings and cash flow gives businesses the financial ability to increase investment spending.	Business conditions at peak, with healthy cash flow. Interest rates tend to be higher—aimed at reducing overheating and encouraging investment slowdown.	Companies experience fall in demand, profits, and cash flows.
Capital spending	Low but increasing as companies start to enjoy better conditions. Capex focus on efficiency rather than capacity. Upturn most pronounced in orders for light producer equipment. Typically, the orders initially reinstated are for equipment with a high rate of obsolescence, such as software, systems, and technological hardware.	Customer orders and capacity utilization increase. Companies start to focus on capacity expansion. The composition of the economy's capacity may not be optimal for the current structure of demand, necessitating spending on new types of equipment. Orders precede actual shipments, so orders for capital equipment are a widely watched indicator of the future direction of capital spending.	New orders intended to increase capacity may be an early indicator of the late stage of the expansion phase. Companies continue to place new orders as they operate at or near capacity.	New orders halted, and some existing orders canceled (no need to expand). Initial cutbacks may be sharp and exaggerate the economy's downturn. As the general cyclical bust matures, cutbacks in spending on heavy equipment further intensify the contraction. Maintenance scaled back.
Examples	Software, systems, and hardware (high rates of obsolescence) orders placed or reinstated first.	Heavy and complex equipment, warehouses, and factories. A company may need warehouse space in locations different from where existing facilities are.	Fiber-optic overinvestment in late 1990s that peaked with the "technology, media, telecoms bubble."	Technology and light equipment with short lead times get cut first. Cuts in construction and heavy equipment follow.

EXAMPLE 3

Capital Spending

1 Levels of capacity utilization are one of the factors that determine companies' aggregate need for additional capital expenditure. Which of the following is another factor that impacts the capital expenditure decision?

   A The rate of unemployment.

   B The composition of the economy's capacity in relation to how it can satisfy demand.

   C The ability to re-instate orders canceled during the contraction stage.

2 Orders for technology and light equipment decline before construction orders in a contraction because:

   A businesses are uncertain about cyclical directions.

   B they are easier to cancel than large construction contracts.

   C businesses value light equipment less than structures and heavy machinery.

Solution to 1:

B is correct. The composition of the current productive capacity may not be optimal of the current structure of demand. C is incorrect because the ability to re-instate canceled orders is a matter that is relevant once the decision to increase capital expenditure is made.

Solution to 2:

B is correct. Because it usually takes much longer time to plan and complete large construction projects than for equipment orders, construction projects may be less influenced by business cycles.

4.3 Fluctuations in Inventory Levels

The aggregate size of inventories is small relative to the size of the economy, but their accumulation and cutbacks by businesses can occur with substantial speed and frequency. Changing inventories reflect differences between the growth (or decline) in sales and the growth (or decline) in production. A key indicator in this area is the inventory–sales ratio that measures the inventories available for sale to the level of sales. Analysts pay attention to inventories to gauge the position of the economy in the cycle. Exhibit 9 shows how production, sales, and inventories typically move through the phases of a cycle.

Exhibit 9 Inventories throughout the Cycle

Phase of the Cycle	Recovery	Expansion	Slowdown	Contraction
Sales and production	Sales decline slows. Sales subsequently recover. Production upturn follows but lags behind sales growth. Over time, production approaches normal levels as excess inventories from the downturn are cleared.	Sales increase. Production rises fast to keep up with sales growth and to replenish inventories of finished products. This increases the demand for intermediate products. "Inventory rebuilding or restocking stage."	Sales slow faster than production; inventories increase. Economic slowdown leads to production cutbacks and order cancellations.	Businesses produce at rates below the sales volumes necessary to dispose of unwanted inventories.
Inventory–sales ratio	Begins to fall as sales recovery outpaces production.	Ratio stable.	Ratio increases. Signals weakening economy.	Ratio begins to fall back to normal.

EXAMPLE 4

Inventory Fluctuation

1 Although a small part of the overall economy, changes in inventories can influence economic growth measures significantly because they:

- **A** reflect general business sentiment.
- **B** tend to move forcefully up or down.
- **C** determine the availability of goods for sale.

2 Inventories tend to rise when:

- **A** inventory–sales ratios are low.
- **B** inventory–sales ratios are high.
- **C** economic activity begins to rebound.

3 Inventories will often fall early in a recovery because:

- **A** businesses need profit.
- **B** sales outstrip production.
- **C** businesses ramp up production because of increased economic activity.

Solution to 1:

B is correct. Inventory levels can fluctuate dramatically over the business cycle.

Solution to 2:

A is correct. When the economy starts to recover, sales of inventories can outpace production, which results in low inventory–sales ratios. Companies then need to accumulate more inventories to restore the ratio to normal level. C is incorrect because in the early stages of a recovery, inventories are likely to fall as sales increase faster than production.

Solution to 3:

B is correct. The companies are slow to increase production in the early recovery phase because they first want to confirm the recession is over. Increasing output also takes time after the downsizing during the recession.

5 CONSUMER BEHAVIOR

Households represent the largest single sector of most economies. As a result, patterns of household consumption determine overall economic direction more than any other sector. For this reason, the patterns of household consumption are important to a range of practitioners, not just to analysts focusing on companies serving consumers.

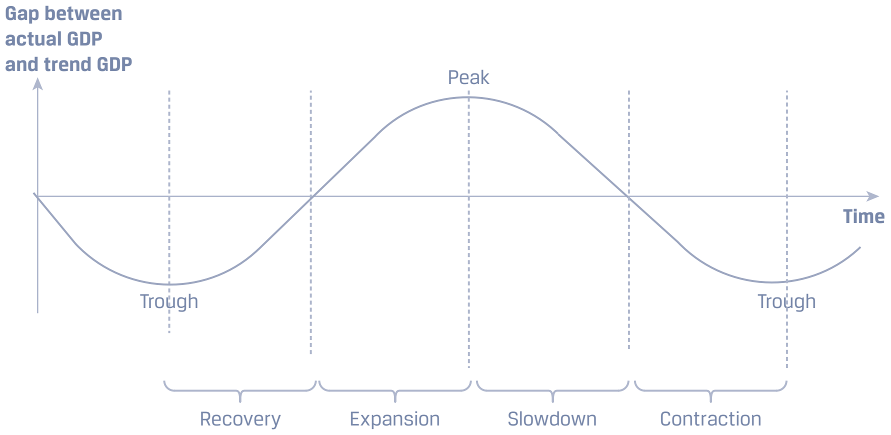
5.1 Consumer Confidence

Consumer confidence plays a significant role in spending decisions and reflects expectations of future incomes and employment prospects. Beyond direct observations of consumer spending and its mix, practitioners can also gauge future directions of spending by analyzing measures of consumer confidence or sentiment to ascertain how aggressive consumers may be in their spending. Usually, such information is in the form of surveys intended to provide practitioners with a general guide to trends.

5.2 Measures of Consumption

Two primary measures of household consumption are retail sales (physical as well as online) and broad-based indicators of consumer spending that also include purchases outside purely retail establishments, such as utilities and household services. Some additional measures make finer distinctions, such as tracking spending, both real and nominal, of a specific group of consumer products. Exhibit 10 shows how consumer spending changes through the economic cycle, and it splits consumer spending into three parts: (1) durable goods, (2) non-durable goods, and (3) services. Spending on durables is the most cyclical part, while spending on non-durables is the least cyclical. Spending on services, which include both more and less cyclical sub-components, in aggregate fits between the durable and non-durable goods categories. It is also worth noting that consumer spending, including that on durables, while cyclical, is less cyclical than investment spending by firms.

Exhibit 10 Consumer Behavior



Phase of the Cycle	Recovery	Expansion	Slowdown	Contraction
Incomes, employment, and confidence	Unemployment remains above average. Layoffs slow. Businesses rely on overtime before moving to hiring. Consumer confidence starts improving.	Hiring restarts. Unemployment rate stabilizes and starts falling. Consumers experience rising incomes, healthy employment prospects, and greater confidence.	Businesses continue hiring but at a slower pace. Unemployment rate continues to fall. Incomes are still growing. Consumers remain confident.	Businesses first cut hours of overtime prior to freezing hiring and starting layoffs. Employment levels decline, and consumer confidence weakens.
Spending on consumer durables (autos, motorcycles, appliances, furniture)	Spending limited as households postpone spending.	Spending increases.	Spending above average.	Purchases postponed; spending decreasing.
Consumer non-durables (i.e., medicines, food, household products)	Spending shows little change through the cycle.			
Services (entertainment, outdoor eating, communications, personal services)	Spending below average.	Spending increases.	Spending above average.	Spending declines.

Analysts often compare trends in durable purchases with those in the other categories to understand the economy's progress through the cycle. Because durable goods purchases usually replace items with longer useful lives, households usually postpone such purchases during economic downturns. Therefore, a weakness in durables spending may be an early indication of general economic weakness, and an increase in such spending may signal a more general cyclical recovery.

5.3 Income Growth

Growth in income is normally a good indicator of consumption prospects, and, albeit not as timely as consumer surveys, household income figures are widely available in most countries. Especially relevant is after-tax income, known as disposable income. Some analysts chart consumer spending based on a concept termed permanent income. Permanent income excludes temporary income (windfall or sudden, unexpected) and unsustainable losses and gains and tries to capture the income flow on which households believe they can rely. The basic level of consumption reflects this notion of permanent income. However, spending on durables tends to rise and fall with disposable income, regardless of the source, not just permanent income.

5.4 Saving Rates

Consumer spending patterns frequently diverge from trends in income, no matter how income is measured. An analysis of the saving rates can assist practitioners in this regard. Changes in saving rates can capture consumers' intent to reduce spending out of current income. The saving rate may also reflect future income uncertainties perceived by consumers (precautionary savings). Therefore, a higher saving rate may indicate consumers' ability to spend despite possible lower income in the future. A rise in the saving rate, usually measured as a percentage of income, may indicate caution among households and signal economic weakening. At the same time, the greater the stock of savings in the household sector and the wider the gap between ongoing income and spending, the greater the capacity for households to increase their spending. So, although unusually high savings may at first say something negative about the cyclical outlook, they point longer-term to the potential for recovery.

EXAMPLE 5

Consumer Behavior

1 Durable goods have the most pronounced cyclical behavior because:
  - **A** they have a longer useful life.
  - **B** their purchase cannot be delayed.
  - **C** they are needed more than non-durable goods or services.

2 Permanent income provides a better guide to:
  - **A** saving rates.
  - **B** spending on services.
  - **C** spending on durable goods.

Solution to 1:

A is correct. Durable goods are usually big ticket items, the life span of which can be extended with repairs and without incurring high replacement costs. So, consumers tend to delay replacement when the economic outlook is not favorable.

Solution to 2:

B is correct. Households adjust consumption of discretionary goods and services based on the perceived permanent income level rather than temporary earning fluctuations. Saving rates and durable goods consumption are more related to the short-term uncertainties caused by recessions.

6 HOUSING SECTOR BEHAVIOR

Housing sector activity includes new and existing home sales and residential construction activity. It is a small part of the overall economy compared to consumer spending, but because housing activity experiences dramatic swings, it often counts more in overall economic movements than the sector's relatively small size might suggest. In many countries, such as the United States and the United Kingdom, changing property values are closely related to consumer wealth and confidence.

6.1 Available Statistics

Many economies offer statistics on a range of housing activities, including the inventory of unsold homes on the market and average or median price of homes (sometimes recorded by type of housing unit and sometimes as the price per square foot or square meter). The relationships in this area typically follow fairly regular cyclical patterns.

6.2 Sensitivity to Interest Rates and Relationship to Credit Cycle

Because many home buyers finance their purchase with a loan (called mortgage in some countries), the sector is especially sensitive to interest rates. Home buying and consequently construction activity expand in response to lower loan interest rates and contract in response to higher loan interest rates.

Beyond such interest rate effects, housing also follows its own internal cycle, seen as part of the credit cycle. When housing prices are low relative to average incomes, and especially when mortgage rates are also low, the cost of owning a house falls and demand for housing increases. Often indicators of the cost of owning a house are available to compare household incomes with the cost of supporting an average house, both its price and the expense of a typical mortgage. Commonly, housing prices and mortgage rates rise disproportionately as expansionary cycles mature, bringing on an increase in relative housing costs even as household incomes rise. The resulting slowdown of house sales can lead to a cyclical downturn first in buying and then, as the inventory of unsold houses builds, in actual construction activity.

These links, clear as they are, are far from mechanical. If housing prices have risen rapidly in the recent past, for instance, many people will buy to gain exposure to the expected further price gains even as the purchase in other respects becomes harder to rationalize. Such behavior can extend the cycle upward and may later result in a more severe correction. This result occurs because "late buying" activity invites overbuilding. The large inventory of unsold homes eventually puts downward pressure on real estate prices, catching late buyers, who have stretched their resources. This pattern occurred in many countries during the 2008–2009 global financial crisis.

6.3 The Role of Demographics

Cyclical behavior in housing occurs around the long0run growth trend in housing determined by demographics, such as family and household formation. Where data on family formation are unavailable, analysts can gauge the rate of family formation by assessing the age structure of the population, as household formation commonly occurs in the 25–40-year-old age group. Adjusted for older people who are vacating existing homes, such calculations serve as an indicator of underlying, longer-term, secular housing demand.

6.4 Impact on the Economic Cycle

Although such measures have little to do with business cycles, they do offer a gauge, along with affordability, of how quickly the housing market can correct excess and return to growth. As some faster-growing developing economies experience urbanization (people moving from the countryside to cities), they witness fast growth in demand for housing units. Such housing demand may quickly reverse any cyclical weakness in the economy, more so than in such economies as Italy or Japan where net new family formation is relatively slight.

EXAMPLE 6

Housing Sector Behavior

1 Housing is more sensitive than other sectors of the economy to:
- **A** interest rates.
- **B** permanent income.
- **C** government spending.

2 Apart from questions of affordability, house buying is affected by:
- **A** the rate of family formation.
- **B** expectation of housing price increases.
- **C** both the rate of family formation and expectation of housing price increases.

Solution to 1:

A is correct. Because real estate purchases are usually financed with mortgage loans, interest rate changes directly influence the monthly payment amounts.

Solution to 2:

C is correct. Family formation constitutes the actual need for housing, whereas buying on the expectation of housing price increases reflects the fact that real estate has investment value.

7 EXTERNAL TRADE SECTOR BEHAVIOR

The external trade sector varies tremendously in size and importance from one economy to another. In such places as Singapore or smaller open economies, such as Lithuania, where almost all inputs are imported and the bulk of the economy's output finds its way to the export market, trade (the sum of both exports and imports) easily exceeds its GDP. External conditions, especially the stage of the cycle in key export regions and in primary trading partners, play a crucial role. In other places, such as Brazil or the United States, external trade assumes a much smaller part of GDP. Since the 1970s, global trade has risen substantially, so the business cycles of the large economies in the world can be more easily transmitted to other economies.

7.1 Cyclical Fluctuations of Imports and Exports

Typically, imports rise, all else equal, with the pace of domestic GDP growth because rising domestic demand increases purchases of goods and services, by consumers and business, from abroad. Thus, imports respond to the domestic cycle. Exports are more dependent on cycles in the rest of the world. If these external cycles are strong, all else equal, exports will grow even if the domestic economy should experience a decline in growth. To understand the impact of exports, financial analysts need to understand the strength of the major trading partners of the economy under consideration. The net effect of trade may offset cyclical weakness and, depending on the importance of exports to the economy, could erase it altogether. For these reasons, such differences can mean the pattern of external trade balances is entirely different from the rest of the domestic economic cycle. Exhibit 11 provides a short summary.

Exhibit 11 External Trade

Phase of the Cycle (domestic economy)	Recovery	Expansion	Slowdown	Contraction
Exports	Driven by external demand			
Imports (*assuming exchange rate remains unchanged)	Imports below average, start to increase.	Imports increase.	Imports peak and start to decline.	Imports in decline to below-average levels.

7.2 The Role of the Exchange Rate

The currency exchange rate also has an independent effect that can move trade in directions strikingly different from the domestic economic cycle. When a nation's currency appreciates (the currency gains in strength relative to other currencies), foreign goods seem cheaper than domestic goods to the domestic population, prompting, all else equal, a relative rise in imports. At the same time, such currency appreciation makes that nation's exports more expensive in global markets and may reduce exports. Of course, currency depreciation has the opposite effect. Although currency moves may be volatile and on occasion extreme, they have a significant effect on trade and the balance of payments only when they cumulate in a single direction for some time. Moves from one month or quarter to the next, however great, have a minimal effect until they persist. Thus, cumulative currency movements that take place over a period of years will have an impact on trade flows that will persist even if the currency subsequently moves in the opposite direction for a temporary period.

7.3 Overall Effect on Exports and Imports

Financial analysts need to consider a wide range of variables, both in the domestic economy and abroad, to assess relative GDP growth rates and then factor in currency considerations to ascertain whether they reinforce or counteract other cyclical forces. Generally, GDP growth differentials in global economic growth rates between countries have the most immediate and straightforward effects: Domestic changes in economic activity raise or reduce imports, and foreign economic activity changes raise or reduce exports. Currency moves have a more complex and, despite the interim short-term currency moves, a more gradual effect.

EXAMPLE 7

External Trade

1 Imports most clearly respond to:
   - A the level of exports.
   - B domestic industrial policy.
   - C domestic GDP growth rate.

2 Exports generally respond to the:

A government expenditure
B global GDP growth rates.
C domestic GDP growth rates.

Solution to 1:

C is correct. As a part of aggregate demand, imports reflect the domestic needs for foreign goods, which vary together with domestic economic growth.

Solution to 2:

B is correct. Exports reflect the foreign demands on domestic output, which depend on the conditions of the global economy.

EXAMPLE 8 INVESTMENT IMPLICATIONS

A junior portfolio manager is reviewing portfolio holdings. He focuses on three specific companies that he was tasked with reviewing. He has just received an analysis of the stage of the business cycle from a team of the firm's economists. To set the context of his analysis, he needs to understand the behavior of the three companies as they have exposures to different parts of the economy. One is focused on consumer goods (electronics and small vehicles) sold on the local market; one is focused on equipment sold to domestic companies; and one makes household and personal care consumer products, which it sells both domestically and abroad.

Which of the these three companies' financial performance fluctuates with the business cycle of their domestic economy the *most*?

A Consumer goods company
B Business equipment company
C Producer of household and personal care products

Solution:

B is correct. Business investment fluctuates more than consumer spending and exports. A and C are incorrect because although consumer goods companies are impacted by the cycle, they are impacted less than companies exposed to business investment.

8 THEORETICAL CONSIDERATIONS

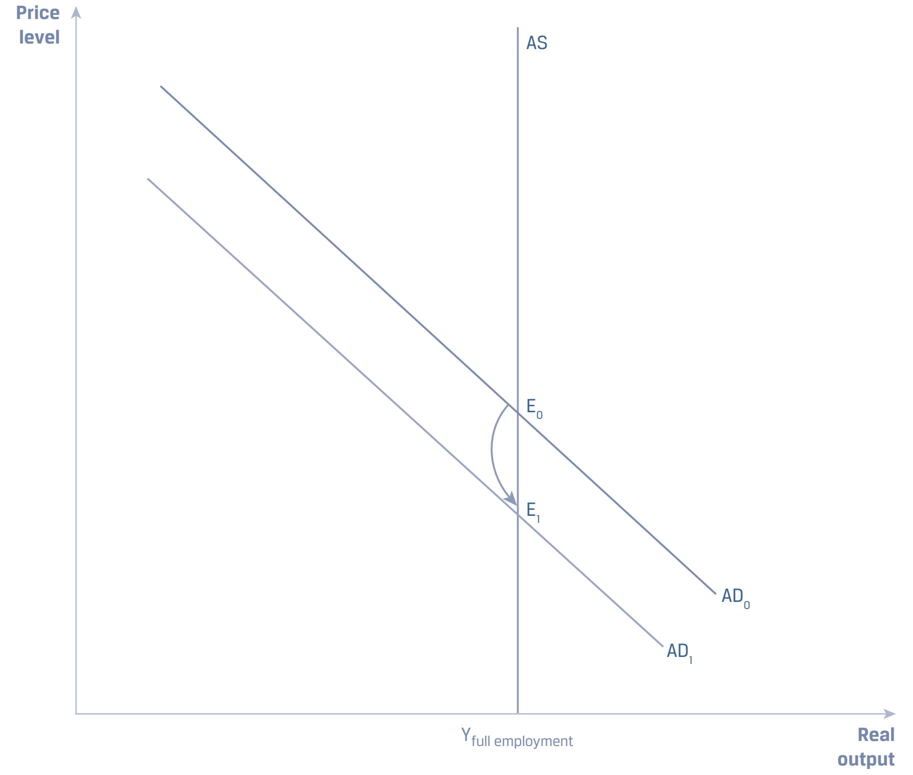
d describe theories of the business cycle

In this section we review the main theories that try to explain causes of business cycles and mechanisms through which the economy returns (or does not return) to the longer-term growth path.

8.1 Historical Context

Modern mainstream macroeconomists use the aggregate demand/aggregate supply (AD/AS) framework to synthesize the different schools of thought that have developed over the decades. In all of the key theories, movements in either of these curves create business cycles over time. Until the 1930s, the classical view has been that the aggregate supply curve is very steep (see Exhibit 12), prices and wages adjust quickly, and therefore recessions are mild and only short-lived. There was no need for government policy to intervene in the economy to try to smooth the business cycle as full employment level of output would be quickly restored. On the contrary, a policy of active demand management might, because of reaction lags that we will explore later, actually trigger business cycles rather than smooth them.

Exhibit 12 Classical Aggregate Supply Curve



The depth and severity of the Great Depression in the 1930s triggered the development of so-called Keynesian theories, in which fiscal policies play an important role in stabilizing the aggregate demand curve. Because prices and wages are seen as sticky and do not adjust, particularly on the way down (shown in Exhibit 13), the economy may experience a significant output gap without intervention.

Rising inflation in the late 1960s and the subsequent stagflation periods triggered by the oil price shocks in the 1970s were an incubator for another approach: monetarism. This school of thought highlights the importance of monetary policy and advocates stable money growth to keep the position of the aggregate demand curve stable, inflation under control, and reduce business cycle volatility.

Since the mid-1980s, we witnessed the emergence of micro-founded theories and the role of rational expectations that spurred the development of neoclassical (within which are Real Business Cycle, RBC, theorists) and new-Keynesian types of macroeconomic models to explain short-term fluctuations. Despite the “neo” and “new” in front of the classical and Keynesian theories, the fundamental differences remain the same: Classical-oriented economists have more faith in the working of the "invisible hand" (i.e., the proper functioning of markets), while Keynesian-oriented economists see rigidities and market failures as more prevalent and the government in a position to actively intervene.

We will now explain the two broad groups of theories through this "non-intervention" vs "intervention" lens and describe each in terms of a) the role of demand and supply shocks; b) the completeness of markets and thereby the stickiness of prices; c) the importance of asymmetric information problems; and d) the role of policy lags.

8.2 Neoclassical Economics

Neoclassical economists, but in particular the so-called Real Business Cycle (RBC) advocates, stress the importance of movements in the supply curve. Technological progress, natural disasters, changes in relative prices of key inputs, other supply-side constraints, or changes in firm expectations lead to changes in the position of the supply curve. RBC models of the business cycle conclude that expansions and contractions represent efficient operation of the economy in response to external real shocks; the level of economic activity at any time is consistent with maximizing expected utility. For that reason, the policy recommendation of RBC theorists, and neoclassical economists in general, is for governments *not* to intervene in the economy with discretionary fiscal and monetary policy. This is stressed even further by arguing that monetary and fiscal policy will affect the economy with substantial delay, referred to as "lags": There is a lag between when policy makers realize what is happening with the economy and when they develop a plan and implement it. Finally, it takes time for policy to actually have an impact on the economy. It is possible that any policy action will impact the economy too late, and instead of bringing output closer to the long-term path, it may amplify the next cycle.

In a more mainstream context, reducing economic volatility is translated into stressing the importance of expectation and information management. Removing asymmetric information problems should speed up price adjustments. If all participants in a market are well informed, market forces can work efficiently. Information problems, however, create inefficiencies. If a firm cannot distinguish between an increase in the relative price of the good it is offering and general inflation, it might make inefficient production decisions. If the price increase in the market arises due to an increase in demand, then the firm should increase production. However, if price increases merely reflect general inflation, the firm should not increase production. Removing such information problems reduces economic volatility. Similarly, fostering agents in the economy to form precise and well-anchored expectations about inflation, for example, can lead to self-fulfilling outcomes and stabilize the position of the supply curve.

Neoclassical economists would also argue that policies should not prevent Schumpeterian "creative destruction" from taking place. The term "creative destruction" describes the disruptive process of transformation that accompanies innovation, originally described by Joseph Schumpeter in the 1950s. When an inventor presents a new product or a new, better way to produce an existing good or service (e.g., radio frequency identification tracking of inventories), then the entrepreneur that introduces the new discovery will likely have bigger profits and may drive the existing producers out of business. Therefore, innovations can generate movements in the aggregate supply curve that policy makers should not try to correct.

8.3 The Austrian School

The so-called Austrian school, another "non-intervention" theory, shares many views of the neoclassical economists. In contrast to the neoclassical school, however, it focuses on the role of money, which they believe does more than facilitate the exchange of goods and services. Austrian economists argue that low interest rates and excessive credit growth during boom times result in over-investment in projects with low returns, causing failure and a move into recession. They argue that following a period of overinvestment, companies will realize that they have overinvested, resulting in a sudden drop in investment with a negative shock for aggregate demand. Because the Austrian school sees business cycles arising from government (and central bank) policies as the cause of overinvestment and subsequent failure, they suggest that policy makers should rarely intervene in the economy.

These classical views stress that markets are in general flexible and prices will easily adjust to changing situations. This makes the aggregate supply curve very steep, even in the short run (see Exhibit 12). Policy should be designed to allow markets to function and prices to adjust with flexibility.

8.4 Monetarism

For monetarists, other than maintaining steady growth of the money supply, government should play only a very limited role in the economy. They reject active management of aggregate demand. Fiscal and monetary policy should be clear and consistent over time so all economic agents can forecast government actions. In this way, the uncertainty of economic fluctuations would not be increased by any uncertainty about the timing and magnitude of economic policies and their lagged effects.

According to the monetarist school, business cycles may occur both because of exogenous shocks and because of government intervention. So, it is better to let aggregate demand and supply find their own equilibrium than to risk causing further economic fluctuations. In contrast to neoclassical economists, a key part of monetarist thought is that the money supply needs to continue to grow at a moderate rate. If it falls, as occurred in the 1930s, the economic downturn could be severe; whereas if money grows too fast, inflation will follow.

Monetarists also see business cycles largely stemming from movements in the aggregate demand curve. But their conclusions differ from Keynesian interventionism, which they reject, for three main reasons:

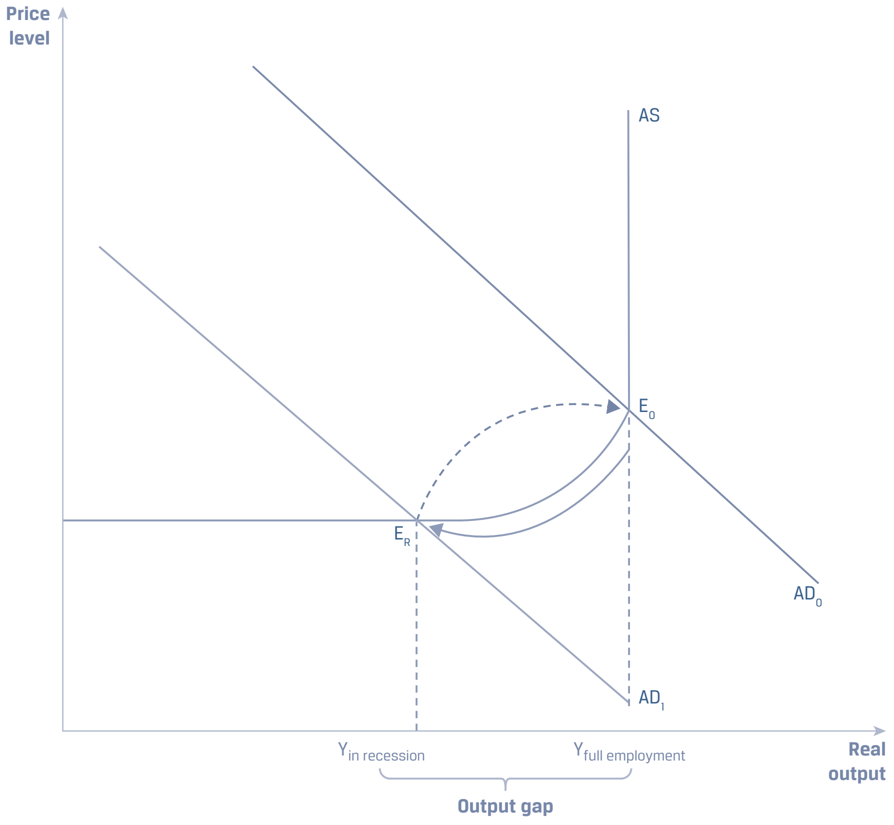
1. Unlike the monetarist model, the Keynesian model does not recognize the supreme importance of the money supply. If the money supply grows too fast, there will be an unsustainable boom, and if it grows too slowly, there will be a recession.
2. Monetarists consider the long-term costs of government intervention (e.g., growing government debt and a high cost of interest on this debt).
3. The timing of a government's economic policy responses is uncertain, and the stimulative effects of a fiscal expansion may take effect after the crisis was over, thus causing more harm than good.

8.5 Keynesianism

Keynesians see an active role for the government in managing aggregate demand. Keynesian-oriented theories give limited importance to the supply curve. Instead of thinking about it as being rather steep, these theories see prices being sticky in the short run, leading to a relatively flat AS curve (see the horizontal section of the AS curve in Exhibit 13). It is argued that nominal wages, as one of the most important prices in the economy, are generally rigid and not downward flexible; workers do not want to see their nominal compensation decrease. Movements in the AD curve, therefore, have large consequences on the overall output of the economy.

But even if workers agreed to accept lower salaries in times of recession, this situation might exacerbate the crisis. Lower wage expectations would hurt confidence, and workers would cut back on their spending—reducing aggregate demand, causing further weakening of confidence, and leading to contractions in demand for goods and services. These contractions would start with the more expensive items, such as durable goods, and move in a "domino effect" through the economy (the downward spiral of the aggregate demand curve continuously shifting left). Exhibit 13 shows how a decrease in aggregate demand (from AD_0 to AD_1) may result in an output gap and an equilibrium in recession (ER) from which it may take a long time to recover and restore full employment levels of output (at equilibrium E_0).

Exhibit 13 Keynesian AS/AD Framework



Furthermore, there could be circumstances in which lower interest rates would not reignite growth because business confidence or "animal spirit" (consumer's lack of confidence) was too low. Therefore, Keynesian economists advocate government intervention in the form of fiscal policy. While Keynes accepted the possibility that markets would reach the equilibrium envisioned by neoclassical economists over the long run, he famously quipped that "in the long run, we are all dead"—that is, the human suffering is excessive while waiting for all shocks to be absorbed and for the economy to return to equilibrium.

In Keynesians' opinion, when crises occur, the government should intervene to keep capital and labor employed by deliberately running a larger fiscal deficit. This intervention would limit the damage caused by major recessions. Although this concept continues to be a highly politically charged debate, many economists agree that government expenditure can limit the negative effect of major economic crises in the short term. The practical criticisms that are often expressed about Keynesian fiscal policy are:

1. Fiscal deficits mean higher government debt that needs to be serviced and repaid eventually.
2. Keynesian cyclical policies are focused on the short term. In the long run, the economy may come back and the presence of the expansionary policy may cause it to "overheat"—that is, to have unsustainably fast economic growth, which causes inflation and other problems. This result is because of the typical lags involved in expansionary policy taking effect on the economy.
3. Fiscal policy takes time to implement. Quite often, by the time stimulatory fiscal policy kicks in, the economy is already recovering.

THE PERSPECTIVE OF HYMAN MINSKY

A different view of business cycles came from Hyman Minsky. Minsky believed that excesses in financial markets exacerbate economic fluctuations. For example, a rapid growth of credit, often given to risky ventures in the late expansion phase of the cycle, will be followed by a "credit crunch" during the down-swing phase. Minsky traced excesses to a type of complacency in which people underestimate the risk of events that have not occurred in a while. Therefore, if the economy has been in a long expansion, people may think that the market works very well and that the expansion will last forever—that is, extrapolating past experiences. In this sense, Minsky's theories could be seen as a precursor of behavioral finance, which is the branch of finance that studies how cognitive and emotional biases, such as overconfidence and short memory, induce investors to be overconfident and make suboptimal choices.

The term **Minsky moment** has been coined for a point in the business cycle when, after individuals become overextended in borrowing to finance speculative investments, people start realizing that something is likely to go wrong and a panic ensues leading to asset sell-offs. The subprime crisis that affected many industrialized countries starting in 2008 has been represented as a "Minsky moment" (see McCulley 2009) because it came after years in which risk premiums (e.g., the differentials, or spreads, between very risky bonds and very safe bonds) were at historically low levels. Typically, low risk premiums suggest that no adverse events are expected; in other words, investors believe that because the economy and the markets have been enjoying a protracted expansion, there is no reason to worry about the future. As a consequence, many market observers suggest that business cycles are being tamed. This kind of view of the world leads people to underestimate risk—for example, by not doing the appropriate diligent research before granting a loan or before purchasing a security (i.e., complacency).

The "Minsky moment" has been compared with a cartoon in which a cartoon character walks over a cliff without realizing it. When he looks down and sees that he is walking on thin air, he panics and falls to the bottom of the canyon—just like the world economy in 2008.

8.6 Modern Approach to Business Cycles

Today's mainstream economics can be described as a synthesis of three macroeconomic theories: neoclassical economics (including RBC theories), Keynesian economics, and monetarism. As the different models do not necessarily rule each other out, mainstream economists in practice use an analytical framework that encompasses all three broad theories.

The debate about business cycles often receives a politically partisan treatment in the press because some experts and commentators are generally against government intervention in the economy (for example, because it may lead to large fiscal deficits) and others are in favor (for example, because it may alleviate the effects of a large economic shock). Financial analysts should base investment decisions on analysis and try to put personal biases aside.

However, there is little doubt that central banks try to manage the business cycle by moving the aggregate demand curve through raising interest rates when the economy is growing rapidly and inflation accelerates and cutting rates when the economy is weak. In the 2008–2009 downturn, when official interest rates first approached zero, central bankers extended the range of their actions and tools to influence aggregate demand and stimulate the economy through the use of unconventional policies that will be covered later.

EXAMPLE 9

Theories of Business Cycle

1 The main difference between neoclassical (RBC) and Keynesian models is that the neoclassical models:

  A focus on the role of money in the economy.

  B focus on the aggregate supply curve in explaining the causes of economic cycles.

  C assume that prices adjust slowly, whereas Keynesians assume that prices adjust quickly.

2 Which of the following theories would advocate government intervention to stimulate aggregate demand during economic contractions?

  A The Austrian school

  B Keynesians

  C Real Business Cycle theorists

Solution to 1:

B is correct. Neoclassical (including RBC) economists suggest that business cycles are caused by factors that shift the aggregate supply curve. A is incorrect as monetarists focus on the role of money and the growth in money supply in the economy. C is incorrect because Keynesians believe that prices adjust slowly.

Solution to 2:

B is correct. Keynesians would advocate government intervention to stimulate aggregate demand. A and C are incorrect as advocates of the Austrian school and the RBC theorists would discourage government intervention.

ECONOMIC INDICATORS

9

e interpret a set of economic indicators, and describe their uses and limitations

Economic indicators are variables that provide information on the state of the overall economy. They are statistics or data readings that reflect economic circumstances of a country, group of countries, region, or sector. Economic indicators are used by policy makers and analysts to understand and assess the existing condition of the economy and its position in the cycle. They can also be used to help predict or confirm the turning points in the cycle. Such knowledge allows analysts to better predict the financial and market performance of stocks and bonds of issuers operating in different sectors of the economy with different sensitivity to the economic cycle.

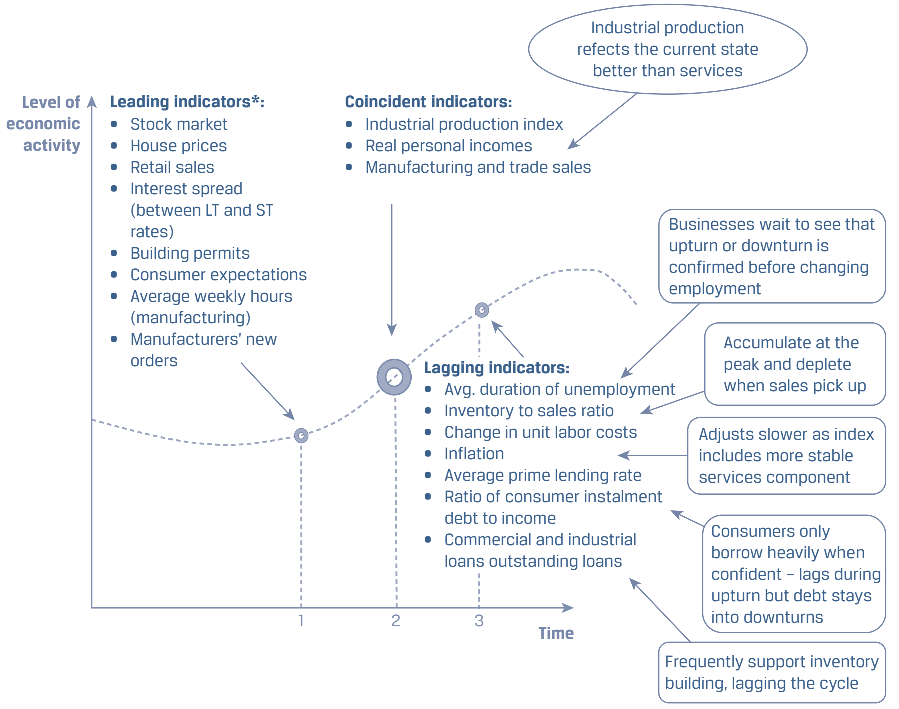
9.1 Types of Indicators

Economic indicators are often classified according to whether they lag, lead, or coincide with changes in an economy's growth.

- **Leading economic indicators** have turning points that usually precede those of the overall economy. They are believed to have value for predicting the economy's future state, usually near-term.
- **Coincident economic indicators** have turning points that are usually close to those of the overall economy. They are believed to have value for identifying the economy's present state.
- **Lagging economic indicators** have turning points that take place later than those of the overall economy. They are believed to have value in identifying the economy's past condition and only change after a trend has been established.

Exhibit 14 provides an illustration of several leading, lagging, and coincident indicators. The leading indicators observed at a point in time labeled as time "1" indicate the direction of the of the activity (output) at a future point in time, such as time "2." The lagging indicators, released around time "3," refer to and help confirm what the state of the economy was at time "2."

Exhibit 14 Types of Economic Indicators



Industrial production refects the current state better than services

Leading indicators*:
- Stock market
- House prices
- Retail sales
- Interest spread (between LT and ST rates)
- Building permits
- Consumer expectations
- Average weekly hours (manufacturing)
- Manufacturers' new orders

Coincident indicators:
- Industrial production index
- Real personal incomes
- Manufacturing and trade sales

Lagging indicators:
- Avg. duration of unemployment
- Inventory to sales ratio
- Change in unit labor costs
- Inflation
- Average prime lending rate
- Ratio of consumer instalment debt to income
- Commercial and industrial loans outstanding loans

Businesses wait to see that upturn or downturn is confirmed before changing employment

Accumulate at the peak and deplete when sales pick up

Adjusts slower as index includes more stable services component

Consumers only borrow heavily when confident – lags during upturn but debt stays into downturns

Frequently support inventory building, lagging the cycle

---

* Leading indicators will be explored in the subsequent section.

9.2 Composite Indicators

An economic indicator either consists of a single variable, like industrial production or the total value of outstanding building permits, or can be a composite of different variables that all tend to move together. The latter are regularly labelled composite indicators. Traditionally, most composite indicators to measure the cyclical state of the economy consist of up to a dozen handpicked variables published by organizations like the OECD or national research institutes. The exact variables combined into these composites vary from one economy to the other. In each case, however, they bring together various economic and financial measures that have displayed a consistently leading, coincident, or lagging relationship to that economy's general cycle.

9.3 Leading Indicators

In the United States, for instance, The Conference Board, a US industry research organization, publishes a composite leading indicator known as the Conference Board Leading Economic Index (LEI) that consists of 10 component parts (it uses the classical business cycle as the underlying concept). Exhibit 15 presents the 10 components used in the LEI. In addition to naming the indicators, it also offers a general description of why each measure is leading the business cycle.

Exhibit 15 Index of Leading Economic Indicators—United States

Leading indicators	Reason for use
Average weekly hours, manufacturing	Businesses will cut overtime before laying off workers in a downturn and increase it before rehiring in a cyclical upturn. Moves up and down before the general economy.
Average weekly initial claims for unemployment insurance	A very sensitive test of initial layoffs and rehiring.
Manufacturers' new orders for consumer goods and materials	Businesses cannot wait too long to meet demand without ordering. Orders tend to lead at upturns and downturns & captures business sentiment.
ISM new order index *Survey based* *The Institute of Supply Management (ISM) polls its members to build indexes of manufacturing orders, output, employment, pricing, and comparable gauges for services.* *A diffusion index usually measures the percentage of components in a series that are rising in the same period. It indicates how widespread a particular movement in the trend is among the individual components.*	Reflects the month on month change in new orders for final sales. Decline of new orders can signal weak demand and can lead to recession.
Manufacturers' new orders for non-defense capital goods excluding aircraft	Captures business expectations and offers first signal of movement up or down. Important sector.
Building permits for new private housing units	Signals new construction activity as permits required before new building can begin.
S&P 500 Index	Stocks tends to anticipate economic turning points; useful early signal.
Leading Credit Index *Aggregates the information from six leading financial indicators, which reflect the strength of the financial system to endure stress.*	A vulnerable financial system can amplify the effects of negative shocks, causing widespread recessions.
Interest rate spread between 10-year treasury yields and overnight borrowing rates (federal funds rate) *Inversion of the yield curve occurs when ST interest rate exceed LT rates – meaning that ST rates are expected to fall and activity is expected to weaken.*	LT (10 or 30 year) bond yields express market expectations about the direction of short-term interest rates. As rates ultimately follow the economic cycle up and down, a wider spread, by anticipating short rate increases, also anticipates an economic upswing and vice versa.
Average consumer expectations for business conditions *Survey based*	Optimism tends to increase spending. Provides early insight into the direction ahead for the whole economy.

9.4 Using Economic Indicators

Exhibit 16 shows a simplified process that an analyst could use to identify business cycle phases. The conclusions can then be used to make investment decisions—for example, to decide in what sectors companies are likely to see improving or deteriorating cash flows, which could impact the investment performance of the equity and debt securities issued by the companies. Note that the order of the steps shown does not have to follow this particular sequence.

Exhibit 16 Use of Statistics to Identify Business Cycle Phase

Step 1

Data release: Analyst notes an increase in the reported level of consumer instalment debt to income.

Analysis: The above indicator normally lags cyclical upturns.

Possible conclusion: Initial evidence that an upturn has been underway.

Step 2

Data release: Industrial Production Index and non-farm payrolls (employees on non-agricultural payrolls) are moving higher.

Analysis: These are coincident indicators that suggest activity is picking up.

Possible conclusion: Further evidence that expansion is underway.

Step 3

Observation: Equity market index has been trending higher. Equity index is a leading indicator. Analyst checks the aggregate LEI Index.

Analysis: If the aggregate LEI is moving higher too, evidence that recovery is underway. -> Confirmation that output is moving higher.

Or

Aggregate LEI is not moving higher. Analyst cannot draw conclusions about recovery.

9.5 Other Composite Leading Indicators

For about 30 countries and several aggregates, such as the EU and G–7, the Organisation for Economic Co-Operation and Development (OECD) calculates OECD Composite Leading Indicator (CLI), which gauge the state of the business cycle in the economy using the growth cycle concept. One of the interesting features of OECD CLI is that the underlying methodology is consistent across several countries. It can therefore be compared more easily to see how each region is faring. Exhibit 17 shows the eight components of CLI used by the OECD. As is usually the case with leading indicators, some are based on surveys while others are based on reported market or economic data.

Exhibit 17 OECD Euro Area CLI Components

OECD CLI

Composite indicator of economic tendency survey results → Economic sentiment index

Residential building permits

Capital goods orders

Euro Stoxx Equity Index

M2 money supply

An interest rate spread

Survey – based → EurozoneEuro area Manufacturing Purchasing Managers Index (PMI)

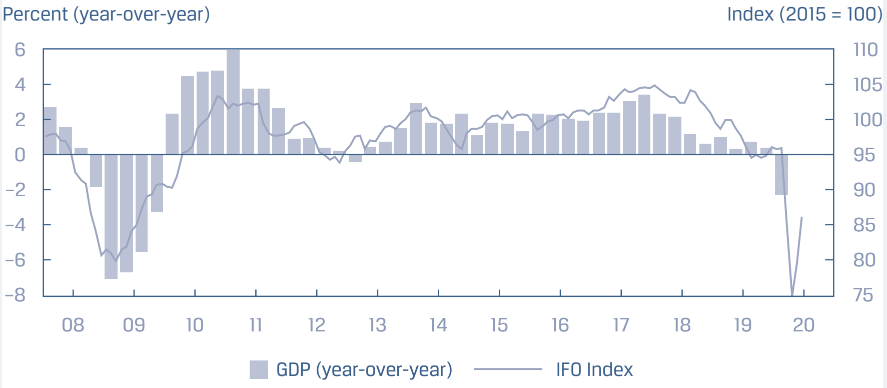
Survey – based → EurozoneEuro area Service Sector Future Business Activity Expectations Index

} Composite leading indicator

The parallels between many of these components and those used in the United States are clear, but the Euro area has a services component in its business activity measures that the United States lacks; the Euro area forgoes many of the overtime and employment gauges that the United States includes. The OECD CLI for Japan is again similar, but it does include labor market indicators (unlike the Eura area) and adds a measure of business failures not included in the other two.

The German ifo survey is a widely used index capturing business climate in Germany and is published monthly. Exhibit 18 shows how the index moved ahead of quarterly-reported YoY changes in German GDP. It also shows an uptick in 2018 despite the GDP growth downturn. These indicators are useful, but they are not foolproof.

Exhibit 18 Business climate survey and GDP



Sources: The ifo Institute and the Federal Statistical Office of Germany.

9.6 Surveys

The composite indicators for region- or country-specific business cycles often make use of economic tendency surveys. These monthly or quarterly surveys carried out by central banks, research institutes, statistical offices, and trade associations are conducted among either businesses, consumers, or experts. The mostly qualitative questions posed are often on the state of their finances, level of activity, and their confidence in the future.

CIRET (a forum for leading economists and institutions that conducts and analyzes business and/or consumer survey data), the United Nations, the OECD, and the European Commission are at different regional levels harmonizing and exchanging knowledge regarding these business tendency and consumer surveys (United Nations 2015). The Directorate-General for Economic and Financial Affairs (DG ECFIN) publishes fully harmonized results for different sectors of the EU member and applicant economies. The Bank of Japan carries out similar surveys and releases the findings in what is called the "Tankan Report." These diverse sources multiply within and across economies. Over the last decade, so-called "purchasing managers" indexes along similar lines, albeit with often much smaller survey samples, have been introduced in a wide range of countries and regions, including Europe and China.

These economic tendency surveys are often aggregated into or are a part of composite indicators. IHS Markit publishes a global Purchasing Managers' Index (PMI) indicator using results of its national business surveys (see Exhibit 17). A global consumer confidence indicator is published by The Conference Board. Some other institutions collect national survey data and calculate supra-national results. For example, the OECD publishes a business confidence indicator and a consumer confidence indicator for the OECD aggregate in its report "Main Economic Indicators." In addition, the European Commission calculates various survey-based indicators for the EU, reflecting the situation in Europe. Based upon economic tendency surveys and using the growth rate cycle concept, KOF and FGV publish a coincident and a leading indicator for the world economy called the Global Economic Barometers. These incorporate hundreds of survey variables from all over the world.

9.7 The Use of Big Data in Economic Indicators

The vast increase in this information and academic developments regarding the use of big(ger) data have in recent years increased the number of variables that go into these composite indicators. For instance, using a statistical technique called "principal components analysis," the Federal Reserve Bank of Chicago computes the Chicago Fed National Activity Index (CFNAI) using 85 monthly macroeconomic series. These series cover industrial production, personal income, capital utilization, employment by sectors, housing starts, retail sales, and so on. Principal components analysis "extracts" the underlying trend that is common to most of these variables, thus distilling the essence of the US business cycle. Similarly, the Bank of Italy in conjunction with the Centre for Economic Policy Research (CEPR) produces the EuroCOIN statistic, which is also based on principal component analysis. There are more than 100 macroeconomic series included in €-coin. €-coin also includes data derived from surveys, interest rates, and other financial variables.

9.8 Nowcasting

Policy makers and market practitioners use real-time monitoring of economic and financial variables to continuously assess current conditions. To overcome the publication delays (for example, GDP numbers are published with a substantial delay) and forecast the "present," they make use of a large variety of data—such as financial market transactions, data from the usage of the large amounts of timely internet searches, and electronic payment data—to provide estimates for key low frequency (monthly or quarterly) economic indicators. What they produce is a nowcast, an estimate of the current state. We refer to the process of producing such an estimate as nowcasting. It can be applied to various macroeconomic variables of interest, such as GDP growth, inflation, or unemployment.

9.9 GDPNow

Nowcasts are produced by a number of entities in investment banking and asset management for their internal or client use, but they are also published by institutions, such as the Atlanta Fed (the Federal Reserve Bank of Atlanta, one of the 12 Federal Reserve Banks in the United States). According to the Atlanta Fed, "GDPNow" is "best viewed as a running estimate of real GDP growth based on available data for the current measured quarter." The objective is to forecast GDP for the current quarter (which will not be released until after quarter-end) in real time based on data as they are released throughout the quarter. To do this, the Atlanta Fed attempts to use the same methodology and data as will be used by the US Bureau of Economic Analysis (BEA) to estimate GDP, replacing data that have not yet been released with forecasts based on the data already observed. As the quarter progresses, more of the actual data will have been observed. GDPNow should, at least on average, converge to what will be released by the BEA as their "advance" estimate of quarterly GDP about four weeks after quarter-end.

DIFFUSION INDEX OF ECONOMIC INDICATORS

In the United States, The Conference Board also compiles a monthly diffusion index of the leading, lagging, and coincident indicators. The **diffusion index** reflects the proportion of the index's components that are moving in a pattern consistent with the overall index. Analysts often rely on these diffusion indexes to provide a measure of the breadth of the change in a composite index.

For example, The Conference Board tracks the growth of each of the 10 constituents of its Leading Economic Index, assigning a value of 1.0 to each indicator that rises by more than 0.05% during the monthly measurement period, a value of 0.5 for each component indicator that changes by less than 0.05%, and a value of 0 for each component indicator that falls by more than 0.05%. These assigned values, which of course differ in other indexes in other countries, are then summed and divided by 10 (the number of components). To make the overall measure resemble the more familiar indexes, the Board multiplies the result by 100.

A simple numerical example will help explain. Say, for ease of exposition, the indicator has only four component parts: stock prices, money growth, orders, and consumer confidence. In one month, stock prices rise 2.0%, money growth rises 1.0%, orders are flat, and consumer confidence falls by 0.6%. Using The Conference Board's assigned values, these would contribute respectively: 1.0 + 1.0 + 0.5 + 0 to create a numerator of 2.5. When divided by four (the number of components) and multiplied by 100, it generates an indicator of 62.5 for that month.

Assume that the following month stock prices fall 0.8%, money grows by 0.5%, orders pick up 0.5%, and consumer confidence grows 3.5%. Applying the appropriate values, the components would add to 0 + 1.0 + 1.0 + 1.0 = 3.0. Divided by the number of components and multiplied by 100, it yields an index value of 75. The 20.0% increase in the index value means more components of the composite index are rising. Given this result, an analyst can be more confident that the higher composite index value actually represents broader movements in the economy. In general, a diffusion index does not reflect outliers in any component (like a straight arithmetic mean would do) but instead tries to capture the overall change common to all components.

EXAMPLE 10

Economic Indicators

1 Leading, lagging, and coincident indicators are:

- **A** the same worldwide.
- **B** based on historical cyclical observations.
- **C** based on Keynesian and/or Monetarist theory.

2 A diffusion index:

- **A** measures growth.
- **B** reflects the consensus change in economic indicators.
- **C** is roughly analogous to the indexes used to measure industrial production.

3 In the morning business news, a financial analyst, Kevin Durbin, learned that average hourly earnings had increased last month. The most appropriate action for Durbin is to:

- **A** call his clients to inform them of a good trading opportunity today.
- **B** examine other leading indicators to see any confirmation of a possible turning point for the economy.
- **C** use the news in his research report as a confirmation for his belief that the economy has recovered from a recession.

4 The following table shows the trends in various economic indicators in the two most recent quarters:

Economic Indicator	Trend
Interest rate spread between long-term government bonds and overnight borrowing rate	Narrowing
New orders for capital goods	Declining
Residential building permits	Declining
Employees on non-agricultural payrolls	Turned from rising to falling
Manufacturing and trade sales	Stable
Average duration of unemployment	Small decline
Change in unit labor costs	Rising

Given the information, this economy is *most likely* experiencing a:

- **A** continuing recession.
- **B** peak in the business cycle.
- **C** strong recovery out of a trough.

5 The indicator indexes created by various organizations or research agencies:

- **A** include only leading indicators to compute their value.
- **B** are highly reliable signals on the phase of business cycles.
- **C** evolve over time in terms of composition and computation formula.

6 Which one of the following trends in various economic indicators is *most* consistent with a recovery from a recession?

- **A** A declining inventory–sales ratio and stable industrial production index

B A rising broad stock market index and unit labor costs turning from increasing to decreasing

C A decrease in average weekly initial claims for unemployment insurance and an increase in aggregate real personal income

7 Which of the following statements gives the *best* description of nowcasting?

A It is a method used to forecast future trends in economic variables based on their past and current values.

B It is a method used for real-time monitoring of economic and financial variables to continuously assess current conditions and provide an estimate of the current state.

C It is a technique used to study past relationships between variables to explain which ones have explained the path of a particular variable of interest.

Solution to 1:

B is correct. The recognition of economic indicators is based on empirical observations for an economy.

Solution to 2:

B is correct. The diffusion indexes are constructed to reflect the common trends embedded in the movements of all the indicators included in such an index.

Solution to 3:

B is correct. Financial analysts need to synthesize the information from various indicators in order to gather a reliable reading of the economic trends.

Solution to 4:

B is correct. The first three indicators are leading indicators, and all of them are indicating an impending recession, which means the economy has reached the peak in this cycle. Non-agricultural payrolls and manufacturing and trade sales are coincident indicators. The trends in these two variables further indicate that the economy may begin to decline. The trends in the last two indicators—both lagging indicators—indicate that the economy may either continue to grow or it may be close to a peak. Aggregating the signals given by all three groups of economic indicators, it appears the economy may be near the peak of a business cycle.

Solution to 5:

C is correct. The indicator indexes are constantly updated for their composition and methodology based on the accumulation of empirical knowledge, and they can certainly include more than just leading indicators.

Solution to 6:

C is correct. The improving leading indicator, average weekly initial claims for unemployment insurance, and the improving coincident indicator, aggregate real personal income, are most consistent with an economic recovery. Even though a declining inventory-to-sales ratio, a lagging indicator, is consistent with an early recovery, the coincident indicator, the stable industrial production index, does not support that conclusion. Although a rising stock market index can signal economic expansion, the lagging indicator, the unit labor costs, has peaked, which is more consistent with a recession.

Solution to 7:

B is correct. Nowcasting involves the use of techniques to estimate the present state. A is incorrect because nowcasting aims to estimate the present, not forecast the future. C is incorrect because the focus of nowcasting is to estimate the current, present value of a variable, such as GDP.

10 UNEMPLOYMENT

f describe types of unemployment, and compare measures of unemployment

Many governments and central banks have economic policy objectives related to limiting the rate at which citizens are unemployed and containing price inflation (i.e., preserving the purchasing power of a domestic currency). The relationships of these variables to the business cycle are discussed in the following sections. In general, unemployment is at its highest somewhat after the recovery has started and is at its lowest somewhat after the peak of the economy.

10.1 Unemployment

An overheated economy often leads to inflation when unemployment is very low. Workers ask for higher wages because they expect prices of goods and services to keep going up; at the same time, they have market power against employers because there are few available workers to be hired. The upward pressure on wages coupled with the impact of wage escalator clauses (automatic increases in wages as the consumer price index grows) trigger a price–wage inflationary spiral. This issue was a particular problem in industrialized countries during the 1960s and 1970s and to a degree remains relevant today.

A key aspect in this process is inflation expectations. Because inflation expectations are high, the request for higher wages is stronger, which induces employers to increase prices in advance to keep their profit margins stable. This avalanche process grows with time. Central banks act, sometimes drastically, to slow down the economy and reset inflationary expectations throughout the economy at a low level so that if everyone expects low inflation, the inflationary spiral itself will stop. These actions may trigger a deep recession. Therefore, whenever a financial analyst sees signs of a price–wage spiral in the making, a reasonable response would be to consider the effect of both high inflation and sharp tightening of monetary policy.

This example shows that measures of labor market conditions are important in assessing whether an economy is at risk of cyclical downturn.

The following are definitions of a few terms used to summarize the state of the labor market:

- **Employed**: The number of people with a job. This figure normally does not include people working in the informal sector (e.g., unlicensed cab drivers, illegal workers).
- **Labor force**: The number of people who either have a job or are actively looking for a job. This number excludes retirees, children, stay-at-home parents, full-time students, and other categories of people who are neither employed nor actively seeking employment.
- **Unemployed**: People who are actively seeking employment but are currently without a job. Some special subcategories include:

  - **Long-term unemployed**: People who have been out of work for a long time (more than three to four months in many countries) but are still looking for a job.
  - **Frictionally unemployed**: People who are not working at the time of filling out the statistical survey because they are taking time to search for a job that matches their skills, interests, and other preferences better than what is currently available, or people who have left one job and are about to start another job. The frictionally unemployed includes people who have voluntarily left their previous positions to change their jobs—in other words, they are "between jobs"—and those new entrants or re-entrants into the labor force who have not yet found work. Frictional unemployment is short-term and transitory in nature.
- **Unemployment rate**: The ratio of unemployed to labor force.
- **Participation rate** (or activity ratio): The ratio of labor force to total population of working age (i.e., those between 16 and 64 years of age).
- **Underemployed**: A person who has a job but either has the qualifications to work at a significantly higher-paying job or works part-time but desires a full-time position. For example, a lawyer who is out of work and takes a job in a bookstore could call herself underemployed. This lawyer would not count as unemployed for the computation of the unemployment rate (she does have a job, even if it may not be her highest-paying job). Although the unemployment rate statistic is criticized for not taking the issue of underemployment into account, it may be difficult to classify whether a person is truly underemployed—for example, the lawyer may find legal work too stressful and prefers working at the bookstore.
- **Discouraged worker**: A person who has stopped looking for a job. Perhaps because of a weak economy, the discouraged worker has given up seeking employment. Discouraged workers are statistically outside the labor force (similar to children and retirees), which means they are not counted in the official unemployment rate. During prolonged recessions, the unemployment rate may actually decrease because many discouraged workers stop seeking work. It is important to observe the participation rate together with the unemployment rate to understand if unemployment is decreasing because of an improved economy or because of an increase in discouraged workers. Discouraged workers and underemployed people may be considered examples of "hidden unemployment."
- **Voluntarily unemployed**: A person voluntarily outside the labor force, such as a jobless worker refusing an available vacancy for which the wage is lower than their threshold or those who retired early.

10.1.1 *The Unemployment Rate*

The unemployment rate is certainly the most quoted measure of unemployment; it attempts to measure those people who have no work but would work if they could find it, generally stated as a percentage of the overall labor force. In the United States, the indicator emerges from a monthly survey of households by the US Bureau of Labor Statistics. The survey asks how many household members have jobs and how many of working age do not have jobs but are seeking work. Other statistical bureaus rely on other sources for the calculation, using claims for unemployment assistance, for instance, or their equivalent. Some statistical bureaus measure the labor force simply as those of working age, regardless of whether they are ready or willing to work. These differences can make precise international comparisons problematic. One solution is to use the International Labour Organization (ILO) statistics that try to estimate on a consistent basis. As indicated earlier, some statistical agencies add perspective with other measures—for example, what proportion of those who have ceased work are discouraged, underemployed, or have opted out of the labor force for other reasons or are working part-time.

Although these various unemployment measures provide insight to the state of the economy, they are inaccurate in pointing to cyclical directions for two primary reasons, both of which make unemployment a lagging economic indicator of the business cycle.

One reason is that the unemployment rate tends to point to a past economic condition—that is, it lags the cycle—because the labor force expands and declines in response to the economic environment. Compounding the inaccuracy, discouraged workers cease searching for work when times get hard, reducing the number typically counted as unemployed and making the job market look stronger than it really is. Conversely, when the job market picks up, these people return to the search; however, because they seldom find work immediately, they at least initially raise the calculation of those unemployed, giving the false impression of the lack of recovery in the jobs market, when, in fact, it is the improvement that brought these people back into the labor force in the first place. Sometimes this cyclical flow of new job seekers is so great that the unemployment rate actually rises even as the economic recovery gains momentum. Those agencies that measure the labor force in terms of the working-age population avoid this bias because the working-age population measure is unaffected by economic conditions in the labor market. But this approach introduces biases of its own, such as counting as unemployed those people who have severe disabilities and could never seek work.

Another reason the unemployment indicator tends to lag the cycle comes from the typical reluctance of businesses to lay off people. The reluctance may stem from a desire to retain good workers for the long run, or it may just reflect constraints written into labor contracts that make layoffs expensive. The reluctance makes the various measures of unemployment rise more slowly as the economy slides into recession than they otherwise might. Then as the recovery develops, a business waits to hire until it has fully employed the workers it has kept on the payroll during the recession; this delay causes decreases in the unemployment rate to lag in the cyclical recovery, sometimes for a long time.

10.1.2 *Overall Payroll Employment and Productivity Indicators*

To get a better picture of the employment cycle, practitioners often rely on more straightforward measures of payroll growth. By measuring the size of payrolls, practitioners sidestep such issues as the ebb and flow of discouraged workers. These statistics, however, have biases of their own. It is hard, for instance, to count employment in smaller businesses, which may be significant drivers of employment growth. Still, there is a clear indication of economic trouble when payrolls shrink and a clear indication of recovery when they rise.

The examination of other measures can also assist in understanding the employment situation and its use in determining cyclical directions. Two additional measures are hours worked, especially overtime, and the use of temporary workers. A business does not want to make mistakes with full-time staff, either hiring or firing. Thus, at the first signs of economic weakness, managers cut back hours, especially overtime. Such movements can simply reflect minor month-to-month production shifts, but if followed by cutbacks in part-time and temporary staff, the picture gives a strong signal of economic weakness, especially if confirmed by other independent indicators. Similarly, on the cyclical upswing, a business turns first to increases in overtime and hours. If a business then increases temporary staffing, it gives a good signal of economic recovery long before any movement in rehiring full-time staff again, especially if confirmed by independent cyclical indicators.

Productivity measures also offer insight into this cyclical process. Because productivity is usually measured by dividing output by hours worked, a business's tendency to keep workers on the payroll even as output falls usually prompts a reduction in measured productivity. If measures are available promptly enough, this sign of cyclical weakness might precede even the change in hours. This drop in productivity precedes any change in full-time payrolls. Productivity also responds promptly when business conditions improve and the business first begins to utilize its underemployed workers, which occurs earlier than any upturn in full-time payrolls.

On a more fundamental level, productivity can also pick up in response to technological breakthroughs or improved training techniques. As already mentioned, such changes affect potential GDP. If strong enough, they can negatively affect employment trends, keeping them slower than they would be otherwise by relieving the need for additional staff to increase production. But these influences usually unfold over decades and mean little to cyclical considerations, which, at most, unfold over years. What is more, there are few statistical indicators to gauge the onset of technological change, confining analysts to the use of anecdotal evidence or occasional longitudinal studies.

EXAMPLE 11

Analyzing Unemployment

1. At the peak of the business cycle, if the unemployment rate is low, the majority of the unemployed are *most likely*:
   - **A** discouraged workers.
   - **B** long-term unemployed.
   - **C** frictionally unemployed.
2. As an economy starts to recover from a trough in the business cycle, the unemployment rate is *most likely* to:
   - **A** continue to rise with a decline in the number of discouraged workers.
   - **B** start to decline with an increase in the number of discouraged workers.
   - **C** continue to rise with an increase in the number of discouraged workers.
3. An analyst observes that the unemployment rate is high and rising, whereas productivity and hours worked have declined. The analyst is *most likely* to conclude that the labor market is signaling the:
   - **A** end of a recession.
   - **B** deepening of a recession.
   - **C** peak of the business cycle.

Solution to 1:

C is correct. At the peak of a business cycle, the labor market is usually tight, and people become unemployed largely because they are either "between jobs" or they have entered or reentered the labor force but have not yet found work.

Solution to 2:

A is correct. As the economy starts to recover, discouraged workers return to the labor force and start looking for jobs, which increases both the number of unemployed and the size of the labor force. The unemployment rate rises because the rise in the unemployed population is proportionately larger than the increase in the size of the labor force. B and C are incorrect because an increase in the number of discouraged workers typically occurs when the economy is contracting.

Solution to 3:

B is correct. High and rising unemployment, declining hours worked, and falling productivity are all signs of a weak economy getting weaker. When the economy first slows down, businesses cut back employees' hours. As the recession deepens, they then lay off employees, leading to a higher unemployment rate. Yet, because workforce turnover is costly for businesses, the scale of the layoff can be less than the decline in output, resulting in a decline in productivity. A is incorrect because toward the end of a recession, businesses are hesitant to increase hiring and instead use more overtime, increasing both productivity and the hours worked. C is incorrect because at the peak of a business cycle, the unemployment rate is usually low and the level of hours worked is high.

11

INFLATION

g explain inflation, hyperinflation, disinflation, and deflation

The overall price level changes at varying rates during different phases of a business cycle. Thus, when studying business cycles, it is important to understand this phenomenon. In general, the inflation rate is pro-cyclical (that is, it goes up and down *with* the cycle), but with a lag of a year or more.

Inflation refers to a sustained rise in the overall level of prices in an economy. Economists use various price indexes to measure the overall price level, also called the aggregate price level. The inflation rate is the percentage change in a price index—that is, the speed of overall price level movements. Investors follow the inflation rate closely not only because it can help to infer the state of the economy, but also because an unexpected change may result in a change in monetary policy, which can have a large and immediate impact on asset prices. In developing countries, very high inflation rates can lead to social unrest or even shifts of political power, which constitutes political risk for investments in those economies.

Central banks, the monetary authority in most economies, monitor the domestic inflation rates closely when conducting monetary policy. Monetary policy determines interest rates and the available quantities of money and loans in an economy. A high inflation rate combined with fast economic growth and low unemployment usually indicates the economy is overheating, which may trigger some policy movements to cool it down. However, if a high inflation rate is combined with a high level of unemployment and a slowdown of the economy—an economic state known as stagflation (stagnation plus inflation)—the economy will typically be left to correct itself because no short-term economic policy is deemed effective.

11.1 Deflation, Hyperinflation, and Disinflation

Various terms are related to the levels and changes of the inflation rate:

- **Deflation**: A sustained decrease in aggregate price level, which corresponds to a negative inflation rate—that is, an inflation rate of less than 0%.

- **Hyperinflation**: An extremely fast increase in aggregate price level, which corresponds to an extremely high inflation rate—for example, 500% to 1,000% per year.
- **Disinflation**: A decline in the inflation rate, such as from around 15% to 20% to 5% or 6%. Disinflation is very different from deflation because even after a period of disinflation, the inflation rate remains positive and the aggregate price level keeps rising (although at a slower speed).

Inflation means that the same amount of money can purchase less real goods or services in the future. So, the value of money or the purchasing power of money decreases in an inflationary environment. When deflation occurs, the value of money actually increases. Because most debt contracts are written in fixed monetary amounts, the liability of a borrower also rises in real terms during deflation. As the price level falls, the revenue of a typical company also falls during a recession. Facing increasing real debt, a company that is short of cash usually cuts its spending, investment, and workforce sharply. Less spending and high unemployment further exacerbate the economic contraction. To avoid getting too close to deflation, the consensus on the preferred inflation rate is around 2% per year for developed economies. Deflation occurred in the United States during the Great Depression and briefly during the recession following the global financial crisis of 2008–2009. Since the late 1990s, Japan has experienced several episodes of deflation.

Hyperinflation usually occurs when large scale government spending is not backed by real tax revenue and the monetary authority accommodates government spending by increasing the money supply. Hyperinflation may also be caused by the shortage of supply created during or after a war, economic regime transition, or prolonged economic distress of an economy caused by political instability. During hyperinflation, people are eager to change their cash into real goods because prices are rising very fast. As a result, money changes hands at an extremely high frequency. The government also has to print more money to support its increased spending. As more cash chases a limited supply of goods and services, the rate of price increases accelerates. After World War I, a famous case of hyperinflation occurred in Germany from 1923 to 1924. During the peak of this episode, prices doubled every 3.7 days. After World War II, Hungary experienced a severe hyperinflation during which prices doubled every 15.6 hours at its peak in 1946. In 1993, the inflation rate in Ukraine peaked at 10,155% per year. In January 1994, the *monthly* inflation rate peaked at 313 million% in Yugoslavia. The most recent hyperinflation in Zimbabwe reached a peak of *monthly* inflation at 79.6 billion% in the middle of November 2008. Because the basic cause for hyperinflation is too much money in circulation, regaining control of the money supply is the key to ending hyperinflation.

Exhibit 19 shows recent episodes of disinflation in selected countries around the world. The first episode happened during the early 1980s. Because of the two oil crises in the 1970s, many countries around the world were experiencing high levels of inflation. In Exhibit 19, the annual inflation rates in most countries around 1980 ranged between 10% and 20%. Even though this level is still far from hyperinflation, it generated social pressure against inflationary monetary policy. At the cost of a severe recession early in the 1980s, these countries brought inflation rates down to around 5%, on average, by 1985. In the early 1990s, inflationary experiences varied widely in world markets as some countries entered recessions, such as the United States and the United Kingdom, and others boomed. However, from the beginning to the end of the 1990s, there was a broad-based decline in inflation rates; in some countries, annual inflation rates were below 2% by the end of the decade. In many countries, the decline in inflation was attributed to high productivity growth rates.

Exhibit 19 Two Episodes of Disinflation around the World Annual Inflation Rates

	First Episode					Second Episode			
Year	**1979**	**1980**	**1983**	**1984**	**1985**	**1990**	**1991**	**1998**	**1999**
Country									
Australia	9.1	10.2	10.1	3.9	6.7	7.3	3.2	0.9	1.5
Canada	9.1	10.1	5.9	4.3	4.0	4.8	5.6	1.0	1.7
Finland	7.5	11.6	8.4	7.1	5.2	6.1	4.3	1.4	1.2
France	10.6	13.6	9.5	7.7	5.8	3.2	3.2	0.6	0.5
Germany	4.0	5.4	3.3	2.4	2.1	2.7	4.0	1.0	0.6
Italy	14.8	21.1	14.6	10.8	9.2	6.5	6.3	2.0	1.7
Japan	3.7	7.8	1.9	2.3	2.0	3.1	3.3	0.7	−0.3
South Korea	18.3	28.7	3.4	2.3	2.5	8.6	9.3	7.5	0.8
Spain	15.7	15.6	12.2	11.3	8.8	6.7	5.9	1.8	2.3
Sweden	7.2	13.7	8.9	8.0	7.4	10.4	9.4	−0.3	0.5
United Kingdom	13.4	18.0	4.6	5.0	6.1	7.0	7.5	1.6	1.3
United States	11.3	13.5	3.2	4.3	3.5	5.4	4.2	1.6	2.2
Average	10.4	14.1	7.2	5.8	5.3	6.0	5.5	1.6	1.2
G–7 Countries	9.6	12.5	4.6	4.6	3.9	4.8	4.4	1.3	1.4

Source: The Organisation for Economic Co-Operation and Development (OECD).

11.2 Measuring Inflation: The Construction of Price Indexes

h explain the construction of indexes used to measure inflation

Because the inflation rate is measured as the percentage change of a price index, it is important to understand how a price index is constructed so that the inflation rate derived from that index can be accurately interpreted. A price index represents the average prices of a basket of goods and services, and various methods can be used to average the different prices. Exhibit 20 shows a simple example of the change of a consumption basket over time.

Exhibit 20 Consumption Basket and Prices over Two Months

Time	**January 2019**		**February 2019**	
Goods	**Quantity**	**Price**	**Quantity**	**Price**
Rice	50 kg	¥3/kg	70 kg	¥4/kg
Gasoline	70 liters	¥4.4/liter	60 liters	¥4.5/liter

For January 2019, the total value of the consumption basket is:

Value of rice + Value of gasoline = (50 × 3) + (70 × 4.4) = ¥458.

A price index uses the relative weight of a good in a basket to weight the price in the index. Therefore, the same consumption basket in February 2019 is worth:

Value of rice + Value of gasoline = (50 × 4) + (70 × 4.5) = ¥515.

The price index in the base period is usually set to 100. So, if the price index in January 2019 is 100, then the price index in February 2019 is

$$\text{Price index in February 2019} = \frac{515}{458} \times 100 = ¥112.45 \text{ and}$$

$$\text{Inflation rate} = \frac{112.45}{100} - 1 = 0.1245 = 12.45\%.$$

A price index created by holding the composition of the consumption basket constant is called a Laspeyres index. Most price indexes around the world are Laspeyres indexes because the survey data on the consumption basket are only available with a lag. In many countries, the basket is updated every five years. Because most price indexes are created to measure the cost of living, however, simply using a fixed basket of goods and services has three serious biases:

- Substitution bias: As the price of one good or service rises, people may substitute it with other goods or services that have a lower price. This substitution will result in an upward bias in the measured inflation rate based on a Laspeyres index.
- Quality bias: As the quality of the same product improves over time, it satisfies people's needs and wants better. One such example is the quality of cars. Over the years, the prices of cars have been rising while the safety and reliability of cars have been enhanced. If not adjusted for quality, the measured inflation rate will experience another upward bias.
- New product bias: New products are frequently introduced, but a fixed basket of goods and services will not include them. In general, this situation again creates an upward bias in the inflation rate.

It is relatively easy to resolve the quality bias and new product bias. Many countries adjust for the quality of the products in a basket, a practice called hedonic pricing. New products can be introduced into the basket over time. The substitution bias can be somewhat resolved by using the chained price index formula. One such example is the Fisher index, which is the geometric mean of the Laspeyres index and the Paasche index. The latter is an index formula using the current composition of the basket. Using the consumption basket for February 2019 in Exhibit 20, the value of the Paasche index is

$$\begin{aligned}\text{Paasche Index}_{02/2019} = I_P &= \frac{(70 \times 4) + (60 \times 4.5)}{(70 \times 3) + (60 \times 4.4)} \times 100 \\ &= \frac{550}{474} \times 100 = 116.03\end{aligned}$$

The value of the Fisher index is

$$\text{Fisher Index}_{02/2019} = \sqrt{I_P \times I_L} = \sqrt{116.03 \times 112.45} = 114.23,$$

where I_L is the Laspeyres index.

11.3 Price Indexes and Their Usage

i. compare inflation measures, including their uses and limitations

Most countries use a consumer price index (CPI) specific to the domestic economy to track inflation. Exhibit 21 shows the different weights for various categories of goods and services in the consumer price indexes of different countries.

Exhibit 21 The Consumption Basket of Different Consumer Price Indexes

Country	China	India	India	Germany	United States	United States
Name of Index	CPI	CPI(UNME)	CPI(Urban)[c]	HICP	CPI-U	PCE
Year[a]	2016[b]	1984/85	2012	2017/18	2017/18	2009
Category (%):						
Food and Beverage	30	47.1	37.7	14	15.3	13.4
Housing and Utility	21	21.9	27.2	31.7	37.9	17.3
Furniture	6	2	3.9	5	4.2	3.4
Apparel	7.5	7	5.6	4.5	3.3	3.3
Medical Care	8	2.5	4.8	4.4	7.7	16.9
Transportation and Communication	13	5.2	9.7	16.5	15.3	9.3[d]
Education and Recreation	10.5	6.8	7.6	12.3	9.1	8.9[e]
Others	4	7.5	3.5	11.6	7.2	27.6

[a] The base year of the weights, where it is appropriate.
[b] Weights for China are not released publicly. Imputed numbers given as of 2018 are from the Reserve Bank of Australia.
[c] India redefined the CPI bundle in 2012 for urban consumers.
[d] Includes only transportation expenditures by consumers.
[e] Includes only recreational expenditures by consumers.
Sources: Government websites and authors' calculations.

As shown in Exhibit 21, in different countries the consumer price indexes have different names and different weights on various categories of goods and services. For example, food weights are higher in the CPI for China and India but less for the developed countries; a greater proportion of income of the average consumer goes to food in the developing countries of China and India than in the developed countries shown in Exhibit 21. For India, the weights across categories change dramatically over time as the country has developed. Weights across the categories change across time in developed countries too, but these changes tend to be much smaller. The scope of the index is also different among countries. For China and Germany, the surveys used to collect data for CPI cover both urban and rural areas. The CPI for the United States covers only urban areas using a household survey, which is why it is called the CPI-U. On the other hand, using business surveys, the personal consumption expenditure (PCE) price index covers all personal consumption in the United States.

The producer price index (PPI) is another important inflation measure. The PPI reflects the price changes experienced by domestic producers in a country. Because price increases may eventually pass through to consumers, the PPI can influence the future CPI. The items in the PPI include fuels, farm products (such as grains and meat), machinery and equipment, chemical products (such as drugs and paints), transportation equipment, metals, pulp and paper, and so on. These products are usually further grouped by stage-of-processing categories: crude materials, intermediate materials, and finished goods. Similar to the CPI, scope and weights vary among countries. The differences in the weights can be much more dramatic for the PPI than for the CPI because different countries may specialize in different industries. In some countries, the PPI is called the wholesale price index (WPI).

As an important inflation indicator, many economic activities are indexed to a certain price index. For example, in the United States, Treasury Inflation-Protected Securities (TIPS) adjusts the bond's principal according to the US CPI-U index. The terms of labor contracts and commercial real estate leases may adjust periodically according to the CPI. Recurring payments in business contracts can be linked to the PPI or its sub-indexes for a particular category of products.

Central banks usually use a consumer price index to monitor inflation. For example, the European Central Bank (ECB), the central bank for the European Union (EU), focuses on the Harmonised Index of Consumer Prices (HICP). Each member country in the EU first reports their own individual HICP, and then Eurostat, the statistical office for the EU, aggregates the country-level HICPs with country weights. But there are exceptions. The Reserve Bank of India follows the inflation in India using a WPI. Because food items only represent about 27% in the India WPI (much lower than the 70% in the India rural CPI), the rural CPIs can rise faster than the WPI when there is high food price inflation. Besides the weight differences, the wholesale prices in the WPI also understate market prices because they do not consider retail margins (markups). The choice of inflation indicator may also change over time. The central bank of the United States, known as the Federal Reserve Board (the Fed), once focused on the CPI-U produced by the Bureau of Labor Statistics under the US Department of Labor. Because the CPI-U is a Laspeyres index and has the previously discussed upward biases, however, the Fed switched in 2000 to the PCE index, a Fisher index produced by the Bureau of Economic Analysis under the US Department of Commerce. The PCE index also has the advantage that it covers the complete range of consumer spending rather than just a basket.

HEADLINE AND CORE INFLATION

Headline inflation refers to the inflation rate calculated based on the price index that includes all goods and services in an economy. **Core inflation** usually refers to the inflation rate calculated based on a price index of goods and services except food and energy. Policy makers often choose to focus on the core inflation rate when reading the trend in the economy and making economic policies As they try to avoid overreaction to short-term fluctuations in food and energy prices that may not have a significant impact on future headline inflation.

The ultimate goal for policy makers is to control headline inflation, which reflects the actual cost of living. The fluctuations in the prices of food and energy are often the result of short-term changes in supply and demand. These changes in the prices of energy, particularly oil, are internationally determined and not necessarily reflective of the domestic business cycle. These imbalances may not persist, or even if some changes are permanent, the economy may be able to absorb them over time. These possibilities make headline inflation a noisy predictor. The core inflation rate may be a better signal of the trend in domestically driven inflation. To the extent that some trends in the headline inflation rate are permanent, policy makers need to pay attention to these as well.

Besides tracking inflation, financial analysts also use the price index to deflate GDP (i.e., to eliminate the price effect in nominal GDP data so as to identify trends in real economic growth). Many countries publish a particular price index, called the GDP deflator, for that purpose. Sub-indexes are also commonly available and may prove more valuable to an analyst with an interest in a particular industry or company.

SUB-INDEXES AND RELATIVE PRICES

As mentioned previously, a sub-index refers to the price index for a particular category of goods or services. **Relative price** is the price of a specific good or service in comparison with those of other goods and services. Good examples for relative prices include the prices for food and energy. The movements in a sub-index or a relative price may be difficult to detect in the headline inflation rate. Because macroeconomic policy decision makers rely heavily on the headline inflation rate, they may not be aware of price movements at the sub-index level. These price movements, however, can be very useful for analyzing the prospects of an industry or a company. For example, if the producer price index for the machinery used by an industry rises quickly, the allowable capital depreciation permitted by the existing tax code may not generate sufficient tax benefits for the companies in that industry to meet future replacement expenses. The future profitability of the industry may decline for this reason. The decline in prices for flat screen televisions provides an example of relative price movements. The price drop for these TVs may help to lower inflation pressure but can hurt manufacturers' profits.

EXAMPLE 12

Inflation

1 Which of the following statements regarding the movements of overall price levels is *most* accurate?

A Disinflation means that the overall price level declines.

B Deflation occurs when the inflation rate turns negative.

C When the price of chicken rises, the inflation rate will increase.

2 Deflation can exacerbate a recession because firms may reduce their investments and hiring when:

A the slower pace of inflation lowers aggregate demand.

B their revenues decline but their debt burden rises in real terms.

C the prices of their products continue to fall because of intense competition.

3 Which one of the following economic phenomena related to inflation cannot be determined by using observations of the inflation rate alone?

A Deflation

B Stagflation

C Hyperinflation

4 If a price index is calculated based on a fixed basket of goods, in an inflationary environment the inflation rate calculated based on this index over time will:

A overstate the actual cost of living.

B understate the actual cost of living.

C track the actual cost of living quite closely.

5 To adjust nominal economic growth for general price level changes in a country, an analyst would prefer to use:

A the CPI.

B the GDP deflator.

C the Personal Consumption Expenditures (PCE) index.

6 To estimate the trends in sales and production costs of a given industry, an analyst would prefer to collect data on:

A the sub-indexes of the wholesale price index (WPI) for that industry.

B the sub-indexes of both the CPI and WPI that are relevant to the industry.

- **C** the sub-indexes of the CPI relevant to the output and inputs of that industry.

7 Compared with core inflation, headline inflation:

- **A** has an upward bias.
- **B** is more subject to short-term market conditions.
- **C** can more accurately predict future inflation.

Solution to 1:

B is correct. When the inflation rate falls below zero—that is, the overall price level declines—the economy is experiencing deflation. A is incorrect, because disinflation indicates that the overall price level is rising but at a slower pace. C is incorrect because inflation measures are designed to reflect changes in the overall price level. Consumption baskets in modern economies usually contain a large number of goods and services, thus the price of a particular product usually cannot significantly influence the overall price level.

Solution to 2:

B is correct. As the prices of the output of firms fall, the firms receive lower revenues. Because the nominal amount of debt that firms carry is usually fixed, lower general price levels leads to higher debt balances in real terms. These two forces push firms closer to default, so they may scale back spending on investments and labor, which further lowers the aggregate demand and pushes the general price level even lower. In macroeconomic analysis, it is usually the changes in aggregate demand that influence inflation instead of the reverse causality. Furthermore, neither inflation fluctuations nor aggregate demand shifts explain the potential damaging effect of deflation. Price decline attributable to the competitive environment is a microeconomic phenomenon that is not sufficient to explain the macroeconomic impact of deflation.

Solution to 3:

B is correct. A high inflation rate alone does not indicate stagflation, which happens if high unemployment occurs together with high inflation.

Solution to 4:

A is correct. Upward biases, such as the substitution bias or quality bias, will overstate the actual cost of living.

Solution to 5:

B is correct. The GDP deflator reflects the prices of the goods and services produced domestically. Both the CPI and PCE indexes are constructed using consumption baskets, and the components of a consumption basket can be very different from the components of output of that same country.

Solution to 6:

B is correct. A sub-index of the CPI reflects the market price changes of the products of an industry, whereas a sub-index of the WPI reflects the price changes of the inputs of an industry. The different composition of outputs and inputs of an industry need to be appropriately accounted for when selecting a price series. Furthermore, the WPI may not take the markups set by the industry into account.

Solution to 7:

B is correct. Headline inflation is heavily influenced by food and energy price fluctuations, which are affected by short-term supply and demand changes in these markets. These market conditions may not persist. It is also possible for an economy to absorb the price changes so that they will not have long-lasting impact on the headline inflation rate. This means headline inflation contains a great deal of noise and is not a reliable predictor of future inflation trends. The biases in various inflation measures are inherent in the index construction methodology and are not related to the price movements of the goods and services.

11.4 Explaining Inflation

j contrast cost-push and demand-pull inflation

Economists describe two types of inflation: cost-push, in which rising costs, usually wages, compel businesses to raise prices generally; and demand-pull, in which increasing demand raise prices generally, which then are reflected in a business's costs as workers demand wage hikes to catch up with the rising cost of living. Whatever the sequence by which prices and costs rise in an economy, the fundamental cause is the same: excessive demands—either for raw materials, finished goods, or labor—that outstrip the economy's ability to respond. The initial signs appear in the areas with the greatest constraints: the labor market, the commodity market, or some area of final output. Even before examining particular cost and price measures, when considering inflation, practitioners look to indicators that might reveal when the economy faces such constraints.

11.4.1 *Cost-Push Inflation*

Considering cost-push inflation, analysts may look at commodity prices because commodities are an input to production. But because wages are the single biggest cost to businesses, practitioners focus most particularly on wage-push inflation, which is tied to the labor market. Because the object is to gauge demand for labor relative to capacity, the unemployment rate is key, as well as measures of the number of workers available to meet the economy's expanding needs. Obviously, the higher the unemployment rate, the lower the likelihood that shortages will develop in labor markets; whereas the lower the unemployment rate, the greater the likelihood shortages will drive up wages. Because the unemployment rate generally only counts people who are looking for work, some practitioners argue that it fails to account for the economy's full labor potential. They state that a tight labor market will bring people out in search of work and ease any potential wage strains. To account for this issue and to modify the unemployment rate indicator, these practitioners also look at the participation rate of people in the labor force, arguing that it gives a fuller and more accurate picture of potential than the unemployment rate.

Analysis in this area recognizes that not all labor is alike. Structural factors related to training deficiencies, cultural patterns in all or some of the population, inefficiencies in the labor market, and the like can mean that the economy will effectively face labor shortages long before the unemployment rate reaches very low figures. This effective unemployment rate, below which pressure emerges in labor markets, is frequently referred to as the **non-accelerating inflation rate of unemployment** (NAIRU) or, drawing on the work of the Nobel Prize winner Milton Friedman, the **natural rate of unemployment** (NARU). Of course, these rates vary from one economy to another and over time in a single economy. It is this rate rather than full employment that determines when an economy will experience bottlenecks in the labor market and wage-push inflationary pressures.

Take, for example, the technology sector. It has grown so rapidly in some economies that training in the labor force cannot keep up with demand. This sector can, therefore, face shortages of trained workers and attendant wage pressures even though the economy as a whole seems to have considerable slack in the overall labor market. Until training (supply) catches up with demand, that economy may experience wage and inflation pressure at rates of unemployment that in other places and circumstances might suggest ample slack in the labor market and much less wage-push pressure.

Assessments of wage-push inflation also consider direct observations of wage trends that, when they accelerate, might force businesses to raise prices (initiating the wage-price spiral mentioned earlier in this reading). Statistical agencies provide a wide array of wage-cost indicators, such as hourly wage gauges, weekly earnings, and overall labor costs, including the outlays for benefits. Some of these indicators include the effects of special overtime pay or bonuses, and others do not. And although these measures give an idea of the cost to businesses and hence the kind of wage-push inflationary pressure, a complete picture only emerges when practitioners examine such trends alongside productivity measures.

Productivity, or output per hour, is an essential part of wage-push inflation analysis because the output available from each worker determines the number of units over which businesses can spread the cost of worker compensation. The greater each worker's output per hour, the lower the price businesses need to charge for each unit of output to cover hourly labor costs. And by extension, the faster output per hour grows, the faster labor compensation can expand without putting undue pressure on businesses' costs per unit of output. The equation for this **unit labor cost** (ULC) indicator, as it is called, is as follows:

$$\text{ULC} = W/O,$$

where

ULC = unit labor costs
O = output per hour per worker
W = total labor compensation per hour per worker

Many factors can affect labor productivity across time and between economies. The cyclical swings have already been described, as have the effects of technology and training. The pace of development also tends to increase worker productivity because the more sophisticated equipment, systems, and technologies workers have at their disposal, the higher their output per hour. Whatever causes the productivity growth, if it fails to keep up with worker compensation, unit costs to a business rise, and as a business tries to protect its profit margins, prices generally come under increasing upward pressure. Generally, this situation occurs because heavy demand for labor relative to available labor resources has pushed up compensation faster than productivity. Practitioners use a variety of indicators to identify cost- or wage-push inflationary pressure.

EXAMPLE 13

Unemployment Too High

Which of the following is **not** a problem with the NARU and NAIRU?

A They only work in monetarist models.

B They may change over time given changes in technology and economic structure.

C They do not account for bottlenecks in segments of the labor market (e.g., college graduates).

Solution:

A is correct. The NARU and NAIRU are the unemployment rates at which the inflation rate will not rise because of a shortage of labor. This concept does not tie to a particular school of macroeconomic models.

11.4.2 *Demand-Pull Inflation*

The search for indicators from the demand-pull side of the inflation question brings practitioners back to the relationship between actual and potential real GDP and industrial capacity utilization. The higher the rate of capacity utilization or the closer actual GDP is to potential, the more likely an economy will suffer shortages, bottlenecks, a general inability to satisfy demand, and hence, price increases. The more an economy operates below its potential or the lower the rate of capacity utilization, the less such supply pressure will exist and the greater likelihood of a slowdown in inflation, or outright deflation. In addition to these macro indicators, practitioners will also look for signs of inflationary pressure in commodity prices—in part because they are a cost to business, but more as a general sign of excess demand. For an individual economy, observations of commodity prices could be misleading because commodities trade in a global market and accordingly reflect global economic conditions more than those in an individual economy.

Taking a different perspective, Monetarists contend that inflation is fundamentally a monetary phenomenon. A surplus of money, they argue, will inflate the money price of everything in the economy. Stated in terms of straightforward supply and demand relationships, a surplus of money would bring down its value just as a surplus in any market would bring down the price of the product in excess. Because the price of money is stated in terms of the products it can buy, its declining value would have an expression in generally higher prices—that is, in inflation. This Monetarist argument, as it is called, finds a simpler expression in the old saying: "Inflation results when too much money chases too few goods." Although it seems distant from other explanations of inflation, in practice it is not that distinct. Excess money causes inflationary pressure by increasing liquidity, which ultimately causes a rapid rise in demand. In this sense, the Monetarist argument is a special case under the more general heading of demand-pull concepts of inflation. The practical distinction between the monetarist and other approaches is in identifying the initial cause of the demand excess.

Practitioners can track this effect by examining various money supply indicators, usually provided by the central bank. To detect inflationary or deflationary pressure, practitioners note acceleration or deceleration in monetary growth based on past trends. Obviously in the absence of a special explanation, acceleration signals the potential for inflationary pressure. In applying this approach, practitioners also compare monetary growth with the growth of the nominal economy, represented by nominal GDP. If monetary growth is outpacing the growth of the nominal economy, there is deemed to be inflationary potential. This is especially the case if monetary growth has also accelerated from its trend. There is a disinflationary or deflationary potential if monetary growth lags the economy's rate of expansion, especially if it has also decelerated from its trend.

INFLATION AND VELOCITY

Some practitioners view the likelihood of inflationary pressure from the vantage point of the ratio of nominal GDP to money supply, commonly called the "velocity of money." If this ratio remains stable around a constant or a historical trend, they see reason to look for relative price stability. If velocity falls, it could suggest a surplus of money that might have inflationary potential, but much depends on why it has declined. If velocity has fallen because a cyclical correction has brought down the GDP numerator relative to the money denominator, then practitioners view prospects as more likely to lead to a cyclical upswing to reestablish the former relationship than inflationary pressure. If velocity has fallen, however, because of an increase in the money denominator, then inflationary pressure becomes more likely. If velocity rises, financial analysts might be concerned about a shortage of money in the economy and disinflation or deflation.

The 2008–2009 global recession and financial crisis offers an extreme example of these velocity ambiguities. As the global economy slipped into recession, which held back the GDP numerator in velocity measures, central banks, most notably the Federal Reserve in the United States, tried to help financial institutions cope by injecting huge amounts of money into their respective financial systems, raising the velocity denominator. Velocity measures plummeted accordingly. The expectation is that subsequent GDP growth as economies and financial markets heal will bring velocity back to a more normal level and trend. That said, the fear is that the monetary surge will, over the very long run, lead to inflation. For policy makers, this situation has created a very difficult policy choice. On the one side, they need to sustain the supply of money to help their respective economies cope with the aftereffects of the financial crisis. On the other side, they need ultimately to withdraw any monetary excess to preclude potential inflationary pressures.

11.4.3 *Inflation Expectations*

Beyond demand-pull, monetary, and cost-push inflation considerations, practitioners also need to account for the effect of inflation expectations. Once inflation becomes embedded in an economy, businesses, workers, consumers, and economic actors of every kind begin to expect it and build those expectations into their actions. This reaction, in turn, creates an inflationary momentum of its own. Such expectations give inflation something of a self-sustaining character and cause it to persist in an economy even after its initial cause has disappeared. High inflation rates persisted in the 1970s and early 1980s in Europe and the United States on the basis of expectations—even after these economies had sunk into recession. The resulting slow or negative economic growth combined with high unemployment and rising inflation was termed “stagflation.”

Measuring inflation expectations is not easy. Some practitioners gauge expectations by relying on past inflation trends and on the assumption that market participants largely extrapolate their past experiences. In some markets, surveys of inflation expectations are available, although these are often biased by the way the questions are asked. Another indicator of inflation expectations becomes available when governments issue bonds, such as Treasury Inflation-Protected Securities (TIPS), that adjust in various ways to compensate holders for inflation. By comparing the interest available on these bonds with other government bonds that do not offer such inflation-linked adjustments, practitioners can gauge the general level of inflation expectations among market participants and factor it into their own inflation forecasts and strategies.

For example, if today’s yield on the 10-year nominal bond of a certain country is 3.5% and the yield on the 10-year inflation-protected bond of the same country is 1.5%, we infer that the market is pricing in a 3.5% – 1.5% = 2% average annual inflation over the next 10 years. However, this calculation needs to be treated with caution because the market for inflation-linked bonds is relatively small and thus yields can be influenced by other factors, such as the very strong demand from US pension funds seeking to match their liabilities.

EXAMPLE 14

Inflation

1 To examine whether there is inflationary pressure caused by rising costs, an analyst will *most likely* gather data on:

A the growth rates of money supply and nominal GDP.

B the unemployment rate, the NAIRU, and productivity growth.

C commodity prices, past inflation trends, and expected inflation surveys.

2 The most recent macroeconomic data for an economy are given in the following table:

Variable	**Value**
Hourly wage growth rate	3.4%
Unit labor cost growth rate	–0.25%
Nominal GDP growth rate	3.4%
Money supply growth rate	6.7%
Implied inflation rate from government issued inflation-linked securities	2.2%

Based on the information in the table, an analyst will conclude that current inflation pressure in this economy is *most likely* caused by:

A rising wages.

B rising inflation expectations.

C bottlenecks in increasing supply to satisfy demand.

3 Cost-push inflation *most likely* occurs when:

A unemployment rates are low.

B unemployment rates are high.

C unemployment is either high or low.

4 Unit labor costs measure:

A hourly wage rates.

B total labor compensation per hour.

C a combination of hourly wages and output.

5 Demand-pull inflation:

A is a discredited concept.

B depends on the movements in commodity prices.

C reflects the state of economic activity relative to potential.

6 Monetarists believe inflation:

A reflects the growth of money.

B is driven by the level of interest rates.

C is largely a cost-push phenomenon.

7 The inflationary potential of a particular inflation rate depends on the economy's NAIRU or NARU, which, in turn, depends in part on:

A the intensity of past cyclical swings.

B the bargaining power of trade unions.

C the skill set of the labor force relative to the economy's industrial mix.

8 Which of the following is *not* a problem with the NARU and NAIRU?

A They are not observable directly.

B They work only in monetarist models.

C They change over time given changes in technology and economic structure.

Solution to 1:

B is correct. Comparing the current unemployment rate with the NAIRU and productivity growth with the wage growth can help an analyst determine whether inflation may be rising because of higher costs (cost-push inflation). Comparing the monetary growth with nominal GDP growth is helpful to determine whether high demand is creating inflationary pressure (demand-pull inflation). Commodity price increases could be an indicator of either cost-push or demand-pull inflation but may contain limited information in some situations. Past inflation trends and surveys on inflation expectations can help to gauge expected inflation rates; inflation expectations can be a driver of inflation even in the absence of the original underlying cause.

Solution to 2:

C is correct. The table shows that growth in the money supply has outpaced nominal GDP growth, which can result in too much money chasing too few goods. In other words, inflation pressure results from demand beyond the economy's current capacity to produce. Although the wage rate is rising, the negative unit labor cost growth rate indicates an increase in productivity. Thus, it is unlikely the economy will experience cost-push inflation. The implied inflation rate is very modest, which is unlikely to lead to a rising inflation rate.

Solution to 3:

A is correct. When unemployment is below the NAIRU, there is a shortage of labor that pushes up labor cost.

Solution to 4:

C is correct. Unit labor costs reflect the labor cost in each unit of output.

Solution to 5:

C is correct. When the economy is operating above its potential capacity allowed by the resources available, inflation will start to rise.

Solution to 6:

A is correct. Monetarists emphasize the role of money growth in determining the inflation rate, especially in the long run. As Milton Friedman famously put it: "Inflation is always and everywhere a monetary phenomenon."

Solution to 7:

C is correct. If the skill set of a large part of the labor force cannot satisfy the hiring need from the employers, the NAIRU of such an economy can be quite high.

Solution to 8:

B is correct. The NAIRU or NARU reflects the potential of an economy and thus cannot be directly observed from the economic data. They also change over time depending on technological progress and social factors.

SUMMARY

- Business cycles are recurrent expansions and contractions in economic activity affecting broad segments of the economy.

- Business cycles are a fundamental feature of market economies, but their amplitude and/or length vary considerably.
- Business cycles can be split into many different phases. The investment industry typically refers to four phases of the cycle: recovery, expansion, slowdown, and contraction, with the peak in output occurring during the slowdown phase and the trough in output occurring in the recovery phase.
- Classical cycle refers to fluctuations in the level of economic activity when measured by GDP in volume terms.
- Growth cycle refers to fluctuations in economic activity around the long-term potential trend growth level with focus on how much actual economic activity is below or above trend growth in economic activity.
- Growth rate cycle refers to fluctuations in the growth rate of economic activity.
- Credit cycles describe the changing availability—and pricing—of credit. When the economy is strong or improving, the willingness of lenders to extend credit, and on favorable terms, is high. When the economy is weak or weakening, lenders pull back, or “tighten” credit, by making it less available and more expensive, contributing to asset values and further economic weakness and higher defaults.
- Credit cycles are relevant due to the importance of credit in the financing of construction and the purchase of property.
- The extent of business cycle fluctuations and the duration of recessions and recoveries are often shaped by linkages between business and credit cycles.
- Key economic variables change through the business cycle. The level of business investment shows significant changes over the cycle.
- Employment levels follow the cycle with a delay as companies initially use overtime before hiring after the onset of recovery and then reduce overtime before reducing employment as the economy passes its peak and enters contraction.
- Consumer spending, the largest component of output, follows cyclical patterns as workers make decisions based on their levels of income, wage growth, and employment outlook.
- Spending on consumer durables is the most cyclical consumer activity, followed by services and consumer staples, which are less affected by the cyclicality of the economy.
- The size of inventories is small relative to the size of the economy, but they have a much greater effect on economic growth than justified by their relatively small aggregate size relative to the economy as a whole.
- Inventory–sales ratio measures the inventories available for sale to the level of sales. Analysts pay attention to inventories to gauge the position of the economy in the cycle.
- Neoclassical and Real Business Cycle (RBC) theories focus on fluctuations of aggregate supply (AS). If AS shifts left because of an input price increase or right because of a price decrease or technical progress, the economy will gradually converge to its new equilibrium. Government intervention is generally not necessary because it may exacerbate the fluctuation or delay the convergence to equilibrium.
- Keynesian theories focus on fluctuations of aggregate demand (AD). If AD shifts left, Keynesians advocate government intervention to restore full employment and avoid a deflationary spiral.

- Monetarists argue that the timing of the impact from government policies is uncertain, and it is generally better to let the economy find its new equilibrium unassisted while ensuring that the money supply is kept growing at an even pace.
- Economic indicators are statistics on macroeconomic variables that help in understanding which stage of the business cycle an economy is at. Of particular importance are the leading indicators, which suggest where the economy is likely to be in the near future. No economic indicator is perfect, and many of these statistics are subject to periodic revisions.
- Leading economic indicators have turning points that usually precede those of the overall economy. They include survey-based indicators, such as (ISM), new orders, and average consumer expectations. They also include average weekly hours, initial claims for unemployment insurance, new building permits, the stock market index, and the difference between yields on short-term and long-term bonds.
- Coincident economic indicators have turning points that are usually close to those of the overall economy. They are believed to have value for identifying the economy's present state. They include industrial production indexes, manufacturing and trade sales indexes, aggregate real personal income, and non-agricultural employment.
- Lagging economic indicators have turning points that take place later than those of the overall economy; they change after a trend has been established. They include average duration of unemployment, inventory–sales ratio, change in unit labor costs, average bank prime lending rate, commercial and industrial loans outstanding, ratio of consumer installment debt to income, and change in consumer price index for services.
- Policy makers and market practitioners use real-time monitoring of economic and financial variables to continuously assess current conditions and produce a nowcast. Nowcasting produces an estimate of the present state of the economy. It is useful because the actual data on such measures as GDP are only published with delay, after the end of the time period under consideration.
- Unemployed are those people who are actively seeking employment but are currently without a job.
- Long-term unemployed are those who have been out of work for a long time (more than three to four months in many countries) but are still looking for a job.
- Frictionally unemployed are those people who are not working at the time of filling out the statistical survey because they are taking time to search for a job that matches their skills, interests, and other preferences better than what is currently available, or people who have left one job and are about to start another job.
- There are different types of inflation. Hyperinflation indicates a high (e.g., 100% annual) and increasing rate of inflation; deflation indicates a negative inflation rate (prices decrease); imported inflation is associated with increasing cost of inputs that come from abroad; demand inflation is caused by constraints in production that prevent companies from making as many goods as the market demands (it is sometimes called wartime inflation, because in times of war, goods tend to be rationed).
- Inflation is measured by many indexes. Consumer price indexes reflect the prices of a basket of goods and services that is typically purchased by a normal household. Producer price indexes measure the cost of a basket of raw materials, intermediate inputs, and finished products. GDP deflators measure the

price of the basket of goods and services produced within an economy in a given year. Core indexes exclude volatile items, such as agricultural products and energy, whose prices tend to vary more than other goods.

- Price levels are affected by real factors and monetary factors. Real factors include aggregate supply (an increase in supply leads to lower prices) and aggregate demand (an increase in demand leads to higher prices). Monetary factors include the supply of money (i.e., more money circulating, if the economy is in equilibrium, will lead to higher prices) and the velocity of money (i.e., higher velocity, if the economy is in equilibrium, will lead to higher prices).
- Economists describe two types of inflation: cost-push, in which rising costs, usually wages, compel businesses to raise prices generally; and demand-pull, in which increasing demand raises prices generally, which then are reflected in a business's costs as workers demand wake hikes to catch up with the rising cost of living.

REFERENCES

Burns, Wesley Clair, and Arthur F. Mitchell. 1946. Measuring Business Cycles. National Bureau of Economic Research.

Claessens, Stijn, M. Ayhan Kose, and Marco E. Terrones. 2012. "How Do Business and Financial Cycles Interact?" *Journal of International Economics* 87 (1): 178–90.

McCulley, Paul. 2009. "The Shadow Banking System and Hyman Minsky's Economic Journey." In Insights into the Global Financial Crisis. Edited by Laurence B. Siegel. Charlottesville, VA: Research Foundation of CFA Institute.

Schüler, Y. S., P. Hiebert, and T. A. Peltonen. 2017. "*Coherent Financial Cycles for G-7 Countries: Why Extending Credit Can Be an Asset.*" Working Paper No. 43, European Systemic Risk Board (May).

United Nations. 2015. *Handbook on Economic Tendency Surveys.* United Nations Statistical Papers Series M, no. 96. https://unstats.un.org/unsd/nationalaccount/docs/ETS_Handbook_wCover.pdf.

PRACTICE PROBLEMS

1 The characteristic business cycle patterns of trough, expansion, peak, and contraction are:

A periodic.

B recurrent.

C of similar duration.

2 During the contraction phase of a business cycle, it is *most likely* that:

A inflation indicators are stable.

B aggregate economic activity relative to potential output is decreasing.

C investor preference for government securities declines.

3 An economic peak is *most* closely associated with:

A accelerating inflation.

B stable unemployment.

C declining capital spending.

4 With which sector of the economy would analysts most commonly associate credit cycles?

A Exports

B Construction and purchases of property

C Food retail

5 The reason analysts follow developments in the availability of credit is that:

A loose private sector credit may contribute to the extent of asset price and real estate bubbles and subsequent crises.

B loose credit helps reduce the extent of asset price and real estate bubbles.

C credit cycles are of same length and depth as business cycles.

6 Based on typical labor utilization patterns across the business cycle, productivity (output per hours worked) is *most likely* to be highest:

A at the peak of a boom.

B into a maturing expansion

C at the bottom of a recession.

7 As the expansion phase of the business cycle advances from early stage to late stage, businesses *most likely* experience a decrease in:

A labor costs.

B capital investment.

C availability of qualified workers.

8 An analyst writes in an economic report that the current phase of the business cycle is characterized by accelerating inflationary pressures and borrowing by companies. The analyst is *most likely* referring to the:

A peak of the business cycle.

B contraction phase of the business cycle.

C early expansion phase of the business cycle.

9 In a recession, companies are *most likely* to adjust their stock of physical capital by:

© 2020 CFA Institute. All rights reserved.

A selling it at fire sale prices.

B not maintaining equipment.

C quickly canceling orders for new construction equipment.

10 The inventory–sales ratio is *most likely* to be rising:

A as a contraction unfolds.

B partially into a recovery.

C near the top of an economic cycle.

11 A decrease in a country's total imports is *most likely* caused by:

A an increase in the pace of domestic GDP growth.

B a cyclical downturn in the economies of primary trading partners.

C persistent currency depreciation relative to primary trading partners.

12 Monetarists favor a limited role for the government because they argue:

A government policy responses may lag.

B firms take time to adjust to systemic shocks to the economy.

C resource use is efficient with marginal revenue and cost equal.

13 A national government responds to a severe recession by funding numerous infrastructure projects using deficit spending. Which school of economic thought is *most* consistent with such action?

A Keynesian

B Monetarist

C Neoclassical

14 According to Real Business Cycle models, an economic contraction is *most likely* caused by:

A sticky wages.

B rising energy prices.

C a contraction in the money supply.

15 The Austrian economic school attributes the primary cause of the business cycle to:

A misguided government intervention.

B the creative destruction of technological progress.

C sticky price and wage expectations that exaggerate trends.

16 According to the Austrian school, the *most* appropriate government response to an economic recession is to:

A allow the market to adjust naturally.

B maintain steady growth in the money supply.

C decrease the market rate of interest below its natural value.

17 Which of the following statements is the *best* description of the characteristics of economic indicators?

A Leading indicators are important because they track the entire economy.

B Lagging indicators, in measuring past conditions, do not require revisions.

C A combination of leading and coincident indicators can offer effective forecasts.

18 When the spread between 10-year US Treasury yields and the federal funds rate narrows and at the same time the prime rate stays unchanged, this mix of indicators *most likely* forecasts future economic:

A growth.

B decline.

C stability.

19 Current economic statistics indicating little change in services inflation, rising residential building permits, and increasing average duration of unemployment are *best* interpreted as:

A conflicting evidence about the direction of the economy.

B evidence that a cyclical upturn is expected to occur in the future.

C evidence that a cyclical downturn is expected to occur in the future.

20 If relative to prior values of their respective indicators, the inventory–sales ratio has risen, unit labor cost is stable, and real personal income has decreased, it is *most likely* that a peak in the business cycle:

A has occurred.

B is just about to occur.

C will occur sometime in the future.

21 When aggregate real personal income, industrial output, and the S&P 500 Index all increase in a given period, it is *most accurate* to conclude that a cyclical upturn is:

A occurring.

B about to end.

C about to begin.

22 Which of the following is *most likely* to increase after an increase in aggregate real personal income?

A Equity prices

B Building permits for new private housing units

C The ratio of consumer installment debt to income

23 Which of the following indicators is *most* appropriate in predicting a turning point in the economy?

A The Industrial Production Index

B The average bank prime lending rate

C Average weekly hours, manufacturing

24 The unemployment rate is considered a lagging indicator because:

A new job types must be defined to count their workers.

B multi-worker households change jobs at a slower pace.

C businesses are slow to hire and fire due to related costs.

25 A decrease in both the labor force participation ratio and the unemployment rate is *most likely* caused by:

A an increase in discouraged workers.

B an increase in underemployed workers.

C a decrease in voluntarily unemployed persons.

26 During an economic recovery, a lagging unemployment rate is *most likely* attributable to:

A businesses quickly rehiring workers.

B new job seekers entering the labor force.

C underemployed workers transitioning to higher-paying jobs.

27 The treasury manager of a large company has recently left his position to accept a promotion with a competitor six months from now. A statistical employment survey conducted now should categorize the status of the former treasury manager as:

A underemployed.

B voluntarily unemployed.

C frictionally unemployed.

28 The discouraged worker category is defined to include people who:

A are overqualified for their job.

B could look for a job but choose not to.

C currently look for work without finding it.

29 The category of persons who would be *most likely* to be harmed by an increase in the rate of inflation is:

A homeowners with fixed 30-year mortgages.

B retirees relying on a fixed annuity payment.

C workers employed under contracts with escalator clauses.

30 The term that describes when inflation declines but nonetheless remains at a positive level is:

A deflation.

B stagflation.

C disinflation.

31 Deflation is *most likely* to be associated with:

A a shortage of government revenue.

B substantial macroeconomic contraction.

C explicit monetary policy to combat inflation.

32 The *least likely* consequence of a period of hyperinflation is the:

A reduced velocity of money.

B increased supply of money.

C possibility of social unrest.

The following information relates to Questions 33–34

Exhibit 1 Consumption Baskets and Prices over Two Months

	November 2010		December 2010	
Goods	**Quantity**	**Price**	**Quantity**	**Price**
Sugar	70 kg	€0.90/kg	120 kg	€1.00 /kg
Cotton	60 kg	€0.60/kg	50 kg	€0.80/kg

33 Assuming the base period for 2010 consumption is November and the initial price index is set at 100, then the inflation rate after calculating the December price index as a Laspeyres index is *closest to*:

A 19.2%.

B 36.4%.

C 61.6%.

34 For the December consumption basket in Exhibit 1, the value of the Paasche index is *closest to*:

A 116.

B 148.

C 160.

---

35 A central bank will *most likely* allow the economy to self-correct in periods of:

A high inflation, fast economic growth, and low unemployment.

B low inflation, slow economic growth, and high unemployment.

C high inflation, slow economic growth, and high unemployment.

36 Disinflation is *best* described as a:

A decline in price levels.

B negative inflation rate.

C decline in the inflation rate.

37 The characteristic of national consumer price indexes that is *most likely* shared across major economies worldwide is:

A the geographic areas covered in their surveys.

B the weights they place on covered goods and services.

C their use in the determination of macroeconomic policy.

38 Of the following statements regarding the Producer Price Index (PPI), which is the *least likely*? The PPI:

A can influence the future CPI.

B category weights can vary more widely than analogous CPI terms.

C is used more frequently than CPI as a benchmark for adjusting labor contract payments.

39 The following presents selected commodity price data for July–August 20X0:

	July 20X0		August 20X0	
Goods	**Quantity**	**Price**	**Quantity**	**Price**
Milk	18 liters	€1.00/liter	17 liters	€1.00/liter
Orange juice	6 liters	€2.00/liter	4 liters	€2.50/liter

Given the consumption basket and prices presented, which type of price index will result in the highest calculated inflation rate over a two-month time period?

A One that uses a current consumption basket

B One that uses a constant consumption basket

C One reflecting substitutions made by consumers over time

40 The inflation rate *most likely* relied on to determine public economic policy is:

A core inflation.

B headline inflation.

C index of food and energy prices.

41 What is the *most* important effect of labor productivity in a cost-push inflation scenario?

A Rising productivity indicates a strong economy and a bias towards inflation.

B The productivity level determines the economy's status relative to its "natural rate of unemployment."

C As productivity growth proportionately exceeds wage increases, product price increases are less likely.

42 A product is part of a price index based on a fixed consumption basket. If, over time, the product's quality improves while its price stays constant, the measured inflation rate is *most likely*:

A unaffected.

B biased upward.

C biased downward.

43 A price index of goods and services that excludes food and energy is *most likely* used to calculate:

A core inflation.

B the GDP deflator.

C headline inflation.

44 An economist expects the following:

- The decline in the unemployment rate will result in higher revenues for home retailers.
- A tighter labor market will put upward pressure on wages, compelling home retailers to raise prices.

Which type of inflation *best* corresponds to the economist's expectations?

A Stagflation

B Cost-push inflation

C Demand-pull inflation

SOLUTIONS

1 B is correct. The stages of the business cycle occur repeatedly over time.

2 B is correct. The net trend during contraction is negative.

3 A is correct. Inflation is rising at peaks.

4 B is correct. Credit cycles are associated with availability of credit, which is important in the financing of construction and the purchase of property.

5 A is correct. Studies have shown that loose credit conditions contribute to the extent of asset price and real estate bubbles that tend to be followed by crises.

6 C is correct. At the end of a recession, firms will run "lean production" to generate maximum output with the fewest number of workers.

7 C is correct. When an economy's expansion is well established, businesses often have difficulty finding qualified workers.

8 A is correct. Accelerating inflation and rapidly expanding capital expenditures typically characterize the peak of the business cycle. During such times, many businesses finance their capital expenditures with debt to expand their production capacity.

9 B is correct. Physical capital adjustments to downturns come through aging of equipment plus lack of maintenance.

10 C is correct. Near the top of a cycle, sales begin to slow before production is cut, leading to an increase in inventories relative to sales.

11 C is correct. When a nation's currency depreciates, domestic goods seem cheaper than foreign goods, placing downward pressure on demand for imports. When the depreciation persists for some time, the country's total imports are likely to decrease.

12 A is correct. Monetarists caution that policy effects can occur long after the need for which they were implemented is no longer an issue.

13 A is correct. Keynesian economics is based on government intervention in the form of fiscal policy. The national government responds to the recession by using deficit spending to fund infrastructure projects.

14 B is correct. Real Business Cycle models conclude that expansions and contractions of the economy are responses to external shocks, such as supply shocks arising from advances in technology or changes in the relative prices of inputs (e.g., energy prices). An increase in energy prices shifts short-run aggregate supply to the left, resulting in higher prices and lower GDP.

15 A is correct. Austrian economists see monetary policy mistakes as leading to booms and busts.

16 A is correct. Austrian economists advocate limited government intervention in the economy. They advise that the best thing to do in a recession is to allow the necessary market adjustment to take place.

17 C is correct. While no single indicator is definitive, a mix of them—which can be affected by various economic determinants—can offer the strongest signal of performance.

18 B is correct. The narrowing spread of this leading indicator foretells a drop in short-term rates and a fall in economic activity. The prime rate is a lagging indicator and typically moves after the economy turns.

19 B is correct. Rising building permits—a leading indicator—indicate that an upturn is expected to occur or continue. Increasing average duration of unemployment—a lagging indicator—indicates that a downturn has occurred, whereas the lack of any change in services inflation—also a lagging indicator—is neither negative nor positive for the direction of the economy. Taken together, these statistics indicate that a cyclical upturn may be expected to occur.

20 A is correct. Both inventory–sales and unit labor costs are lagging indicators that decline somewhat after a peak. Real personal income is a coincident indicator that by its decline shows a slowdown in business activity.

21 A is correct. Aggregate real personal income and industrial output are coincident indicators, whereas the S&P 500 is a leading indicator. An increase in aggregate personal income and industrial output signals that an expansion is occurring, whereas an increase in the S&P 500 signals that an expansion will occur or is expected to continue. Taken together, these statistics indicate that a cyclical upturn is occurring.

22 C is correct. Aggregate real personal income is a coincident indicator of the business cycle, and the ratio of consumer installment debt to income is a lagging indicator. Increases in the ratio of consumer installment debt follow increases in average aggregate income during the typical business cycle.

23 C is correct. Leading economic indicators have turning points that usually precede those of the overall economy. Average weekly hours, manufacturing is a leading economic indicator. The Industrial Production Index is a coincident economic indicator, and the average bank prime lending rate is a lagging economic indicator.

24 C is correct. This effect makes unemployment rise more slowly as recessions start and fall more slowly as recoveries begin.

25 A is correct. Discouraged workers have given up seeking employment and are statistically outside the labor force. Therefore, an increase in discouraged workers will decrease the labor force and thus the labor participation ratio, which is the ratio of labor force to total working age population. Additionally, an increase in discouraged workers will decrease the unemployment rate because discouraged workers are not counted in the official unemployment rate.

26 B is correct. In an economic recovery, new job seekers return to the labor force, and because they seldom find work immediately, their return may initially raise the unemployment rate.

27 C is correct. Frictionally unemployed people are not working at the time of the employment survey but have recently left one job and are about to start another job. The frictionally unemployed have a job waiting for them and are not 100% unemployed, it is just that they have not started the new job yet. Although the treasury manager has left his current employment, he has accepted a new position at another firm that starts in six months.

28 B is correct. Discouraged workers are defined as persons who have stopped looking for work and are outside the labor force.

29 B is correct. With inflation, a fixed amount of money buys fewer goods and services, thus reducing purchasing power.

30 C is correct. Disinflation is known as a reduction of inflation from a higher to lower, but still above zero, level.

31 B is correct. Deflation is connected to a vicious cycle of reduced spending and higher unemployment.

32 A is correct. In hyperinflation, consumers accelerate their spending to beat price increases, and money circulates more rapidly.

33 A is correct. The Laspeyres index is calculated with these inputs:

- November consumption bundle: 70 × 0.9 + 60 × 0.6 = 99
- December consumption bundle: 70 × 1 + 60 × 0.8 = 118
- December price index: (118/99) × 100 = 119.19
- Inflation rate: (119.19/100) – 1 = 0.1919 = 19.19%

34 A is correct. The Paasche index uses the current product mix of consumption combined with the variation of prices. For December, its value is

$$(120 \times 1 + 50 \times 0.8)/(120 \times 0.9 + 50 \times 0.6) = (160/138) \times 100 = 115.9.$$

35 C is correct. This scenario is often referred to as stagflation. Here, the economy is likely to be left to self-correct because no short-term economic policy is thought to be effective.

36 C is correct. Disinflation is a decline in the inflation rate—for example, from 7% to 4%.

37 C is correct. Central banks typically use consumer price indexes to monitor inflation and evaluate their monetary policies.

38 C is correct. The CPI is typically used for this purpose, while the PPI is more closely connected to business contracts.

39 B is correct. The inflation rate calculated by using a constant consumption basket (the Laspeyres index) is 10%, derived as follows:

$$\begin{aligned} \text{July 20X0 consumption basket} &= (18 \times €1) + (6 \times €2) = €30 \\ \text{August 20X0 consumption basket} &= (18 \times €1) + (6 \times €2.5) = €33 \\ \text{Value of the Laspeyres index (IL)} &= (€33/€30) \times 100 = €110 \\ \text{Inflation rate} &= (110/100) - 1 = 0.10 = 10\% \end{aligned}$$

The inflation rate calculated using a current consumption basket (the Paasche index) is 8%, derived as follows:

$$\begin{aligned} \text{July 20X0 consumption basket} &= (17 \times €1) + (4 \times €2) = €25 \\ \text{August 20X0 consumption basket} &= (17 \times €1) + (4 \times €2.5) = €27 \\ \text{Value of the Paasche index (IP)} &= (€27/€25) \times 100 = €108 \\ \text{Inflation rate} &= (108/100) - 1 = 0.08 = 8\% \end{aligned}$$

The inflation rate calculated by "chaining" the monthly prices of consumption baskets as they change over time (the Fisher index) is derived as follows:

$$\text{Value of the Fisher index} = \sqrt{I_P \times I_L}$$

$$\text{Value of the Fisher Index} = \sqrt{€110 \times €108} = €108.99$$

$$\text{Inflation rate} = (108.99/100) - 1 = 0.0899 = 8.99\%$$

40 A is correct. Core inflation is less volatile since it excludes food and energy prices and therefore will not be as likely to lead to policy overreactions when serving as a target.

41 C is correct. For productivity, or output per hour, the faster that it can grow, the further that wages can rise without putting pressure on business costs per unit of output.

42 B is correct. As the quality of a product improves, it satisfies people's needs and wants better. The measured inflation rate is skewed higher than otherwise unless an adjustment is made for the increase in the quality of the good. Even if the good's price had increased over time, the improvements in quality would still bias the measured inflation rate upward.

43 A is correct. A price index of goods and services that excludes food and energy is used to calculate core inflation. Policy makers often use core inflation when reading the trend in the economy and making economic policies. The reason is because policy makers are trying to avoid overreaction to short-term fluctuations in prices as a result of short-term changes in supply and demand.

44 B is correct. Cost-push inflation refers to the situation in which rising costs, usually wages, compel businesses to raise prices.

ECONOMICS
STUDY SESSION
4

Economics (2)

This study session begins with monetary and fiscal policy, including their use by central banks and governments. Economics in a global context is then introduced. Next follows a discussion on the flows of goods and services and physical and financial capital that occur across national borders. Highlighted in the discussion are the relationships between different types of flows and the benefits of trade to trade partners. Finally, given that operations and investments in global markets involve foreign exchange (currency) risk, the session concludes with an overview of currency market fundamentals.

READING ASSIGNMENTS

Reading 12	Monetary and Fiscal Policy by Andrew Clare, PhD, and Stephen Thomas, PhD
Reading 13	International Trade and Capital Flows by Usha Nair-Reichert, PhD, and Daniel Robert Witschi, PhD, CFA
Reading 14	Currency Exchange Rates by William A. Barker, PhD, CFA, Paul D. McNelis, and Jerry Nickelsburg, PhD

© 2021 CFA Institute. All rights reserved.

READING

12

Monetary and Fiscal Policy

by Andrew Clare, PhD, and Stephen Thomas, PhD

Andrew Clare, PhD, and Stephen Thomas, PhD, are at Cass Business School (United Kingdom).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** compare monetary and fiscal policy;
☐	**b.** describe functions and definitions of money;
☐	**c.** explain the money creation process;
☐	**d.** describe theories of the demand for and supply of money;
☐	**e.** describe the Fisher effect;
☐	**f.** describe roles and objectives of central banks;
☐	**g.** contrast the costs of expected and unexpected inflation;
☐	**h.** describe tools used to implement monetary policy;
☐	**i.** describe the monetary transmission mechanism;
☐	**j.** describe qualities of effective central banks;
☐	**k.** explain the relationships between monetary policy and economic growth, inflation, interest, and exchange rates;
☐	**l.** contrast the use of inflation, interest rate, and exchange rate targeting by central banks;
☐	**m.** determine whether a monetary policy is expansionary or contractionary;
☐	**n.** describe limitations of monetary policy;
☐	**o.** describe roles and objectives of fiscal policy;
☐	**p.** describe tools of fiscal policy, including their advantages and disadvantages;
☐	**q.** describe the arguments about whether the size of a national debt relative to GDP matters;

(continued)

© 2019 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	The candidate should be able to:
☐	**r.** explain the implementation of fiscal policy and difficulties of implementation;
☐	**s.** determine whether a fiscal policy is expansionary or contractionary;
☐	**t.** explain the interaction of monetary and fiscal policy.

1 INTRODUCTION TO MONETARY AND FISCAL POLICY

a compare monetary and fiscal policy;

The economic decisions of households can have a significant impact on an economy. For example, a decision on the part of households to consume more and to save less can lead to an increase in employment, investment, and ultimately profits. Equally, the investment decisions made by corporations can have an important impact on the real economy and on corporate profits. But individual corporations can rarely affect large economies on their own; the decisions of a single household concerning consumption will have a negligible impact on the wider economy.

By contrast, the decisions made by governments can have an enormous impact on even the largest and most developed of economies for two main reasons. First, the public sectors of most developed economies normally employ a significant proportion of the population, and they are usually responsible for a significant proportion of spending in an economy. Second, governments are also the largest borrowers in world debt markets. Exhibit 1 gives some idea of the scale of government borrowing and spending.

Exhibit 1

Panel A. Central Government Debt to GDP, 2017

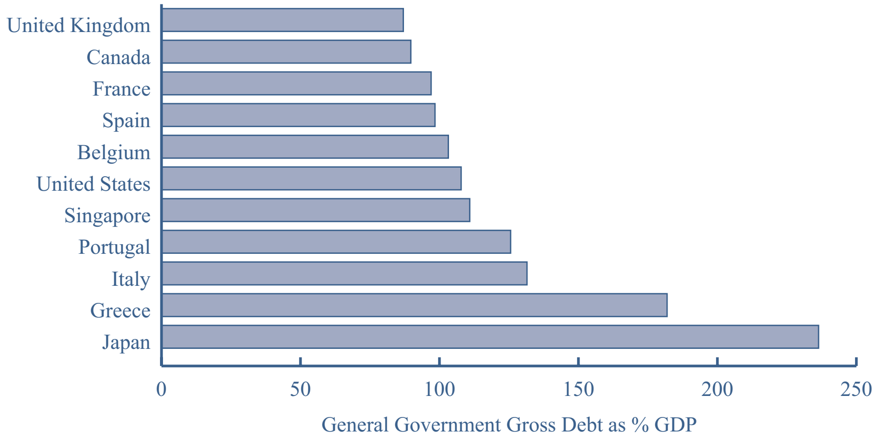
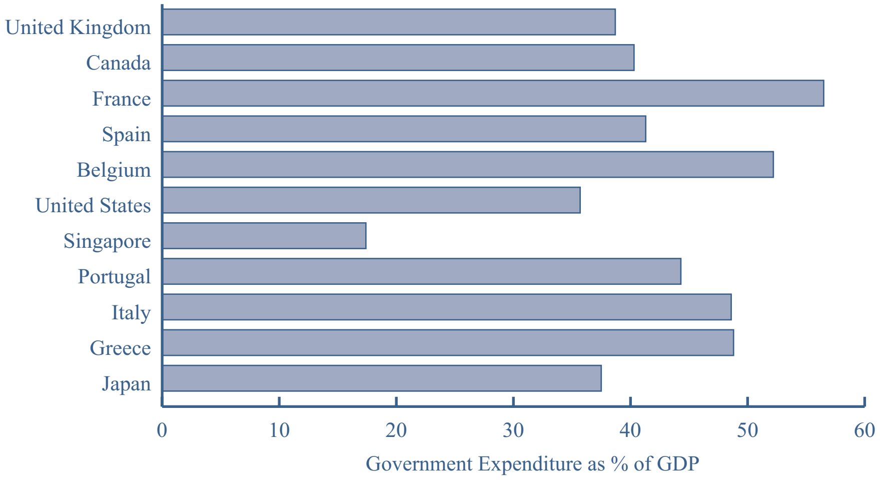


Exhibit 1 (Continued)

Panel B. Public Sector Spending to GDP, 2017



Source: IMF, World Economic Outlook Database, April 2018.

Government policy is ultimately expressed through its borrowing and spending activities. In this reading, we identify and discuss two types of government policy that can affect the macroeconomy and financial markets: monetary policy and fiscal policy.

Monetary policy refers to central bank activities that are directed toward influencing the quantity of money and credit in an economy.[1] By contrast, **fiscal policy** refers to the government's decisions about taxation and spending. Both monetary and fiscal policies are used to regulate economic activity over time. They can be used to accelerate growth when an economy starts to slow or to moderate growth and activity when an economy starts to overheat. In addition, fiscal policy can be used to redistribute income and wealth.

The overarching goal of both monetary and fiscal policy is normally the creation of an economic environment where growth is stable and positive and inflation is stable and low. Crucially, the aim is therefore to steer the underlying economy so that it does not experience economic booms that may be followed by extended periods of low or negative growth and high levels of unemployment. In such a stable economic environment, householders can feel secure in their consumption and saving decisions, while corporations can concentrate on their investment decisions, on making their regular coupon payments to their bond holders and on making profits for their shareholders.

The challenges to achieving this overarching goal are many. Not only are economies frequently buffeted by shocks (such as oil price jumps), but some economists believe that natural cycles in the economy also exist. Moreover, there are plenty of examples from history where government policies—either monetary, fiscal, or both—have exacerbated an economic expansion that eventually led to damaging consequences for the real economy, for financial markets, and for investors.

The balance of the reading is organized as follows. Sections 2–9 provide an introduction to monetary policy and related topics. Sections 10–13 present fiscal policy. The interactions between monetary policy and fiscal policy are the subject of Section 14. A summary and practice problems conclude the reading.

---

[1] Central banks can implement monetary policy almost completely independent of government interference and influence at one end of the scale, or simply as the agent of the government at the other end of the scale.

1.1 Monetary Policy

As stated above, monetary policy refers to government or central bank activities that are directed toward influencing the quantity of money and credit in an economy. Before we can begin to understand how monetary policy is implemented, we must examine the functions and role of **money**. We can then explore the special role that **central banks** play in today's economies.

EXAMPLE 1

Monetary and Fiscal Policy

1 Which of the following statements *best* describes monetary policy? Monetary policy:
   - **A** involves the setting of medium-term targets for broad money aggregates.
   - **B** involves the manipulation by a central bank of the government's budget deficit.
   - **C** seeks to influence the macro economy by influencing the quantity of money and credit in the economy.

2 Which of the following statements *best* describes fiscal policy? Fiscal policy:
   - **A** is used by governments to redistribute wealth and incomes.
   - **B** is the attempt by governments to balance their budgets from one year to the next.
   - **C** involves the use of government spending and taxation to influence economy activity.

Solution to 1:

C is correct. Choice A is incorrect because, although the setting of targets for monetary aggregates is a possible *tool* of monetary policy, monetary policy itself is concerned with influencing the overall, or macro, economy.

Solution to 2:

C is correct. Note that governments may wish to use fiscal policy to redistribute incomes and balance their budgets, but the overriding goal of fiscal policy is usually to influence a broader range of economic activity.

2

MONEY: FUNCTIONS, CREATION, AND DEFINITION

b describe functions and definitions of money;

c explain the money creation process;

To understand the nature, role, and development of money in modern economies, it is useful to think about a world without money—where to purchase any good or service, an individual would have to "pay" with another good or service. An economy where such economic agents as households, corporations, and even governments pay for goods and services in this way is known as a **barter economy**. There are many drawbacks to such an economy. First, the exchange of goods for other goods (or services) would require both economic agents in the transaction to want what the other is selling. This means that there has to be a **double coincidence of wants**. It might also be impossible to undertake transactions where the goods are indivisible—that is, where one agent wishes to buy a certain amount of another's goods, but that agent only has one indivisible unit of another good that is worth more than the good that the agent is trying to buy. Another problem occurs if economic agents do not wish to exchange all of their goods on other goods and services. This may not be a problem, however, when the goods they have to sell can be stored safely so that they retain their value for the future. But if these goods are perishable, they will not be able to store value for their owner. Finally, in a barter economy, there are many measures of value: the price of oranges in terms of pears; of pears in terms of bread; of bread in terms of milk; or of milk in terms of oranges. A barter economy has no common measure of value that would make multiple transactions simple.

2.1 The Functions of Money

The most generic definition of money is that it is any generally accepted medium of exchange. A **medium of exchange** is any asset that can be used to purchase goods and services or to repay debts. Money can thus eliminate the debilitating double coincidence of the "wants" problem that exists in a barter economy. When this medium of exchange exists, a farmer wishing to sell wheat for wine does not need to identify a wine producer in search of wheat. Instead, he can sell wheat to those who want wheat in exchange for money. The farmer can then exchange this money for wine with a wine producer, who in turn can exchange that money for the goods or services that she wants.

However, for money to act as this liberating medium of exchange, it must possess certain qualities. It must:

i. be readily acceptable,
ii. have a known value,
iii. be easily divisible,
iv. have a high value relative to its weight, and
v. be difficult to counterfeit.

Qualities (i) and (ii) are closely related; the medium of exchange will only be acceptable if it has a known value. If the medium of exchange has quality (iii), then it can be used to purchase items of relatively little value and of relatively large value with equal ease. Having a high value relative to its weight is a practical convenience, meaning that people can carry around sufficient wealth for their transaction needs. Finally, if the medium of exchange can be counterfeited easily, then it would soon cease to have a value and would not be readily acceptable as a means of effecting transactions; in other words, it would not satisfy qualities (i) and (ii).

Given the qualities that money needs to have, it is clear why precious metals (particularly gold and silver) often fulfilled the role of medium of exchange in early societies, and as recently as the early part of the twentieth century. Precious metals were acceptable as a medium of exchange because they had a known value, were easily divisible, had a high value relative to their weight, and could not be easily counterfeited.

Thus, precious metals were capable of acting as a medium of exchange. But they also fulfilled two other useful functions that are essential for the characteristics of money. In a barter economy, it is difficult to store wealth from one year to the next when one's produce is perishable, or indeed, if it requires large warehouses in which to store it. Because precious metals like gold had a high value relative to their bulk and were not perishable, they could act as a **store of wealth**. However, their ability to act as a store of wealth not only depended on the fact that they did not perish physically over time, but also on the belief that others would always value precious metals. The value from year to year of precious metals depended on people's continued demand for them in ornaments, jewellery, and so on. For example, people were willing to use gold as a store of wealth because they believed that it would remain highly valued. However, if gold became less valuable to people relative to other goods and services year after year it would not be able to fulfill its role as a **store of value**, and as such might also lose its status as a medium of exchange.

Another important characteristic of money is that it can be used as a universal unit of account. As such, it can create a single unitary **measure of value** for all goods and services. In an economy where gold and silver are the accepted medium of exchange, all prices, debts, and wealth can be recorded in terms of their gold or silver coin exchange value. Money, in its role as a unit of account, drastically reduces the number of prices in an economy compared to barter, which requires that prices be established for a good in terms of all other goods for which it might be exchanged.

In summary, money fulfills three important functions, it:

- acts as a medium of exchange;
- provides individuals with a way of storing wealth; and
- provides society with a convenient measure of value and unit of account.

2.2 Paper Money and the Money Creation Process

Although precious metals like gold and silver fulfilled the required functions of money relatively well for many years, and although carrying gold coins around was easier than carrying around one's physical produce, it was not necessarily a safe way to conduct business.

A crucial development in the history of money was the **promissory note**. The process began when individuals began leaving their excess gold with goldsmiths, who would look after it for them. In turn the goldsmiths would give the depositors a receipt, stating how much gold they had deposited. Eventually these receipts were traded directly for goods and services, rather than there being a physical transfer of gold from the goods buyer to the goods seller. Of course, both the buyer and seller had to trust the goldsmith because the goldsmith had all the gold and the goldsmith's customers had only pieces of paper. These depository receipts represented a promise to pay a certain amount of gold on demand. This paper money therefore became a proxy for the precious metals on which they were based, that is, they were directly related to a physical commodity. Many of these early goldsmiths evolved into banks, taking in excess wealth and in turn issuing promissory notes that could be used in commerce.

In taking in other people's gold and issuing depository receipts and later promissory notes, it became clear to the goldsmiths and early banks that not all the gold that they held in their vaults would be withdrawn at any one time. Individuals were willing to buy and sell goods and services with the promissory notes, but the majority of the gold that backed the notes just sat in the vaults—although its ownership would change with the flow of commerce over time. A certain proportion of the gold that was not being withdrawn and used directly for commerce could therefore be lent to others at a rate of interest. By doing this, the early banks created money.

The process of **money creation** is a crucial concept for understanding the role that money plays in an economy. Its potency depends on the amount of money that banks keep in reserve to meet the withdrawals of its customers. This practice of lending customers' money to others on the assumption that not all customers will want all of their money back at any one time is known as **fractional reserve banking**.

We can illustrate how it works through a simple example. Suppose that the bankers in an economy come to the view that they need to retain only 10 percent of any money deposited with them. This is known as the **reserve requirement**.[2] Now consider what happens when a customer deposits €100 in the First Bank of Nations. This deposit changes the balance sheet of First Bank of Nations, as shown in Exhibit 2, and it represents a liability to the bank because it is effectively loaned to the bank by the customer. By lending 90 percent of this deposit to another customer the bank has two types of assets: (1) the bank's reserves of €10, and (2) the loan equivalent to €90. Notice that the balance sheet still balances; €100 worth of assets and €100 worth of liabilities are on the balance sheet.

Now suppose that the recipient of the loan of €90 uses this money to purchase some goods of this value and the seller of the goods deposits this €90 in another bank, the Second Bank of Nations. The Second Bank of Nations goes through the same process; it retains €9 in reserve and loans 90 percent of the deposit (€81) to another customer. This customer in turn spends €81 on some goods or services. The recipient of this money deposits it at the Third Bank of Nations, and so on. This example shows how money is created when a bank makes a loan.

Exhibit 2 Money Creation via Fractional Reserve Banking

First Bank of Nations

Assets		**Liabilities**	
Reserves	€10	Deposits	€100
Loans	€90		

Second Bank of Nations

Assets		**Liabilities**	
Reserves	€9	Deposits	€90
Loans	€81		

Third Bank of Nations

Assets		**Liabilities**	
Reserves	€8.1	Deposits	€81
Loans	€72.9		

This process continues until there is no more money left to be deposited and loaned out. The total amount of money 'created' from this one deposit of €100 can be calculated as:

$$\text{New deposit/Reserve requirement} = €100/0.10 = €1{,}000 \quad (1)$$

It is the sum of all the deposits now in the banking system. You should also note that the original deposit of €100, via the practice of reserve banking, was the catalyst for €1,000 worth of economic transactions. That is not to say that economic growth would be zero without this process, but instead that it can be an important component in economic activity.

---

[2] This is an example of a *voluntary* reserve requirement because it is self-imposed.

The amount of money that the banking system creates through the practice of fractional reserve banking is a function of 1 divided by the reserve requirement, a quantity known as the **money multiplier**.[3] In the case just examined, the money multiplier is 1/0.10 = 10. Equation 1 implies that the smaller the reserve requirement, the greater the money multiplier effect.

In our simplistic example, we assumed that the banks themselves set their own reserve requirements. However, in some economies, the central bank sets the reserve requirement, which is a potential means of affecting money growth. In any case, a prudent bank would be wise to have sufficient reserves such that the withdrawal demands of their depositors can be met in stressful economic and credit market conditions.

Later, when we discuss central banks and central bank policy, we will see how central banks can use the mechanism just described to affect the money supply. Specifically, the central bank could, by purchasing €100 in government securities credited to the bank account of the seller, seek to initiate an increase in the money supply. The central bank may also lend reserves directly to banks, creating excess reserves (relative to any imposed or self-imposed reserve requirement) that can support new loans and money expansion.

EXAMPLE 2

Money and Money Creation

1 To fulfill its role as a medium of exchange, money should:

- **A** be a conservative investment.
- **B** have a low value relative to its weight.
- **C** be easily divisible and a good store of value.

2 If the reserve requirement for banks in an economy is 5 percent, how much money could be created with the deposit of an additional £100 into a deposit account?

- **A** £500
- **B** £1,900
- **C** £2,000

3 Which of the following functions does money normally fulfill for a society? It:

- **A** acts as a medium of exchange only.
- **B** provides economic agents with a means of storing wealth only.
- **C** provides society with a unit of account, acts as a medium of exchange, and acts as a store of wealth.

Solution to 1:

C is correct. Money needs to have a known value and be easily divisible. It should also be readily acceptable, difficult to counterfeit, and have a high value relative to its weight.

---

[3] This quantity, known as the simple money multiplier, represents a maximum expansion. To the extent that banks hold excess reserves or that money loaned out is not re-deposited, the money expansion would be less. More complex multipliers incorporating such factors are developed in more advanced texts.

Solution to 2:

C is correct. To calculate the increase in money from an additional deposit in the banking system, use the following expression: new deposit/reserve requirement.

Solution to 3:

C is correct. Money needs to be able to fulfill the functions of acting as a unit of account, a medium of exchange, and a means of storing wealth.

2.3 Definitions of Money

The process of money creation raises a fundamental issue: What is money? In an economy with money but without promissory notes and fractional reserve banking, money is relatively easy to define: Money is the total amount of gold and silver coins in circulation, or their equivalent. The money creation process above, however, indicates that a broader definition of money might encompass all the notes and coins in circulation *plus* all bank deposits.

More generally, we might define money as any medium that can be used to purchase goods and services. Notes and coins can be used to fulfill this purpose, and yet such currency is not the only means of purchasing goods and services. Personal cheques can be written based on a bank chequing account, while debit cards can be used for the same purpose. But what about time deposits or savings accounts? Nowadays transfers can be made relatively easily from a savings account to a current account; therefore, these savings accounts might also be considered as part of the stock of money. Credit cards are also used to pay for goods and services; however, there is an important difference between credit card payments and those made by cheques and debit cards. Unlike a cheque or debit card payment, a credit card payment involves a deferred payment. Basically, the greater the complexity of any financial system, the harder it is to define money.

The monetary authorities in most modern economies produce a range of measures of money (see Exhibit 3). But generally speaking, the money stock consists of notes and coins in circulation, plus the deposits in banks and other financial institutions that can be readily used to make purchases of goods and services in the economy. In this regard, economists often speak of the rate of growth of **narrow money** and/or **broad money**. By narrow money, they generally mean the notes and coins in circulation in an economy, plus other very highly liquid deposits. Broad money encompasses narrow money but also includes the entire range of liquid assets that can be used to make purchases.

Because financial systems, practice, and institutions vary from economy to economy, so do definitions of money; thus, it is difficult to make international comparisons. Still, most central banks produce both a narrow and broad measure of money, plus some intermediate ones too. Exhibit 3 shows the money definitions in four economies.

Exhibit 3 Definitions of Money

Money Measures in the United States

The US Federal Reserve produces two measures of money. The first is M1, which comprises notes and coins in circulation, travelers' cheques of non-bank issuers, demand deposits at commercial banks, plus other deposits on which cheques can be written. M2 is the broadest measure of money currently produced by the Federal Reserve and includes M1, plus savings and money market deposits, time deposit accounts of less than $100,000, plus other balances in retail money market and mutual funds.

(continued)

Exhibit 3 (Continued)

Money Measures in the Eurozone

The European Central Bank (ECB) produces three measures of euro area money supply. The narrowest is M1. M1 comprises notes and coins in circulation, plus all overnight deposits. M2 is a broader definition of euro area money that includes M1, plus deposits redeemable with notice up to three months and deposits with maturity up to two years. Finally, the euro area's broadest definition of money is **M3**, which includes M2, plus repurchase agreements, money market fund units, and debt securities with up to two years maturity.

Money Measures in Japan

The Bank of Japan calculates three measures of money. M1 is the narrowest measure and consists of cash currency in circulation. M2 incorporates M1 but also includes certificates of deposit (CDs). The broadest measure, M3, incorporates M2, plus deposits held at post offices, plus other savings and deposits with financial institutions. There is also a "broad measure of liquidity" that encompasses M3 as well as a range of other liquid assets, such as government bonds and commercial paper.

Money Measures in the United Kingdom

The United Kingdom produces a set of four measures of the money stock. **M0** is the narrowest measure and comprises notes and coins held outside the Bank of England, plus Bankers' deposits at the Bank of England. **M2** includes M0, plus (effectively) all retail bank deposits. **M4** includes M2, plus wholesale bank and building society deposits and also certificates of deposit. Finally, the Bank of England produces another measure called **M3H**, which is a measure created to be comparable with money definitions in the EU (see above). M3H includes M4, plus UK residents' and corporations' foreign currency deposits in banks and building societies.

3 MONEY: QUANTITY THEORY, SUPPLY AND DEMAND, FISHER EFFECT

d describe theories of the demand for and supply of money;

e describe the Fisher effect;

The previous section of this reading shows that there are many definitions of money. In this section, we explore the important relationship between money and the price level. This relationship is best expressed in the **quantity theory of money**, which asserts that total spending (in money terms) is proportional to the quantity of money. The theory can be explained in terms of Equation 2, known as the **quantity equation of exchange**:

$$M \times V = P \times Y \qquad (2)$$

where M is the quantity of money, V is the velocity of circulation of money (the average number of times in a given period that a unit of currency changes hands), P is the average price level, and Y is real output. The expression is really just an accounting identity. Effectively, it says that over a given period, the amount of money used to purchase all goods and services in an economy, $M \times V$, is equal to monetary value of this output, $P \times Y$. If the velocity of money is approximately constant—which is an assumption of quantity theory—then spending $P \times Y$ is approximately proportional to M. The quantity equation can also be used to explain a consequence of **money neutrality**. If money neutrality holds, then an increase in the money supply, M, will not affect Y, real output, or the speed with which money changed hands, V, because if real output is unaffected, there would be no need for money to change hands more rapidly.[4] However, it will cause the aggregate price level, P, to rise.

The simple quantity theory gave rise to the equally simple idea that the price level, or at least the rate of inflation, could be controlled by manipulating the rate of growth of the money supply. Economists who believe this are referred to as **monetarists**. They argue that there is a causal relationship running from money growth to inflation. In the past, some governments have tried to apply this logic in their efforts to control inflation, most notably and unsuccessfully the United Kingdom's government in 1979 (see Example 5). However, it is possible that causality runs the other way—that is, from real activity to the money supply. This means that the quantity of money in circulation is determined by the level of economic activity, rather than vice versa.

3.1 The Demand for Money

The amount of wealth that the citizens of an economy choose to hold in the form of money—as opposed to bonds or equities—is known as the demand for money. There are three basic motives for holding money:

- transactions-related;
- precautionary; and
- speculative.

Money balances that are held to finance transactions are referred to as **transactions money balances**. The size of the transactions balances will tend to increase with the average value of transactions in an economy. Generally speaking, as gross domestic product (GDP) grows over time, transactions balances will also tend to grow; however, the ratio of transactions balances to GDP remains fairly stable over time.

As the name suggests, **precautionary money balances** are held to provide a buffer against unforeseen events that might require money. These balances will tend to be larger for individuals or organizations that enter into a high level of transactions over time. In other words, a precautionary buffer of $100 for a company that regularly enters into transactions worth millions of dollars might be considered rather small. When we extend this logic to the overall economy, we can see that these precautionary balances will also tend to rise with the volume and value of transactions in the economy, and therefore, GDP as well.

Finally, the **speculative demand for money** (sometimes called the **portfolio demand for money**) relates to the demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments (e.g., bonds). **Speculative money balances** consist of monies held in anticipation that other assets will decline in value. But in choosing to hold speculative money balances rather than bonds, investors give up the return that could be earned from the bond or other financial assets. Therefore, the speculative demand for money will tend to fall as the returns available on other financial assets rises. However, it will tend to rise as the perceived risk in other financial instruments rises. In equilibrium, individuals will tend to increase their holdings of money relative to riskier assets until the marginal benefit of having a lower risk portfolio of wealth is equal to the marginal cost of giving up a unit of expected return on these riskier assets. In aggregate then, speculative balances will tend to be inversely related to the expected return on other financial assets and directly related to the perceived risk of other financial assets.

---

[4] Note that the full version of the quantity theory of money uses the symbol T rather than Y to indicate transactions because money is used not just for buying goods and services but also for financial transactions. We will return to this point in the discussion of quantitative easing.

EXAMPLE 3

Money

1 The transactions demand for money refers to the demand to hold money:

- **A** as a buffer against unforeseen events.
- **B** to use in the purchase of goods and services.
- **C** based on the opportunity or risks available on other financial instruments.

2 The speculative demand for money will tend to:

- **A** fall as the perceived risk on other assets rises.
- **B** rise as the expected returns on other assets fall.
- **C** be inversely related to the transactions demand for money.

3 What is the difference between narrow and broad money? Broad money:

- **A** is limited to those liquid assets most commonly used to make purchases.
- **B** can be used to purchase a wider range of goods and services than narrow money.
- **C** encompasses narrow money and refers to the stock of the entire range of liquid assets that can be used to make purchases.

Solution to 1:

B is correct. The transactions demand for money refers to the amount of money that economic agents wish to hold to pay for goods and services.

Solution to 2:

B is correct. If the expected return on other assets falls, then the opportunity cost of holding money also falls and can, in turn, lead to an increase in the speculative demand for money.

Solution to 3:

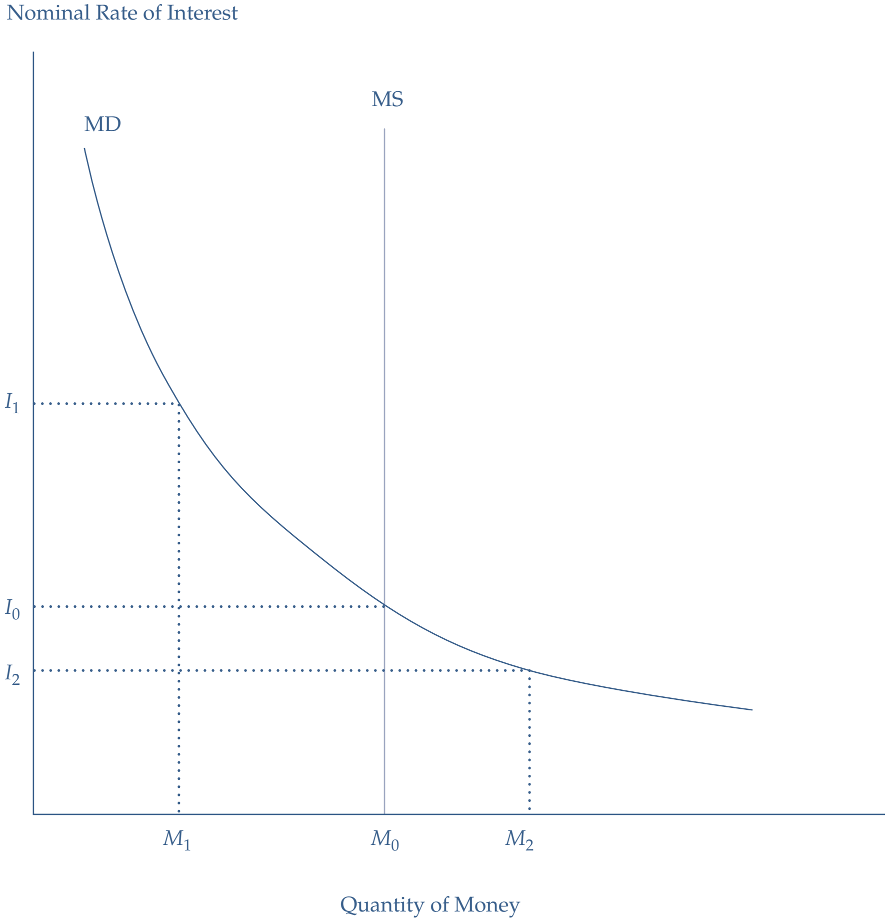
C is correct. This is the definition of broad money. Broad money encompasses narrow money.

3.2 The Supply and Demand for Money

We have now discussed definitions of money, its relationship with the aggregate price level, and the demand for it. We now discuss the interaction between the supply of and demand for money.

As with most other markets, the supply of money and the demand to hold it will interact to produce an equilibrium price for money. In this market, the price of money is the nominal interest rate that could be earned by lending it to others. Exhibit 4 shows the supply and demand curves for money. The vertical scale represents the rate of interest; the horizontal scale plots the quantity of nominal money in the economy. The supply curve (MS) is vertical because we assume that there is a fixed nominal amount of money circulating at any one time. The demand curve (MD) is downward sloping because as interest rates rise, the speculative demand for money falls. The supply and demand for money are both satisfied at an equilibrium interest rate of I_0. I_0 is the rate of interest at which no excess money balances exist.

Exhibit 4 The Supply and Demand for Money



To see why I_0 is the equilibrium rate of interest where there are no excess money balances, consider the following. If the interest rate on bonds were I_1 instead of I_0, there would be excess supply of money ($M_0 - M_1$). Economic agents would seek to buy bonds with their excess money balances, which would force the price of bonds up and the interest rate back down to I_0. Similarly, if bonds offered a rate of interest, I_2, there would be an excess demand for money ($M_2 - M_0$). Corporations and individuals would seek to sell bonds so that individuals could increase their money holdings, but in doing so, the price of bonds would fall and the interest rate offered on them would rise until it reached I_0. Interest rates effectively adjust to bring the market into equilibrium ("clear the market"). In this simple example, we have also assumed that the supply of money and bonds is fixed as economic agents readjust their holdings. In practice, this may not be true, but the dynamics of the adjustment process described here essentially still hold.

Exhibit 4 also reemphasises the relationship between the supply of money and the aggregate price level, which we first encountered when discussing the quantity theory of money. Suppose that the central bank increases the supply of money from

M_0 to M_2, so that the vertical supply curve shifts to the right. Because the increase in the supply of money makes it more plentiful and hence less valuable, its price (the interest rate) falls as the price level rises.

This all sounds very simple, but in practice the effects of an increase in the money supply are more complex. The initial increase in the money supply will create excess supply of cash. People and companies could get rid of the excess by loaning the money to others by buying bonds, as implied above, but they might also deposit it in a bank or simply use it to buy goods and services. But an economy's capacity to produce goods and services depends on the availability of real things: notably, natural resources, capital, and labour—that is, factors of production supplied either directly or indirectly by households. Increasing the money supply does not change the availability of these real things. Thus, some economists believe that the long-run impact of an exogenous increase in the supply of money is an increase in the aggregate price level.

This phenomenon—whereby an increase in the money supply is thought in the long run simply to lead to an increase in the price level while leaving real variables like output and employment unaffected—is known as **money neutrality**. To see why in the long run money should have a neutral effect on real things, consider the following simple example.

Suppose the government declared today that 1kg would henceforth be referred to as 2kg and that 1.5kg would be referred to as 3kg. In other words, suppose that they halved the "value" of a kilogram. Would anything real have changed? A 1kg bag of sugar would not have changed physically, although it would be relabelled as a 2kg bag of sugar. However, there might be some short-run effects; confused people might buy too little sugar, and some people might go on crash diets! But ultimately people would adjust. In the long run, the change wouldn't matter. There is a clear parallel here with the theory of money neutrality. Doubling the prices of everything—halving the value of a currency—does not change anything real. This is because, like kilograms, money is a unit of account. However, halving the value of a currency could affect real things in the short run.

There are two points worth making with regard to money neutrality. First, although the simple kilogram analogy above does suggest that money should not affect real things in the long run, as the British economist Keynes said: "*In the long run we are all dead*!" In practice, it is very difficult for economists to be sure that money neutrality holds in the long run. And second, we must assume that monetary authorities do believe that the money supply can affect real things in the short run. If they did not, then there would be almost no point to monetary policy.

3.3 The Fisher Effect

The **Fisher effect** is directly related to the concept of money neutrality. Named after the economist Irving Fisher, the Fisher effect states that the real rate of interest in an economy is stable over time so that changes in nominal interest rates are the result of changes in expected inflation. Thus, the nominal interest rate (R_{nom}) in an economy is the sum of the required real rate of interest (R_{real}) and the expected rate of inflation (π^e) over any given time horizon:

$$R_{\text{nom}} = R_{\text{real}} + \pi^e \tag{3}$$

According to money neutrality, over the long term the money supply and/or the growth rate in money should not affect R_{real} but will affect inflation and inflation expectations.

The Fisher effect also demonstrates that embedded in every nominal interest rate is an expectation of future inflation. Suppose that 12-month US government T-bills offered a yield equal to 4 percent over the year. Suppose also that T-bill investors wished to earn a real rate of interest of 2 percent and expected inflation to be 2 percent over the next year. In this case, the return of 4 percent would be sufficient to deliver the investors' desired real return of 2 percent (so long as inflation did not exceed 2 percent). Now suppose that investors changed their view about future inflation and instead expected it to equal 3 percent over the next 12 months. To compensate them for the higher expected inflation, the T-bill rate would have to rise to 5 percent, thereby preserving the required 2 percent real return.

There is one caveat to this example. Investors can never be sure about future values of such economic variables as inflation and real growth. To compensate them for this uncertainty, they require a **risk premium**. The greater the uncertainty, the greater the required risk premium. So, all nominal interest rates are comprised of three components:

- a required real return;
- a component compensating investors for expected inflation; and
- a risk premium to compensate them for uncertainty.

EXAMPLE 4

Interest Rates and the Supply of Money

1 According to the quantity equation of exchange, an increase in the money supply can lead to an:

   A increase in the aggregate price level, regardless of changes in the velocity of circulation of money.

   B increase in the aggregate price level as long as the velocity of circulation of money rises sufficiently to offset the increase in the money supply.

   C increase in the aggregate price level as long as the velocity of circulation of money does not fall sufficiently to offset the increase in the money supply and real output is unchanged.

2 The nominal interest rate comprises a real rate of interest:

   A plus a risk premium only.

   B plus a premium for expected inflation only.

   C compensation for both expected inflation and risk.

3 An expansion in the money supply would *most likely*:

   A lead to a decline in nominal interest rates.

   B lead to an increase in nominal interest rates.

   C reduce the equilibrium amount of money that economic agents would wish to hold.

Solution to 1:

C is correct. If the velocity of circulation of money does not change with an increase in the money supply and real output is fixed, then the aggregate price level should increase. If the velocity of circulation of money falls sufficiently, or if real output rises sufficiently, then the increase in money may have no impact on prices.

Solution to 2:

C is correct. Investors demand a real rate of interest and compensation for expected inflation and a risk premium to compensate them for uncertainty.

Solution to 3:

A is correct. Increasing the supply of money, all other things being equal, will reduce its "price," that is, the interest rate on money balances.

EXAMPLE 5

Mrs. Thatcher's Monetary Experiment

The Background

Over the 1970s, the United Kingdom had one of the worst inflation records of any developed economy. Retail price inflation averaged 12.6 percent over that decade and peaked at 26.9 percent in August 1975. Over this period, then-Prime Minister Margaret Thatcher and her advisers had become convinced that inflation could not be controlled by the income and price policies used in the United Kingdom in the past. Instead, they believed that inflation could be tamed by controlling the rate of growth of the money supply. Mrs. Thatcher's first administration took power in May 1979 with the intention of pursuing a monetarist agenda—that is, a macroeconomic policy that would be underpinned by targets for money supply growth.

The Medium-Term Financial Strategy

Targets for monetary growth were set for a definition of the money supply known as Sterling M3 (£M3), which was to be kept in the range of 7–11 percent for the period 1980–1981 and then gradually reduced to within 4–8 percent by 1983–1984. This set of targets was known as the Medium Term Financial Strategy (MTFS). The idea was simple: Control the rate of growth of the money supply, and the rate of growth of prices (i.e., inflation) would remain under control too. The instrument of control was the Bank of England's policy interest rate that would be set to achieve the desired rate of growth of the money supply. This was a macroeconomic policy built, however imperfectly, on an interpretation of the quantity theory of money.

The theory was simple, but the practice proved to be less so. Over the first two and a half years of the MTFS, £M3 overshot its target by 100 percent. The inability of the monetary authorities to control the rate of growth of the broad money supply was largely caused by Thatcher's abolition of exchange controls in 1979. By abolishing these controls, there was a significant increase in foreign exchange business that came into the British banking system, which changed the velocity of money and therefore meant that the relationship between broad money and nominal incomes had changed fundamentally.[5]

Despite the inability to control the money supply, in 1983 the Thatcher administration reasserted its confidence in the policy and published a further set of monetary targets for several years ahead. However, the persistent failure to meet these targets, too, eventually led to the abandonment of any type of monetary targeting by the summer of 1985.

The experience of the UK monetary authorities over this period emphasizes how unstable the relationship between money and the policy interest rate could be along with the relationship between money and aggregate demand—particularly in an economy that is experiencing rapid financial innovation, as the UK economy was following the abolition of exchange controls and the introduction of greater competition within the banking industry.

---

[5] See Goodhart (1989) for a discussion.

Today the Bank of England is responsible for the operation and implementation of monetary policy in the United Kingdom. The trends in money supply are watched very carefully, but they are not the subject of targets, per se.

ROLES OF CENTRAL BANKS & OBJECTIVES OF MONETARY POLICY

4

f describe roles and objectives of central banks;

Central banks play several key roles in modern economies. Generally, a central bank is the monopoly supplier of the currency, the banker to the government and the bankers' bank, the lender of last resort, the regulator and supervisor of the payments system, the conductor of monetary policy, and the supervisor of the banking system. Let us examine these roles in turn.

In its earliest form, money could be exchanged for a pre-specified precious commodity, usually gold, and promissory notes were issued by many private banks. Today, however, state-owned institutions—usually central banks—are designated in law as being the monopoly suppliers of a currency. Initially, these monopolists supplied money that could be converted into a pre-specified amount of gold; they adhered to a **gold standard**. For example, up until 1931, bank notes issued by Britain's central bank, the Bank of England, could be redeemed at the bank for a pre-specified amount of gold. But Britain, like most other major economies, abandoned this convertibility principle in the first half of the twentieth century. Money in all major economies today is not convertible by law into anything else, but it is, in law, **legal tender**. This means that it must be accepted when offered in exchange for goods and services. Money that is not convertible into any other commodity is known as **fiat money**. Fiat money derives its value via government decree and because people accept it for payment of goods and services and for debt repayment.

As long as fiat money is acceptable to everyone as a medium of exchange, and it holds its value over time, then it will also be able to serve as a unit of account. However, once an economy has moved to a system of fiat money, the role of the supplier of that money becomes even more crucial because they could, for example, expand the supply of this money indefinitely should they wish to do so. Central banks therefore play a crucial role in modern economies as the suppliers and guardians of the value of their fiat currencies and as institutions charged with the role of maintaining confidence in their currencies. As the monopoly suppliers of an economy's currency, central banks are at the centre of economic life. As such, they assume other roles in addition to being the suppliers and guardians of the value of their currencies.

Most central banks act as the banker to the government and to other banks. They also act as a **lender of last resort** to banks. Because the central bank effectively has the capacity to print money, it is in the position to be able to supply the funds to banks that are facing a damaging shortage. The facts that economic agents know that the central bank stands ready to provide the liquidity required by any of the banks under its jurisdiction and that they trust government bank deposit insurance help to prevent bank runs in the first place. However, the recent financial crisis has shown that this knowledge is not always sufficient to deter a bank run.

EXAMPLE 6

The Northern Rock Bank Run

In the latter part of the summer of 2007, the fall in US house prices and the related implosion of the US sub-prime mortgage market became the catalyst for a global liquidity crisis. Banks began to hoard cash and refused to lend to other banks at anything other than extremely punitive interest rates through the interbank market. This caused severe difficulties for a UK mortgage bank, Northern Rock. Northern Rock's mortgage book had expanded rapidly in the preceding years as it borrowed aggressively from the money markets. It is now clear that this expansion was at the expense of loan quality. The then UK regulatory authority, the Financial Services Authority (FSA),[6] later reported in 2008 that Northern Rock's lending practices did not pay due regard to either the credit quality of the mortgagees or the values of the properties on which the mortgages were secured. Being at the worst end of banking practice, and relying heavily on international capital markets for its funding, Northern Rock was therefore very susceptible to a global reduction in liquidity. As the liquidity crisis took hold, Northern Rock found that it could not replace its maturing money market borrowings. On 12 September 2007, in desperate need of liquidity, Northern Rock's board approached the UK central bank to ask for the necessary funds.

However, the news of Northern Rock's perilous liquidity position became known by the public and, more pertinently, by Northern Rock's retail depositors. On 14 September, having heard the news, queues began to form outside Northern Rock branches as depositors tried to withdraw their savings. On that day, it was estimated that Northern Rock depositors withdrew around £1bn, representing 5 percent of Northern Rock's deposits. Further panic ensued as investors in "internet only" Northern Rock accounts could not withdraw their money because of the collapse of Northern Rock's website. A further £1bn was withdrawn over the next two days.

Northern Rock's share price dropped rapidly, as did the share prices of other similar UK banks. The crisis therefore threatened to engulf more than one bank. To prevent contagion, the Chancellor of the Exchequer announced on 17 September that the UK government would guarantee all Northern Rock deposits. This announcement was enough to stabilize the situation, and given that lending to Northern Rock was now just like lending to the government, deposits actually started to rise again.

Eventually Northern Rock was nationalized by the UK government, with the hope that at some time in the future it could be privatized once its balance sheet had been repaired.

Central banks are also often charged by the government to supervise the banking system, or at least to supervise those banks that they license to accept deposits. However, in some countries, this role is undertaken by a separate authority. In other countries, the central bank can be jointly responsible with another body for the supervision of its banks.

Exhibit 5 lists the banking supervisors in the G-10 countries; central banks are underlined. As the exhibit shows, most but not all bank systems have a single supervisor, which is not necessarily a central bank. A few countries, such as Germany and the United States, have more than one supervisor.

---

[6] In 2013, the Financial Services Authority was replaced by two new regulatory authorities, the Financial Conduct Authority (FCA) and the Prudential Regulation Authority (PRA).

Exhibit 5 Banking Supervision in the G10

Country	Institution(s)
Belgium	Banking and Finance Commission
Canada	Office of the Superintendent of Financial Institutions
France	Commission Bancaire
Germany	Federal Banking Supervisory Office; Deutsche Bundesbank
Italy	Bank of Italy
Japan	Financial Services Agency
Netherlands	Bank of Netherlands
Sweden	Swedish Financial Supervisory Authority
Switzerland	Federal Commission
United Kingdom	Bank of England
United States	Office of the Comptroller of the Currency; Federal Reserve; Federal Deposit Insurance Corporation

The United Kingdom is an interesting case study in this regard. Until May 1997, the Bank of England had statutory responsibility for banking supervision in the United Kingdom. In May 1997, banking supervision was removed from the Bank of England and assigned to a new agency, the Financial Services Authority (FSA). However, the removal of responsibility for banking supervision from the central bank was seen by some as being a contributory factor in the run on the mortgage bank Northern Rock, and generally as a contributory factor in the recent banking crisis. Because of this perceived weakness in the separation of the central bank from banking supervision, the Bank of England regained responsibility for banking supervision and regulation in 2013.

Perhaps the least appreciated role of a central bank is its role in the **payments system**. Central banks are usually asked to oversee, regulate, and set standards for a country's payments system. Every day millions of financial transactions take place in a modern economy. For the system to work properly, procedures must be robust and standardized. The central bank will usually oversee the payments system and will also be responsible for the successful introduction of any new processes. Given the international nature of finance, the central bank will also be responsible for coordinating payments systems internationally with other central banks.

Most central banks will also be responsible for managing their country's **foreign currency reserves** and also its gold reserves. With regard to the latter, even though countries abandoned the gold standard in the early part of the twentieth century, the world's central bankers still hold large quantities of gold. As such, if central banks were to decide to sell significant proportions of their gold reserves, it could potentially depress gold prices.

Finally, central banks are usually responsible for the operation of a country's **monetary policy**. This is arguably the highest profile role that these important organizations assume. Recall that monetary policy refers to central bank activities that are directed toward influencing the quantity of money and credit in an economy. As the monopoly supplier of a country's currency, central banks are in the ideal position to implement and/or determine monetary policy.

To summarise, central banks assume a range of roles and responsibilities. They do not all assume responsibility for the supervision of the banks, but all of the other roles listed below are normally assumed by the central bank:

- Monopoly supplier of the currency;

- Banker to the government and the bankers' bank;
- Lender of last resort;
- Regulator and supervisor of the payments system;
- Conductor of monetary policy; and
- Supervisor of the banking system.

4.1 The Objectives of Monetary Policy

Central banks fulfill a variety of important roles, but for what overarching purpose? A brief perusal of the websites of the world's central banks will reveal a wide range of explanations of their objectives. Their objectives are clearly related to their roles, and so there is frequent mention of objectives related to the stability of the financial system and to the payments systems. Some central banks are charged with doing all they can to maintain full employment and output. But some also have related but less tangible roles, like "maintaining confidence in the financial system," or even to "promote understanding of the financial sector." But there is one overarching objective that most seem to acknowledge explicitly, and that is the objective of maintaining **price stability**.

So, although central banks usually have to perform many roles, most specify an overarching objective. Exhibit 6 lists what we might call the primary objective(s) of a number of central banks, from both developed and developing economies.

Exhibit 6 The Objectives of Central Banks

The Central Bank of Brazil

Its "institutional mission" is to "ensure the stability of the currency's purchasing power and a solid and efficient financial system."

The European Central Bank

"[T]o maintain price stability is the primary objective of the Euro system and of the single monetary policy for which it is responsible. This is laid down in the Treaty on the Functioning of the European Union, Article 127 (1)."

"Without prejudice to the objective of price stability", the euro system will also "support the general economic policies in the Community with a view to contributing to the achievement of the objectives of the Community." These include a "high level of employment" and "sustainable and non-inflationary growth."

The US Federal Reserve

"The Federal Reserve sets the nation's monetary policy to promote the objectives of maximum employment, stable prices, and moderate long-term interest rates."

The Reserve Bank of Australia

"It is the duty of the Reserve Bank Board, within the limits of its powers, to ensure that the monetary and banking policy of the Bank is directed to the greatest advantage of the people of Australia and that the powers of the Bank ... are exercised in such a manner as, in the opinion of the Reserve Bank Board, will best contribute to:

a the stability of the currency of Australia;
b the maintenance of full employment in Australia; and
c the economic prosperity and welfare of the people of Australia."

Exhibit 6 (Continued)

The Bank of Korea

"The primary purpose of the Bank, as prescribed by the Bank of Korea Act of 1962, is the pursuit of price stability."

Source: Central bank websites found at http://www.bis.org/cbanks.htm.

EXAMPLE 7

Central Banks

1 A central bank is normally *not* the:

  A lender of last resort.

  B banker to the government and banks.

  C body that sets tax rates on interest on savings.

2 Which of the following *best* describes the overarching, long-run objective of most central banks?

  A Price stability

  B Fast economic growth

  C Current account surplus

Solution to 1:

C is correct. A central bank is normally the lender of last resort and the banker to the banks and government, but the determination of all tax rates is normally the preserve of the government and is a fiscal policy issue.

Solution to 2:

A is correct. Central banks normally have a variety of objectives, but the overriding one is nearly always price stability.

As we have already discussed, one of the essential features of a monetary system is that the medium of exchange should have a relatively stable value from one period to the next. Arguably then, the overarching goal of most central banks in maintaining price stability is the associated goal of controlling inflation. But before we explore the tools central banks use to control inflation, we should first consider the potential costs of inflation. In other words, we should ask why it is that central bankers believe that it is so important to control a nominal variable.

THE COSTS OF INFLATION

5

g contrast the costs of expected and unexpected inflation;

Huge efforts have been put into controlling inflation since the major economies experienced such high levels of inflation in the 1970s. From the early 1970s then, inflation has been seen as a very bad thing. But why? What are the costs of inflation? The debate around the "costs" of inflation really centers on the distinction between

expected inflation and **unexpected inflation**. Expected inflation is clearly the level of inflation that economic agents expect in the future. Unexpected inflation can be defined as the level of inflation that we experience that is either below or above that which we expected; it is the component of inflation that is a surprise.

At a micro level, high inflation means that businesses constantly have to change the advertised prices of their goods and services. These are known as **menu costs**. There also exists what economists refer to as "shoe leather" costs of inflation. In times of high inflation, people would naturally tend to hold less cash and would therefore wear out their shoe leather (or more likely the engines of their cars) in making frequent trips to the bank to withdraw cash. But these are relatively old arguments, used to demonstrate that inflation is bad. In a modern economy, with the internet and with transactions becoming increasingly cashless, these costs associated with inflation will be lower today than they may have been in the past.

To demonstrate the potentially more significant costs of inflation, consider the following. Imagine a world where inflation is high but where all prices (including asset prices) in an economy are perfectly indexed to inflation, and that technology has eliminated the issues surrounding the menu and shoe leather costs of inflation. In such a world, would economic agents care about inflation? Probably not. If the average price of goods and services rose by 10 percent, people's salaries (and all other prices) would rise by the same amount, which would therefore make economic agents indifferent to the rise in prices.

In practice, however, all prices, wages, salaries, rents, and so forth are not indexed, in which case economic agents would certainly need to think about inflation more carefully. But what if inflation in this world where prices are no longer perfectly indexed is high, but perfectly predictable? In this alternative, imaginary world, economic agents would have to think about inflation, but not too hard as long as they were capable of calculating the impact of the known inflation rate on all future prices. So, if everyone knew that inflation was going to be 10 percent over the next year, then everyone could bargain for a 10 percent increase in their salaries to accommodate this, and companies could plan to put up the prices of their goods and services by 10 percent. In this world, an expectation of 10 percent inflation would become a self-fulfilling prophecy.

However, economic agents would worry about inflation in a world where all prices were not indexed and, crucially, where inflation was high and unpredictable. In fact, this is a crude description the inflationary backdrop in many developed economies over the 1970s and 1980s, including the United States, France, the United Kingdom, Italy, and Canada.

Arguably it is **unexpected inflation** that is most costly. Inflation that is fully anticipated can be factored into wage negotiations and priced into business and financial contracts. But when inflation turns out to be higher than is anticipated, then borrowers benefit at the expense of lenders because the real value of their borrowing declines. Conversely, when inflation is lower than is anticipated, lenders benefit at the expense of borrowers because the real value of the payment on debts rises. Furthermore, if inflation is very uncertain or very volatile, then lenders will ask for a premium to compensate them for this uncertainty. As a result, the costs of borrowing will be higher than would otherwise have been the case. Higher borrowing costs could in turn reduce economic activity, for example, by discouraging investment.

It is also possible that **inflation uncertainty** can exacerbate the economic cycle. Inflation uncertainty is the degree to which economic agents view future rates of inflation as hard to forecast. Take for example the case of an imaginary television manufacturer. Suppose one day that the manufacturer looks out at the market for televisions and sees that the market price of televisions has risen by 10 percent. Armed with this information, the manufacturer assumes that there has been an increase in demand for televisions or maybe a reduction in supply. So, to take advantage of the new, higher prices, the manufacturer extends the factory, employs more workers, and begins to produce more televisions.

Having now increased the output of the factory, the manufacturer then attempts to sell the extra televisions that the factory has produced. But to its horror, the manufacturer finds out that there is no extra demand for televisions. Instead, the 10 percent rise in television prices was caused by a generalized 10 percent increase in all consumer prices across the economy. The manufacturer realizes that it has surplus stock, surplus factory capacity, and too many workers. So, it cuts back on production, lays off some of the workforce, and realizes that it won't need to invest in new plant or machinery for a long time.

This example emphasizes the potentially destabilizing impact of unexpected inflation. It demonstrates how unanticipated inflation can reduce the information content of market prices for economic agents. If we scale this example up, it should not be too difficult to imagine how unanticipated increases or decreases in the general price level could help to exacerbate—and in some extreme cases cause—economic booms and busts.

Over the last two to three decades the consensus among economists has been that unanticipated and high levels of inflation can have an impact on real things like employment, investment and profits, and therefore that controlling inflation should be one of the main goals of macroeconomic policy. In summary:

Expected inflation can give rise to:

- menu costs and
- shoe leather costs.

Unanticipated (unexpected) inflation can in addition:

- lead to inequitable transfers of wealth between borrowers and lenders (including losses to savings);
- give rise to risk premia in borrowing rates and the prices of other assets; and
- reduce the information content of market prices.

MONETARY POLICY TOOLS[7]

6

h describe tools used to implement monetary policy;

i describe the monetary transmission mechanism;

k explain the relationships between monetary policy and economic growth, inflation, interest, and exchange rates;

Central banks have three primary tools available to them: open market operations, the refinancing rate, and reserve requirements.

---

[7] Monetary policy tools and operations often vary considerably from economy to economy. We have tried to describe the generics of the process here. For a more-detailed review of monetary operations across the world, see Gray and Talbot (2006).

6.1 Open Market Operations

One of the most direct ways for a central bank to increase or reduce the amount of money in circulation is via **open market operations**. Open market operations involve the purchase and sale of government bonds from and to commercial banks and/or designated market makers. For example, when the central bank buys government bonds from commercial banks, this increases the reserves of private sector banks on the asset side of their balance sheets. If banks then use these surplus reserves by increasing lending to corporations and households, then via the money multiplier process explained in Section 2, broad money growth expands. Similarly, the central bank can sell government bonds to commercial banks. By doing this, the reserves of commercial banks decline, reducing their capacity to make loans (i.e., create credit) to households and corporations and thus causing broad money growth to decline through the money multiplier mechanism. In using open market operations, the central bank may target a desired level of commercial bank reserves or a desired interest rate for these reserves.

6.2 The Central Bank's Policy Rate

The most obvious expression of a central bank's intentions and views comes via the interest rate it sets. The name of the **official interest rate** (or **official policy rate** or just **policy rate**) varies from central bank to central bank, but its purpose is to influence short- and long-term interest rates and ultimately real economic activity.

The interest rate that a central bank sets and that it announces publicly is normally the rate at which it is willing to lend money to the commercial banks (although practices do vary from country to country). This policy rate can be achieved by using short-term collateralized lending rates, known as repo rates. For example, if the central bank wishes to increase the supply of money, it might buy bonds (usually government bonds) from the banks, with an agreement to sell them back at some time in the future. This transaction is known as a **repurchase agreement**. Normally, the maturity of repo agreements ranges from overnight to two weeks. In effect, this represents a secured loan to the banks, and the lender (in this case the central bank) earns the repo rate.

Suppose that a central bank announces an increase in its official interest rate. Commercial banks would normally increase their **base rates** at the same time. A commercial bank's base rate is the reference rate on which it bases lending rates to all other customers. For example, large corporate clients might pay the base rate plus 1 percent on their borrowing from a bank, while the same bank might lend money to a small corporate client at the base rate plus 3 percent. But why would commercial banks immediately increase their base or reference rates just because the central bank's refinancing rate had increased?

The answer is that commercial banks would not want to have lent at a rate of interest that would be lower than they might be charged by the central bank. Effectively, the central bank can force commercial banks to borrow from it at this rate because it can conduct open market operations that create a shortage of money, forcing the banks to sell bonds to it with a pre-agreed repurchase price (i.e., do a repurchase agreement). The repo rate would be such that the central bank earned the official refinancing rate on the transactions.

The name of each central bank's official refinancing rate varies. The Bank of England's refinancing rate is the **two-week repo rate**. In other words, the Bank of England fixes the rate at which it is willing to lend two-week money to the banking sector. The ECB's official policy rate is known as the **refinancing rate** and defines the rate at which it is willing to lend short-term money to the euro area banking sector.

The corresponding rate in the United States is the discount rate, which is the rate for member banks borrowing directly from the Federal Reserve System. But the most important interest rate used in US monetary policy is the **federal funds rate**. The federal funds rate (or **fed funds rate**) is the interbank lending rate on overnight borrowings of reserves. The Federal Open Market Committee (FOMC) seeks to move this rate to a target level by reducing or adding reserves to the banking system by means of open market operations. The level of the rate is reviewed by the FOMC at its meetings held every six weeks (although the target can be changed between meetings, if necessary).

Through the setting of a policy rate, a central bank can manipulate the amount of money in the money markets. Generally speaking, the higher the policy rate, the higher the potential penalty that banks will have to pay to the central bank if they run short of liquidity, the greater will be their willingness to reduce lending, and the more likely that broad money growth will shrink.

6.3 Reserve Requirements

The third primary way in which central banks can limit or increase the supply of money in an economy is via their **reserve requirements**. We have already seen that the money creation process is more powerful the lower the percentage reserve requirement of banks. So, a central bank could restrict money creation by raising the reserve requirements of banks. However, this policy tool is not used much nowadays in developed economies. Indeed, some central banks, such as the Bank of England, do not even set minimum reserve requirements for the banks under their jurisdiction anymore. Changing reserve requirements frequently is disruptive for banks. For example, if a central bank increased the reserve requirements, a bank that was short on reserves might have to cease its lending activities until it had built up the necessary reserves, because deposits would be unlikely to rise quickly enough for the bank to build its reserves in this way. However, reserve requirements are still actively used in many developing countries to control lending and remain a potential policy tool for those central banks that do not currently use it.

To summarize, central banks can manipulate the money supply in one of three ways:

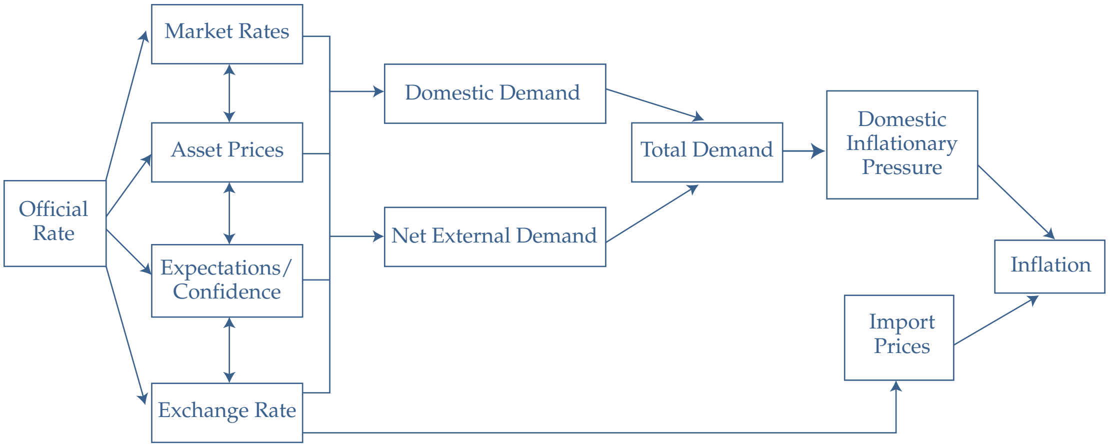
- open market operations;
- its official policy rate and associated actions in the repo market; and
- manipulation of official reserve requirements.

6.4 The Transmission Mechanism

The overarching goal of a central bank is to maintain price stability. We demonstrated above how a central bank can manipulate the money supply and growth of the money supply. We also indicated how policy rates set and targeted by the central banks are usually very short term in nature; often they target overnight interest rates. However, most businesses and individuals in the real economy borrow and lend over much longer time frames than this. It may not be obvious, then, how changing short-term interest rates can influence the real economy, particularly if money neutrality holds in the long run. The fact that central bankers believe that they can affect real economic variables, in particular economic growth, by influencing broad money growth suggests that they believe that money is not neutral—at least not in the short run.

Exhibit 7 presents a stylized representation of the **monetary transmission mechanism**. This is the process whereby a central bank's interest rate gets transmitted through the economy and ultimately affects the rate of increase of prices—that is, inflation.

Exhibit 7 A Stylized Representation of the Monetary Transmission Mechanism



Source: Bank of England.

Suppose that a central bank announces an increase in its official interest rate. The implementation of the policy may begin to work through the economy via four interrelated channels. Those channels include bank lending rates, asset prices, agents' expectations, and exchange rates. First, as described above, the base rates of commercial banks and interbank rates should rise in response to the increase in the official rate. Banks would, in turn, increase the cost of borrowing for individuals and companies over both short- and long-term horizons. Businesses and consumers would then tend to borrow less as interest rates rise. An increase in short-term interest rates could also cause the price of such assets as bonds or the value of capital projects to fall as the discount rate for future cash flows rises.

Market participants would then come to the view that higher interest rates will lead to slower economic growth, reduced profits, and reduced borrowing to finance asset purchases. Exporters' profits might decline if the rise in interest rates causes the country's exchange rate to appreciate, because this would make domestic exports more expensive to overseas buyers and dampen demand to purchase them. The fall in asset prices as well as an increase in prices would reduce household financial wealth and therefore lead to a reduction in consumption growth. Expectations regarding interest rates can play a significant role in the economy. Often companies and individuals will make investment and purchasing decisions based on their interest rate expectations, extrapolated from recent events. If the central bank's interest rate move is widely expected to be followed by other interest rate increases, investors and companies will act accordingly. Consumption, borrowing, and asset prices may all decline as a result of the revision in expectations.

There is a whole range of interconnected ways in which a rise in the central bank's policy rate can reduce real domestic demand and net external demand (that is, the difference between export and import consumption). Weaker total demand would tend to put downward pressure on the rate of domestic inflation—as would a stronger currency, which would reduce the prices of imports. Taken together, these might begin to put downward pressure on the overall measure of inflation.

To summarize, the central bank's policy rate works through the economy via any one, and often all, of the following interconnected channels:

- Short-term interest rates;
- Changes in the values of key asset prices;
- The exchange rate; and

- The expectations of economic agents.

EXAMPLE 8

Central Bank Tools

1. Which of the following variables are *most likely* to be affected by a change in a central bank's policy rate?
   - **A** Asset prices only
   - **B** Expectations about future interest rates only
   - **C** Both asset prices and expectations about future interest rates
2. Which of the following does a central bank seek to influence directly via the setting of its official interest rate?
   - **A** Inflation expectations
   - **B** Import prices
   - **C** Domestic inflation

Solution to 1:

C is correct. The price of equities, for example, might be affected by the expectation of future policy interest rate changes. In other words, a rate change may be taken as a signal of the future stance of monetary policy—contractionary or expansionary.

Solution to 2:

A is correct. By setting its official interest rate, a central bank could expect to have a direct influence on inflation expectations—as well as on other market interest rates, asset prices, and the exchange rate (where this is freely floating). If it can influence these factors, it might ultimately hope to influence import prices (via changes in the exchange rate) and also domestically generated inflation (via its impact on domestic and/or external demand). The problem is that the workings of the transmission mechanism—from the official interest rate to inflation—are complex and can change over time.

INFLATION TARGETING

7

j describe qualities of effective central banks;

k explain the relationships between monetary policy and economic growth, inflation, interest, and exchange rates;

l contrast the use of inflation, interest rate, and exchange rate targeting by central banks;

Over the 1990s, a consensus began to build among both central bankers and politicians that the best way to control inflation and thereby maintain price stability was to target a certain level of inflation and to ensure that this target was met by monitoring a wide range of monetary, financial, and real economic variables. Nowadays, inflation-targeting frameworks are the cornerstone of monetary policy and macroeconomic policy in many economies. Exhibit 8 shows the growth in the number of inflation-targeting monetary policy regimes over time.

The inflation-targeting framework that is now commonly practiced was pioneered in New Zealand. In 1988, the New Zealand Minister of Finance, Roger Douglas, announced that economic policy would focus on bringing inflation down from the prevailing level of around 6.0 percent to a target range of 0 to 2 percent. This goal was given legal status by the Reserve Bank of New Zealand Act 1989. As part of the Act, the Reserve Bank of New Zealand (RBNZ) was given the role of pursuing this target. The bank was given **operational independence**; it was free to set interest rates in the way that it thought would best meet the inflation target. Although the RBNZ had independent control of monetary policy, it was still accountable to the government and was charged with communicating its decisions in a clear and transparent way. As Exhibit 8 shows, the New Zealand model was widely copied.

Exhibit 8 The Progressive Adoption of Inflation Targeting by Central Banks

1989	New Zealand				
1990	Chile	Canada			
1991	Israel	United Kingdom			
1992	Sweden	Finland	Australia		
1995	Spain				
1998	Czech Republic	South Korea	Poland		
1999	Mexico	Brazil	Colombia	ECB	
2000	South Africa	Thailand			
2001	Iceland	Norway	Hungary	Peru	Philippines
2005	Guatemala	Indonesia	Romania		
2006	Turkey	Serbia			
2007	Ghana				

Note: Spain and Finland later joined the EMU.
Sources: For 2001 and earlier, Truman (2003). For 2002 to 2007, Roger (2010).

Although these inflation-targeting regimes vary a little from economy to economy, their success is thought to depend on three key concepts: central bank independence, credibility, and transparency.

7.1 Central Bank Independence[8]

In most cases, the central bank that is charged with targeting inflation has a degree of independence from its government. This independence is thought to be important. It is conceivable that politicians could announce an inflation target and direct the central bank to set interest rates accordingly. Indeed, this was the process adopted in the United Kingdom between 1994 and 1997. But politicians have a constant eye on re-election and might be tempted, for example, to keep rates "too low" in the lead up to an election in the hope that this might help their re-election prospects. As a consequence, this might lead to higher inflation. Thus, it is now widely believed that monetary policy decisions should rest in the hands of an organization that is remote from the electoral process. The central bank is the natural candidate to be the monopoly supplier of a currency.

---

[8] For information about the degree of independence of any central bank, the roles that it assumes in an economy, and the framework in which it operates, analysts should go to a central bank's website. A list of central bank websites can be found at http://www.bis.org/cbanks.htm.

However, there are degrees of independence. For example, the head of the central bank is nearly always chosen by government officials. The Chairman of the US Federal Reserve's Board of Governors is appointed by the President of the United States of America; the Head of the ECB is chosen by the committee of Euro area finance ministers; while the Governor of the Bank of England is chosen by the Chancellor of the Exchequer. So, in practice, separating control from political influence completely is probably an impossible (although a desirable) goal.

There are further degrees of independence. Some central banks are both operationally and **target independent**. This means that they not only decide the level of interest rates, but they also determine the definition of inflation that they target, the rate of inflation that they target, and the horizon over which the target is to be achieved. The ECB has independence of this kind. By contrast, other central banks—including those in New Zealand, Sweden, and the United Kingdom—are tasked to hit a definition and level of inflation determined by the government. These central banks are therefore only operationally independent.

7.2 Credibility

The independence of the central bank and public confidence in it are key in the design of an inflation-targeting regime.

To illustrate the role of credibility, suppose that instead of the central bank, the government assumes the role of targeting inflation but the government is heavily indebted. Given that higher inflation reduces the real value of debt, the government would have an incentive to avoid reaching the inflation target or to set a high inflation target such that price stability and confidence in the currency could be endangered. As a result, few would believe the government was really intent on controlling inflation; thus, the government would lack credibility. Many governments have very large levels of debt, especially since the 2008–2009 global financial crisis. In such a situation, economic agents might expect a high level of inflation, regardless of the actual, stated target. The target might have little credibility if the organization's likelihood of sticking to it is in doubt.

If a respected central bank assumes the inflation-targeting role and if economic agents believe that the central bank will hit its target, the belief itself could become self-fulfilling. If everyone believes that the central bank will hit an inflation target of 2 percent next year, this expectation might be built into wage claims and other nominal contracts that would make it hit the 2 percent target. It is for this reason that central bankers pay a great deal of attention to inflation expectations. If these expectations were to rise rapidly, perhaps following a rapid increase in oil prices, unchecked expectations could get embedded into wage claims and eventually cause inflation to rise.

7.3 Transparency

One way of establishing credibility is for a central bank to be transparent in its decision making. Many, if not all, independent inflation-targeting central banks produce a quarterly assessment of their economies. These **Inflation reports**, as they are usually known, give central banks' views on the range of indicators that they watch when they come to their (usually) monthly interest rate decision. They will consider and outline their views on the following subjects, usually in this order:

- Broad money aggregates and credit conditions;
- Conditions in financial markets;

- Developments in the real economy (e.g., the labour market); and
- Evolution of prices.

Consideration of all of these important components of an economy is then usually followed by a forecast of growth and inflation over a medium-term horizon, usually two years.

By explaining their views on the economy and by being transparent in decision making, the independent, inflation-targeting central banks seek to gain reputation and credibility, making it easier to influence inflation expectations and hence ultimately easier to meet the inflation target.

The Target

Whether the target is set by the central bank or by the government for the central bank to hit, the level of the target and the horizon over which the target is to be hit is a crucial consideration in all inflation-targeting frameworks.

Exhibit 9 A Range of Inflation Targets

Country/Region	
Australia	Australian Federal Reserve's target is inflation between 2.0% and 3.0%.
Canada	Bank of Canada's target is CPI inflation within the 1.0% and 3.0% range.
Euro-area	ECB's target is CPI inflation close to, but below, a ceiling of 2%.
South Korea	Bank of Korea's target for 2010–2012 is CPI inflation within ±1.0 percentage of 2.0%.
New Zealand	Reserve Bank of New Zealand's target is to keep future inflation between 1.0% and 3.0% with a focus on the average future inflation rate near 2.0%.
Sweden	Riksbank's target is CPI inflation within ±1.0 percentage point of 2.0%.
United Kingdom	Bank of England's target is CPI inflation within ±1.0 percentage point of 2.0%.

Source: Central bank websites (http://www.bis.org/cbanks.htm).

Exhibit 9 shows that many central banks in developed economies target an inflation rate of 2 percent based on a consumer price index. Given that the operation of monetary policy is both art and science, the banks are normally allowed a range around the central target of +1 percent or –1 percent. For example, with a 2 percent target, they would be tasked to keep inflation between 1 percent and 3 percent. But why target 2 percent and not 0 percent?

The answer is that aiming to hit 0 percent could result in negative inflation, known as **deflation**. One of the limitations of monetary policy that we discuss below is its ability or inability to deal with periods of deflation. If deflation is something to be avoided, why not target 10 percent? The answer to this question is that levels of inflation that high would not be consistent with price stability; such a high inflation rate would further tend to be associated with high inflation volatility and uncertainty. Central bankers seem to agree that 2 percent is far enough away from the risks of deflation and low enough not to lead to destabilizing inflation shocks.

Finally, we should keep in mind that the headline inflation rate that is announced in most economies every month, and which is the central bank's target, is a measure of how much a basket of goods and services has risen over the previous twelve months. It is history. Furthermore, interest rate changes made today will take some time to have their full effect on the real economy as they make their way through the monetary transmission mechanism. It is for these two reasons that inflation targeters do not target current inflation but instead usually focus on inflation two years ahead.

Although inflation-targeting mandates may vary from country to country, they have common elements: the specification of an explicit inflation target, with permissible bounds, and a requirement that the central bank should be transparent in its objectives and policy actions. This is all usually laid out in legislation that imposes statutory obligations on the central bank. As mentioned earlier, New Zealand pioneered the inflation-targeting approach to monetary policy that has since been copied widely. Below is New Zealand's Policy Targets Agreement, which specifies the inflation-targeting mandate of its central bank, the Reserve Bank of New Zealand.

Exhibit 10 New Zealand's Policy Targets Agreement

"This agreement between the Minister of Finance and the Governor of the Reserve Bank of New Zealand (the Bank) is made under section 9 of the Reserve Bank of New Zealand Act 1989 (the Act). The Minister and the Governor agree as follows:

1 Price stability

- **a** Under Section 8 of the Act the Reserve Bank is required to conduct monetary policy with the goal of maintaining a stable general level of prices.
- **b** The Government's economic objective is to promote a growing, open and competitive economy as the best means of delivering permanently higher incomes and living standards for New Zealanders. Price stability plays an important part in supporting this objective.

2 Policy target

- **a** In pursuing the objective of a stable general level of prices, the Bank shall monitor prices as measured by a range of price indexes. The price stability target will be defined in terms of the All Groups Consumers Price Index (CPI), as published by Statistics New Zealand.
- **b** For the purpose of this agreement, the policy target shall be to keep future CPI inflation outcomes between 1 per cent and 3 per cent on average over the medium term.

3 Inflation variations around target

- **a** For a variety of reasons, the actual annual rate of CPI inflation will vary around the medium-term trend of inflation, which is the focus of the policy target. Amongst these reasons, there is a range of events whose impact would normally be temporary. Such events include, for example, shifts in the aggregate price level as a result of exceptional movements in the prices of commodities traded in world markets, changes in indirect taxes,[9] significant government policy changes that directly affect prices, or a natural disaster affecting a major part of the economy.

(continued)

---

[9] "Indirect taxes" refer to such taxes as sales taxes and value-added taxes that are levied on goods and services rather than directly on individuals and companies.

Exhibit 10 (Continued)

  - **b** When disturbances of the kind described in clause 3(a) arise, the Bank will respond consistent with meeting its medium-term target.

4 Communication, implementation and accountability

  - **a** On occasions when the annual rate of inflation is outside the medium-term target range, or when such occasions are projected, the Bank shall explain in Policy Statements made under section 15 of the Act why such outcomes have occurred, or are projected to occur, and what measures it has taken, or proposes to take, to ensure that inflation outcomes remain consistent with the medium-term target.
  - **b** In pursuing its price stability objective, the Bank shall implement monetary policy in a sustainable, consistent and transparent manner and shall seek to avoid unnecessary instability in output, interest rates and the exchange rate.
  - **c** The Bank shall be fully accountable for its judgments and actions in implementing monetary policy."

Source: http://www.rbnz.govt.nz/.

To summarize, an inflation-targeting framework normally has the following set of features:

- An independent and credible central bank;
- A commitment to transparency;
- A decision-making framework that considers a wide range of economic and financial market indicators; and
- A clear, symmetric and forward-looking medium-term inflation target, sufficiently above 0 percent to avoid the risk of deflation but low enough to ensure a significant degree of price stability.

Indeed, independence, credibility, and transparency are arguably the crucial ingredients for an effective central bank, whether they target inflation or not.

The Main Exceptions to the Inflation-Targeting Rule

Although the practice of inflation targeting is widespread, there are two prominent central banks that have not adopted a formal inflation target along the lines of the New Zealand model: the Bank of Japan and the US Federal Reserve System.

The Bank of Japan

> Japan's central bank, the Bank of Japan (BoJ), does not target an explicit measure of inflation. Japan's government and its monetary authorities have been trying to combat deflation for much of the last two decades. However, despite their efforts—including the outright printing of money—inflation has remained very weak. Inflation targeting is seen very much as a way of combating and controlling inflation; as such, it would seem to have no place in an economy that suffers from persistent deflation.
>
> Some economists have argued, however, that an inflation target is exactly what the Japanese economy needs. By announcing that positive inflation of say 3 percent is desired by the central bank, this might become a self-fulfilling prophecy if Japanese consumers and companies factor this target into nominal

wage and price contracts. But for economic agents to believe that the target will be achieved, they have to believe that the central bank is capable of achieving it. Given that the BoJ has failed to engineer persistent, positive inflation, it is debatable how much credibility Japanese households and corporations would afford such an inflation-targeting policy.

The US Federal Reserve System

It is perhaps rather ironic that the world's most influential central bank, the US Federal Reserve, which controls the supply of the world's de facto reserve currency, the US dollar, does not have an explicit inflation target. However, it is felt that the single-minded pursuit of inflation might not be compatible with the Fed's statutory goal as laid out in the Federal Reserve Act, which charges the Fed's board to:

> "promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates."

In other words, it has been argued that inflation targeting might compromise the goal of "maximum employment." In practice, however, the Fed has indicated that it sees core inflation measured by the personal consumption expenditure (PCE) deflator of about, or just below, 2 percent as being compatible with "stable prices." Financial markets therefore watch this US inflation gauge very carefully in order to try and anticipate the rate actions of the Fed.

Monetary Policy in Developing Countries

Developing economies often face significant impediments to the successful operation of any monetary policy—that is, the achievement of price stability. These include:

- the absence of a sufficiently liquid government bond market and developed interbank market through which monetary policy can be conducted;
- a rapidly changing economy, making it difficult to understand what the neutral rate might be and what the equilibrium relationship between monetary aggregates and the real economy might be;
- rapid financial innovation that frequently changes the definition of the money supply;
- a poor track record in controlling inflation in the past, making monetary policy intentions less credible; and
- an unwillingness of governments to grant genuine independence to the central bank.

Taken together, any or all of these impediments might call into question the effectiveness of any developing economy's monetary policy framework, making any related monetary policy goals difficult to achieve.

EXAMPLE 9

Central Bank Effectiveness

1 The reason some inflation-targeting banks may target low inflation and not 0 percent inflation is *best* described by which of the following statements?

A Some inflation is viewed as being good for an economy.

- **B** Targeting zero percent inflation runs a higher risk of a deflationary outcome.
- **C** It is very difficult to eliminate all inflation from a modern economy.

2 The degree of credibility that a central bank is afforded by economic agents is important because:

- **A** they are the lender of last resort.
- **B** their targets can become self-fulfilling prophecies.
- **C** they are the monopolistic suppliers of the currency.

Solution to 1:

B is correct. Inflation targeting is art, not science. Sometimes inflation will be above target and sometimes below. Were central banks to target zero percent, then inflation would almost certainly be negative on some occasions. If a deflationary mindset then sets in among economic agents, it might be difficult for the central bank to respond to this because they cannot cut interest rates much below zero.

Solution to 2:

B is correct. If a central bank operates within an inflation-targeting regime and if economic agents believe that it will achieve its target, this expectation will become embedded into wage negotiations, for example, and become a self-fulfilling prophecy. Also, banks need to be confident that the central bank will lend them money when all other sources are closed to them; otherwise, they might curtail their lending drastically, leading to a commensurate reduction in money and economic activity.

8 EXCHANGE RATE TARGETING

k explain the relationships between monetary policy and economic growth, inflation, interest, and exchange rates;

l contrast the use of inflation, interest rate, and exchange rate targeting by central banks;

Many developing economies choose to operate monetary policy by targeting their currency's exchange rate, rather than an explicit level of domestic inflation. Such targeting involves setting a fixed level or band of values for the exchange rate against a major currency, with the central bank supporting the target by buying and selling the national currency in foreign exchange markets. There are recent examples of developed economies using such an approach. In the 1980s, following the failure of its policy of trying to control UK inflation by setting medium-term goals for money supply growth (see Example 5), the UK government decided to operate monetary policy such that the sterling's exchange rate equalled a pre-determined value in terms of German deutschemarks. The basic idea is that by tying a domestic economy's currency to that of an economy with a good track record on inflation, the domestic economy would effectively "import" the inflation experience of the low inflation economy.

Suppose that a developing country wished to maintain the value of its currency against the US dollar. The government and/or central bank would announce the currency exchange rate that they wished to target. To simplify matters, let us assume that the domestic inflation rates are very similar in both countries and that the monetary authorities of the developing economy have set an exchange rate target that is consistent with relative price levels in the two economies. Under these (admittedly unlikely) circumstances, in the absence of shocks, there would be no reason for the exchange rate to deviate significantly from this target level. So as long as domestic inflation closely mirrors US inflation, the exchange rate should remain close to its target (or within a target band). It is in this sense that a successful exchange rate policy imports the inflation of the foreign economy.

Now suppose that economic activity in the developing economy starts to rise rapidly and that domestic inflation in the developing economy rises above the level in the United States. With a freely floating exchange rate regime, the currency of the developing economy would start to fall against the dollar. To arrest this fall, and to protect the exchange rate target, the developing economy's monetary authority sells foreign currency reserves and buys its own currency. This has the effect of reducing the domestic money supply and increasing short-term interest rates. The developing economy experiences a monetary policy tightening which, if expected to bring down inflation, will cause its exchange rate to rise against the dollar.

By contrast, in a scenario in which inflation in the developing country fell relative to the United States, the central bank would need to sell the domestic currency to support the target, tending to increase the domestic money supply and reduce the rate of interest.

In practice, the interventions of the developing economy central bank will simply stabilize the value of its currency, with many frequent adjustments. But this simplistic example should demonstrate one very important fact: *When the central bank or monetary authority chooses to target an exchange rate, interest rates and conditions in the domestic economy must adapt to accommodate this target and domestic interest rates and money supply can become more volatile.*

The monetary authority's commitment to and ability to support the exchange rate target must be credible for exchange rate targeting to be successful. If that is not the case, then speculators may trade against the monetary authority. Speculative attacks forced sterling out of the European Exchange Rate Mechanism in 1992. The fixed exchange rate regime was abandoned and the United Kingdom allowed its currency to float freely. Eventually, the UK government adopted a formal inflation target in 1997. Similarly, in the Asian financial crisis of 1997–1998, Thailand's central bank tried to defend the Thai baht against speculative attacks for much of the first half of 1997 but then revealed at the beginning of July that it had no reserves left. The subsequent devaluation triggered a debt crisis for banks and companies that had borrowed in foreign currency, and contagion spread throughout Asia.

Despite these risks, many currencies are pegged to other currencies, most notably the US dollar. Exhibit 11 shows a list of some of the currencies that were pegged to (fixed against) the US dollar at the end of 2018. Other currencies operate under a "managed exchange rate policy," where they are allowed to fluctuate within a range that is maintained by a monetary authority via market intervention. Dollarization occurs when a country adopts the US dollar as their functional currency. This is stronger than pegging to the dollar because under dollarization the US dollar replaces the previous national currency. Exhibit 11 breaks out countries that peg their currency to the dollar and those that have adopted the US dollar as their currency.

Exhibit 11 Select Currencies Pegged to the US Dollar, as of December 2018

Pegged to USD

- Bermuda
- Bahamas
- Lebanon
- Hong Kong SAR
- Saudi Arabia
- Qatar
- United Arab Emirates

(continued)

Exhibit 11 (Continued)

Dollarized

- Panama (1904)
- Ecuador (2000)
- East Timor (2001)
- El Salvador (2000)
- Caribbean Netherlands (2011)

EXAMPLE 10

Exchange Rate Targeting

1 When the central bank chooses to target a specific value for its exchange rate:

   A it must also target domestic inflation.

   B it must also set targets for broad money growth.

   C conditions in the domestic economy must adapt to accommodate this target.

2 With regard to monetary policy, what is the hoped for benefit of adopting an exchange rate target?

   A Freedom to pursue redistributive fiscal policy

   B Freedom to set interest rates according to domestic conditions

   C To “import” the inflation experience of the economy whose currency is being targeted

3 Which of the following is *least* likely to be an impediment to the successful implementation of monetary policy in developing economies?

   A Fiscal deficits

   B Rapid financial innovation

   C Absence of a liquid government bond market

Solution to 1:

C is correct. The adoption of an exchange rate target requires that the central bank set interest rates to achieve this target. If the target comes under pressure, domestic interest rates may have to rise, regardless of domestic conditions. It may have a “target” level of inflation in mind as well as “targets” for broad money growth, but as long as it targets the exchange rate, domestic inflation and broad money trends must simply be allowed to evolve.

Solution to 2:

C is correct. Note that interest rates have to be set to achieve this target and are therefore subordinate to the exchange rate target and partially dependent on economic conditions in the foreign economy.

Solution to 3:

A is correct. Note that the absence of a liquid government bond market through which a central bank can enact open market operations and/or repo transactions will inhibit the implementation of monetary policy—as would rapid financial innovation because such innovation can change the relationship between money and economic activity. Fiscal deficits, on the other hand, are not normally an impediment to the implementation of monetary policy, although they could be if they were perceived to be unsustainable.

MONETARY POLICIES: CONTRACTIONARY, EXPANSIONARY, LIMITATIONS

9

- **m** determine whether a monetary policy is expansionary or contractionary;
- **n** describe limitations of monetary policy;

Most central banks will adjust liquidity conditions by adjusting their official policy rate.[10] When they believe that economic activity is likely to lead to an increase in inflation, they might increase interest rates, thereby reducing liquidity. In these cases, market analysts describe such actions as **contractionary** because the policy is designed to cause the rate of growth of the money supply and the real economy to contract (see Exhibit 7 for the possible transmission mechanism here). Conversely, when the economy is slowing and inflation and monetary trends are weakening, central banks may increase liquidity by cutting their target rate. In these circumstances, monetary policy is said to be **expansionary**.

Thus, when policy rates are high, monetary policy may be described as contractionary; when low, they may be described as expansionary. But what are they "high" and "low" in comparison to?

The **neutral rate of interest** is often taken as the point of comparison. One way of characterizing the neutral rate is to say that it is that rate of interest that neither spurs on nor slows down the underlying economy. As such, when policy rates are above the neutral rate, monetary policy is contractionary; when they are below the neutral rate, monetary policy is expansionary. The neutral rate should correspond to the average policy rate over a business cycle.

However, economists' views of the neutral rate for any given economy might differ, and therefore, their view of whether monetary policy is contractionary, neutral, or expansionary might differ too. What economists do agree on is that the neutral policy rate for any economy comprises two components:

- Real trend rate of growth of the underlying economy, and
- Long-run expected inflation.

The real trend rate of growth of an economy is also difficult to discern, but it corresponds to that rate of economic growth that is achievable in the long run that gives rise to stable inflation. If we are thinking about an economy with a credible inflation-targeting regime, where the inflation target is say 2 percent per year and where an analyst believes that the economy can grow sustainably over the long term at a rate of 2.5 percent per year, then they might also estimate the neutral rate to be:

$$\text{Neutral rate} = \text{Trend growth} + \text{Inflation target} = 2.5\% + 2\% = 4.5\% \quad (4)$$

The analyst would therefore describe the central bank's monetary policy as being contractionary when its policy rate is above 4.5 percent and expansionary when it is below this level.

---

[10] Although, if they have reduced their policy rate to 0 percent, to increase liquidity further they have to resort to less-conventional monetary policy measures.

In practice, central banks often indicate what they believe to be the neutral rate of interest for their economy too. But determining this “neutral rate” is more art than science. For example, many analysts have recently revised down their estimates of trend growth for many western countries following the collapse of the credit bubble, because in many cases, the governments and private individuals of these economies are now being forced to reduce consumption levels and pay down their debts.

9.1 What’s the Source of the Shock to the Inflation Rate?

An important aspect of monetary policy for those charged with its conduct is the determination of the source of any shock to the inflation rate. Suppose that the monetary authority sees that inflation is rising beyond its target, or simply in a way that threatens price stability. If this rise was caused by an increase in the confidence of consumers and business leaders, which in turn has led to increases in consumption and investment growth rates, then we could think of it as being a **demand shock**. In this instance, it might be appropriate to tighten monetary policy in order to bring the inflationary pressures generated by these domestic demand pressures under control.

However, suppose instead that the rise in inflation was caused by a rise in the price of oil (for the sake of argument). In this case, the economy is facing a **supply shock**, and raising interest rates might make a bad situation worse. Consumers are already facing an increase in the cost of fuel prices that might cause profits and consumption to fall and eventually unemployment to rise. Putting up interest rates in this instance might simply exacerbate the oil price-induced downturn, which might ultimately cause inflation to fall sharply.

It is important, then, for the monetary authority to try to identify the source of the shock before engineering a contractionary or expansionary monetary policy phase.

9.2 Limitations of Monetary Policy

The limitations of monetary policy include problems in the transmission mechanism and the relative ineffectiveness of interest rate adjustment as a policy tool in deflationary environments.

9.2.1 *Problems in the Monetary Transmission Mechanism*

In Exhibit 7, we presented a stylized representation of the monetary policy transmission mechanism, including the channels of bank lending rates, asset prices, expectations, and exchange rates. The implication of the diagram is that there are channels through which the actions of the central bank or monetary authority are transmitted to both the nominal and real economy. However, there may be some occasions when the will of the monetary authority is not transmitted seamlessly through the economy.

Suppose that a central bank raises interest rates because it is concerned about the strength of underlying inflationary pressures. Long-term interest rates are influenced by the path of expected short-term interest rates, so the outcome of the rate hike will depend on market expectations. Suppose that bond market participants think that short-term rates are already too high, that the monetary authorities are risking a recession, and that the central bank will likely undershoot its inflation target. This fall in inflation expectations could cause long-term interest rates to fall. That would make long-term borrowing cheaper for companies and households, which could in turn stimulate economic activity rather than cause it to contract.

Arguably, the more credible the monetary authority, the more stable the long end of the yield curve; moreover, the monetary authority will be more confident that its “policy message” will be transmitted throughout the economy. A term recently used in the marketplace is **bond market vigilantes**. These “vigilantes” are bond market participants who might reduce their demand for long-term bonds, thus pushing up their yields, if they believe that the monetary authority is losing its grip on inflation. That yield increase could act as a brake on any loose monetary policy stance. Conversely, the vigilantes may push long-term rates down by increasing their demand for long-dated government bonds if they expect that tight monetary policy is likely to cause a sharp slowdown in the economy, thereby loosening monetary conditions for long-term borrowers in the economy.

A credible monetary policy framework and authority will tend not to require the vigilantes to do the work for it.

In very extreme instances, there may be occasions where the demand for money becomes infinitely elastic—that is, where the demand curve is horizontal and individuals are willing to hold additional money balances without any change in the interest rate—so that further injections of money into the economy will not serve to further lower interest rates or affect real activity. This is known as a **liquidity trap**. In this extreme circumstance, monetary policy can become completely ineffective. The economic conditions for a liquidity trap are associated with the phenomenon of **deflation**.

9.2.2 *Interest Rate Adjustment in a Deflationary Environment and Quantitative Easing as a Response*

Deflation is a pervasive and persistent fall in a general price index and is more difficult for conventional monetary policy to deal with than inflation. This is because cutting nominal interest rates much below zero to stimulate the economy is difficult.[11] It is at this point that the economic conditions for a liquidity trap arise.

Deflation raises the real value of debt, while the persistent fall in prices can encourage consumers to put off consumption today, leading to a fall in demand that leads to further deflationary pressure. Thus a deflationary "trap" can develop, which is characterized by weak consumption growth, falling prices, and increases in real debt levels. Japan eventually found itself in such a position following the collapse of its property bubble in the early 1990s.

If conventional monetary policy—the adjustment of short-term interest rates—is no longer capable of stimulating the economy once the zero or even negative nominal interest rate bound has been reached, is monetary policy useless?

In the aftermath of the collapse of the high-tech bubble in November 2002, Federal Reserve Governor (now Chairman) Ben Bernanke gave a speech entitled "Deflation: Making Sure 'It' Doesn't Happen Here." In this speech, Bernanke stated that inflation was always and everywhere a monetary phenomenon, and he expressed great confidence that by expanding the money supply by various means (including dropping it out of a helicopter on the population below), the Federal Reserve as the monopoly supplier of money could always engineer positive inflation in the US economy. He said:

> I am confident that the Fed would take whatever means necessary to prevent significant deflation in the United States and, moreover, that the US central bank, in cooperation with other parts of the government as needed, has sufficient policy instruments to ensure that any deflation that might occur would be both mild and brief.

Following the collapse of the credit bubble in 2008, a number of governments along with their central banks cut rates to (near) zero, including those in the United States and the United Kingdom. However, there was concern that the underlying economies might not respond to this drastic monetary medicine, mainly because the related banking crisis had caused banks to reduce their lending drastically. In order to kick start the process, both the Federal Reserve and the Bank of England effectively printed money and pumped it in to their respective economies. This "unconventional" approach to monetary policy, known as **quantitative easing** (QE), is operationally similar to open market purchase operations but conducted on a much larger scale.

---

[11] Interest rates were cut to below zero in several European countries in 2014 and subsequently in Japan in 2016.

The additional reserves created by central banks in a policy of quantitative easing can be used to buy any assets. The Bank of England chose to buy **gilts** (bonds issued by the UK government), where the focus was on gilts with three to five years maturity. The idea was that this additional reserve would kick-start lending, causing broad money growth to expand, which would eventually lead to an increase in real economic activity. But there is no guarantee that banks will respond in this way. In a difficult economic climate, it may be better to hold excess reserves rather than to lend to households and businesses that may default.

In the United States, the formal plan for QE mainly involved the purchase of mortgage bonds issued or guaranteed by Freddie Mac and Fannie Mae. Part of the intention was to push down mortgage rates to support the US housing market, as well as to increase the growth rate of broad money. Before implementing this formal program, the Federal Reserve intervened in several other markets that were failing for lack of liquidity, including interbank markets and the commercial paper market. These interventions had a similar effect on the Federal Reserve's balance sheet and the money supply as the later QE program.

This first round of QE by the Federal Reserve was then followed by a further round of QE, known as QE2. In November 2010, the Federal Reserve judged that the US economy had not responded sufficiently to the first round of QE (QE1). The Fed announced that it would create $600 billion and use this money to purchase long-dated US Treasuries in equal tranches over the following eight months. The purpose of QE2 was to ensure that long bond yields remained low in order to encourage businesses and households to borrow for investment and consumption purposes, respectively.

The final round of QE, known as QE3, was implemented in September 2012 to provide $40 billion per month to purchase agency mortgage-backed securities "until the labor market improved substantially." QE3 lasted until December 2013, when the Federal Reserve announced it was tapering back on these purchases. These purchases, and quantitative easing, ended 10 months later in October 2014.

As long as central banks have the appropriate authority from the government, they can purchase any assets in a quantitative easing program. But the risks involved in purchasing assets with credit risk should be clear. In the end, the central bank is just a special bank. If it accumulates bad assets that then turn out to create losses, it could face a fatal loss of confidence in its main product: fiat money.

9.2.3 *Limitations of Monetary Policy: Summary*

The ultimate problem for monetary authorities as they try to manipulate the supply of money in order to influence the real economy is that they cannot control the amount of money that households and corporations put in banks on deposit, nor can they easily control the willingness of banks to create money by expanding credit. Taken together, this also means that they cannot always control the money supply. Therefore, there are definite limits to the power of monetary policy.

EXAMPLE 11

The Limits of Monetary Policy: The Case of Japan

The Background

Between the 1950s and 1980s, Japan's economy achieved faster real growth than any other G7 economy. But the terrific success of the economy sowed the seeds of the problems that were to follow. The very high real growth rates achieved by Japan over four decades became built in to asset prices, particularly equity and commercial property prices. Toward the end of the 1980s, asset prices rose to even higher levels when the Bank of Japan followed a very easy monetary policy as it tried to prevent the Japanese yen from appreciating too much against the US dollar. However, when interest rates went up in 1989–1990 and the economy slowed, investors eventually came to believe that the growth assumptions that were built in to asset prices and other aspects of the Japanese economy were unrealistic. This realization caused Japanese asset prices to collapse. For example, the Nikkei 225 stock market index reached 38,915 in 1989; by the end of March 2003, it had fallen by 80 percent to 7,972. The collapse in asset prices caused wealth to decline dramatically. Consumer confidence understandably fell sharply too, and consumption growth slowed. Corporate spending also fell, while bank lending contracted sharply in the weak economic climate. Although many of these phenomena are apparent in all recessions, the situation was made worse when deflation set in. In an environment when prices are falling, consumers may put off discretionary spending today until tomorrow; by doing this, however, they exacerbate the deflationary environment. Deflation also raises the real value of debts; as deflation takes hold, borrowers find the real value of their debts rising and may try to increase their savings accordingly. Once again, such actions exacerbate the recessionary conditions.

The Monetary Policy Response

Faced with such a downturn, the conventional monetary policy response is to cut interest rates to try to stimulate real economic activity. The Japanese central bank, the Bank of Japan, cut rates from 8 percent in 1990 to 1 percent by 1996. By February 2001, the Japanese policy rate was cut to zero where it stayed.

Once rates are at or near zero, there are two broad approaches suggested by theory, though the two are usually complementary. First, the central bank can try to convince markets that interest rates will remain low for a long time, even after the economy and inflation pick up. This will tend to lower interest rates along the yield curve. Second, the central bank can try to increase the money supply by purchasing assets from the private sector, so-called quantitative easing. The Bank of Japan (BoJ) did both in 2001. It embarked on a program of quantitative easing supplemented by an explicit promise not to raise short-term interest rates until deflation had given way to inflation.

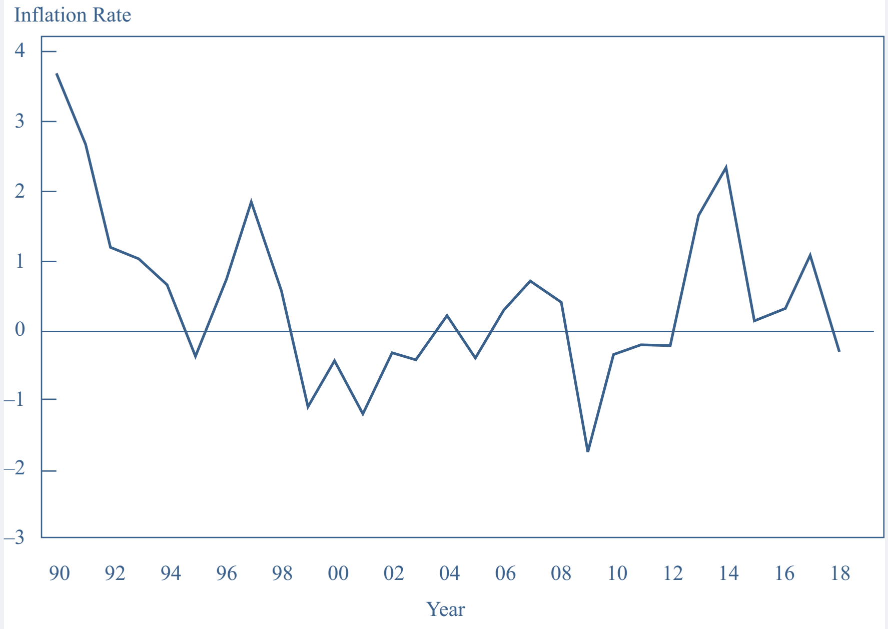
Quantitative easing simply involves the printing of money by the central bank. In practice, this involved the BoJ using open market operations to add reserves to the banking system through the direct purchase of government securities in the open market.

The reserve levels became the new target. The BoJ's monetary policy committee determined the level of reserves and the quantity of bond purchases that should be undertaken, rather than voting on the policy rate.

The success of this policy is difficult to judge. As the chart below shows, although deflation turned to inflation for a while, it returned to deflation in 2008–2009 when the Japanese economy suffered a sharp recession along with much of the rest of the world. At that time, having reversed its QE policy during 2004–2008 by reducing its bond holdings, the Bank of Japan began to buy again.

The Bank of Japan ramped up its asset purchases starting in 2013, when other central banks began to unwind their QE programs. At the beginning of 2013, BoJ Assets to Japanese GDP were approximately 30%. By mid-2018, BoJ assets to Japanese GDP were almost 100%! Economists debate the point, but arguably, even the Bank of Japan's much larger program of QE has not been able to eliminate deflation. The Japanese experience suggests that there may be limits to the power of monetary policy.

Exhibit 12 Inflation and Deflation in Japan



Source: www.statbureau.org/en/japan/inflation-tables.

EXAMPLE 12

Evaluating Monetary Policy

1 If an economy's trend GDP growth rate is 3 percent and its central bank has a 2 percent inflation target, which policy rate is *most consistent* with an expansionary monetary policy?

A 4 percent

B 5 percent

C 6 percent

2 An increase in a central bank's policy rate might be expected to reduce inflationary pressures by:

   A reducing consumer demand.

   B reducing the foreign exchange value of the currency.

   C driving up asset prices leading to an increase in personal sector wealth.

3 Which of the following statements *best* describes a fundamental limitation of monetary policy? Monetary policy is limited because central bankers:

   A cannot control the inflation rate perfectly.

   B are appointed by politicians and are therefore never truly independent.

   C cannot control the amount of money that economic agents put in banks, nor the willingness of banks to make loans.

Solution to 1:

A is correct. The neutral rate of interest, which in this example is 5 percent, is considered to be that rate of interest that neither spurs on nor slows down the underlying economy. As such, when policy rates are above the neutral rate, monetary policy is contractionary; when they are below the neutral rate, monetary policy is expansionary. It comprises two components: the real trend rate of growth of the underlying economy (in this example, 3 percent) and long-run expected inflation (in this example, 2 percent).

Solution to 2:

A is correct. If an increase in the central bank's policy rate is successfully transmitted via the money markets to other parts of the financial sector, consumer demand might decline as the rate of interest on mortgages and other credit rises. This decline in consumer demand should, all other things being equal and amongst other affects, lead to a reduction in upward pressure on consumer prices.

Solution to 3:

C is correct. Central bankers do not control the decisions of individuals and banks that can influence the money creation process.

ROLES AND OBJECTIVES OF FISCAL POLICY

10

o describe roles and objectives of fiscal policy;

The second set of tools used for influencing economic activity consists of the tools associated with fiscal policy. These involve the use of government spending and changing tax revenue to affect a number of aspects of the economy:

- Overall level of aggregate demand in an economy and hence the level of economic activity.
- Distribution of income and wealth among different segments of the population.
- Allocation of resources between different sectors and economic agents.

Often, a discussion of fiscal policy focuses on the impact of changes in the difference between government spending and revenue on the aggregate economy, rather than on the actual levels of spending and revenue themselves.

10.1 Roles and Objectives of Fiscal Policy

A primary aim for fiscal policy is to help manage the economy through its influence on aggregate national output, that is, real GDP.

10.1.1 *Fiscal Policy and Aggregate Demand*

Aggregate demand is the amount companies and households plan to spend. We can consider a number of ways that fiscal policy can influence aggregate demand. For example, an **expansionary** policy could take one or more of the following forms:

- Cuts in personal income tax raise disposable income with the objective of boosting aggregate demand.
- Cuts in sales (indirect) taxes to lower prices which raises real incomes with the objective of raising consumer demand.
- Cuts in corporation (company) taxes to boost business profits, which may raise capital spending.
- Cuts in tax rates on personal savings to raise disposable income for those with savings, with the objective of raising consumer demand.
- New public spending on social goods and infrastructure, such as hospitals and schools, boosting personal incomes with the objective of raising aggregate demand.

We must stress, however, that the reliability and magnitude of these relationships will vary over time and from country to country. For example, in a recession with rising unemployment, it is not always the case that cuts in income taxes will raise consumer spending because consumers may wish to raise their precautionary (rainy day) saving in anticipation of further deterioration in the economy. Indeed, in very general terms economists are often divided into two camps regarding the workings of fiscal policy: **Keynesians** believe that fiscal policy can have powerful effects on aggregate demand, output, and employment when there is substantial spare capacity in an economy. **Monetarists** believe that fiscal changes only have a temporary effect on aggregate demand and that monetary policy is a more effective tool for restraining or boosting inflationary pressures. Monetarists tend not to advocate using monetary policy for countercyclical adjustment of aggregate demand. This intellectual division will naturally be reflected in economists' divergent views on the efficacy of the large fiscal expansions observed in many countries following the credit crisis of 2008, along with differing views on the possible impact of quantitative easing.

10.1.2 *Government Receipts and Expenditure in Major Economies*

In Exhibit 13, we present the total government revenues as a percentage of GDP for some major economies. This is the share of a country's output that is gathered by the government through taxes and such related items as fees, charges, fines, and capital transfers. It is often considered as a summary measure of the extent to which a government is involved both directly and indirectly in the economic activity of a country.

Taxes are formally defined as compulsory, unrequited payments to the general government (they are unrequited in the sense that benefits provided by a government to taxpayers are usually not related to payments). Exhibit 13 contains taxes on incomes and profits, social security contributions, indirect taxes on goods and services, employment taxes, and taxes on the ownership and transfer of property.

Exhibit 13 General Government Revenues as Percent of GDP

	1995	2000	2005	2008	2010	2015
Australia	34.5	36.1	36.5	35.3	32.4	34.9
Germany	45.1	46.4	43.6	43.8	43.0	44.5
Japan	31.2	31.4	31.7	34.4	30.6	35.7
United Kingdom	38.2	40.3	40.8	42.2	38.2	38.0
United States	33.8	35.4	33.0	32.3	30.9	33.4
OECD	37.9	39.0	37.7	37.9	39.8	40.9

Source: Organisation for Economic Co-Operation and Development (OECD).

Taxes on income and profits have been fairly constant for the Organisation for Economic Co-Operation and Development (OECD) countries overall at around 12.5–13 percent of GDP since the mid-1990s, while taxes on goods and services have been steady at about 11 percent of GDP for that period. Variations between countries can be substantial; taxes on goods and services are around 5 percent of GDP for the United States and Japan but over 16 percent for Denmark.

Exhibit 14 shows the percentage of GDP represented by government expenditure in a variety of major economies over time. Generally, these have been fairly constant since 1995, though Germany had a particularly high number at the start of the period because of reunification costs. The impacts of governments' fiscal stimulus programs in the face of the 2008–2009 financial crisis show up as significant increases in government expenditures in Exhibit 14 and increases in government deficits in Exhibit 15 between 2008 and 2010.

Exhibit 14 General Government Expenditures as Percent of GDP

	1995	2000	2005	2008	2010	2015
Australia	38.2	35.2	34.8	34.3	34.4	36.2
Germany	54.8	45.1	46.9	43.8	47.3	43.9
Japan	36.0	39.0	38.4	37.1	39.6	39.4
United Kingdom	44.1	36.6	44.0	47.5	47.6	42.2
United States	37.1	33.9	36.2	38.8	42.9	37.6
OECD	42.7	38.7	40.5	41.4	45.2	41.8

Source: OECD.

Clearly, the possibility that fiscal policy can influence output means that it may be an important tool for **economic stabilization**. In a recession, governments can raise spending (**expansionary fiscal policy**) in an attempt to raise employment and output. In boom times—when an economy has full employment and wages and prices are rising too fast—then government spending may be reduced and taxes raised (**contractionary fiscal policy**).

Hence, a key concept is the **budget surplus/deficit**, which is the difference between government revenue and expenditure for a fixed period of time, such as a fiscal or calendar year. Government revenue includes tax revenues net of transfer payments; government spending includes interest payments on the government debt. Analysts often focus on changes in the budget surplus or deficit from year to year as indicators of whether the fiscal policy is getting tighter or looser. An increase in a budget surplus would be associated with contractionary fiscal policy, while a rise in a deficit is an expansionary fiscal policy. Of course, over the course of a business cycle the budget surplus will vary automatically in a countercyclical way. For example, as an economy slows and unemployment rises, government spending on social insurance and unemployment benefits will also rise and add to aggregate demand. This is known as an **automatic stabilizer**. Similarly, if boom conditions ensue and employment and incomes are high, then progressive income and profit taxes are rising and also act as automatic stabilizers increasing budget surplus or reducing budget deficit. The great advantage of automatic stabilizers is that they are indeed automatic, not requiring the identification of shocks to which policymakers must consider a response. By reducing the responsiveness of the economy to shocks, these automatic stabilizers reduce output fluctuations. Automatic stabilizers should be distinguished from discretionary fiscal policies, such as changes in government spending or tax rates, which are actively used to stabilize aggregate demand. If government spending and revenues are equal, then the budget is **balanced**.

Exhibit 15 General Government Net Borrowing or Lending as Percent of GDP

	1995	2000	2005	2008	2010	2015
Australia	−3.7	0.9	1.7	−3.8	−4.4	−2.2
Germany	−9.7	1.3	−3.3	−0.2	−4.2	0.8
Japan	−4.7	−7.6	−6.7	−4.1	−9.1	−3.6
United Kingdom	−5.8	3.7	−3.3	−5.1	−9.4	−4.2
United States	−3.3	1.5	−3.3	−7.0	−12.0	−4.2
OECD	−4.8	0.2	−2.7	−1.5	−5.1	−1.9

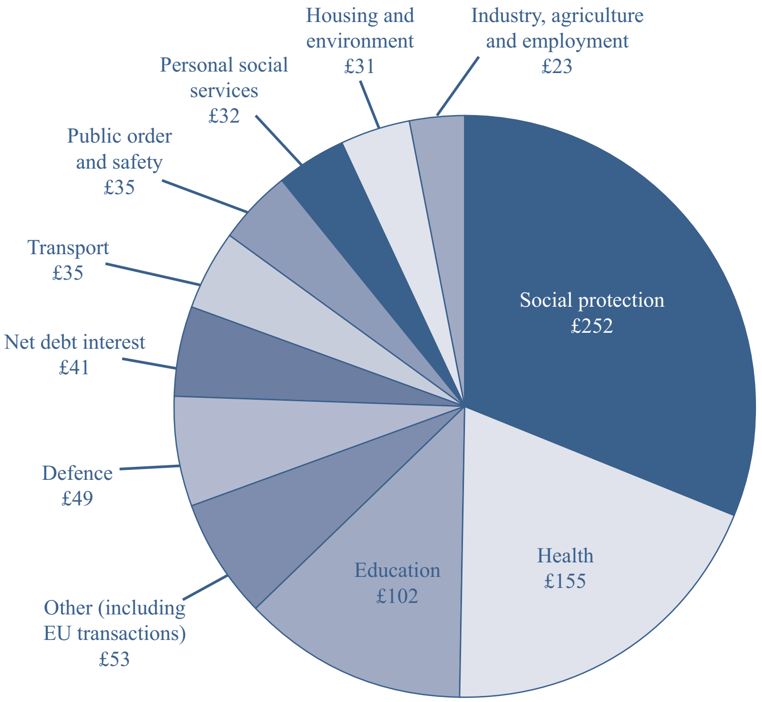
Source: OECD.

EXAMPLE 13

Sources and Uses of Government Cash Flows: The Case of the United Kingdom

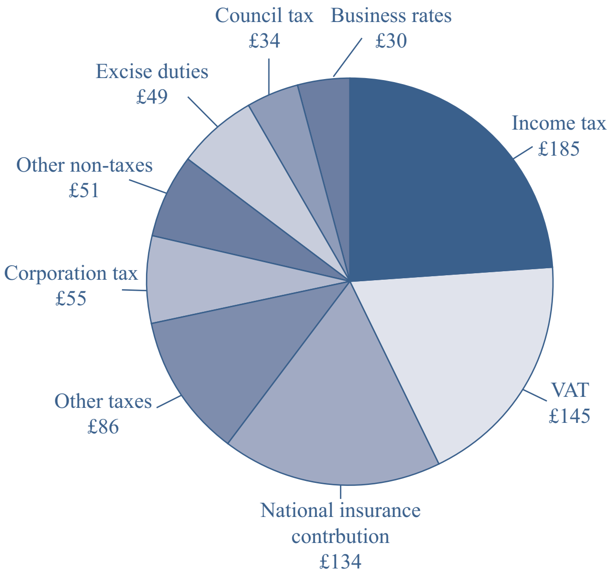
The precise components of revenue and expenditure will of course vary over time and between countries. But, as an example of the breakdown of expenditure and revenue, in Exhibits 16 and 17 we have presented the budget projections of the United Kingdom for 2018/2019. The budget projected that total spending would come to £808bn, while total revenue would only be £769bn. The government was therefore forecasting a budget shortfall of £39bn for the fiscal year, meaning that it had an associated need to borrow £39bn from the private sector in the United Kingdom or the private and public sectors of other economies.

Exhibit 16 **Where Does the Money Go? The United Kingdom, 2018–2019**



Note: All values are in billions of pounds.
Source: HM Treasury, United Kingdom.

Exhibit 17 Where Does the Money Come From? The United Kingdom, 2018–2019



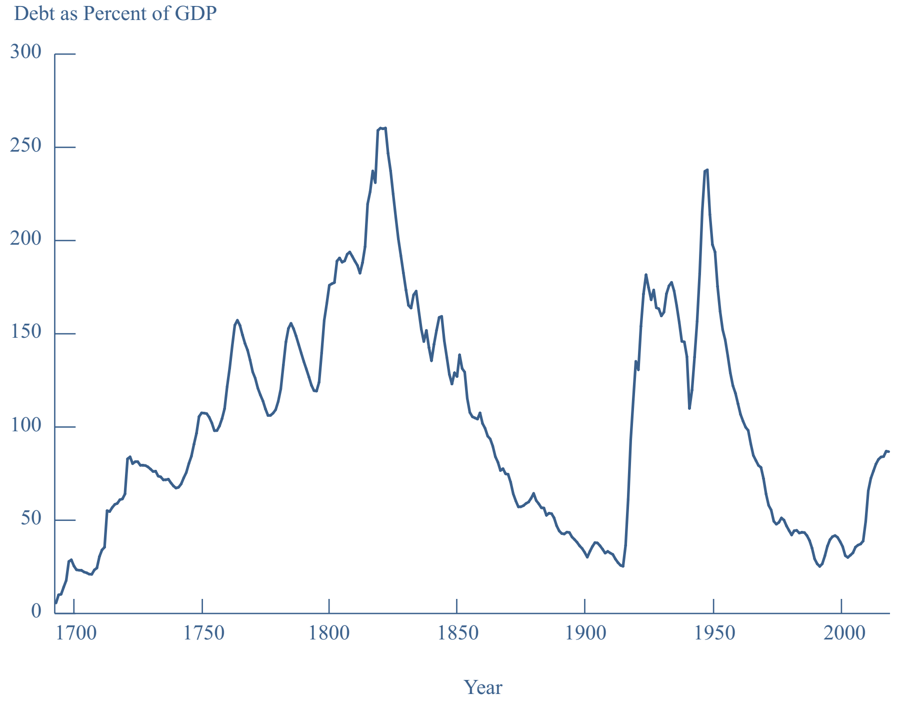
Note: All values are in billions of pounds.
Source: HM Treasury, United Kingdom.

11 DEFICITS AND THE NATIONAL DEBT

q describe the arguments about whether the size of a national debt relative to GDP matters;

Government deficits are the difference between government revenues and expenditures over a period of calendar time, usually a year. Government (or national) debt is the accumulation over time of these deficits. Government deficits are financed by borrowing from the private sector, often via private pension and insurance fund portfolio investments. We saw above that governments are more likely to have deficits than surpluses over long periods of time. As a result, there may exist a large stock of outstanding government debt owned by the private sector. This will vary as the business cycle ebbs and flows. Exhibit 18 shows the time path of the ratio of public debt to GDP for the United Kingdom over several hundred years. It can be clearly seen that the major cause of fluctuations in that ratio through history has been the financing of wars, in particular the Napoleonic Wars of 1799–1815 and the First and Second World Wars of 1914–1918 and 1939–1945.

Exhibit 18 UK National Debt as Percent of GDP (1692–2018)



Source: http://ukpublicspending.co.uk.

With the onset of the credit crisis of 2008, governments actively sought to stimulate their economies through increased expenditures without raising taxes and revenues. This led to increased borrowing, shown in Exhibits 15 and 19, which has become a concern in the financial markets in 2010 for such countries as Greece. Indeed, between 2008 and 2009, central government debt rose from $1.2 trillion to $1.6 trillion in the United Kingdom and from $5.8 trillion to $7.5 trillion for the United States.[12] The fiscal expansion by governments in the face of the financial crisis seems to have significantly raised the General Government Debt to GDP ratio over the long term for many countries, as illustrated in Exhibit 19.

Exhibit 19 General Government Debt as Percent of GDP

	1995	2000	2005	2008	2010	2015
Australia	57.3	41.1	30.0	30.0	41.9	64.1
Germany	54.1	59.5	70.1	68.1	84.5	78.9
Japan	94.7	142.6	176.2	181.6	207.5	237.4
United Kingdom	51.4	48.7	51.3	63.3	88.8	111.7
United States	83.2	61.7	79.0	93.2	117.0	125.3
OECD	65.8	59.9	59.5	60.8	73.0	85.3

Source: www.oecd.org.

---

[12] www.oecd.org.

Ultimately, if the ratio of debt to GDP rises beyond a certain unknown point, then the solvency of the country comes into question. An additional indicator for potential insolvency is the ratio of interest rate payments to GDP, which is shown for some major economies in Exhibit 20. These represent payments required of governments to service their debts as a percentage of national output and as such reflect both the size of debts and the interest charged on them. Such ratios could rise rapidly with the growing debt ratios of 2009 and 2010, particularly if the interest rates on the debt were to rise from the historically low levels.

Exhibit 20 General Government Net Debt Interest Payments as Percent of GDP

	1995	2000	2005	2008	2010	2015
Australia	3.5	1.7	1.0	–0.5	0.0	0.3
Germany	2.9	2.7	2.4	2.3	2.1	0.9
Japan	1.3	1.5	0.8	0.3	0.6	0.4
United Kingdom	3.1	2.4	1.8	1.7	2.6	2.0
United States	3.5	2.5	1.8	2.6	2.9	2.8
OECD	3.6	2.5	1.8	1.9	2.1	1.9

Source: OECD.

Governments' spending was far in excess of revenues following the credit crisis of 2007–2010 as governments tried to stimulate their economies; this level of spending raised concerns in some quarters about the scale of governmental debt accumulation. Exhibit 19 shows that government debt relative to GDP for the OECD countries overall rose from 59.5 percent in 2005 to 85.3 percent in 2015. In Japan, where fiscal spending has been used to stimulate the economy from the early 1990s, the ratio has risen from 94.7 percent in 1995 to 237.4 percent in 2015. If an economy grows in real terms, so do the real tax revenues and hence the ability to service a growing real debt at constant tax rate levels. However, if the real growth in the economy is lower than the real interest rate on the debt, then the debt ratio will worsen even though the economy is growing because the debt burden (i.e., the real interest rate times the debt) grows faster than the economy. Hence, an important issue for governments and their creditors is whether their additional spending leads to sufficiently higher tax revenues to pay the interest on the debt used to finance the extra spending.

However, within a national economy, the real value of the outstanding debt will fall if the overall price level rises (i.e., inflation, and hence a rise in nominal GDP even if real GDP is static) and thus the ratio of debt to GDP may not be rising. But if the general price level falls (i.e., deflation), then the ratio may stay elevated for longer. If net interest payments rise rapidly and investors lose confidence in a government's ability to honour its debts, then financing costs may escalate even more quickly and make the situation unstable.

Should we be concerned about the size of a national debt (relative to GDP)? There are strong arguments both for and against:

> The arguments against being concerned about national debt (relative to GDP) are as follows:
>
> - The scale of the problem may be overstated because the debt is owed internally to fellow citizens. This is certainly the case in Japan and South Korea, where 93 percent is owned by local residents. Canada is similar with 90 percent owned by residents. However, other countries have a much lower

percentage owned internally. In the United States and United Kingdom, the figures are 53 percent and 73 percent, respectively, while Italy has only 46 percent owned by local residents.[13]

- A proportion of the money borrowed may have been used for capital investment projects or enhancing human capital (e.g., training, education); these should lead to raised future output and tax revenues.
- Large fiscal deficits require tax changes which may actually reduce distortions caused by existing tax structures.
- Deficits may have no net impact because the private sector may act to offset fiscal deficits by increasing saving in anticipation of future increased taxes. This argument is known as "Ricardian equivalence" and is discussed in more detail later.
- If there is unemployment in an economy, then the debt is not diverting activity away from productive uses (and indeed the debt could be associated with an increase in employment).

The arguments in favour of being concerned are:

- High levels of debt to GDP may lead to higher tax rates in the search for higher tax revenues. This may lead to disincentives to economic activity as the higher marginal tax rates reduce labour effort and entrepreneurial activity, leading to lower growth in the long run.
- If markets lose confidence in a government, then the central bank may have to print money to finance a government deficit. This may lead ultimately to high inflation, as evidenced by the economic history of Germany in the 1920s and more recently in Zimbabwe.
- Government borrowing may divert private sector investment from taking place (an effect known as **crowding out**); if there is a limited amount of savings to be spent on investment, then larger government demands will lead to higher interest rates and lower private sector investing.

An important distinction to make is between long- and short-run effects. Over short periods of time (say, a few years), crowding out may have little effect. If it lasts for a longer time, however, then capital accumulation in an economy may be damaged. Similarly, tax distortions may not be too serious over the short-term but will have a more substantial impact over many years.

EXAMPLE 14

Types of Fiscal Policies

1. Which of the following is *not* associated with an expansionary fiscal policy?
   - **A** A rise in capital gains taxes
   - **B** Cuts in personal income taxes
   - **C** New capital spending by the government on road building
2. Fiscal expansions will *most likely* have the most impact on aggregate output when the economy is in which of the following states?
   - **A** Full employment

---

[13] These data come from the Bank for International Settlements (BIS), IMF, and central bank websites. All figures are as of 2018.

B Near full employment

C Considerable unemployment

3 Which one of the following is *most likely* a reason to *not* use fiscal deficits as an expansionary tool?

A They may crowd out private investment.

B They may facilitate tax changes to reduce distortions in an economy.

C They may stimulate employment when there is substantial unemployment in an economy.

Solution to 1:

A is correct. A rise in capital gains taxes reduces income available for spending and hence reduces aggregate demand, other things being equal. Cutting income tax raises disposable income, while new road building raises employment and incomes; in both cases, aggregate demand rises and hence policy is expansionary.

Solution to 2:

C is correct. When an economy is close to full employment a fiscal expansion raising aggregate demand can have little impact on output because there are few spare unused resources (e.g., labour or idle factories); instead, there will be upward pressure on prices (i.e., inflation).

Solution to 3:

A is correct. A frequent argument against raises in fiscal deficits is that the additional borrowing to fund the deficit in financial markets will displace private sector borrowing for investment (i.e., "crowd it out").

12 FISCAL POLICY TOOLS

p describe tools of fiscal policy, including their advantages and disadvantages;

We now look at the nature of the fiscal tools available to a government. Government spending can take a variety of forms:

- **Transfer payments** are welfare payments made through the social security system and, depending on the country, comprise payments for state pensions, housing benefits, tax credits and income support for poorer families, child benefits, unemployment benefits, and job search allowances. Transfer payments exist to provide a basic minimum level of income for low-income households, and they also provide a means by which a government can change the overall income distribution in a society. Note that these payments are not included in the definition of GDP because they do not reflect a reward to a factor of production for economic activity. Also, they are not considered to be part of general government spending on goods and services.
- **Current government spending** involves spending on goods and services that are provided on a regular, recurring basis—including health, education, and defense. Clearly, such spending will have a big impact on a country's skill level and overall labour productivity.
- **Capital expenditure** includes infrastructure spending on roads, hospitals, prisons, and schools. This investment spending will add to a nation's capital stock and affect productive potential for an economy.

Government spending can be justified on both economic and social grounds:

- To provide such services as defense that benefit all citizens equally.
- For infrastructure capital spending (e.g., roads) to help a country's economic growth.
- To guarantee a minimum level of income for poorer people and hence redistribute income and wealth (e.g., welfare and related benefits).
- To influence a government's economic objectives of low inflation and high employment and growth (e.g., management of aggregate demand).
- To subsidize the development of innovative and high-risk new products or markets (e.g., alternative energy sources).

Government revenues can take several forms:

- **Direct taxes** are levied on income, wealth, and corporate profits and include capital gains taxes, national insurance (or labour) taxes, and corporate taxes. They may also include a local income or property tax for both individuals and businesses. Inheritance tax on a deceased's estate will have both revenue-raising and wealth-redistribution aspects.
- **Indirect taxes** are taxes on spending on a variety of goods and services in an economy—such as the excise duties on fuel, alcohol, and tobacco as well as sales (or value-added tax)—and often exclude health and education products on social grounds. In addition, taxes on gambling may also be considered to have a social aspect in deterring such activity, while fuel duties will have an environmental purpose in making fuel consumption and hence travel more expensive.

Taxes can be justified both in terms of raising revenues to finance expenditures and in terms of income and wealth redistribution policies. Economists typically consider four desirable attributes of a tax policy:

- **Simplicity**: This refers to ease of compliance by the taxpayer and enforcement by the revenue authorities. The final liability should be certain and not easily manipulated.
- **Efficiency**: Taxation should interfere as little as possible in the choices individuals make in the market place. Taxes affect behaviour and should, in general, discourage work and investment as little as possible. A major philosophical issue among economists is whether tax policy should deliberately deviate from efficiency to promote "good" economic activities, such as savings, and discourage harmful ones, such as tobacco consumption. Although most would accept a limited role in guiding consumer choices, some will question if policymakers are equipped to decide on such objectives and whether there will be unwanted ancillary effects, such as giving tax breaks for saving among people who already save and whose behaviour does not change.
- **Fairness**: This refers to the fact that people in similar situations should pay the same taxes ("horizontal equity") and that richer people should pay more taxes ("vertical equity"). Of course, the concept of fairness is really subjective. Still, most would agree that income tax rates should be progressive—that is, that households and corporations should pay proportionately more as their incomes rise. However, some people advocate "flat" tax rates, whereby all should pay the same proportion of taxable income.
- **Revenue sufficiency**: Although revenue sufficiency may seem obvious as a criterion for tax policy, there may be a conflict with fairness and efficiency. For example, one may believe that increasing income tax rates to reduce fiscal deficits reduces labour effort and that tax rate increases are thus an inefficient policy tool.

EXAMPLE 15

Some Issues with Tax Policy

1. *Incentives*. Some economists believe that income taxes reduce the incentive to work, save, and invest and that the overall tax burden has become excessive. These ideas are often associated with supply-side economics and the US economist Arthur Laffer. A variety of income tax cuts and simplifications have taken place in the United States since 1981, and although there is substantial controversy, some claim that work effort did rise (although tax cuts had little impact on savings). Similarly, some found that business investment did rise, while others claimed it was independent of such cuts.
2. *Fairness*. How do we judge the fairness of the tax system? One way is to calibrate the tax burden falling on different groups of people ranked by their income and to assess how changes in taxes affect these groups. Of course, this imposes huge data demands on investigators and must be considered incomplete. In the United States, it has been found that the federal system is indeed highly progressive. Many countries use such methods to analyze the impact of tax changes on different income groups when they announce their annual fiscal policy plans.
3. *Tax reform*. There is continuous debate on reforming tax policy. Should there be a flat-rate tax on labour income? Should all investment be immediately deducted for corporate taxes? Should more revenue be sourced from consumption taxes? Should taxes be indexed to inflation? Should dividends be taxed when profits have already been subject to tax? Should estates be taxed at all? Many of these issues are raised in the context of their impact on economic growth.

EXAMPLE 16

Fiscal Tools

1. Which of the following is *not* a tool of fiscal policy?
   - **A** A rise in social transfer payments
   - **B** The purchase of new equipment for the armed forces
   - **C** An increase in deposit requirements for the buying of houses
2. Which of the following is not an indirect tax?
   - **A** Excise duty
   - **B** Value-added Tax
   - **C** Employment taxes
3. Which of the following statements is *most* accurate?
   - **A** Direct taxes are useful for discouraging alcohol consumption.
   - **B** Because indirect taxes cannot be changed quickly, they are of no use in fiscal policy.
   - **C** Government capital spending decisions are slow to plan, implement, and execute and hence are of little use for short-term economic stabilization.

Solution to 1:

C is correct. Rises in deposit requirements for house purchases are intended to reduce the demand for credit for house purchases and hence would be considered a tool of monetary policy. This is a policy used actively in several countries, and is under consideration by regulators in other countries to constrain house price inflation.

Solution to 2:

C is correct. Both excise duty and VAT are applied to prices, whereas taxes on employment apply to labour income and hence are not indirect taxes.

Solution to 3:

C is correct. Capital spending is much slower to implement than changes in indirect taxes; and indirect taxes affect alcohol consumption more directly than direct taxes.

12.1 The Advantages and Disadvantages of Using the Different Tools of Fiscal Policy

The different tools used to expedite fiscal policy as a means to try to put or keep an economy on a path of positive, stable growth with low inflation have both advantages and disadvantages:

Advantages:

- Indirect taxes can be adjusted almost immediately after they are announced and can influence spending behaviour instantly and generate revenue for the government at little or no cost to the government.
- Social policies, such as discouraging alcohol or tobacco use, can be adjusted almost instantly by raising such taxes.

Disadvantages:

- Direct taxes are more difficult to change without considerable notice, often many months, because payroll computer systems will have to be adjusted (although the announcement itself may well have a powerful effect on spending behaviour more immediately). The same may be said for welfare and other social transfers.
- Capital spending plans take longer to formulate and implement, typically over a period of years. For example, building a road or hospital requires detailed planning, legal permissions, and implementation. This is often a valid criticism of an active fiscal policy and was widely heard during the US fiscal stimulus in 2009–2010. On the other hand, such policies add to the productive potential of an economy, unlike a change in personal or indirect taxes. Of course, the slower the impact of a fiscal change, the more likely other exogenous changes will already be influencing the economy before the fiscal change kicks in.

The above-mentioned tools may also have expectational effects at least as powerful as the direct effects. The announcement of future income tax rises a year ahead could potentially lead to reduced consumption immediately. Such delayed tax rises were a feature of UK fiscal policy of 2009–2010; however, the evidence is anecdotal because spending behaviour changed little until the delayed tax changes actually came into force.

We may also consider the relative potency of the different fiscal tools. Direct government spending has a far bigger impact on aggregate spending and output than income tax cuts or transfer increases; however, if the latter are directed at the poorest in society (basically, those who spend all their income), then this will give a relatively strong boost. Further discussion and examples of these comparisons are given in section 14 below on the interaction between monetary and fiscal policy.

12.2 Modeling the Impact of Taxes and Government Spending: The Fiscal Multiplier

The conventional macroeconomic model has government spending, G, adding directly to aggregate demand, AD, and reducing it via taxes, T; these comprise both indirect taxes on expenditures and direct taxes on factor incomes. Further government spending is increased via the payment of transfer benefits, B, such as social security payments. Hence, the net impact of the government sector on aggregate demand is:

$$G - T + B = \text{Budget surplus OR deficit} \tag{5}$$

Net taxes (NT; taxes less transfers) reduce disposable income (YD) available to individuals relative to national income or output (Y) as follows:

$$YD = Y - NT = (1 - t)\ Y \tag{6}$$

where t is the **net tax rate**. Net taxes are often assumed to be proportional to national income, Y, and hence total tax revenue from net taxes is tY. If t = 20% or 0.2, then for every $1 rise in national income, net tax revenue will rise by 20 cents and household disposable income will rise by 80 cents.

The **fiscal multiplier** is important in macroeconomics because it tells us how much output changes as exogenous changes occur in government spending or taxation. The recipients of the increase in government spending will typically save a proportion $1 - c$ of each additional dollar of disposable income, where c is the **marginal propensity to consume** (MPC) this additional income. Ignoring income taxes, we can see that $\$c$ will, in turn, be spent by these recipients on more goods and services. The recipients of this $\$c$ will themselves spend a proportion c of this additional income (i.e., $\$c \times c$, or c-squared). This process continues with income and spending growing at a constant rate of c as it passes from hand to hand through the economy. This is the familiar geometric progression with constant factor c, where $0 < c < 1$. The sum of this geometric series is $1/(1 - c)$.

We define s as the **marginal propensity to save** (MPS), the amount saved out of an additional dollar of disposable income. Because $c + s = 1$, hence $s = 1 - c$.

Exhibit 21 Disposable Income, Saving, and the MPC

Income	Income tax	Disposable income	Consumption	Saving
$100	$20	$80	$72	$8

In Exhibit 21, the MPC out of disposable income is 90% or 0.9 (72/80). The MPS is therefore 1 – 0.9 or 0.1

For every dollar of new (additional) spending, total incomes and spending rises by $\$1/(1 - c)$. And because $0 < c < 1$, this must be > 1; this is the multiplier. If $c = 0.9$ (or individuals spend 90 percent of additions to income), then the multiplier = $1/(1 - 0.9) = 10$.

A formal definition of the multiplier would be the ratio of the change in equilibrium output to the change in autonomous spending that caused the change. This is a monetary measure, but because prices are assumed to be constant in this analysis, real and monetary amounts are identical. Given that fiscal policy is about changes in government spending, G, net taxes, NT, and tax rates, t, we can see that the multiplier is an important tool for calibrating the possible impact of policy changes on output. How can we introduce tax changes into the multiplier concept? We do this by introducing the idea of disposable income, YD, defined as income less income taxes net of transfers, $Y - NT$.

Households spend a proportion c of disposable income, YD, that is, cYD or $c(Y - NT)$ or $c(1 - t)Y$. The **marginal propensity to consume** in the presence of taxes is then $c(1 - t)$. If the government increases spending, say on road building, by an amount, G, then disposable income rises by $(1 - t)G$ and consumer spending by $c(1 - t)G$. Provided there are unused sources of capital and labour in the economy, this leads to a rise in aggregate demand and output; the recipients of this extra consumption spending will have $(1 - t)c(1 - t)G$ extra disposable income available and will spend c of it. This cumulative extra spending and income will continue to spread through the economy at a decreasing rate as $0 < c(1 - t) < 1$. The overall final impact on aggregate demand and output will effectively be the sum of this decreasing geometric series with common ratio $c(1 - t)$, and this sums to $1/[1 - c(1 - t)]$. This is known as the **fiscal multiplier** and is very relevant to studies of fiscal policy as changes in G or tax rates will affect output in an economy through the value of the multiplier.

For example, if the tax rate is 20 percent, or 0.2, and the marginal propensity to spend is 90 percent, or 0.9, then the fiscal multiplier will be: $1/[1 - 0.9(1 - 0.2)]$ or $1/0.28 = 3.57$. In other words, if the government raises G by \$1 billion, total incomes and spending rise by \$3.57 billion.

Discretionary fiscal policy (see below) will involve changes in these variables with a view to influencing Y.

12.3 The Balanced Budget Multiplier

If a government increases G by the same amount as it raises taxes, the aggregate output actually rises. Why is this?

It is because the marginal propensity to spend out of disposable income is less than 1, and hence for every dollar less in YD, spending only falls \$$c$. Hence, aggregate spending falls less than the tax rise by a factor of c. A balanced budget leads to a rise in output, which in turn leads to further rises in output and incomes via the multiplier effect.

Suppose an economy has an equilibrium output or income level of \$1,000 consisting of \$900 of consumption and \$100 of investment spending, which is fixed and not related to income. If government spending is set at \$200, financed by a tax rate of 20 percent (giving tax revenue of \$200), what will happen to output? First, additional government spending of \$200 will raise output by that amount; but will taxes of \$200 reduce output by a similar amount? Not if the MPC is less than 1; suppose it is 0.9, and hence spending will only fall by 90 percent of \$200, or \$180. The initial impact of the balanced fiscal package on aggregate demand will be to raise it by \$200–\$180 = \$20. This additional output will, in turn, lead to further increases in income and output through the multiplier effect.

Even though the above policy involved a combination of government spending and tax increases that initially left the government's budget deficit/surplus unchanged, the induced rise in output will lead to further tax revenue increases and a further change in the budget position. Could the government adjust the initial change in spending to offset exactly the eventual total change in tax revenues? The answer is "yes," and we can ask what will be the effect on output of this genuinely balanced budget change? This balanced budget multiplier always takes the value unity.

EXAMPLE 17

Government Debt, Deficits, and Ricardo

The total stock of government debt is the outstanding stock of IOUs issued by a government and not yet repaid. They are issued when the government has insufficient tax revenues to meet expenditures and has to borrow from the public. The size of the outstanding debt equals the cumulative quantity of net borrowing it has done, and the fiscal or budget deficit is added in the current period to the outstanding stock of debt. If the outstanding stock of debt falls, we have a negative deficit or a surplus.

If a government reduces taxation by $10 billion one year and replaces that revenue with borrowing of $10 billion from the public, will it have any real impact on the economy? The important issue here is how people perceive that action: Do they recognize what will happen over time as interest and bond principal have to be repaid out of future taxes? If so, they may think of the bond finance as equivalent to delayed taxation finance; thus, the reduction in current taxation will have no impact on spending because individuals save more in anticipation of higher future taxes to repay the bond. This is called **Ricardian equivalence** after the economist David Ricardo. If people do not correctly anticipate all the future taxes required to repay the additional government debt, then they feel wealthier when the debt is issued and may increase their spending, adding to aggregate demand.

Whether Ricardian equivalence holds in practice is ultimately an empirical issue and is difficult to calibrate conclusively given the number of things that are changing at any time in a modern economy.

13 FISCAL POLICY IMPLEMENTATION

r explain the implementation of fiscal policy and difficulties of implementation;

s determine whether a fiscal policy is expansionary or contractionary;

In the following, we discuss major issues in fiscal policy implementation.

13.1 Deficits and the Fiscal Stance

An important question is the extent to which the budget is a useful measure of the government's fiscal stance. Does the size of the deficit actually indicate whether fiscal policy is **expansionary** or **contractionary**? Clearly, such a question is important for economic policymakers insofar as the deficit can change for reasons unrelated to actual fiscal policy changes. For example, the **automatic stabilizers** mentioned earlier will lead to changes in the budget deficit unrelated to fiscal policy changes; a recession will cause tax revenues to fall and the budget deficit to rise. An observer may conclude that fiscal policy has been loosened and is expansionary and that no further government action is required.

To this end, economists often look at the **structural (or cyclically adjusted) budget deficit** as an indicator of the fiscal stance. This is defined as the deficit that would exist *if the economy was at full employment (or full potential output)*. Hence, if we consider a period of relatively high unemployment, such as 2009–2010 with around 9–10 percent of the workforce out of work in the United States and Europe, then the budget deficits in those countries would be expected to be reduced substantially if the economies returned to full employment. At this level, tax revenues would be higher and social transfers lower. Recent data for major countries are given in Exhibit 22, where negative numbers refer to deficits and positive numbers are surpluses.

Exhibit 22 General Government Cyclically Adjusted Balances as Percent of GDP

	1995	2000	2005	2008	2010	2015
Australia	−3.1	0.9	2.0	−0.4	−3.8	−0.1
Germany	−9.5	0.9	−2.6	−0.8	−3.3	0.7
Japan	−4.6	−6.4	−4.1	−4.0	−8.2	−3.6
United Kingdom	−5.6	0.8	−4.5	−5.6	−7.6	−4.3
United States	−2.9	−0.4	−5.4	−7.1	−10.0	−3.5
OECD	−4.6	−1.2	−3.6	−4.5	−6.9	−2.0

Source: OECD Economic Outlook, Volume 2018 Issue 1.

A further reason why actual government deficits may *not* be a good measure of fiscal stance is the distinction between real and nominal interest rates and the role of inflation adjustment when applied to budget deficits. Although national economic statistics treat the cash interest payments on debt as government expenditure it makes more sense to consider only the inflation-adjusted (or real) interest payments because the real value of the outstanding debt is being eroded by inflation. Automatic stabilizers—such as income tax, VAT, and social benefits—are important because as output and employment fall and reduce tax revenues, so *net* tax revenues also fall as unemployment benefits rise. This acts as a fiscal stimulus and serves to reduce the size of the multiplier, dampening the output response of whatever caused the fall in output in the first place. By their very nature, automatic stabilizers do not require policy changes; no policymaker has to decide that an economic shock has occurred and how to respond. Hence, the responsiveness of the economy to shocks is automatically reduced, as are movements in employment and output.

In addition to these automatic adjustments, governments also use discretionary fiscal adjustments to influence aggregate demand. These will involve tax changes and/or spending cuts or increases usually with the aim of stabilizing the economy. A natural question is why fiscal policy cannot stabilize aggregate demand completely, hence ensuring full employment at all times.

13.2 Difficulties in Executing Fiscal Policy

Fiscal policy cannot stabilize aggregate demand completely because the difficulties in executing fiscal policy cannot be completely overcome.

First, the policymaker does not have complete information on how the economy functions. It may take several months for policymakers to realize that an economy is slowing, because data appear with a considerable time lag and even then are subject to substantial revision. This is often called the **recognition lag** and has been likened to the problem of driving with the rear view mirror. Then, when policy changes are finally decided on, they may take many months to implement. This is the **action lag**. If a government decides to raise spending on capital projects to increase employment and incomes, for example, these may take many months to plan and put into action. Finally, the result of these actions on the economy will take additional time to become evident; this is the **impact lag**. These types of policy lags also occur in the case of discretionary monetary policy.

A second aspect of time in this process is the uncertainty of where the economy is heading independently of these policy changes. For example, a stimulus may occur simultaneously with a surprise rise in investment spending or in the demand for a country's exports just as discretionary government spending starts to rise. Macroeconomic forecasting models do not generally have a good track record for accuracy and hence cannot be relied on to aid the policy-making process in this context. In addition, when discretionary fiscal adjustments are announced (or are already underway), private sector behaviour may well change leading to rises in consumption or investment, both of which will reinforce the effects of a rise in government expenditure. Again, this will make it difficult to calibrate the required fiscal adjustment to secure full employment.

There are wider macroeconomic issues also involved here.

- If the government is concerned with both unemployment *and* inflation in an economy, then raising aggregate demand toward the full employment level may also lead to a tightening labour market and rising wages and prices. The policymaker may be reluctant to further fine tune fiscal policy in an uncertain world because it might induce inflation.
- If the budget deficit is already large relative to GDP and further fiscal stimulus is required, then the necessary increase in the deficit may be considered unacceptable by the financial markets when government funding is raised, leading to higher interest rates on government debt and political pressure to tackle the deficit.
- Of course, all this presupposes that we know the level of full employment, which is difficult to measure accurately. Fiscal expansion raises demand, but what if we are already at full employment, which will be changing as productive capacity changes and workers' willingness to work at various wage levels changes?
- If unused resources reflect a low supply of labour or other factors rather than a shortage of demand, then discretionary fiscal policy will not add to demand and will be ineffective, raising the risk of inflationary pressures in the economy.
- The issue of crowding out may occur: If the government borrows from a limited pool of savings, the competition for funds with the private sector may crowd out private firms with subsequent less investing and economic growth. In addition, the cost of borrowing may rise, leading to the cancellation of potentially profitable opportunities. This concept is the subject of continuing empirical debate and investigation.

EXAMPLE 18

Evaluating Fiscal Policy

1 Which of the following statements is *least* accurate?

A The economic data available to policymakers have a considerable time lag.

B Economic models always offer an unambiguous guide to the future path of the economy.

C Surprise changes in exogenous economic variables make it difficult to use fiscal policy as a stabilization tool.

2 Which of the following statements is *least* accurate?

A Discretionary fiscal changes are aimed at stabilizing an economy.

B In the context of implementing fiscal policy, the recognition lag is often referred to as "driving in the rear view mirror."

C Automatic fiscal stabilizers include new plans for additional road building by the government.

3 Which of the following statements regarding a fiscal stimulus is *most* accurate?

A Accommodative monetary policy reduces the impact of a fiscal stimulus.

B Different statistical models will predict different impacts for a fiscal stimulus.

C It is always possible to predict precisely the impact of a fiscal stimulus on employment.

4 Which of the following statements is *most* accurate?

A An increase in the budget deficit is always expansionary.

B An increase in government spending is always expansionary.

C The structural deficit is always larger than the deficit below full employment.

5 Crowding out refers to a:

A fall in interest rates that reduces private investment.

B rise in private investment that reduces private consumption.

C rise in government borrowing that reduces the ability of the private sector to access investment funds.

6 A contractionary fiscal policy will always involve which of the following?

A A balanced budget

B A reduction in government spending

C A fall in the budget deficit or rise in the surplus

7 Which one of the following statements is *most* accurate?

A Ricardian equivalence refers to individuals having no idea of future tax liabilities.

B If there is high unemployment in an economy, then easy monetary and fiscal policies should lead to an expansion in aggregate demand.

C Governments do not allow political pressures to influence fiscal policies but do allow voters to affect monetary policies.

Solution to 1:

B is correct. Economic forecasts from models will always have an element of uncertainty attached to them and thus are not unambiguous or precise in their prescriptions. Once a fiscal policy decision has been made and implemented, unforeseen changes in other variables may affect the economy in ways that would lead to changes in the fiscal policy if we had perfect foresight. Note that it is true that official economic data may be available with substantial time lags, making fiscal judgements more difficult.

Solution to 2:

C is correct. New plans for road building are discretionary and not automatic.

Solution to 3:

B is correct. Different models embrace differing views on how the economy works, including differing views on the impact of fiscal stimuli.

Solution to 4:

A is correct. Note that increases in government spending may be accompanied by even bigger rises in tax receipts and hence may not be expansionary.

Solution to 5:

C is correct. A fall in interest rates is likely to lead to a rise in investment. Crowding out refers to government borrowing that reduces the ability of the private sector to invest.

Solution to 6:

C is correct. Note that a reduction in government spending could be accompanied by an even bigger fall in taxation, making it be expansionary.

Solution to 7:

B is correct. Note that governments often allow pressure groups to affect fiscal policy and that Ricardian equivalence involves individuals correctly anticipating future taxes, so A and C are not correct choices.

14

THE RELATIONSHIP BETWEEN MONETARY AND FISCAL POLICY

t explain the interaction of monetary and fiscal policy.

Both monetary and fiscal policies can be used to try and influence the macroeconomy. But the impact of monetary policy on aggregate demand may differ depending on the fiscal policy stance. Conversely, the impact of fiscal policy might vary under various alternative monetary policy conditions. Clearly, policymakers need to understand this interaction. For example, they need to consider the impact of changes to the budget when monetary policy is accommodative as opposed to when it is restrictive: Can we expect the same impact on aggregate demand in both situations?

Although both fiscal and monetary policy can alter aggregate demand, they do so through differing channels with differing impact on the composition of aggregate demand. The two policies are not interchangeable. Consider the following cases in which the assumption is made that *wages and prices are rigid*:

- *Easy fiscal policy/tight monetary policy*: If taxes are cut or government spending rises, the expansionary fiscal policy will lead to a rise in aggregate output. If this is accompanied by a reduction in money supply to offset the fiscal expansion, then interest rates will rise and have a negative effect on private sector demand. We have higher output and higher interest rates, and government spending will be a larger proportion of overall national income.
- *Tight fiscal policy/easy monetary policy*: If a fiscal contraction is accompanied by expansionary monetary policy and low interest rates, then the private sector will be stimulated and will rise as a share of GDP, while the public sector will shrink.

- *Easy monetary policy/easy fiscal policy*: If both fiscal and monetary policy are easy, then the joint impact will be highly expansionary—leading to a rise in aggregate demand, lower interest rates (at least if the monetary impact is larger), and growing private and public sectors.
- *Tight monetary policy/tight fiscal policy*: Interest rates rise (at least if the monetary impact on interest rates is larger) and reduce private demand. At the same time, higher taxes and falling government spending lead to a drop in aggregate demand from both public and private sectors.

14.1 Factors Influencing the Mix of Fiscal and Monetary Policy

Although governments are concerned about stabilizing the level of aggregate demand at close to the full employment level, they are also concerned with the growth of potential output. To this end, encouraging private investment will be important. It may best be achieved by accommodative monetary policy with low interest rates and a tight fiscal policy to ensure free resources for a growing private sector.

At other times, the lack of a good quality, trained workforce—or perhaps a modern capital infrastructure—will be seen as an impediment to growth; thus, an expansion in government spending in these areas may be seen as a high priority. If taxes are not raised to pay for this, then the fiscal stance will be expansionary. If a loose monetary policy is chosen to accompany this expansionary spending, then it is *possible* that inflation may be induced. Of course, it is an open question as to whether policymakers can judge the appropriate levels of interest rates or fiscal spending levels.

Clearly, the mix of policies will be heavily influenced by the political context. A weak government may raise spending to accommodate the demands of competing vested interests (e.g., subsidies to particular sectors, such as agriculture in the EC), and thus a restrictive monetary policy may be needed to hold back the possibly inflationary growth in aggregate demand through raised interest rates and less credit availability.

Both fiscal and monetary policies suffer from lack of precise knowledge of where the economy is today, because data appear initially subject to revision and with a time lag. However, fiscal policy suffers from two further issues with regard to its use in the short run.

As we saw earlier, it is difficult to implement quickly because spending on capital projects takes time to plan, procure, and put into practice. In addition, it is politically easier to loosen fiscal policy than to tighten it; in many cases, automatic stabilizers are the source of fiscal tightening, because tax rates are not changing and political opposition is muted. Similarly, the independence of many central banks means that decisions on raising interest rates are outside the hands of politicians and thus can be taken more easily.

The interaction between monetary and fiscal policies was also implicitly evident in our discussion of Ricardian equivalence because if tax cuts have no impact on private spending as individuals anticipate future higher taxes, then clearly this may lead policymakers to favour monetary tools.

Ultimately, the interaction of monetary and fiscal policies in practice is an empirical question, which we touched on earlier. In their detailed research paper using the IMF'S Global Integrated Monetary and Fiscal Model (IMF 2009), IMF researchers examined four forms of coordinated global fiscal loosening over a two-year period, which will be reversed gradually after the two years are completed. These are:

- an increase in social transfers to all households,
- a decrease in tax on labour income,
- a rise in government investment expenditure, and
- a rise in transfers to the poorest in society.

The two types of monetary policy responses considered are:

- no monetary accommodation, so rising aggregate demand leads to higher interest rates immediately; or
- interest rates are kept unchanged (accommodative policy) for the two years.

The following important policy conclusions from this study emphasize the role of policy interactions:

- *No monetary accommodation*: Government spending increases have a much bigger effect (six times bigger) on GDP than similar size social transfers because the latter are not considered permanent, although real interest rates rise as monetary authorities react to rises in aggregate demand and inflation. Targeted social transfers to the poorest citizens have double the effect of the non-targeted transfers, while labour tax reductions have a slightly bigger impact than the latter.
- *Monetary accommodation:* Except for the case of the cut in labour taxes, fiscal multipliers are now much larger than when there is no monetary accommodation. The cumulative multiplier (i.e., the cumulative effect on real GDP over the two years divided by the percentage of GDP, which is a fiscal stimulus) is now 3.9 for government expenditure compared to 1.6 with no monetary accommodation. The corresponding numbers for targeted social transfer payments are 0.5 without monetary accommodation and 1.7 with it. The larger multiplier effects with monetary accommodation result from rises in aggregate demand and inflation, leading to falls in real interest rates and additional private sector spending (e.g., on investment goods). Labour tax cuts are less positive.

14.2 Quantitative Easing and Policy Interaction

What about the scenario of zero interest rates and deflation? Fiscal stimulus should still raise demand and inflation, lowering real interest rates and stimulating private sector demand. We saw earlier that quantitative easing has been a feature of major economies during 2009–2010. This involves the purchase of government or private securities by the central bank from individuals, institutions, or banks and substituting central bank balances for those securities. The ultimate aim is that recipients will subsequently increase expenditures, lending or borrowing in the face of raised cash balances and lower interest rates.

If the central bank purchases government securities on a large scale, it is effectively funding the budget deficit and the independence of monetary policy is an illusion. This so-called "printing of money" is feared by many economists as the monetization of the government deficit. Note that it is unrelated to the conventional inflation target of central banks, such as the Bank of England. Some economists question whether an independent central bank should engage in such activity.

14.3 The Importance of Credibility and Commitment

The IMF model implies that if governments run persistently high budget deficits, real interest rates rise and crowd out private investment, reducing each country's productive potential. As individuals realize that deficits will persist, inflation expectations and longer-term interest rates rise: This reduces the effect of the stimulus by half.

Further, if there is a real lack of commitment to fiscal discipline over the longer term, (e.g., because of aging populations) and the ratio of government debt to GDP rose by 10 percentage points permanently in the United States alone, then world real interest rates would rise by 0.14 percent—leading to a 0.6 percent permanent fall in world GDP.

EXAMPLE 19

Interactions of Monetary and Fiscal Policy

1 In a world where Ricardian equivalence holds, governments would *most likely* prefer to use monetary rather than fiscal policy because under Ricardian equivalence:

- **A** real interest rates have a more powerful effect on the real economy.
- **B** the transmission mechanism of monetary policy is better understood.
- **C** the future impact of fiscal policy changes are fully discounted by economic agents.

2 If fiscal policy is easy and monetary policy tight, then:

- **A** interest rates would tend to fall, reinforcing the fiscal policy stance.
- **B** the government sector would tend to shrink as a proportion of total GDP.
- **C** the government sector would tend to expand as a proportion of total GDP.

3 Which of the following has the greatest impact on aggregate demand according to an IMF study? A 1 percent of GDP stimulus in:

- **A** government spending.
- **B** rise in transfer benefits.
- **C** cut in labour income tax across all income levels.

Solution to 1:

C is correct. If Ricardian equivalence holds, then economic agents anticipate that the consequence of any current tax cut will be future tax rises, which leads them to increase their saving in anticipation of this so that the tax cut has little effect on consumption and investment decisions. Governments would be forced to use monetary policy to affect the real economy on the assumption that money neutrality did not hold in the short term.

Solution to 2:

C is correct. With a tight monetary policy, real interest rates should rise and reduce private sector activity, which could be at least partially offset by an expansion in government activity via the loosening of fiscal policy. The net effect, however, would be an expansion in the size of the public sector relative to the private sector.

Solution to 3:

A is correct. The study clearly showed that direct spending by the government leads to a larger impact on GDP than changes in taxes or benefits.

SUMMARY

In this reading, we have sought to explain the practices of both monetary and fiscal policy. Both can have a significant impact on economic activity, and it is for this reason that financial analysts need to be aware of the tools of both monetary and fiscal policy, the goals of the monetary and fiscal authorities, and most important the monetary and fiscal policy transmission mechanisms.

- Governments can influence the performance of their economies by using combinations of monetary and fiscal policy. Monetary policy refers to central bank activities that are directed toward influencing the quantity of money and credit in an economy. By contrast, fiscal policy refers to the government's decisions about taxation and spending. The two sets of policies affect the economy via different mechanisms.
- Money fulfills three important functions: It acts as a medium of exchange, provides individuals with a way of storing wealth, and provides society with a convenient unit of account. Via the process of fractional reserve banking, the banking system can create money.
- The amount of wealth that the citizens of an economy choose to hold in the form of money—as opposed to, for example, bonds or equities—is known as the demand for money. There are three basic motives for holding money: transactions-related, precautionary, and speculative.
- The addition of 1 unit of additional reserves to a fractional reserve banking system can support an expansion of the money supply by an amount equal to the money multiplier, defined as 1/reserve requirement (stated as a decimal).
- The nominal rate of interest is comprised of three components: a real required rate of return, a component to compensate lenders for future inflation, and a risk premium to compensate lenders for uncertainty (e.g., about the future rate of inflation).
- Central banks take on multiple roles in modern economies. They are usually the monopoly supplier of their currency, the lender of last resort to the banking sector, the government's bank and bank of the banks, and they often supervise banks. Although they may express their objectives in different ways, the overarching objective of most central banks is price stability.
- For a central bank to be able to implement monetary policy objectively, it should have a degree of independence from government, be credible, and be transparent in its goals and objectives.
- The ultimate challenge for central banks as they try to manipulate the supply of money to influence the economy is that they cannot control the amount of money that households and corporations put in banks on deposit, nor can they easily control the willingness of banks to create money by expanding credit. Taken together, this also means that they cannot always control the money supply. Therefore, there are definite limits to the power of monetary policy.
- The concept of money neutrality is usually interpreted as meaning that money cannot influence the real economy in the long run. However, by the setting of its policy rate, a central bank hopes to influence the real economy via the policy rate's impact on other market interest rates, asset prices, the exchange rate, and the expectations of economic agents.

- Inflation targeting is the most common monetary policy—although exchange rate targeting is also used, particularly in developing economies. Quantitative easing attempts to spur aggregate demand by drastically increasing the money supply.
- Fiscal policy involves the use of government spending and revenue raising (taxation) to impact a number of aspects of the economy: the overall level of aggregate demand in an economy and hence the level of economic activity; the distribution of income and wealth among different segments of the population; and hence ultimately the allocation of resources between different sectors and economic agents.
- The tools that governments use in implementing fiscal policy are related to the way in which they raise revenue and the different forms of expenditure. Governments usually raise money via a combination of direct and indirect taxes. Government expenditure can be current on goods and services or can take the form of capital expenditure, for example, on infrastructure projects.
- As economic growth weakens, or when it is in recession, a government can enact an expansionary fiscal policy—for example, by raising expenditure without an offsetting increase in taxation. Conversely, by reducing expenditure and maintaining tax revenues, a contractionary policy might reduce economic activity. Fiscal policy can therefore play an important role in stabilizing an economy.
- Although both fiscal and monetary policy can alter aggregate demand, they work through different channels, the policies are therefore not interchangeable, and they conceivably can work against one another unless the government and central bank coordinate their objectives.

REFERENCES

Goodhart, Charles A.E. 1989. "The Conduct of Monetary Policy." *Economic Journal*, vol. 99, no. 396:293–346.

Gray, Simon, and Nick Talbot. 2006. *Monetary Operations*. London: Bank of England (http://www.bankofengland.co.uk/ccbs/monetary-operations).

IMF (International Monetary Fund). 2009. "The Case for Global Fiscal Stimulus" (March): http://www.imf.org/external/pubs/ft/spn/2009/spn0903.pdf.

Roger, Scott. 2010. "Inflation Targeting Turns 20." *Finance and Development*, vol. 47, no. 1:46–49 (March).

Truman, Edwin. 2003. *Inflation Targeting in the World Economy*. Washington, DC: Institute for International Economics.

PRACTICE PROBLEMS

1 As the reserve requirement increases, the money multiplier:

A increases.

B decreases.

C remains the same.

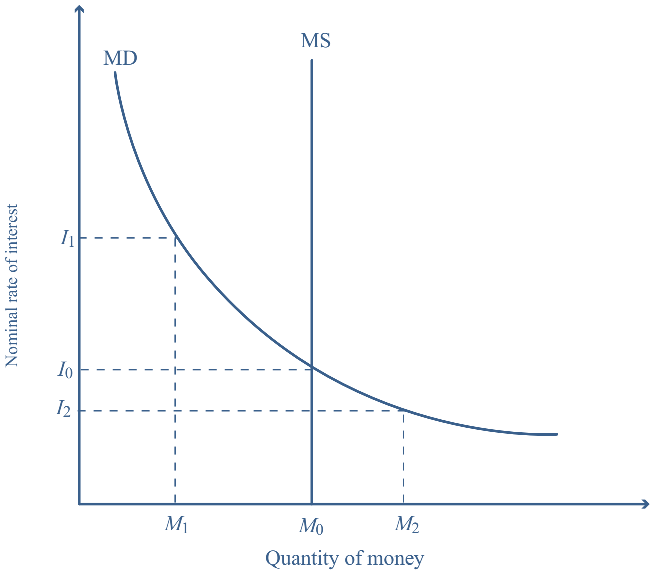
2 Which is the *most* accurate statement regarding the demand for money?

A Precautionary money demand is directly related to GDP.

B Transactions money demand is inversely related to returns on bonds.

C Speculative demand is inversely related to the perceived risk of other assets.

3 The following exhibit shows the supply and demand for money:



There is an excess supply of money when the nominal rate of interest is:

A I_0.

B I_1.

C I_2.

4 According to the theory of money neutrality, money supply growth does *not* affect variables such as real output and employment in:

A the long run.

B the short run.

C the long and short run.

5 Which of the following *best* describes a fundamental assumption when monetary policy is used to influence the economy?

A Financial markets are efficient.

B Money is not neutral in the short run.

C Official rates do not affect exchange rates.

6 Monetarists are *most likely* to believe:

A there is a causal relationship running from inflation to money.

B inflation can be affected by changing the money supply growth rate.

© 2011 CFA Institute. All rights reserved.

C rapid financial innovation in the market increases the effectiveness of monetary policy.

7 The proposition that the real interest rate is relatively stable is *most* closely associated with:

A the Fisher effect.

B money neutrality.

C the quantity theory of money.

8 Which of the following equations is a consequence of the Fisher effect?

A Nominal interest rate = Real interest rate + Expected rate of inflation.

B Real interest rate = Nominal interest rate + Expected rate of inflation.

C Nominal interest rate = Real interest rate + Market risk premium.

9 Central banks would typically be *most* concerned with costs of:

A low levels of inflation that are anticipated.

B moderate levels of inflation that are anticipated.

C moderate levels of inflation that are not anticipated.

10 Monetary policy is *least likely* to include:

A setting an inflation rate target.

B changing an official interest rate.

C enacting a transfer payment program.

11 Which role is a central bank *least likely* to assume?

A Lender of last resort.

B Sole supervisor of banks.

C Supplier of the currency.

12 Which is the *most* accurate statement regarding central banks and monetary policy?

A Central bank activities are typically intended to maintain price stability.

B Monetary policies work through the economy via four independent channels.

C Commercial and interbank interest rates move inversely to official interest rates.

13 When a central bank announces a decrease in its official policy rate, the desired impact is an increase in:

A investment.

B interbank borrowing rates.

C the national currency's value in exchange for other currencies.

14 Which action is a central bank *least likely* to take if it wants to encourage businesses and households to borrow for investment and consumption purposes?

A Sell long-dated government securities.

B Purchase long-dated government treasuries.

C Purchase mortgage bonds or other securities.

15 A central bank that decides the desired levels of interest rates and inflation and the horizon over which the inflation objective is to be achieved is *most* accurately described as being:

A target independent and operationally independent.

B target independent but not operationally independent.

C operationally independent but not target independent.

16 A country that maintains a target exchange rate is *most likely* to have which outcome when its inflation rate rises above the level of the inflation rate in the target country?

A An increase in short-term interest rates.

B An increase in the domestic money supply.

C An increase in its foreign currency reserves.

17 A central bank's repeated open market purchases of government bonds:

A decreases the money supply.

B is prohibited in most countries.

C is consistent with an expansionary monetary policy.

18 In theory, setting the policy rate equal to the neutral interest rate should promote:

A stable inflation.

B balanced budgets.

C greater employment.

19 A prolonged period of an official interest rate very close to zero without an increase in economic growth *most likely* suggests:

A quantitative easing must be limited to be successful.

B there may be limits to the effectiveness of monetary policy.

C targeting reserve levels is more important than targeting interest rates.

20 Raising the reserve requirement is *most likely* an example of which type of monetary policy?

A Neutral.

B Expansionary.

C Contractionary.

21 Which of the following is a limitation on the ability of central banks to stimulate growth in periods of deflation?

A Ricardian equivalence.

B The interaction of monetary and fiscal policy.

C The fact that interest rates cannot fall significantly below zero.

22 The *least likely* limitation to the effectiveness of monetary policy is that central banks cannot:

A accurately determine the neutral rate of interest.

B regulate the willingness of financial institutions to lend.

C control amounts that economic agents deposit into banks.

23 Which of the following is the *most likely* example of a tool of fiscal policy?

A Public financing of a power plant.

B Regulation of the payment system.

C Central bank's purchase of government bonds.

24 The *least likely* goal of a government's fiscal policy is to:

A redistribute income and wealth.

B influence aggregate national output.

C ensure the stability of the purchasing power of its currency.

25 Given an independent central bank, monetary policy actions are *more likely* than fiscal policy actions to be:

A implementable quickly.

B effective when a specific group is targeted.

C effective when combating a deflationary economy.

26 Which statement regarding fiscal policy is *most* accurate?

A To raise business capital spending, personal income taxes should be reduced.

B Cyclically adjusted budget deficits are appropriate indicators of fiscal policy.

C An increase in the budget surplus is associated with expansionary fiscal policy.

27 The *least likely* explanation for why fiscal policy cannot stabilize aggregate demand completely is that:

A private sector behavior changes over time.

B policy changes are implemented very quickly.

C fiscal policy focuses more on inflation than on unemployment.

28 Which of the following *best* represents a contractionary fiscal policy?

A Public spending on a high-speed railway.

B A temporary suspension of payroll taxes.

C A freeze in discretionary government spending.

29 A “pay-as-you-go” rule, which requires that any tax cut or increase in entitlement spending be offset by an increase in other taxes or reduction in other entitlement spending, is an example of which fiscal policy stance?

A Neutral.

B Expansionary.

C Contractionary.

30 Quantitative easing, the purchase of government or private securities by the central banks from individuals and/or institutions, is an example of which monetary policy stance?

A Neutral.

B Expansionary.

C Contractionary.

31 The *most likely* argument against high national debt levels is that:

A the debt is owed internally to fellow citizens.

B they create disincentives for economic activity.

C they may finance investment in physical and human capital.

32 Which statement regarding fiscal deficits is *most* accurate?

A Higher government spending may lead to higher interest rates and lower private sector investing.

B Central bank actions that grow the money supply to address deflationary conditions decrease fiscal deficits.

C According to the Ricardian equivalence, deficits have a multiplicative effect on consumer spending.

33 Which policy alternative is *most likely* to be effective for growing both the public and private sectors?

A Easy fiscal/easy monetary policy.

B Easy fiscal/tight monetary policy.

C Tight fiscal/tight monetary policy.

SOLUTIONS

1 B is correct. There is an inverse relationship between the money multiplier and the reserve requirement. The money multiplier is equal to 1 divided by the reserve requirement.

2 A is correct. Precautionary money demand is directly related to GDP. Precautionary money balances are held to provide a buffer against unforeseen events that might require money. Precautionary balances tend to rise with the volume and value of transactions in the economy, and therefore rise with GDP.

3 B is correct. When the interest rate on bonds is I_1 there is an excess supply of money (equal to $M_0 - M_1 > 0$). Economic agents would seek to buy bonds with their excess money balances, which would force the price of bonds up and the interest rate down to I_0.

4 A is correct. According to the theory of money neutrality, an increase in the money supply ultimately leads to an increase in the price level and leaves real variables unaffected in the long run.

5 B is correct. If money were neutral in the short run, monetary policy would not be effective in influencing the economy.

6 B is correct. By definition, monetarists believe prices may be controlled by manipulating the money supply.

7 A is correct. The Fisher effect is based on the idea that the real interest rate is relatively stable. Changes in the nominal interest rate result from changes in expected inflation.

8 A is correct. The Fisher effect implies that changes in the nominal interest rate reflect changes in expected inflation, which is consistent with Nominal interest rate = Real interest rate + Expected rate of inflation.

9 C is correct. Low levels of inflation has higher economic costs than moderate levels, all else equal; unanticipated inflation has greater costs than anticipated inflation.

10 C is correct. Transfer payment programs represent fiscal, not monetary policy.

11 B is correct. The supervision of banks is not a role that all central banks assume. When it is a central bank's role, responsibility may be shared with one or more entities.

12 A is correct. Central bank activities are typically intended to maintain price stability. Concerning choice B, note that the transmission channels of monetary policy are not independent.

13 A is correct. Investment is expected to move inversely with the official policy rate.

14 A is correct. Such action would tend to constrict the money supply and increase interest rates, all else equal.

15 A is correct. The central bank described is target independent because it set its own targets (e.g., the target inflation rate) and operationally independent because it decides how to achieve its targets (e.g., the time horizon).

16 A is correct. Interest rates are expected to rise to protect the exchange rate target.

17 C is correct. The purchase of government bonds via open market operations increases banking reserves and the money supply; it is consistent with an expansionary monetary policy.

18 A is correct. The neutral rate of interest is that rate of interest that neither stimulates nor slows down the underlying economy. The neutral rate should be consistent with stable long-run inflation.

19 B is correct. A central bank would decrease an official interest rate to stimulate the economy. The setting in which an official interest rate is lowered to zero (or even slightly below zero) without stimulating economic growth suggests that there are limits to monetary policy.

20 C is correct. Raising reserve requirements should slow money supply growth.

21 C is correct. Deflation poses a challenge to conventional monetary policy because once the central bank has cut nominal interest rates to zero (or slightly less than zero) to stimulate the economy, they cannot cut them further.

22 A is correct. The inability to determine exactly the neutral rate of interest does not necessarily limit the power of monetary policy.

23 A is correct. Public financing of a power plant could be described as a fiscal policy tool to stimulate investment.

24 C is correct. Ensuring stable purchasing power is a goal of monetary rather than fiscal policy. Fiscal policy involves the use of government spending and tax revenue to affect the overall level of aggregate demand in an economy and hence the level of economic activity.

25 A is correct. Monetary actions may face fewer delays to taking action than fiscal policy, especially when the central bank is independent.

26 B is correct. Cyclically adjusted budget deficits are appropriate indicators of fiscal policy. These are defined as the deficit that would exist if the economy was at full employment (or full potential output).

27 B is correct. Fiscal policy is subject to recognition, action, and impact lags.

28 C is correct. A freeze in discretionary government spending is an example of a contractionary fiscal policy.

29 A is correct. A "pay-as-you-go" rule is a neutral policy because any increases in spending or reductions in revenues would be offset. Accordingly, there would be no net impact on the budget deficit/surplus.

30 B is correct. Quantitative easing is an example of an expansionary monetary policy stance. It attempts to spur aggregate demand by drastically increasing the money supply.

31 B is correct. The belief is that high levels of debt to GDP may lead to higher future tax rates which may lead to disincentives to economic activity.

32 A is correct. Government borrowing may compete with private sector borrowing for investment purposes.

33 A is correct. If both fiscal and monetary policies are "easy," then the joint impact will be highly expansionary, leading to a rise in aggregate demand, low interest rates, and growing private and public sectors.

READING

13

International Trade and Capital Flows

by Usha Nair-Reichert, PhD, and Daniel Robert Witschi, PhD, CFA

Usha Nair-Reichert, PhD, is at Georgia Institute of Technology (USA). Daniel Robert Witschi, PhD, CFA (Switzerland).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** compare gross domestic product and gross national product;
☐	**b.** describe benefits and costs of international trade;
☐	**c.** contrast comparative advantage and absolute advantage;
☐	**d.** compare the Ricardian and Heckscher–Ohlin models of trade and the source(s) of comparative advantage in each model;
☐	**e.** compare types of trade and capital restrictions and their economic implications;
☐	**f.** explain motivations for and advantages of trading blocs, common markets, and economic unions;
☐	**g.** describe common objectives of capital restrictions imposed by governments;
☐	**h.** describe the balance of payments accounts including their components;
☐	**i.** explain how decisions by consumers, firms, and governments affect the balance of payments;
☐	**j.** describe functions and objectives of the international organizations that facilitate trade, including the World Bank, the International Monetary Fund, and the World Trade Organization.

INTRODUCTION & INTERNATIONAL TRADE-BASIC TERMINOLOGY

1

a compare gross domestic product and gross national product;

© 2019 CFA Institute. All rights reserved.

Global investors must address two fundamentally interrelated questions: where to invest and in what asset classes? Some countries may be attractive from an equity perspective because of their strong economic growth and the profitability of particular domestic sectors or industries. Other countries may be attractive from a fixed income perspective because of their interest rate environment and price stability. To identify markets that are expected to provide attractive investment opportunities, investors must analyze cross-country differences in such factors as expected GDP growth rates, monetary and fiscal policies, trade policies, and competitiveness. From a longer term perspective investors also need to consider such factors as a country's stage of economic and financial market development, demographics, quality and quantity of physical and human capital (accumulated education and training of workers), and its area(s) of comparative advantage.[1]

This reading provides a framework for analyzing a country's trade and capital flows and their economic implications. International trade can facilitate economic growth by increasing the efficiency of resource allocation, providing access to larger capital and product markets, and facilitating specialization based on comparative advantage. The flow of financial capital (funds available for investment) between countries with excess savings and those where financial capital is scarce can increase liquidity, raise output, and lower the cost of capital. From an investment perspective, it is important to understand the complex and dynamic nature of international trade and capital flows because investment opportunities are increasingly exposed to the forces of global competition for markets, capital, and ideas.

This reading is organized as follows. Sections 1–3 define basic terminology used in the reading and describes patterns and trends in international trade and capital flows. It also discusses the benefits of international trade, distinguishes between absolute and comparative advantage, and explains two traditional models of comparative advantage. Sections 6–8 describe trade restrictions and their implications and discuss the motivation for, and advantages of, trade agreements. Sections 9–11 describe the balance of payments and Section 12 discusses the function and objectives of international organizations that facilitate trade. A summary of key points and practice problems conclude the reading.

1.1 International Trade

The following sections describe the role, importance, and possible benefits and costs of international trade. Before beginning those discussions, we define some basic terminology used in this area.

1.1.1 *Basic Terminology*

The aggregate output of a nation over a specified time period is usually measured as its gross domestic product or its gross national product. Gross domestic product (GDP) measures the market value of all final goods and services produced by factors of production (such as labor and capital) located within a country/economy during a given period of time, generally a year or a quarter. Gross national product (GNP), however, measures the market value of all final goods and services produced by factors of production (such as labor and capital) supplied by citizens of a country, regardless of whether such production takes place within the country or outside of the country. The difference between a country's GDP and its GNP is that GDP includes, and GNP excludes, the production of goods and services by foreigners within that country, whereas GNP includes, and GDP excludes, the production of goods and services by its citizens outside of the country. Countries that have large differences between GDP and GNP generally have a large number of citizens who work abroad (for example, Pakistan and Portugal), and/or pay more for the use of foreign-owned capital in domestic production than they earn on the capital they own abroad (for example, Brazil and Canada). Therefore, GDP is more widely used as a measure of economic activity occurring *within* the country, which, in turn, affects employment, growth, and the investment environment.

---

[1] Comparative advantage refers to a country's ability to produce a good at a relatively lower cost than other goods it produces, as compared with another country. It will be more precisely defined and illustrated in Sections 4 and 5.

Imports are goods and services that a domestic economy (i.e., households, firms, and government) purchases from other countries. For example, the US economy imports (purchases) cloth from India and wine from France. **Exports** are goods and services that a domestic economy sells to other countries. For example, South Africa exports (sells) diamonds to the Netherlands, and China exports clothing to the European Union. So how are services imported or exported? If a Greek shipping company transports the wine that the United States imports from France, the United States would classify the cost of shipping as an import of services from Greece and the wine would be classified as an import of goods from France. Similarly, when a British company provides insurance coverage to a South African diamond exporter, Britain would classify the cost of the insurance as an export of services to South Africa. Other examples of services exported/imported include engineering, consulting, and medical services.

The **terms of trade** are defined as the ratio of the price of exports to the price of imports, representing those prices by export and import price indexes, respectively. The terms of trade capture the relative cost of imports in terms of exports. If the prices of exports increase relative to the prices of imports, the terms of trade have improved because the country will be able to purchase more imports with the same amount of exports.[2] For example, when oil prices increased during 2007–2008, major oil exporting countries experienced an improvement in their terms of trade because they had to export less oil in order to purchase the same amount of imported goods. In contrast, if the price of exports decreases relative to the price of imports, the terms of trade have deteriorated because the country will be able to purchase fewer imports with the same amount of exports. Because each country exports and imports a large number of goods and services, the terms of trade of a country are usually measured as an index number (normalized to 100 in some base year) that represents a ratio of the average price of exported goods and services to the average price of imported goods and services. Exhibit 1 shows the terms of trade reported by the World bank, aggregated by their categories for region and income group. A value over (under) 100 indicates that the country, or group of countries, experienced better (worse) terms of trade relative to the base year of 2000.

Exhibit 1 Data on the Barter Terms of Trade for Industrial and Developing Countries (Unit Export Value/Unit Import Value)

	1990	1995	2000	2005	2010	2015
High Income	99.2	110.4	100.0	105.3	115.4	113.3
Low Income	125.9	118.2	100.0	98.0	117.2	113.8

(continued)

---

[2] Although the prices of imports and exports are each stated in currency units, the currency units cancel out when we take the ratio, so the terms of trade reflect the relative price of imports and exports in real (i.e., quantity) terms: units of imports per unit of exports. To see this, note that if one unit of imports costs P_M currency units and one unit of exports is priced at P_X currency units, then the country can buy P_X/P_M (= Terms of trade) units of imports for each unit of exports.

Exhibit 1 (Continued)

	1990	1995	2000	2005	2010	2015
Africa	113.3	108.1	100.0	108.9	132.4	121.9
Asia	115.5	114.7	100.0	99.2	109.6	105.9
Europe	109.1	105.7	100.0	99.8	105.0	101.6
Western Hemisphere	98.7	104.7	100.0	104.9	115.9	107.5
Middle East	92.6	94.9	100.0	125.7	165.2	164.1

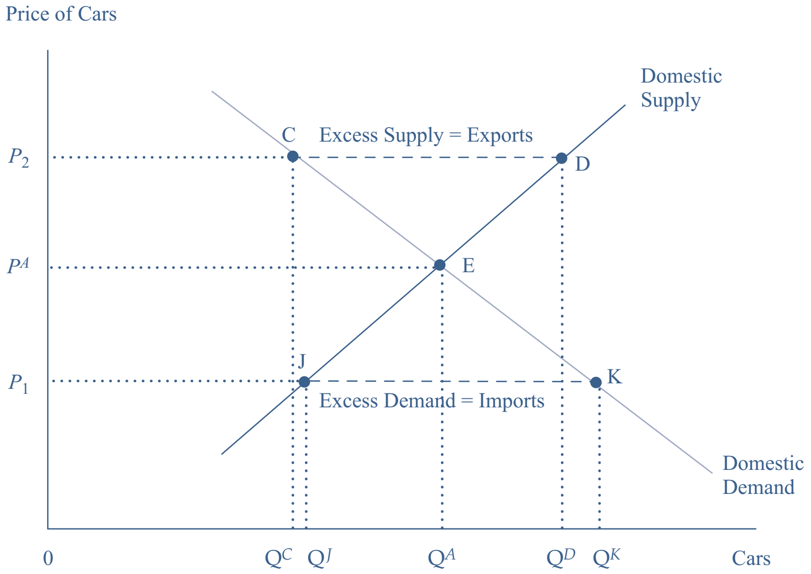
As an example, Exhibit 1 indicates that from 2000 to 2015 both low-income and high-income countries generally experienced an increase in their terms of trade; however, low-income countries experienced more volatility over the period. Looking at the disaggregated data indicates that Africa also experienced improvement in its terms of trade, albeit with volatility, during this period. Countries in Asia and the Western Hemisphere experienced some increase in their terms of trade while those in Europe saw little change. In contrast, countries in the Middle East (which benefited from rising prices of their petroleum exports) experienced a substantial increase in their terms of trade.

Net exports is the difference between the value of a country's exports and the value of its imports (i.e., value of exports minus imports). If the value of exports equals the value of imports, then trade is balanced. If the value of exports is greater (less) than the value of imports, then there is a **trade surplus (deficit)**. When a country has a trade surplus, it lends to foreigners or buys assets from foreigners reflecting the financing needed by foreigners running trade deficits with that country. Similarly, when a country has a trade deficit, it has to borrow from foreigners or sell some of its assets to foreigners. Sections 9–11 on the balance of payments explain these relationships more fully.

Autarky is a state in which a country does not trade with other countries. This means that all goods and services are produced and consumed domestically. The price of a good or service in such an economy is called its **autarkic price**. An autarkic economy is also known as a **closed economy** because it does not trade with other countries. An **open economy**, in contrast, is an economy that trades with other countries. If there are no restrictions on trade, then members of an open economy can buy and sell goods and services at the price prevailing in the world market, the **world price**. An open economy can provide domestic households with a larger variety of goods and services, give domestic companies access to global markets and customers, and offer goods and services that are more competitively priced. In addition, it can offer domestic investors access to foreign capital markets, foreign assets, and greater investment opportunities. For capital intensive industries, such as automobiles and aircraft, manufacturers can take advantage of economies of scale because they have access to a much larger market. **Free trade** occurs when there are no government restrictions on a country's ability to trade. Under free trade, global aggregate demand and supply determine the equilibrium quantity and price of imports and exports. Government policies that impose restrictions on trade, such as tariffs and quotas (discussed later in the reading), are known as **trade protection** and prevent market forces (demand and supply) from determining the equilibrium price and quantity for imports and exports. According to Deardorff, *globalization* refers to the "increasing worldwide integration of markets for goods, services, and capital that began to attract special attention in the late 1990s."[3] It also references "a variety of other changes that were perceived to occur at about the same time, such as an increased role for large corporations (multinational corporations) in the world economy and increased intervention into domestic policies and affairs by international institutions," such as the International Monetary Fund, the World Trade Organization, and the World Bank.

The levels of aggregate demand and supply and the quantities of imports and exports in an economy are related to the concepts of *excess demand* and *excess supply*. Exhibit 2 shows supply and demand curves for cars in the United Kingdom. E is the autarkic equilibrium at price P^A and quantity Q^A, with the quantity of cars demanded equaling the quantity supplied. Now, consider a situation in which the country opens up to trade and the world price is P_1. At this price, the quantity demanded domestically is Q^K while the quantity supplied is Q^J. Hence excess demand is Q^J Q^K. This quantity is satisfied by imports. For example, at a world price of $15,000, the quantity of cars demanded in the United Kingdom might be 2 million and UK production of cars only 1.5 million. As a result, the excess demand of 500,000 would be satisfied by imports. Returning to Exhibit 2, now consider a situation in which the world price is P_2. The quantity demanded is Q^C while the quantity supplied is Q^D. Hence, the domestic excess supply at world price P_2 is Q^C Q^D, which results in exports of Q^C Q^D.

Exhibit 2 Excess Demand, Excess Supply, Imports and Exports



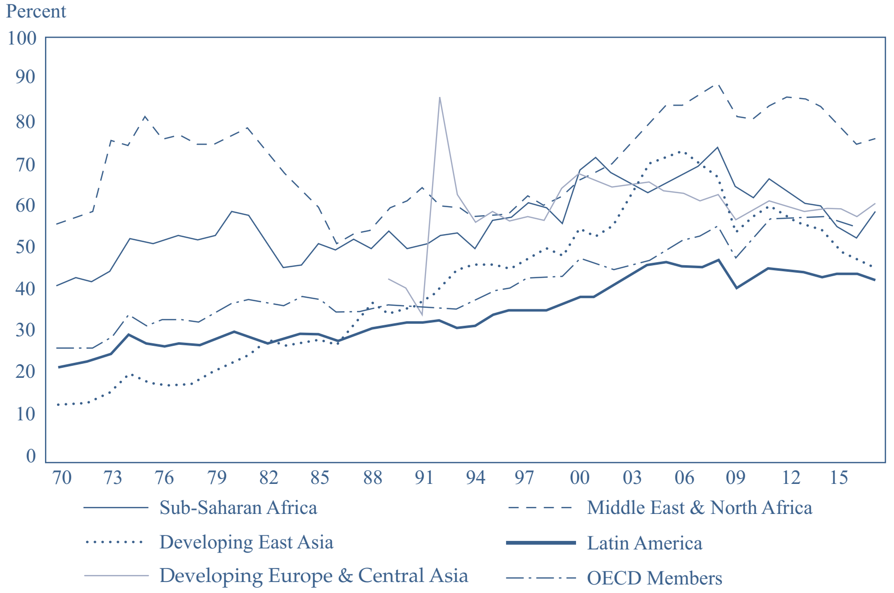
---

[3] Deardorff, Alan. "Deardorff's Glossary of International Economics" (www-personal.umich.edu/~alandear/glossary).

2 PATTERNS AND TRENDS IN INTERNATIONAL TRADE AND CAPITAL FLOWS

The importance of trade in absolute and relative terms (trade-to-GDP ratio) is illustrated in Exhibits 3 through 5. Exhibit 3 shows that trade as a percentage of regional GDP increased in all regions of the world during 1970–2006. Developing countries in Asia had the fastest growth in trade, increasing from less than 20 percent of GDP in 1970 to more than 70 percent of GDP in 2006.

Exhibit 3 Trade in Goods and Services (Percent of Regional GDP)



Note: Developing East Asia and Developing Europe & Central Asia exclude all World Bank designated "High Income" countries in these regions.
Source: World Development Indicators.

Exhibit 4 indicates that trade as a percentage of GDP and the GDP growth rate increased in most regions of the world during 1990–2009. However, data for 2010-2016 indicates a decline that, although consistent with the worldwide economic downturn, varied across country groups. High-income countries that are members of the Organisation for Economic Co-Operation and Development (OECD) experienced a growth rate of 2.4 percent during 2000–2006, but had a growth rate of only 1.1 percent in between 2010-2016. The corresponding numbers for growth in non-OECD high-income countries are 5.0 percent and 2.0 percent, respectively; for lower-middle-income countries, they are 7.7 percent and 3.3 percent, respectively. The 2009 World Development Report affirmed the link between trade and growth and noted evidence that all rich and emerging economies are oriented to being open to trade. More specifically, the report indicated:

> …When exports are concentrated in labor-intensive manufacturing, trade increases the wages for unskilled workers, benefiting poor people. It also encourages macroeconomic stability, again benefiting the poor, who are more likely to be hurt by inflation. And through innovation and factor accumulation, it enhances productivity and thus growth. There may be

> some empirical uncertainty about the strength of trade's relationship with growth. But essentially all rich and emerging economies have a strong trade orientation. (World Bank 2009)

Of course, trade is not the only factor that influences economic growth. Research has also identified such factors as the quality of institutions, infrastructure, and education; economic systems; the degree of development; and global market conditions (World Trade Organization 2008).

Exhibit 4 Trade Openness and GDP Growth

Country Group	Trade as Percent of GDP (averaged over the period) 1980–1989	1990–1999	2000–2009	2010–2016	Average GDP growth (%) 1980–1989	1990–1999	2000–2009	2010–2016
World	75.4	77.9	91.6	94.1	3.1%	2.5%	3.2%	2.1%
High income:								
All	40.7	42.8	53.6	61.3	3.1%	2.5%	1.5%	1.1%
OECD	36.5	39.1	48.7	56.0	3.1%	2.4%	1.3%	1.1%
Non-OECD	126.3	116.5	134.9	138.8	3.9%	3.4%	3.4%	2.0%
Low and middle income:								
All	30.2	42.3	56.3	52.7	3.4%	2.6%	5.5%	3.0%
Middle	30.0	42.2	56.2	52.6	3.4%	2.6%	5.5%	3.0%
Upper middle	28.9	41.5	55.7	51.5	2.1%	2.4%	5.5%	2.9%
Lower middle	33.7	44.7	57.9	56.3	6.0%	3.0%	5.5%	3.3%
Low	52.9	54.1	63.5	66.9	2.6%	1.9%	4.4%	2.3%

Note: Averages indicate the average of the annual data for the period covered.
Source: World Bank.

Exhibit 5 presents trade and foreign direct investment as a percentage of GDP for select countries for 1990-2017. **Foreign direct investment** (FDI) refers to direct investment by a firm in one country (the *source country*) in productive assets in a foreign country (the *host country*). When a firm engages in FDI, it becomes a **multinational corporation** (MNC) operating in more than one country or having subsidiary firms in more than one country. It is important to distinguish FDI from **foreign portfolio investment** (FPI), which refers to shorter-term investment by individuals, firms, and institutional investors (e.g., pension funds) in such foreign financial instruments as foreign stocks and foreign government bonds. Exhibit 5 shows that trade as a percentage of GDP for the world as a whole increased from 39 percent in 1990 to 57 percent in 2010. In Argentina, trade as a percentage of GDP increased from 15 percent in 1990 to 35 percent in 2010, while in India during this same period it increased from 16 percent to almost 50 percent. Among the more advanced economies, trade expanded sharply in Germany (from 46 percent to 87 percent between 1990 and 2017), but in the United States trade expanded more modestly (from 20 percent to 27 percent).

Exhibit 5 Increasing Global Interdependence
FDI and Trade as a percentage of GDP

Country	Type of Flow	1990	2000	2010	2017
World	Trade	38.9	51.3	56.9	56.2*
	FDI: Net Inflows	0.9	4.4	2.7	2.4
	FDI: Net Outflows	1.3	4.1	2.6	2.0
Argentina	Trade	15.0	22.6	35.0	25.0
	FDI: Net Inflows	1.3	3.7	2.7	1.9
	FDI: Net Outflows	0.0	0.3	0.2	0.2
Germany	Trade	46.0	61.4	79.3	86.9
	FDI: Net Inflows	0.2	12.7	2.5	2.1
	FDI: Net Outflows	1.4	5.0	4.3	3.4
India	Trade	15.7	27.2	49.7	40.6
	FDI: Net Inflows	0.1	0.8	1.7	1.5
	FDI: Net Outflows	0.0	0.1	1.0	0.4
United States	Trade	19.8	25.0	28.2	26.6*
	FDI: Net Inflows	0.8	3.4	1.7	1.8
	FDI: Net Outflows	1.0	1.8	2.3	2.2

* Trade figures for 2016.
Source: World Development Indicators.

The increasing importance of multinational corporations is also apparent in Exhibit 5. Net FDI inflows and outflows increased as a percentage of GDP between 1990 and 2000 for each of the countries shown. Trade between multinational firms and their subsidiaries (i.e., intra-firm trade) has become an important part of world trade. For example, 46 percent of US imports occur between related parties (Bernard, Jensen, Redding, and Schott 2010). Globalization of production has increased the productive efficiency of manufacturing firms because they are able to decompose their value chain into individual components or parts, and then outsource their production to different locations where these components can be produced most efficiently.[4] For example, in 2016 Apple's iPhone 6s was manufactured with components sourced from several locations around the world: the camera, display and storage were manufactured in Japan; the RAM and A9 processor were manufactured in Korea; the modem, battery, Wi-Fi module, radio frequency transceiver and chassis were manufactured in China; and much of the hardware and software was designed in the United States while the phone itself was assembled in China.[5] Foreign direct investment and outsourcing have increased business investment in these economies and provided smaller and less developed economies the opportunity to participate in international trade. For example, in 2016 Intel had 10 fabrication plants and 101 assembly and testing sites in 8 countries/regions. These trends indicate the increasing global interdependence of economies, although the degree of interdependence varies. Greater interdependence also means that economies are now more exposed to global competition. As a result, they must be more flexible in their production structure in order to respond effectively to changes in global demand and supply.

---

[4] Hill and Hult (2019) explains the idea of the firm as a value chain: "The operations of the firms can be thought of as a value chain composed of a series of distinct value creation activities including production, marketing and sales, materials management, R&D, human resources, information systems, and firm infrastructure." Production itself can be broken down into distinct components and each component outsourced separately.

[5] "Here's where all the components of your iPhone come from" Skye Gould and Antonio Villas-Boas Apr. 12, 2016 https://www.businessinsider.com/where-iphone-parts-come-from-2016-4

The complexity of trading relationships has also increased with the development of sophisticated global supply chains that include not only final goods but also intermediate goods and services. Increased global interdependence has changed the risk and return profiles of many economies. Economies that have greater international links are more exposed to, and affected by, economic downturns and crises occurring in other parts of the world. The contagion effect of the Asian financial crisis, which began in Thailand in July 1997, spread to many other markets, such as Indonesia, Malaysia, South Korea, Philippines, Hong Kong SAR, Singapore, and Taiwan Region. It even affected Brazil and Russia to some degree, although there is less clarity about the mechanisms by which the crisis spread beyond Asia. Among the outward symptoms of the crisis were exchange rate problems, such as currency speculation and large depreciation of currencies, capital flight, and financial and industrial sector bankruptcies. However, recovery was surprisingly swift and all these economies exhibited positive growth by the second quarter of 1999 (Gerber 2017).

BENEFITS AND COSTS OF INTERNATIONAL TRADE

3

b describe benefits and costs of international trade;

The preceding sections have described the growth of world trade and the increasing interdependence of national economies. Has trade been beneficial? The benefits and costs of international trade have been widely debated. The most compelling arguments supporting international trade are: countries gain from exchange and specialization, industries experience greater economies of scale, households and firms have greater product variety, competition is increased, and resources are allocated more efficiently.

Gains from exchange occur when trade enables each country to receive a higher price for its exports (and greater profit) and/or pay a lower price for imported goods instead of producing these goods domestically at a higher cost (i.e., less efficiently). This exchange, in turn, leads to a more efficient allocation of resources by increasing production of the export good and reducing production of the import good in each country (trading partner). This efficiency allows consumption of a larger bundle of goods, thus increasing overall welfare. The fact that trade increases overall welfare does not, of course, mean that every individual consumer and producer is better off. What it does mean is that the winners could, in theory, compensate the losers and still be better off.

Trade also leads to greater efficiency by fostering specialization based on comparative advantage. Traditional trade models, such as the Ricardian model and the Heckscher–Ohlin model, focus on specialization and trade according to comparative advantage arising from differences in technology and factor endowments, respectively. These models will be discussed in the next section.

Newer models of trade focus on the gains from trade that result from economies of scale, greater product variety, and increased competition. In an open economy, increased competition from foreign firms reduces the monopoly power of domestic firms and forces them to become more efficient, as compared to a closed economy. Industries that exhibit increasing returns to scale (for example, the automobile and steel industries) benefit from increased market size as a country starts trading because the average cost of production declines as output increases in these industries. Monopolistically competitive models of trade have been used to explain why there is significant two-way trade (known as *intra-industry trade*) between countries within the same industry. Intra-industry trade occurs when a country exports and imports goods in the same product category or classification.

In a monopolistically competitive industry, there are many firms; each firm produces a unique or differentiated product, there are no exit or entry barriers, and long-run economic profits are zero. In such a model, even though countries may be similar, they gain from trade because each country focuses on the production and export of one or more varieties of the good and imports other varieties of the good. For example, the European Union exports and imports different types of cars. Consumers gain from having access to a greater variety of final goods. Firms benefit from greater economies of scale because firms both within and outside the EU are able to sell their goods in both markets. Hence, scale economies allow firms to benefit from the larger market size and experience lower average cost of production as a result of trade.

Research suggests that trade liberalization can lead to increased real (that is, inflation-adjusted) GDP although the strength of this relationship is still debated. The positive influence of trade on GDP can arise from more efficient allocation of resources, learning by doing, higher productivity, knowledge spillovers, and trade-induced changes in policies and institutions that affect the incentives for innovation.[6] In industries where there is "learning by doing," such as the semiconductor industry, the cost of production per unit declines as output increases because of expertise and experience acquired in the process of production. Trade can lead to increased exchange of ideas, freer flow of technical expertise, and greater awareness of changing consumer tastes and preferences in global markets. It can also contribute to the development of higher quality and more effective institutions and policies that encourage domestic innovation. For example, Coe and Helpman (1995) show that foreign research and development (R&D) has beneficial effects on domestic productivity. These effects become stronger the more open an economy is to foreign trade. They estimate that about a quarter of the benefits of R&D investment in a G–7 country accrues to their trading partners.[7] Hill (2007) discusses the case of Logitech, a Swiss company that manufactures computer mice. In order to win original equipment manufacturer (OEM) contracts from IBM and Apple, Logitech needed to develop innovative designs and provide high-volume production at a low cost. So in the late 1980s they moved to Taiwan Region, which had a highly qualified labor force, competent parts suppliers, a rapidly expanding local computer industry, and offered Logitech space in a science park at a very competitive rate. Soon thereafter, Logitech was able to secure the Apple contract.

Opponents of free trade point to the potential for greater income inequality and the loss of jobs in developed countries as a result of import competition. As a country moves toward free trade, there will be adjustments in domestic industries that are exporters as well as those that face import competition. Resources (investments) may need to be reallocated into or out of an industry depending on whether that industry is expanding (exporters) or contracting (i.e., facing import competition). As a result of this adjustment process, less-efficient firms may be forced to exit the industry, which may, in turn, lead to higher unemployment and the need for displaced workers to be retrained for jobs in expanding industries. The counter argument is that although there may be short-term and even some medium-term costs, these resources are likely to be more effectively (re-)employed in other industries in the long run. Nonetheless, the adjustment process is virtually certain to impose costs on some groups of stakeholders.

---

[6] "Knowledge spillovers" occur when investments in knowledge creation generate benefits that extend beyond the investing entity and facilitate learning and innovation by other firms or entities.

[7] G–7 countries include Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States.

EXAMPLE 1

Benefits of Trade

Consider two countries that each produce two goods. Suppose the cost of producing cotton relative to lumber is lower in Cottonland than in Lumberland.

1. How would trade between the two countries affect the lumber industry in Lumberland?
2. How would trade between the two countries affect the lumber industry in Cottonland?
3. What would happen to the lumber industry workers in Cottonland in the long run?
4. What is the meaning of the expression "gains from trade"?
5. What are some of the benefits from trade?

Solution to 1:

The lumber industry in Lumberland would benefit from trade. Because the cost of producing lumber relative to producing cotton is lower in Lumberland than in Cottonland (i.e., lumber is relatively cheap in Lumberland), Lumberland will export lumber and the industry will expand.

Solution to 2:

The lumber industry in Cottonland would not benefit from trade, at least in the short run. Because lumber is relatively expensive to produce in Cottonland, the domestic lumber industry will shrink as lumber is imported from Lumberland.

Solution to 3:

The overall welfare effect in both countries is positive. However, in the short run, many lumber producers in Cottonland (and cotton producers in Lumberland) are likely to find themselves without jobs as the lumber industry in Cottonland and the cotton industry in Lumberland contract. Those with skills that are also needed in the other industry may find jobs fairly quickly. Others are likely to do so after some re-training. In the long run, displaced workers should be able to find jobs in the expanding export industry. However, those who remain in the import-competing industry may be permanently worse off because their industry-specific skills are now less valuable. Thus, even in the long run, trade does not necessarily make every stakeholder better off. But the winners could compensate the losers and still be better off, so the overall welfare effect of opening trade is positive.

Solution to 4:

Gains from trade imply that the overall benefits of trade outweigh the losses from trade. It does not mean that all stakeholders (producers, consumers, government) benefit (or benefit equally) from trade.

Solution to 5:

Some of the benefits from trade include: gains from exchange and specialization based on relative cost advantage; gains from economies of scale as the companies add new markets for their products; greater variety of products available to households and firms; greater efficiency from increased competition; and more efficient allocation of resources.

4 COMPARATIVE ADVANTAGE AND THE GAINS FROM TRADE: ABSOLUTE AND COMPARATIVE ADVANTAGE

c contrast comparative advantage and absolute advantage;

Up to this point, we have not been precise about what it means for a country to have a comparative advantage in the production of specific goods and services. In this section, we define comparative advantage, distinguish it from the notion of absolute advantage, and demonstrate the gains from trading in accordance with comparative advantage. We then explain two traditional models of trade—the Ricardian and Heckscher–Ohlin models—and the source of comparative advantage in each model.

4.1 Gains from Trade: Absolute and Comparative Advantage

A country has an **absolute advantage** in producing a good (or service) if it is able to produce that good at a lower cost or use fewer resources in its production than its trading partner. For example, suppose a worker in Brazil can produce either 20 pens or 40 pencils in a day. A worker in Vietnam can produce either 10 pens or 60 pencils. A Vietnamese worker produces 60 pencils a day while a Brazilian worker produces only 40 pencils a day. Hence, Vietnam produces pencils at a lower cost than Brazil, and has an absolute advantage in the production of pencils. Similarly, Brazil produces pens at a lower cost than Vietnam, and hence has an absolute advantage in the production of pens. A country has a **comparative advantage** in producing a good if its opportunity cost of producing that good is less than that of its trading partner. In our example, the opportunity cost of producing an extra pen in Vietnam is 6 pencils. It is the opportunity foregone; namely, the number of pencils Vietnam would have to give up to produce an extra pen. If Brazil does not trade and has to produce both pens and pencils, it will have to give up 2 pencils in order to produce a pen. Similarly, in Vietnam each pen will cost 6 pencils. Hence, the opportunity cost of a pen in Brazil is 2 pencils, whereas in Vietnam it is 6 pencils. Brazil has the lower opportunity cost and thus a comparative advantage in the production of pens. Vietnam has a lower opportunity cost (1 pencil costs ⅙th of a pen) than Brazil (1 pencil costs ½ a pen) in the production of pencils and thus has a comparative advantage in the production of pencils. Example 2 further illustrates these concepts.

EXAMPLE 2

Absolute and Comparative Advantages

Suppose there are only two countries, India and the United Kingdom. India exports cloth to the United Kingdom and imports machinery. The output per worker per day in each country is shown in Exhibit 6:

Exhibit 6 Output per Worker per Day

	Machinery	Cloth (yards)
United Kingdom	**4**	8
India	2	**16**

Based only on the information given, address the following:

1. Which country has an absolute advantage in the production of:
   - **A** machinery?
   - **B** cloth?
2. Do the countries identified in Question 1 as having an absolute advantage in the production of A) machinery and B) cloth, also have a comparative advantage in those areas?

Solution to 1A:

The United Kingdom has an absolute advantage in the production of machinery because it produces more machinery per worker per day than India.

Solution to 1B:

India has an absolute advantage in the production of cloth because it produces more cloth per worker per day than the United Kingdom.

Solution to 2A and 2B:

In both cases, the answer is "yes." In the case of machinery, the opportunity cost of a machine in the United Kingdom is 2 yards of cloth (8 ÷ 4 or 1 machine = 2 yards cloth). This amount is the autarkic price of machines in terms of cloth in the United Kingdom. In India, the opportunity cost of a machine is 8 yards of cloth (16 ÷ 2 or 1 machine = 8 yards cloth). Thus, the United Kingdom has a comparative advantage in producing machines. In contrast, the opportunity cost of a yard of cloth in the United Kingdom and in India is ½ and ⅛ of a machine, respectively. India has a lower opportunity cost (⅛ of a machine) and, therefore, a comparative advantage in the production of cloth.

It is important to note that even if a country does not have an absolute advantage in producing any of the goods, it can still gain from trade by exporting the goods in which it has a comparative advantage. In Example 2, if India could produce only 6 yards of cloth per day instead of 16 yards of cloth, the United Kingdom would have an *absolute* advantage in both machines and cloth. However, India would still have a *comparative* advantage in the production of cloth because the opportunity cost of a yard of cloth in India, ⅓ of a machine in this case, would still be less than the opportunity cost of a yard of cloth in the United Kingdom (½ of a machine as before).

Let us now illustrate the gains from trading according to comparative advantage. In Example 2, if the United Kingdom could sell a machine for more than 2 yards of cloth and if India could purchase a machine for less than 8 yards of cloth, both countries would gain from trade. Although it is not possible to determine the exact world price without additional details regarding demand and supply conditions, both countries would gain from trade as long as the world price for machinery in terms of cloth is between the autarkic prices of the trading partners. In our example, this price corresponds to a price of between 2 and 8 yards of cloth for a machine. *The further away the world price of a good or service is from its autarkic price in a given country, the more that country gains from trade.* For example, if the United Kingdom was able to sell a machine to India for 7 yards of cloth (i.e., closer to India's autarkic price), it would gain 5 yards of cloth per machine sold to India compared with its own autarkic price (with no trade) of 1 machine for 2 yards of cloth. However, if the United Kingdom was able to sell a machine to India for only 3 yards of cloth (closer to the UK autarkic price), it would gain only 1 yard of cloth per machine sold to India compared with its own autarkic price.

Exhibits 7 and 8 provide the production and consumption schedules of both countries at autarky and after trade has commenced. In autarky (Exhibit 7), the United Kingdom produces and consumes 200 machines and 400 yards of cloth (without trade, consumption of each product must equal domestic production). Similarly, India produces 100 machines and 800 yards of cloth in autarky. In a world economy consisting of only these two countries, total output for each commodity is the sum of production in both countries. Therefore, total world output is 300 machines and 1,200 yards of cloth.

Exhibit 7 Production and Consumption in Autarky

	Autarkic Production	Autarkic Consumption
United Kingdom		
Machinery (m)	200	200
Cloth (yards) (c)	400	400
India		
Machinery	100	100
Cloth (yards)	800	800
Total World:		
Machinery	300	300
Cloth (yards)	1200	1200

Now, assume that the United Kingdom and India start trading and that the world price of 1 machine is 4 yards of cloth (1m = 4c). This price is within the range of acceptable world trading prices discussed earlier because this price lies between the autarkic prices of the United Kingdom (1m = 2c) and India (1m = 8c). Exhibit 8 shows that in an open economy, the United Kingdom would specialize in machines and India would specialize in cloth. As a result, the United Kingdom produces 400 machines and no cloth, while India produces 1,600 yards of cloth and no machines. The United Kingdom exports 160 machines to India in exchange for 640 yards of cloth. After trade begins with India, the United Kingdom consumes 240 machines and 640 yards of cloth. Consumption in the United Kingdom increases by 40 machines and 240 yards of cloth. Similarly, India consumes 160 machines and 960 yards of cloth, an increase of 60 machines and 160 yards of cloth. World production and consumption is now 400 machines and 1,600 yards of cloth. Post-trade production and consumption exceeds the autarkic situation by 100 machines and 400 yards of cloth.

Exhibit 8 Gains from Trade

	Post-trade Production	Post-trade Consumption	Change in Consumption (compared with autarky)
UK			
Machinery	400	240	+40
Cloth (yards)	0	640	+240
India			
Machinery	0	160	+60
Cloth (yards)	1600	960	+160
Total World:			

Exhibit 8 (Continued)

	Post-trade Production	Post-trade Consumption	Change in Consumption (compared with autarky)
Machinery	400	400	+100
Cloth (yards)	1600	1600	+400

Exhibit 9 shows a more general case of gains from trade under increasing costs. In Panel A, the curve connecting the *X* and *Y* axes is the UK production possibilities frontier (PPF).[8] That is, it represents the combinations of cloth and machinery that the United Kingdom can produce given its technology and resources (capital and labor). The slope of the PPF at any point is the opportunity cost of one good in terms of the other. The shape of the PPF indicates increasing opportunity cost in terms of machines as more cloth is produced and vice versa. To maximize the value of output, production occurs where the slope of the PPF equals the relative price of the goods. P_A represents the autarkic price line, which is tangent to the PPF at *A*, the autarkic equilibrium. The slope of the autarkic price line represents the opportunity cost before trade. In autarky, the United Kingdom produces and consumes 60 machines and 60 thousand yards cloth, and is on indifference curve *I*.[9] When the United Kingdom starts trading with India, it faces the world price line P^*. This new price line is tangent to the PPF at *B*. The change in relative prices of the goods encourages the United Kingdom to increase the production of the good in which it has comparative advantage (machines) and produce at *B* instead of *A*. We note that at *B* the United Kingdom has increased the production of machines to 120 units and reduced the production of cloth to 30 thousand yards. We also note that trade has expanded the UK consumption possibilities. The United Kingdom consumes at point *E* after trade, exports 80 machines to India and imports 80 thousand yards of cloth from India. Note that *E* is outside the PPF, but on the world price line that is tangent to the PPF at *B*. This line is also the trading possibilities line because trade occurs along this line. The slope of this line is the opportunity cost of a machine in terms of cloth in the world market. The United Kingdom has clearly increased its welfare through trade because it is able to consume at point *E*, which is on a higher indifference curve (III) and thus represents a higher level of welfare compared with the autarkic consumption point *A* on indifference curve *I*.

Panel B shows the corresponding situation for India. When trade opens with the United Kingdom, India shifts production from *A*′ to *B*′, producing more cloth, the good in which it has a comparative advantage, and fewer machines. It now exports 80 thousand yards of cloth to the United Kingdom and imports 80 machines from the United Kingdom. India now consumes at *E*′ which is on the world price line and also on a higher indifference curve, III′, than the autarkic consumption point (*A*′) on indifference curve I′. Thus, by specializing (incompletely, as is typically the case with increasing production costs) in the good in which it has a comparative advantage, each country increases its welfare. We should also note that P^* is the price at which trade is balanced. At this relative world price, the export of cloth from India equals the import of cloth into the United Kingdom (80 thousand yards) and the export of machines from the United Kingdom equals the imports of machines into India (80 machines).

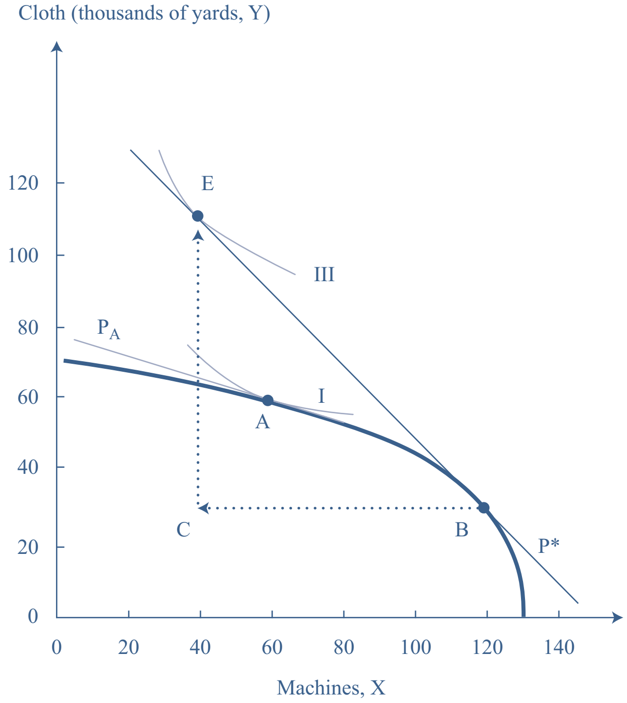
---

[8] Modified from Salvatore (2011).

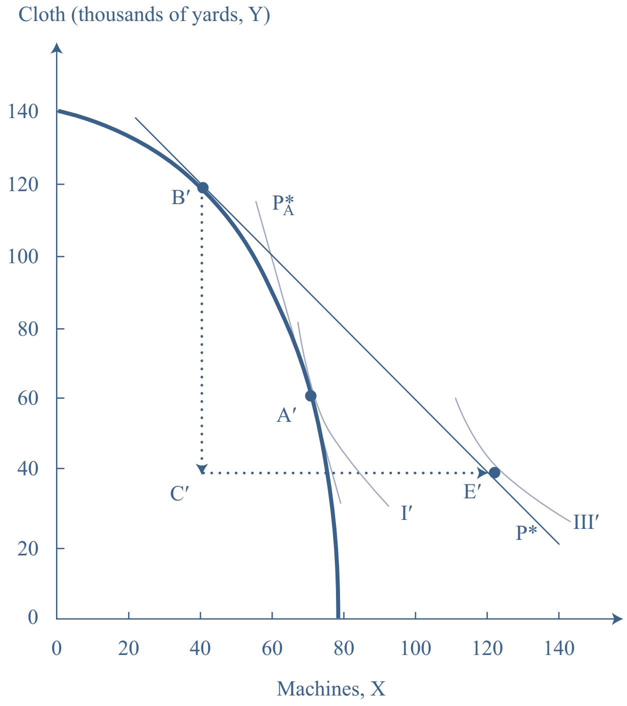
[9] An indifference curve represents the various combinations of goods (machines and cloth) that provide the same level of utility or welfare. Higher indifference curves represent higher levels of utility or welfare.

Exhibit 9 Graphical Depiction of Gains from Trade with Increasing Costs

Panel A. United Kingdom



Panel B. India



A country's comparative advantage can change over time as a result of structural shifts in its domestic economy, shifts in the global economy, the accumulation of physical or human capital, new technology, the discovery of such natural resources as oil, and so on. For example, an increase in skilled labor in China has led several multinational companies to establish R&D facilities in China to benefit from its highly educated workforce.

EXAMPLE 3

Changes in Comparative Advantage

Exhibit 10 shows how the South Korea's comparative advantage changed over time as a result of an export-oriented development strategy it adopted during the 1960s.[10] The challenges of foreign competition created a "virtuous circle" that was self-reinforcing. South Korea's changing comparative advantage was the result of government policy, an increasingly skilled and productive workforce, and proactive firms that learned and adapted new technology.

Exhibit 10 Changes in Structure of South Korea's Exports, 1980–2015 (Percentage Shares)

	1980	1985	1990	1995	2000	2005	2010	2015
Agricultural products	6.8%	4.0%	3.8%	3.2%	2.3%	1.8%	2.0%	2.0%
Fuels and mining products	1.0%	3.1%	1.6%	2.7%	6.1%	6.9%	8.9%	8.1%
Manufactures	69.3%	74.2%	76.9%	81.8%	82.2%	86.9%	86.5%	87.5%
Chemicals	3.3%		3.2%	6.4%	7.3%	9.3%	10.3%	10.9%
Pharmaceuticals			0.1%	0.2%	0.2%	0.2%	0.3%	0.4%
Machinery and transport equipment	15.7%		32.4%	46.9%	53.2%	58.4%	55.5%	57.7%
Office and telecom equipment	7.5%	9.9%	18.2%	23.8%	31.1%	27.9%	20.3%	20.5%
Integrated circuits & electronic components			6.8%	13.9%	13.1%	9.3%	9.1%	10.7%
Textiles	9.8%	6.8%	7.7%	8.8%	6.7%	3.5%	2.3%	2.0%
Clothing	13.1%	11.9%	10.0%	3.5%	2.7%	0.9%	0.3%	0.4%

Source: World Trade Organisation, WTO Statistics Database stat.wto.org/Home/WSDBHome.aspx.

1. How has South Korea's structure of exports changed over time?
2. How did increased foreign competition impact the economy?
3. What were the factors that helped to change South Korea's comparative advantage?

Solution to 1:

In 1980, agriculture and clothing accounted for 6.8 percent and 13 percent of South Korea's exports, respectively. By 2015, the corresponding figures were 2.0 percent and 0.4 percent. In contrast, by 2015 machinery and transport equipment were almost 60% of South Korea's merchandise exports from only about 16% in 1980. Manufactures as a whole were 87.5%, up from 69% in 1980.

---

[10] Wikipedia: Trade Policy of South Korea. In 1962, South Korea first introduced an export promotion policy targeted at labor intensive industries like textiles and clothing. By the 1970s, this plan had shifted focus to heavy industries and chemicals as the main export targets. As South Korea developed in the 1980s and early 1990s, export policies shifted toward consumer products, electronics, and high tech in particular.

Solution to 2:

The challenges of foreign competition created a "virtuous circle" that was self-reinforcing. Success in export markets increased the confidence of South Korean firms and led to greater success in exports through increased productivity, higher-quality products, acquisition of new skills, and adoption of technologies.

Solution to 3:

The factors that helped change South Korea's comparative advantage included government policy, an increasingly skilled and productive workforce, and proactive firms that learned and adapted new technology.

From an investment perspective, it is critical for analysts to be able to examine a country's comparative and absolute advantages and to analyze changes in them. It is also important to understand changes in government policy and regulations, demographics, human capital, demand conditions, and other factors that may influence comparative advantage and production and trade patterns. This information can then be used to identify sectors, industries within those sectors, and companies within those industries that will benefit.

5

RICARDIAN AND HECKSCHER–OHLIN MODELS OF COMPARATIVE ADVANTAGE

d compare the Ricardian and Heckscher–Ohlin models of trade and the source(s) of comparative advantage in each model;

A discussion of absolute and comparative advantage and the gains from specialization would be incomplete without a discussion of two important theories of trade, the Ricardian Model and the Heckscher–Ohlin Model. These models are based on cross-country differences in technology and in factor endowments, respectively. These theoretical models are based on several assumptions, some of which may not be fully satisfied in the real world; nonetheless they provide extremely useful insights into the determinants and patterns of trade.

Adam Smith argued that a country could gain from trade if it had an absolute advantage in the production of a good. David Ricardo extended Smith's idea of the gains from trade by arguing that even if a country did not have an absolute advantage in the production of any good, it could still gain from trade if it had a comparative advantage in the production of a good. In the Ricardian model, labor is the only (variable) factor of production. Differences in labor productivity, reflecting underlying differences in technology, are the source of comparative advantage and hence the key driver of trade in this model. A country with a lower opportunity cost in the production of a good has a comparative advantage in that good and will specialize in its production. In our two-country model, if countries vary in size, the smaller country may specialize completely, but may not be able to meet the total demand for the product. Hence, the larger country may be incompletely specialized, producing and exporting the good in which it has a comparative advantage but still producing (and consuming) some of the good in which it has a comparative disadvantage. It is important to recognize that although differences in technology may be a major source of comparative advantage at a given point in time, other countries can close the technology gap or even gain a technological advantage. The shift of information technology services from developed countries to India is an example of comparative advantage shifting over time.[11] This shift was facilitated by India's growing pool of highly skilled and relatively low-wage labor, the development and growth of its telecommunication infrastructure, and government policies that liberalized trade in the 1990s.

In the Heckscher–Ohlin Model (also known as the factor-proportions theory), both capital and labor are variable factors of production. That is, each good can be produced with varying combinations of labor and capital. According to this model, differences in the relative endowment of these factors are the source of a country's comparative advantage. This model assumes that technology in each industry is the same among countries, but it varies between industries. According to the theory, a country has a comparative advantage in goods whose production is intensive in the factor with which it is relatively abundantly endowed, and would tend to specialize in and export that good. Capital is relatively more (less) abundant in a country if the ratio of its endowment of capital to labor is greater (less) than that of its trading partner.[12] This scenario means a country in which labor is relatively abundant would export relatively labor-intensive goods and import relatively capital-intensive goods. For example, because the manufacture of textiles and clothing is relatively labor intensive, they are exported by such countries as China and India where labor is relatively abundant.

Relative factor intensities in production can be illustrated with the following example. In 2002, capital per worker in the Canadian paper industry was C$118,777, whereas in the clothing manufacturing sector it was C$8,954.[13] These amounts indicate that manufacturing paper is more capital intensive than clothing production. Canada trades with Thailand and, being relatively capital abundant compared with Thailand, it exports relatively capital-intensive paper to Thailand and imports relatively labor-intensive clothing from Thailand.

Because the Heckscher–Ohlin model has two factors of production, labor and capital, (unlike the Ricardian model that has only labor), it allows for the possibility of income redistribution through trade. The demand for an input is referred to as a *derived demand* because it is derived from the demand for the product it is used to produce. As a country opens up to trade, it has a favorable impact on the abundant factor, and a negative impact on the scarce factor. This result is because trade causes output prices to change; more specifically, the price of the export good increases and the price of the import good declines. These price changes affect the demand for factors used to produce the import and export goods, and hence affect the incomes received by each factor of production.

To illustrate this point, consider again the opening of trade between the United Kingdom and India in Exhibit 9. When trade opened, the United Kingdom expanded production of machines—which are assumed to be the capital-intensive industry—and reduced production of clothing. India did the opposite. Machines became more expensive relative to clothing in the United Kingdom (line P^* is steeper than line P^A). The relative price change, along with the shift in output it induces, leads to a redistribution of income from labor to capital in the United Kingdom. The opposite occurs in India—machines become cheaper relative to clothing (line P^* is flatter than $P^{A'}$), production shifts toward clothing, and income is redistributed from capital to labor.

---

[11] According to NASSCOM (India's prominent IT-BPO trade association), Indian firms offer a wide range of information technology services that include consulting, systems integration, IT outsourcing/managed services/hosting services, training, and support/maintenance. See www.nasscom.in.

[12] Alternatively, factor abundance can be defined in terms of the relative factor prices that prevail in autarky. Under this definition, labor is more (less) abundant in a country if the cost of labor relative to the cost of capital is lower (higher) in that country.

[13] Appleyard, Field, and Cobb (2010).

Note that in each country, the relatively cheap good and the relatively cheap factor of production both get more expensive when trade is opened. That raises an interesting question: If free trade equalizes the prices of goods among countries, does it also equalize the prices of the factors of production? In the simple Heckscher–Ohlin world of homogeneous products, homogeneous inputs, and identical technologies among countries, the answer is yes: The absolute and relative factor prices are equalized in both countries if there is free trade. In the real world, we see that factor prices do not converge completely even if there is free trade because several assumptions of the models are not fully satisfied in the real world. Nonetheless, it is important to note that *with international trade factor prices display a tendency to move closer together in the long run.*

Changes in factor endowments can cause changes in the patterns of trade and can create profitable investment opportunities. For example, in 1967 Japan had a comparative advantage in unskilled-labor-intensive goods, such as textiles, apparel, and leather. Meier (1998) notes that by 1980, Japan had greatly increased its skilled labor and consequently had a comparative advantage in skill-intensive products, especially non-electrical machinery.

It is important to note that technological differences, as emphasized in the Ricardian trade model, and differences in factor abundance, as emphasized in the Heckscher–Ohlin model, are both important drivers of trade. They are complementary, not mutually exclusive. Tastes and preferences can also vary among countries and can change over time, leading to changes in trade patterns and trade flows.

6

TRADE AND CAPITAL FLOWS: RESTRICTIONS & AGREEMENTS- TARIFFS, QUOTAS AND EXPORT SUBSIDIES

e compare types of trade and capital restrictions and their economic implications;

Trade restrictions (or trade protection) are government policies that limit the ability of domestic households and firms to trade freely with other countries. Examples of trade restrictions include tariffs, import quotas, voluntary export restraints (VER), subsidies, embargoes, and domestic content requirements. **Tariffs** are taxes that a government levies on imported goods. **Quotas** restrict the quantity of a good that can be imported into a country, generally for a specified period of time. A voluntary export restraint is similar to a quota but is imposed by the exporting country. An **export subsidy** is paid by the government to the firm when it exports a unit of a good that is being subsidized. The goal here is to promote exports, but it reduces welfare by encouraging production and trade that is inconsistent with comparative advantage. **Domestic content provisions** stipulate that some percentage of the value added or components used in production should be of domestic origin. Trade restrictions are imposed by countries for several reasons including protecting established domestic industries from foreign competition, protecting new industries from foreign competition until they mature (infant industry argument), protecting and increasing domestic employment, protecting strategic industries for national security reasons, generating revenues from tariffs (especially for developing countries), and retaliation against trade restrictions imposed by other countries.

Capital restrictions are defined as controls placed on foreigners' ability to own domestic assets and/or domestic residents' ability to own foreign assets. Thus, in contrast with trade restrictions, which limit the openness of goods markets, capital restrictions limit the openness of financial markets. Sections 6 and 7 discuss trade restrictions. Section 8 briefly addresses capital restrictions.

6.1 Tariffs

Tariffs are taxes that a government levies on imported goods.[14] The primary objective of tariffs is to protect domestic industries that produce the same or similar goods. They may also aim to reduce a trade deficit. Tariffs reduce the demand for imported goods by increasing their price above the free trade price. The economic impact of a tariff on imports in a small country is illustrated in Exhibit 11. In this context, a small country is not necessarily small in size, population, or GDP. Instead, a **small country** is one that is a price taker in the world market for a product and cannot influence the world market price. For example, by many measures Brazil is a large country, but it is a price taker in the world market for cars. A large country, however, is a large importer of the product and can exercise some influence on price in the world market. When a large country imposes a tariff, the exporter reduces the price of the good to retain some of the market share it could lose if it did not lower its price. This reduction in price alters the terms of trade and represents a redistribution of income from the exporting country to the importing country. So, in theory it is possible for a large country to increase its welfare by imposing a tariff if 1) its trading partner does not retaliate and 2) the deadweight loss as a result of the tariff (see below) is smaller than the benefit of improving its terms of trade. However, there would still be a net reduction in global welfare—the large country cannot gain by imposing a tariff unless it imposes an even larger loss on its trading partner.

In Exhibit 11, the world price (free trade price) is P^*. Under free trade, domestic supply is Q^1, domestic consumption is Q^4, and imports are Q^1Q^4. After the imposition of a per-unit tariff t, the domestic price increases to P_t, which is the sum of the world price and the per-unit tariff t. At the new domestic price, domestic production increases to Q^2 and domestic consumption declines to Q^3, resulting in a reduction in imports to Q^2Q^3.

The welfare effects can be summarized as follows:

- Consumers suffer a loss of consumer surplus because of the increase in price.[15] This effect is represented by areas A + B + C + D in Exhibit 11.
- Local producers gain producer surplus from a higher price for their output. This effect is represented by area A.
- The government gains tariff revenue on imports Q^2Q^3. This effect is represented by area C.

The net welfare effect is the sum of these three effects. The loss in consumer surplus is greater than the sum of the gain in producer surplus and government revenue and results in a deadweight loss to the country's welfare of B + D.

---

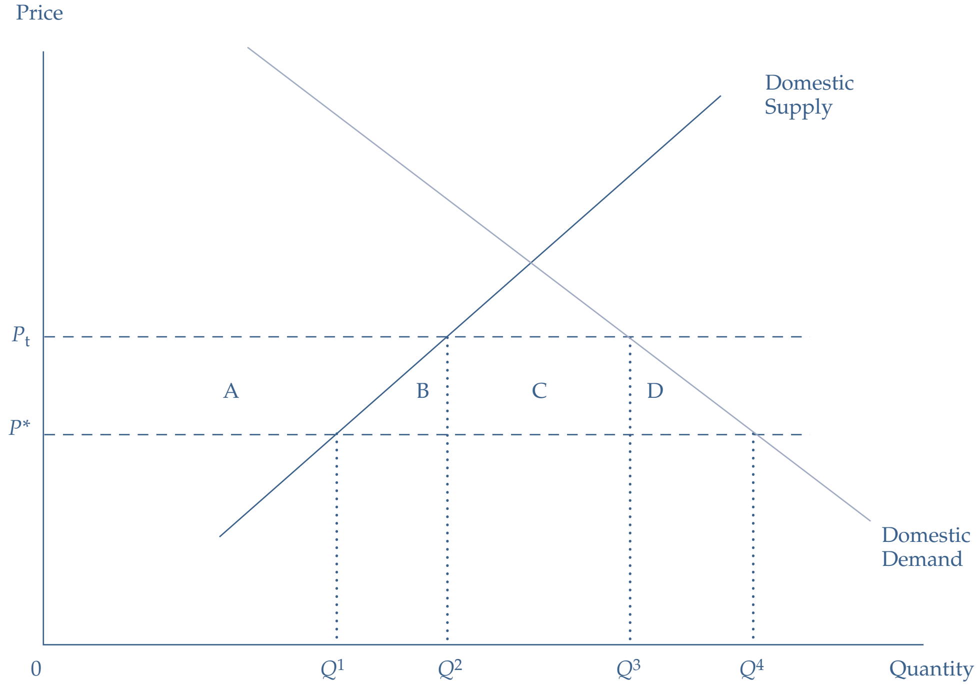
[14] Governments may also impose taxes on exports, although they are less common.

[15] Consumer surplus, producer surplus, and deadweight loss are defined and discussed in the prerequisite reading "Demand and Supply Analysis: Introduction" available online in your Candidate Resources.

Welfare Effects of an Import Tariff or Quota	Importing Country
Consumer surplus	– (A + B + C + D)
Producer surplus	+A
Tariff revenue or Quota rents	+C
National welfare	– *B* – *D*

Tariffs create deadweight loss because they give rise to inefficiencies on both the consumption and production side. B represents inefficiencies in production. Instead of being able to import goods at the world price P^*, tariffs encourage inefficient producers whose cost of production is greater than P^* to enter (or remain in) the market, leading to an inefficient allocation of resources. On the consumption side, tariffs prevent mutually beneficial exchanges from occurring because consumers who were willing to pay more than P^* but less than P_t are now unable to consume the good.

Exhibit 11 Welfare Effects of Tariff and Import Quota



EXAMPLE 4

Analysis of a Tariff

South Africa manufactures 110,000 tons of paper. However, domestic demand for paper is 200,000 tons. The world price for paper is $5 per ton. South Africa will import 90,000 tons of paper from the world market at free trade prices. If the South African government (a small country) decides to impose a tariff of

20 percent on paper imports, the price of imported paper will increase to $6. Domestic production after the imposition of the tariff increases to 130,000 tons, while the quantity demanded declines to 170,000 tons.

1. Calculate the loss in consumer surplus arising from the imposition of the tariff.
2. Calculate the gain in producer surplus arising from the imposition of the tariff.
3. Calculate the gain in government revenue arising from the imposition of the tariff.
4. Calculate the deadweight loss arising from the imposition of the tariff.

Solution to 1:

The loss in consumer surplus = $1 × 170,000 + 1/2 × $1 × 30,000 = $185,000. This calculation is represented by areas A + B + C + D in Exhibit 11.

Solution to 2:

Gain in producer surplus = $1 × 110,000 + 1/2 × ($1 × 20,000) = $120,000; Area A in Exhibit 11.

Solution to 3:

Change in government revenue = $1 × 40,000 = $40,000; Area C in Exhibit 11.

Solution to 4:

Deadweight loss because of the tariff = 1/2 × $1 × 20,000 + 1/2 × $1 × 30,000 = $25,000; Areas B + D in Exhibit 11.

6.2 Quotas

A **quota** restricts the quantity of a good that can be imported into a country, generally for a specified period of time. An **import license** specifies the quantity that can be imported. For example, the European Union operates a system of annual import quotas for steel producers who are not members of the World Trade Organization. The 2010 quota was 0.2 million tons a year for Kazakhstan. In the case of Russia, the 2010 quota of 3.2 million tons per year was a part of an EU–Russia agreement.[16] A key difference between tariffs and quotas is that the government is able to collect the revenue generated from a tariff. This effect is uncertain under a quota. With quotas, foreign producers can often raise the price of their goods and earn greater profits than they would without the quota. These profits are called **quota rents**. In Exhibit 11, if the quota is Q^2Q^3, the equivalent tariff that will restrict imports to Q^2Q^3 is t and the domestic price after the quota is P_t. This is the same as the domestic price after the tariff t was imposed. Area C, however, is now the quota rent or profits that are likely to be captured by the foreign producer rather than tariff revenue that is captured by the domestic government. If the foreign producer or foreign government captures the quota rent, C, then the welfare loss to the importing country, represented by areas B + D + C in Exhibit 11, under a quota is greater than under the equivalent tariff. If the government of the country that imposes the quota can capture the quota rents by auctioning the import licenses for a fee, then the welfare loss under the quota is similar to that of a tariff, represented by areas B + D.

---

[16] For more information, see http://ec.europa.eu/trade/creating-opportunities/economic-sectors/industrial-goods/steel/.

A **voluntary export restraint** (VER) is a trade barrier under which the exporting country agrees to limit its exports of the good to its trading partners to a specific number of units. The main difference between an import quota and a VER is that the former is imposed by the importer, whereas the latter is imposed by the exporter. The VER allows the quota rent resulting from the decrease in trade to be captured by the exporter (or exporting country), whereas in the case of an import quota there is ambiguity regarding who captures the quota rents. Hence, a VER results in welfare loss in the importing country. For example, in 1981 the Japanese government imposed VERs on automobile exports to the United States.

6.3 Export Subsidies

An export subsidy is a payment by the government to a firm for each unit of a good that is exported. Its goal is to stimulate exports. But it interferes with the functioning of the free market and may distort trade away from comparative advantage. Hence, it reduces welfare. *Countervailing duties* are duties that are levied by the importing country against subsidized exports entering the country. As an example, agricultural subsidies in developed countries, notably the EU, have been a contentious issue in trade negotiations with less-developed countries and developed countries that are agricultural exporters, such as New Zealand and Australia.

In the case of an export subsidy, the exporter has the incentive to shift sales from the domestic to the export market because it receives the international price plus the per-unit subsidy for each unit of the good exported. This scenario raises the price in the domestic market by the amount of the subsidy in the small country case (price before subsidy plus subsidy). In the large country case, the world price declines as the large country increases exports. The net welfare effect is negative in both the large and small country cases, with a larger decline in the large country case. This result is because in the large country case, the decline in world prices implies that a part of the subsidy is transferred to the foreign country, unlike in the small country case.

Exhibit 12 summarizes some of these effects.

Exhibit 12

Panel A. Effects of Alternative Trade Policies

	Tariff	Import Quota	Export Subsidy	VER
Impact on	Importing country	Importing country	Exporting country	Importing country
Producer surplus	Increases	Increases	Increases	Increases
Consumer surplus	Decreases	Decreases	Decreases	Decreases
Government revenue	Increases	Mixed (depends on whether the quota rents are captured by the importing country through sale of licenses or by the exporters)	Falls (government spending rises)	No change (rent to foreigners)
National welfare	Decreases in small country Could increase in large country	Decreases in small country Could increase in large country	Decreases	Decreases

Exhibit 12 (Continued)

Panel B. Effects of Alternative Trade Policies on Price, Production, Consumption, and Trade

	Tariff	Import Quota	Export Subsidy	VER
Impact on	Importing country	Importing country	Exporting country	Importing country
Price	Increases	Increases	Increases	Increases
Domestic consumption	Decreases	Decreases	Decreases	Decreases
Domestic production	Increases	Increases	Increases	Increases
Trade	Imports decrease	Imports decrease	Exports increase	Imports decrease

EXAMPLE 5

Tariffs, Quotas, and VERs

Thailand, a small country, has to decide whether to impose a tariff or a quota on the import of computers. You are considering investing in a local firm that is a major importer of computers.

1 What will be the impact of a tariff on prices, quantity produced, and quantity imported in Thailand (the importing country)?

2 If Thailand imposes a tariff, what will the impact be on prices in the exporting country?

3 How would a tariff affect consumer surplus, producer surplus, and government revenue in Thailand?

4 Explain whether the net welfare effect of a tariff is the same as that of a quota.

5 Which policy, a tariff or a quota, would be most beneficial to the local importer in which you may invest and why?

6 If Thailand were to negotiate a VER with the countries from which it imports computers, would this be better or worse than an import quota for the local importing firm in which you may invest? Why?

Solution to 1:

A tariff imposed by a small country, such as Thailand, raises the price of computers in the importing country, reduces the quantity imported, and increases domestic production.

Solution to 2:

A tariff imposed by a small country would not change the price of computers in the exporting country.

Solution to 3:

When a small country imposes a tariff, it reduces consumer surplus, increases producer surplus, and increases government revenue in that country.

Solution to 4:

The quota can lead to a greater welfare loss than a tariff if the quota rents are captured by the foreign government or foreign firms.

Solution to 5:

A tariff will hurt importers because it will reduce their share of the computer market in Thailand. The impact of a quota depends on whether the importers can capture a share of the quota rents. Assuming importers can capture at least part of the rents, they will be better off with a quota.

Solution to 6:

The VER would not be better for the local importer than the import quota and would most likely be worse. Under the VER, all of the quota rents will be captured by the exporting countries whereas with an import quota at least part of the quota rents may be captured by local importers.

It is important to understand existing trade policies and the potential for policy changes that may impact return on investment. Changes in the government's trade policy can affect the pattern and value of trade and may result in changes in industry structure. These changes may have important implications for firm profitability and growth because they can affect the goods a firm can import/export, change demand for its products, impact its pricing policies, and create delays through increased paperwork, procurement of licenses, approvals, and so on. For example, changes in import policies that affect the ability of a firm to import vital inputs for production may increase the cost of production and reduce firm profitability.

7 TRADING BLOCS, COMMON MARKETS, AND ECONOMIC UNIONS

f explain motivations for and advantages of trading blocs, common markets, and economic unions;

There has been a proliferation of trading blocs or regional trading agreements (RTA) in recent years. Important examples of regional integration include the North American Free Trade Agreement (NAFTA) and the European Union (EU). A regional trading bloc is a group of countries that have signed an agreement to reduce and progressively eliminate barriers to trade and movement of factors of production among the members of the bloc. It may or may not have common trade barriers against countries that are not members of the bloc.

There are many different types of regional trading blocs, depending on the level of integration that takes place. **Free trade areas** (FTA) are one of the most prevalent forms of regional integration in which all barriers to the flow of goods and services among members have been eliminated. However, each country maintains its own polices against non-members. The North American Free Trade Agreement (NAFTA) among the United States, Canada, and Mexico is an example of a FTA. A **customs union** extends the FTA by not only allowing free movement of goods and services among members but also creating a common trade policy against non-members. In 1947, Belgium, the Netherlands, and Luxemburg ("Benelux") formed a customs union that became a part of the European Community in 1958. The **common market** is the next level of economic integration that incorporates all aspects of the customs union and extends it by allowing free movement of factors of production among members. The Southern Cone Common Market (MERCOSUR) of Argentina, Brazil, Paraguay, and Uruguay is an example of a common market.[17] An **economic union** requires an even greater degree of integration. It incorporates all aspects of a common market and in addition requires common economic institutions and coordination of economic policies among members. The European Community became the European Union in 1993. If the members of the economic union decide to adopt a common currency, then it is also a **monetary union**. For example, with the adoption of the euro, 19 EU member countries also formed a monetary union.[18]

EXAMPLE 6

Trading Blocs

1 Chile and Australia have a free trade with each other but have separate trade barriers on imports from other countries. Chile and Australia are a part of a(n)

- **A** FTA.
- **B** Economic union.
- **C** Customs union.
- **D** Common market.

2 An RTA that removes all tariffs on imports from member countries, has common external tariffs against all non-members, but does not advance further in deepening economic integration is called a(n)

- **A** FTA.
- **B** Economic union.
- **C** Customs union.
- **D** Common market.

Solution to 1:

A is correct. Chile and Australia do not a have customs union because they do not have a common trade policy with respect to other trade partners (C is incorrect). A common market or a economic union entail even more integration (B and D are incorrect).

Solution to 2:

C is correct. A basic FTA does not entail common external tariffs (A is incorrect), whereas a common market and an economic union entail integration beyond common external tariffs (B and D are incorrect).

Regional integration is popular because eliminating trade and investment barriers among a small group of countries is easier, politically less contentious, and quicker than multilateral trade negotiations under the World Trade Organization (WTO). The

---

[17] For more information, visit the OECD website, http://stats.oecd.org/glossary/.

[18] On 1 January 1999, Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Portugal, and Spain adopted the euro. This adoption meant that these countries had to surrender control over their domestic monetary policy to the European Central Bank. Greece joined in 2001. Euro coins and notes went into circulation on 1 January 2002, and these countries gave up the last vestiges of their national currencies. Other members now include Slovenia (2007), Cyprus (2008), Malta (2008), Slovakia (2009), Estonia (2011), Latvia (2014), and Lithuania (2015). The eurozone (i.e., the monetary union) is only a subset of the EU membership because some EU members, notably the United Kingdom, have not adopted the euro.

WTO is a negotiating forum that deals with the rules of global trade between nations and where member countries can go to sort out trade disputes. The latest rounds of trade negotiations launched by the WTO in 2001 at Doha, Qatar, included several contentious issues of specific concern to developing countries, such as the cost of implementing trade policy reform in developing countries, market access in developed countries for developing countries' agricultural products, and access to affordable pharmaceuticals in developing countries. After nearly a decade of negotiations, very limited progress has been made on the major issues. Hence, it is not surprising to see a renewed interest in bilateral and multilateral trade liberalization on a smaller scale. Policy coordination and harmonization are also easier among a smaller group of countries. Regional integration can be viewed as a movement toward freer trade.

Regional integration results in preferential treatment for members compared with non-members and can lead to changes in the patterns of trade. Member countries move toward freer trade by eliminating or reducing trade barriers against each other, leading to a more efficient allocation of resources. But regional integration may also result in trade and production being shifted from a lower-cost non-member who still faces trade barriers to a higher-cost member who faces no trade barriers. This shift leads to a less-efficient allocation of resources and could reduce welfare. Hence, there are two static effects that are direct results of the formation of the customs union: trade creation and trade diversion.

Trade creation occurs when regional integration results in the replacement of higher-cost domestic production by lower-cost imports from other members. For example, consider two hypothetical countries, Qualor and Vulcan. Qualor produces 10 million shirts annually and imports 2 million shirts from Vulcan, which has a lower cost of production. Qualor has 10 percent tariffs on imports from Vulcan. Qualor and Vulcan then agree to form a customs union. Qualor reduces its production of shirts to 7 million and now imports 11 million shirts from Vulcan. The decline in Qualor's domestic production (from 10 million to 7 million shirts) is replaced by importing 3 million additional shirts from the low-cost producer, Vulcan. This scenario represents trade creation. The rest of the additional imports (6 million shirts) represent increased consumption by Qualor's consumers because the price of shirts declines after formation of the custom union.

Trade diversion occurs when lower-cost imports from nonmember countries are replaced with higher-cost imports from members. In the example in the preceding paragraph, suppose Qualor initially imposes a 10 percent tariff on imports from both Vulcan and Aurelia. Aurelia is the lowest-cost producer of shirts, so Qualor initially imports 2 million shirts from Aurelia instead of from Vulcan. Qualor and Vulcan then form a customs union, which eliminates tariffs on imports from Vulcan but maintains a 10 percent tariff on imports from Aurelia. Now trade diversion could occur if the free trade price on imports from Vulcan is lower than the price on imports from Aurelia. Even though Aurelia is the lowest-cost producer, it may be a higher-priced source of imports because of the tariff. If this is the case, then Qualor will stop importing from Aurelia, a non-member, and divert its imports to Vulcan, a member of the RTA. Both trade creation and trade diversion are possible in an RTA. If trade creation is larger than trade diversion, then the net welfare effect is positive. However, there are concerns that this may not always be the case.

The benefits ascribed to free trade—greater specialization according to comparative advantage, reduction in monopoly power because of foreign competition, economies of scale from larger market size, learning by doing, technology transfer, knowledge spillovers, greater foreign investment, and better quality intermediate inputs at world prices—also apply to regional trading blocs. In addition, fostering greater interdependence among members of the regional trading bloc reduces the potential for conflict. Members of the bloc also have greater bargaining power and political clout in the global economy by acting together instead of as individual countries.

The 2009 World Development Report points to spillover of growth across borders as one of the main benefits of regional integration (Collier and O'Connell 2007). There is evidence of considerable spillovers among OECD countries, which are highly integrated both as a group and within their own geographic regions. The long-run growth of integrated countries is interconnected because members have greater access to each other's markets. Strong growth in any RTA country could have a positive impact on growth in other RTA member countries. RTAs also enhance the benefits of good policy and lead to convergence in living standards. For example, growth spillovers are likely to be much smaller among Sub-Saharan African countries because of a lack of integration arising from deficiencies in RTAs and inadequate levels of transportation and telecommunications infrastructure. Roberts and Deichmann (2008) estimated what the cumulative loss in real GDP between 1970 and 2000 would have been if Switzerland, which is landlocked and fully integrated with both its immediate neighbors and the world economy, had been subject to the same level of spillovers as the Central African Republic. Under such a scenario, Switzerland's GDP per capita in 2000 would have been 9.3 percent lower. The cumulative GDP loss would have been $334 billion (constant US dollars, 2000), which was the equivalent of 162 percent of Switzerland's real GDP in 2000.

Although regional integration has many advantages, it may impose costs on some groups. For example, there was significant concern in the United States that NAFTA and especially low-skilled-labor intensive imports from Mexico could hurt low-skilled workers. Adjustment costs arose as import competition caused inefficient firms to exit the market, and the workers in those firms were at least temporarily unemployed as they sought new jobs. However, the surviving firms experienced an increase in productivity, and US consumers benefited from the increase in product varieties imported from Mexico. Feenstra and Taylor (2008) estimated that the product varieties exported from Mexico to the United States had grown by an average of 2.2 percent a year across all industries. They estimated that NAFTA imposed private costs of nearly $5.4 billion a year in the United States during 1994–2002, but that these costs were offset by an average welfare gain of $5.5 billion a year accruing from increased varieties imported from Mexico. Consumer gains from more varieties of products continued over time as long as the imports continued, while adjustment costs arising from job losses declined over time. In 2003, the gain from increased product varieties from Mexico was $11 billion, far exceeding the adjustment costs of $5.4 billion.[19] Their analysis concluded:

> …Thus the consumer gains from increased product variety, when summed over the years, considerably exceed the private loss from displacement. This outcome is guaranteed to occur because the gains from expanded import varieties occur every year that the imports are available, whereas labor displacement is a temporary phenomenon. (Feenstra and Taylor 2008, p. 208)

It is important to recognize, however, that workers displaced by regional integration may have to bear long-term losses if they are unable to find jobs with wages comparable with the jobs they lost or they remain unemployed for a long period. For example, although import competition was certainly not the only factor that led to a dramatic contraction of the US automobile industry, the impact on employment in that industry is likely to be permanent and many former autoworkers, especially older workers, may never find comparable jobs.

---

[19] Feenstra and Taylor (2008) discuss in their book the data limitations and various assumptions they made in their analysis.

Concerns regarding national sovereignty, especially where big and small nations may be part of the same bloc, have also been an impediment to the formation of FTAs. The proposal for a South Asian regional bloc has faced challenges regarding India's role because it is one of the biggest economies in the region.

Regional integration is important from an investment perspective because it offers new opportunities for trade and investment. The cost of doing business in a large, single, regional market is lower and firms can benefit from economies of scale. However, it is important to note that differences in tastes, culture, and competitive conditions still exist among members of a trading bloc. These differences may limit the potential benefits from investments within the bloc. In addition, depending on the level of integration and the safeguards in place, problems faced by individual member countries in an RTA may quickly spread to other countries in the bloc.

There are at least two challenges in the formation of an RTA and in its potential progression from a free trade area to deeper integration in the form of a customs union, common market, or economic union. First, cultural differences and historical considerations—for example, wars and conflicts—may complicate the social/political process of integration. Second, maintaining a high degree of economic integration limits the extent to which member countries can pursue independent economic and social policies. Free trade and mobility of labor and capital tend to thwart policies aimed at controlling relative prices and/or quantities within a country, while balance of payments and fiscal credibility considerations limit the viability of divergent macroeconomic policies. This situation is especially true in the case of a monetary union because monetary policy is not under the control of individual countries and currency devaluation/revaluation is not available as a tool to correct persistent imbalances.[20] When persistent imbalances do arise, they may lead to a crisis that spills over to other countries facing similar problems. A recent example is the fear of contagion caused by the Greek fiscal crisis in 2010. In May 2010, Standard & Poor's reduced the credit ratings on Greece's government from investment grade to junk status. It also downgraded the government debt of Spain and Portugal. These countries were suffering from a combination of high government deficits and slow GDP growth. The credit downgrades increased fears that Greece, in particular, would default on its debt and cause economic turmoil not only among the healthier countries in the EU but also in the United States and Asia. The EU and the International Monetary Fund (IMF) agreed on a USD145 billion (EUR110 billion) bailout for Greece in May 2010, and provided Ireland with a financing package of about USD113 billion (EUR85 billion) in November 2010. As of late 2010, there were continuing concerns about the financial health of Greece, Ireland, Portugal, and Spain. The EU, which created the European Financial Stability Facility (EFSF) in 2010 to help EU countries in need, has been debating the need for an expansion in the scope and financing capacity of the EFSF.

---

[20] These limitations are inherent in any system with fixed exchange rates and a high degree of capital mobility. They are not unique to a monetary union (i.e., a common currency). For a discussion of currency regimes, see the Level I curriculum reading on Currency Exchange Rates.

EXAMPLE 7

Trade Agreements

Bagopia, Cropland, and Technopia decide to enter into an RTA. In the first stage, they decide to sign a free trade agreement (FTA). After several successful years, they decide that it is time to form a common market.

1. Does an FTA make exporting firms in member countries more attractive as investment options?
2. How does the common market affect firms doing business in these countries compared with an FTA?

Solution to 1:

The first stage, where there is free movement of goods and services among RTA members, is called a free trade area. It makes exporting firms a more attractive investment proposition because they are able to serve markets in member countries without the additional costs imposed by trade barriers.

Solution to 2:

Unlike an FTA, a common market allows for free movement of factors of production, such as labor and capital, among the member economies. Like an FTA, it provides access to a much larger market and free movement of goods and services. But the common market can create more profitable opportunities for firms than an FTA by allowing them to locate production in and purchase components from anywhere in the common market according to comparative advantage.

CAPITAL RESTRICTIONS

8

g describe common objectives of capital restrictions imposed by governments;

There are many reasons for governments to restrict inward and outward flow of capital. For example, the government may want to meet some objective regarding employment or regional development, or it may have a strategic or defense-related objective. Many countries require approval for foreigners to invest in their country and for citizens to invest abroad. Control over inward investment by foreigners results in restrictions on how much can be invested, and on the type of industries in which capital can be invested. For example, such strategic industries as defense and telecommunications are often subject to ownership restrictions. Outflow restrictions can include restrictions on repatriation of capital, interest, profits, royalty payments, and license fees. Citizens are often limited in their ability to invest abroad, especially in foreign exchange–scarce economies, and there can be deadlines for repatriation of income earned from any investments abroad.

Economists consider free movement of financial capital to be beneficial because it allows capital to be invested where it will earn the highest return. Inflows of capital also allow countries to invest in productive capacity at a rate that is higher than could be achieved with domestic savings alone, and it can enable countries to achieve a higher rate of growth. Longer-term investments by foreign firms that establish a presence in the local economy can bring in not only much needed capital but also new technology, skills, and advanced production and management practices as well as create spillover benefits for local firms. Investment by foreign firms can also create a network of local suppliers if they source some of their components locally. Such suppliers may receive advanced training and spillover benefits from a close working relationship with the foreign firms. On the one hand, increased competition from foreign firms in the market may force domestic firms to become more efficient. On the other hand, it is possible that the domestic industry may be hurt because domestic firms that are unable to compete are forced to exit the market.

In times of macroeconomic crisis, capital mobility can result in capital flight out of the country, especially if most of the inflow reflects short-term portfolio flows into stocks, bonds, and other liquid assets rather than foreign direct investment (FDI) in productive assets. In such circumstances, capital restrictions are often used in conjunction with other policy instruments, such as fixed exchange rate targets. Capital restrictions and fixed exchange rate targets are complementary instruments because in a regime of perfect capital mobility, governments cannot achieve domestic and external policy objectives simultaneously using only standard monetary and fiscal policy tools.[21] By limiting the free flow of capital, capital controls provide a way to exercise control over a country's external balance whereas more traditional macro-policy tools are used to address other objectives.

Modern capital controls were developed by the belligerents in World War I as a method to finance the war effort. At the start of the war, all major powers restricted capital outflows (i.e., the purchase of foreign assets or loans abroad). These restrictions raised revenues by keeping capital in the domestic economy, facilitating the taxation of wealth, and producing interest income. Moreover, capital controls helped to maintain a low level of interest rates, reducing the government's borrowing costs on its liabilities. Since WWI, controls on capital outflows have been used similarly in other countries, mostly developing nations, to generate revenue for governments or to permit them to allocate credit in the domestic economy without risking capital flight. In broad terms, a capital restriction is any policy designed to limit or redirect capital flows. Such restrictions may take the form of taxes, price or quantity controls, or outright prohibitions on international trade in assets. Price controls may take the form of special taxes on returns to international investment, taxes on certain types of transactions, or mandatory reserve requirements—that is, a requirement forcing foreign parties wishing to deposit money in a domestic bank account to deposit some percentage of the inflow with the central bank for a minimum period at zero interest. Quantity restrictions on capital flows may include rules imposing ceilings or requiring special authorization for new or existing borrowing from foreign creditors. Or there may be administrative controls on cross-border capital movements in which a government agency must approve transactions for certain types of assets.

Effective implementation of capital restrictions may entail non-trivial administration costs, particularly if the measures have to be broadened to close potential loopholes. There is also the risk that protecting the domestic financial markets by capital restrictions may postpone necessary policy adjustments or impede private-sector adaptation to changing international circumstances. Most importantly, controls may give rise to negative market perceptions, which may, in turn, make it more costly and difficult for the country to access foreign funds.

In a study on the effectiveness of capital controls, the International Monetary Fund considered restrictions on capital outflows and inflows separately.[22] The authors concluded that for restrictions on capital inflows to be effective (i.e., not circumvented), the coverage needs to be comprehensive and the controls need to be implemented forcefully. Considerable administrative costs are incurred in continuously extending, amending, and monitoring compliance with the regulations. Although controls on inflows appeared to be effective in some countries, it was difficult to distinguish the impact of the controls from the impact of other policies, such as strengthening of prudential regulations, increased exchange rate flexibility, and adjustment of monetary policy. In the case of capital outflows, the imposition of controls during episodes of financial crisis seems to have produced mixed results, providing only temporary relief of varying duration to some countries, while successfully shielding others (e.g., Malaysia) and providing them with sufficient time to restructure their economies.

---

[21] The Level I curriculum reading on Currency Exchange Rates provides a concise discussion of the policy implications of capital mobility with fixed versus floating exchange rates.

[22] Ariyoshi, et al. (2000).

EXAMPLE 8

Historical Example—Capital Restrictions: Malaysia's Capital Controls in 1998–2001

After the devaluation of the Thai baht in July 1997, Southeast Asia suffered from significant capital outflows that led to falling local equity and real estate prices and declining exchange rates. To counter the outflows of capital, the IMF urged many of the countries in the region to increase interest rates, thus making their assets more attractive to foreign investors. Higher interest rates, however, weighed heavily on the domestic economies. In response to this dilemma, Malaysia imposed capital controls on 1 September 1998. These controls prohibited transfers between domestic and foreign accounts, eliminated credit facilities to offshore parties, prevented repatriation of investment until 1 September 1999, and fixed the exchange rate of the Malaysian ringgit at 3.8 per US dollar. In February 1999, a system of taxes on capital flows replaced the prohibition on repatriation of capital. Although the details were complex, the net effect was to discourage short-term capital flows while permitting long-term transactions. By imposing capital controls, Malaysia hoped to regain monetary independence, and to be able to cut interest rates without provoking a fall in the value of its currency as investors avoided Malaysian assets. The imposition of outflow controls indeed curtailed speculative capital outflows and allowed interest rates to be reduced substantially. At the same time, under the umbrella of the capital controls, the authorities pursued bank and corporate restructuring and achieved a strong economic recovery in 1999 and 2000. With the restoration of economic and financial stability, administrative controls on portfolio outflows were replaced by a two-tier, price-based exit system in February 1999, which was finally eliminated in May 2001. Although Malaysia's capital controls did contribute to a stabilization of its economy, they came with long-term costs associated with the country's removal from the MSCI developed equity market index, an important benchmark in the institutional asset management industry, and its relegation to the emerging market universe. The Malaysian market was no longer seen as on par with developed equity markets whose institutional and regulatory frameworks provide a higher standard of safety for investors. As a consequence, it became more difficult for Malaysia to attract net long-term capital inflows (Kawai and Takagi 2003).

1. Under what economic circumstances were Malaysia's capital restrictions imposed?
2. What was the ultimate objective of Malaysia's capital restrictions?
3. How successful were the country's capital restrictions?

Solution to 1:

As a result of the Southeast Asian crisis, Malaysia suffered substantial net capital outflows pushing up the domestic interest rate level.

Solution to 2:

The restrictions were designed to limit and redirect capital flows to allow the government to reduce interest rates and pursue bank and corporate restructurings.

Solution to 3:

Although the capital controls helped stabilize Malaysia's economy, they contributed to a change in investors' perception of Malaysian financial markets and removal of the Malaysian equity market from the MSCI benchmark universe of developed equity markets. This situation undermined international demand for Malaysian equities and made it more difficult to attract net long-term capital inflows.

9 BALANCE OF PAYMENTS- ACCOUNTS AND COMPONENTS

h describe the balance of payments accounts including their components;

The **balance of payments** (BOP) is a double-entry bookkeeping system that summarizes a country's economic transactions with the rest of the world for a particular period of time, typically a calendar quarter or year. In this context, a transaction is defined as "an economic flow that reflects the creation, transformation, exchange, or extinction of economic value and involves changes in ownership of goods and/or financial assets, the provision of services, or the provision of labour and capital."[23] In other words, the BOP reflects payments for exports and imports as well as financial transactions and financial transfers. Analyzing the BOP is an important element in assessing a country's macroeconomic environment, its monetary and fiscal policies, and its long-term growth potential. Investors use data on trade and capital flows to evaluate a country's overall level of capital investment, profitability, and risk. The following section describes the balance of payments, the factors that influence it, and its impact on exchange rates, interest rates, and capital market transactions.

9.1 Balance of Payments Accounts

The BOP is a double-entry system in which every transaction involves both a debit and credit. In principle, the sum of all debit entries should equal the sum of all credit entries, and the net balance of all entries on the BOP statement should equal zero. In practice, however, this is rarely the case because the data used to record balance of payments transactions are often derived from different sources.

Debit entries reflect purchases of imported goods and services, purchases of foreign financial assets, payments received for exports, and payments (interest and principal) received from debtors. Credit entries reflect payments for imported goods and services, payments for purchased foreign financial assets, and payments to creditors (see Exhibit 13, Panel A). Put differently, a debit represents an increase in a country's assets (the purchase of foreign assets or the receipt of cash from foreigners) or a decrease in its liabilities (the amount owed to foreigners); a credit represents a decrease in assets (the sale of goods and services to foreigners or the payment of cash to foreigners) or an increase in liabilities (an amount owed to foreigners).

---

[23] IMF Balance of Payments Handbook, chapter II, page 6.

For example, as shown in Panel B of Exhibit 13, on 1 September Country A purchases $1 million of goods from Country B and agrees to pay for these goods on 1 December. On 1 September, Country A would record in its BOP a $1 million debit to reflect the value of the goods purchased (i.e., increase in assets) and $1 million credit to reflect the amount owed to Country B. On 1 December, Country A would record in its BOP a $1 million debit to reflect a decrease in the amount owed (liability) to Country B and $1 million a credit to reflect the actual payment to Country B (decrease in assets).

From Country B's perspective, on 1 September it would record in its BOP a $1 million debit to reflect the amount owed by Country A and a $1 million credit to reflect the sale of goods (exports). On 1 December, Country B would record a $1 million debit to reflect the cash received from Country A, and $1 million credit to reflect the fact that it is no longer owed $1 million by Country A.

Exhibit 13 Basic Entries in a BOP Context

Panel A

DEBITS	CREDITS
Increase in Assets, Decrease in Liabilities	**Decrease in Assets, Increase in Liabilities**
■ Value of imported goods and services	■ Payments for imports of goods and services
■ Purchases of foreign financial assets	■ Payments for foreign financial assets
■ Receipt of payments from foreigners	■ Value of exported goods and services
■ Increase in debt owed by foreigners	■ Payment of debt by foreigners
■ Payment of debt owed to foreigners	■ Increase in debt owed to foreigners

Panel B

Country A	Debits	Credits
1 September	$1 million Goods purchased from Country B *(increase in real assets)*	$1 million Short-term liability for goods purchased from Country B *(increase in financial liabilities)*
1 December	$1 million Elimination of short-term liability for goods purchased from Country B *(decrease in financial liabilities)*	$1 million Payment for goods purchased from Country B *(decrease in financial assets)*

Country B	Debits	Credits
1 September	$1 million Short-term claim for goods delivered to Country A *(increase in financial assets)*	$1 million Goods delivered to Country A *(decrease in real assets)*
1 December	$1 million Receipt of payment for goods delivered to Country A *(increase in financial assets)*	$1 million Elimination of claim for goods delivered to Country A *(decrease in financial assets)*

9.2 Balance of Payment Components

The BOP is composed of the **current account** that measures the flow of goods and services, the **capital account** that measures transfers of capital, and the **financial account** that records investment flows. These accounts are further disaggregated into sub-accounts:

Current Account

The current account can be decomposed into four sub-accounts:

1 **Merchandise trade** consists of all commodities and manufactured goods bought, sold, or given away.

2 **Services** include tourism, transportation, engineering, and business services, such as legal services, management consulting, and accounting. Fees from patents and copyrights on new technology, software, books, and movies are also recorded in the services category.

3 **Income receipts** include income derived from ownership of assets, such as dividends and interest payments; income on foreign investments is included in the current account because that income is compensation for services provided by foreign investments. When a German company builds a plant in China, for instance, the services the plant generates are viewed as a service export from Germany to China equal in value to the profits the plant yields for its German owner.

4 **Unilateral transfers** represent one-way transfers of assets, such as worker remittances from abroad to their home country and foreign direct aid or gifts.

Capital Account

The capital account consists of two sub-accounts:

1 **Capital transfers** include debt forgiveness and migrants' transfers (goods and financial assets belonging to migrants as they leave or enter the country).[24] Capital transfers also include the transfer of title to fixed assets and the transfer of funds linked to the sale or acquisition of fixed assets, gift and inheritance taxes, death duties, uninsured damage to fixed assets, and legacies.

2 **Sales and purchases of non-produced, non-financial assets**, such as the rights to natural resources, and the sale and purchase of intangible assets, such as patents, copyrights, trademarks, franchises, and leases.

Financial Account

The financial account can be broken down in two sub-accounts: financial assets abroad and foreign-owned financial assets within the reporting country.

1 A country's assets abroad are further divided into official reserve assets, government assets, and private assets. These assets include gold, foreign currencies, foreign securities, the government's reserve position in the International Monetary Fund,[25] direct foreign investment, and claims reported by resident banks.

---

[24] Immigrants bring with them goods and financial assets already in their possession. Hence, these goods are imported on grounds other than commercial transactions.

[25] These are in effect official currency reserves held with the International Monetary Fund.

2 Foreign-owned assets in the reporting country are further divided into official assets and other foreign assets. These assets include securities issued by the reporting country's government and private sectors (e.g., bonds, equities, mortgage-backed securities), direct investment, and foreign liabilities reported by the reporting country's banking sector.

EXAMPLE 9

US Current Account Balance

Exhibit 14 shows a simplified version of the US balance of payments for 1970–2009.

Exhibit 14 US International Transactions Accounts Data

(Credits+, Debits–)	1970	1980	1990	2000	2009	2017
Current Account						
Exports of goods and services and income receipts	68,387	344,440	706,975	1,421,515	2,159,000	3,279,190
Exports of goods and services	56,640	271,834	535,233	1,070,597	1,570,797	2,351,072
Income receipts	11,747	72,606	171,742	350,918	588,203	928,118
Imports of goods and services and income payments	–59,901	–333,774	–759,290	–1,779,241	–2,412,489	–3,609,734
Imports of goods and services	–54,386	–291,241	–616,097	–1,449,377	–1,945,705	–2,903,349
Income payments	–5,515	–42,532	–143,192	–329,864	–466,783	–706,385
Unilateral current transfers, net	–6,156	–8,349	–26,654	–58,645	–124,943	–118,597
Capital Account						
Capital account transactions, net	....	....	–7,220	–1	–140	–24,746
Financial Account						
US-owned assets abroad, ex derivatives (increase/financial outflow (–))	–9,337	–86,967	–81,234	–560,523	–140,465	–1,182,749
Foreign-owned assets in the United States, ex derivatives (increase/financial inflow (+))	7,226	62,037	139,357	1,038,224	305,736	1,537,682
Financial derivatives, net	NA	NA	NA	NA	50,804	23,074
Statistical discrepancy (sum of above items with sign reversed)	–219	22,613	28,066	–61,329	162,497	–95,880

Based only on the information given, address the following:

1 Calculate the current account balance for each year.

2 Calculate the financial account balance for each year.

3 Describe the long-term change in the current account balance.

4 Describe the long-term change in the financial account balance.

Solutions to 1 and 2:

(Credits+, Debits–)	1970	1980	1990	2000	2009	2017
Current Account	2,330	2,317	–78,969	–416,371	–378,432	–449,141
Financial Account	–2,111	–24,930	58,123	477,701	216,075	378,007

Solution to 3:

The United States had a current account surplus until 1980. After 1990, the US current account had an increasing deficit as a result of strong import growth.

Solution to 4:

Mirroring the growing US current account deficit, the US financial account, after 1990, registered increasing net capital inflows in similar proportions to the deficit in the current account.

10 PAIRED TRANSACTIONS IN THE BOP BOOKKEEPING SYSTEM

h describe the balance of payments accounts including their components;

The following examples illustrate how some typical cross-border transactions are recorded in the BOP framework outlined previously. They include commercial exports and imports, the receipt of income from foreign investments, loans made to borrowers abroad, and purchases of home-country currency by foreign central banks. Exhibit 15 illustrates the various individual bookkeeping entries from the perspective of an individual country, in this case Germany.

10.1 Commercial Exports: Transactions (ia) and (ib)

A company in Germany sells technology equipment to a South Korean auto manufacturer for a total price of EUR50 million, including freight charges of EUR1 million to be paid within 90 days. The merchandise will be shipped via a German cargo ship. In this case, Germany is exporting two assets: equipment and transportation services. The cargo shipped is viewed as being created in Germany and used by South Korean customers. In return for relinquishing these two assets, Germany acquires a financial asset—the promise by the South Korean manufacturer to pay for the equipment in 90 days.

Germany would record a EUR50 million debit to an account called "private short-term claims" to show an increase in this asset. It would also record a credit of EUR49 million to "goods" and another credit of EUR1 million to "services." Both credit entries are listed in the export category and show the decrease in assets available to German residents. These figures are entered as credits on lines 2 and 3 and as a debit on 19 in Exhibit 15 and are marked with (ia) to identify a typical commercial export transaction. To pay for the technology equipment purchased from Germany, the South Korean auto manufacturer may purchase euros from its local bank (i.e., a EUR demand deposit held by the Korean bank in a German bank) and then transfer them to the German exporter. As a result, German liabilities to South Korean residents (i.e., South Korean private short-term claims) would be debited. The respective entries, marked with (ib) are on lines 19 and 23 in Exhibit 15.

10.2 Commercial Imports: Transaction (ii)

A German utility company imports gas from Russia valued at EUR 45 million (ii), and agrees to pay the Russian company within three months. The imported gas generates a debit on line 6. The obligation pay is recorded as a credit to foreign private short-term claims on line 23.

10.3 Loans to Borrowers Abroad: Transaction (iii)

A German commercial bank purchases EUR 100 million in intermediate-term bonds issued by a Ukrainian steel company. The bonds are denominated in euros, so payment is made in euros (i.e., by transferring EUR demand deposits). A debit entry on line 18 records the increase in German holdings of Ukrainian bonds, and a credit entry on line 23 records the increase in demand deposits held by Ukrainians in German banks.

Exhibit 15 Hypothetical Transactions between German Residents and Foreigners

Item no	Account	Debit	Credit	Balance
		–	+	+/–
1	**Exports of goods and services, income received**			**55**
2	Goods		49 (ia)	49
3	Services		1 (ia)	1
4	Income on residents' investments abroad		5 (v)	5
5	**Imports of goods and services, income paid**			**–45**
6	Goods	45 (ii)		–45
7	Services			
8	Income on foreign investments in home country			
9	**Unilateral transfers**			
10	**Changes in residents' claims on foreigners**			**–105**
11	Official reserve assets			
12	Gold			
13	Foreign currency balances			
14	Other			
15	Government claims			
16	Private claims			
17	Direct investments			
18	Other private long-term claims	100 (iii)		–100
19	Private short-term claims	50 (ia), 5 (v)	50 (ib)	–5
20	**Changes in foreign claims on residents**			**195**
21	Foreign official claims		20 (iv)	20
22	Foreign private long-term claims			
23	Foreign private short-term claims	20 (iv), 50 (ib)	45 (ii), 100 (iii), 100 (vi)	175
24	**Other**	100 (vi)		**–100**
	Total	~~270~~ 370	~~270~~ 370	**0**
	Current Account: (1) + (5) + (9)			**10**

(continued)

Exhibit 15 (Continued)

Item no	Account	Debit	Credit	Balance
		–	+	+/–
	Capital Account: (24)			**–100**
	Financial Account: (10) + (20)			**90**

10.4 Purchases of Home-Country Currency by Foreign Central Banks: Transaction (iv)

Private foreigners may not wish to retain euro balances acquired in earlier transactions. Those who are holding foreign currency, in our example euro claims, typically do so for purposes of financing purchases from Germany (or other euro area member countries). Assume for instance, that Swiss residents attempt to sell EUR20 million in exchange for their native currency, the Swiss franc (CHF), but there is a lack of demand for EUR funds in Switzerland. In such circumstances, the CHF would appreciate against the EUR. To prevent an undesired CHF appreciation, the Swiss National Bank (SNB) might sell CHF in exchange for EUR balances.

Suppose that the Swiss National Bank purchased EUR20 million, typically in the form of a EUR demand deposit held with a German bank, from local commercial banks in Switzerland. The German BOP would register an increase of EUR20 million in German liabilities held by foreign monetary authorities, the Swiss National Bank (line 21), and an equivalent decline in short-term liabilities held by private foreigners (i.e., Swiss private investors, line 23). It may be noteworthy that when the SNB purchases EUR funds from Swiss commercial banks, it also credits them the CHF equivalent of EUR20 million. The SNB's liabilities to Swiss commercial banks arising from this transaction are in fact reserve deposits that Swiss banks can use when they expand their lending business and create new deposits. Currency interventions by central banks, therefore, can contribute to an increase in a country's overall money supply, all else remaining unchanged.

10.5 Receipts of Income from Foreign Investments: Transaction (v)

Each year, residents of Germany receive billions of EUR in interest and dividends from capital invested in foreign securities and other financial claims. German residents receive these payments in return for allowing foreigners to use German capital that otherwise could be put to work in Germany. Foreign residents, in turn, receive similar returns for their investments in Germany. Assume that a German firm has a long-term capital investment in a profitable subsidiary abroad, and that the subsidiary transfers to its German parent EUR5 million in dividends in the form of funds held in a foreign bank. The German firm then has a new (or increased) demand deposit in a foreign bank as compensation for allowing its capital to be used by its subsidiary. A debit entry on line 19 shows German private short-term claims on foreigners have increased by EUR5 million, and a credit entry on line 4 reflects the fact that German residents have given up an asset (the services of capital covered over the period) valued at EUR5 million.

10.6 Purchase of Non-financial Assets: Transaction (vi)

In a move to safeguard its long-term supply of uranium, a German utility company purchases the rights to exploit a uranium mine from the government of Kazakhstan. It agrees to pay within three months. The respective entries are on lines 23 and 24. Because a non-financial, non-produced asset is involved in this transaction, it is recorded in Germany's capital account.

Note that the sum of all BOP entries in Exhibit 15 is 0. Transactions (i)–(iv) produce a current account surplus of EUR10 million, a capital account deficit of EUR100 million, and a financial account surplus of EUR90 million.

Although it is important to understand the detailed structure of official balance of payments accounts as described in the preceding paragraphs, this example is not necessarily how investment professionals think about the balance of payments day-to-day. Practitioners often think of the current account as roughly synonymous with the trade balance (merchandise trade + services) and lump all the financing flows (financial account + capital account) into one category that is usually referred to simply as the "capital account." They then think of the capital account as consisting of two types of flows—portfolio investment flows and foreign direct investment (FDI). The former are shorter-term investments in foreign assets (stocks, bonds, etc.), whereas the latter are long-term investments in production capacity abroad. Although not completely accurate, this way of thinking about the balance of payments focuses attention on the components—trade, portfolio flows, and FDI—that are most sensitive to, and most likely to affect, market conditions, prices of goods and services, asset prices, and exchange rates. In addition, this perspective fits well with the role that the balance of payments plays in the macroeconomy.

NATIONAL ECONOMIC ACCOUNTS AND THE BALANCE OF PAYMENTS

11

i. explain how decisions by consumers, firms, and governments affect the balance of payments;

In a closed economy, all output Y is consumed or invested by the private sector—domestic households and businesses—or purchased by the government. Letting Y denote GDP, C private consumption, I investment, and G government purchases of goods and services, the national income identity for a closed economy is given by:

$$Y = C + I + G \tag{1}$$

Once foreign trade is introduced, however, some output is purchased by foreigners (exports) whereas some domestic spending is used for purchases of foreign goods and services (imports). The national income identity for an open economy is thus

$$Y = C + I + G + X - M \tag{2}$$

where X denotes exports and M denotes imports.

For most countries, exports rarely equal imports. Net exports or the difference between exports and imports ($X - M$) is the equivalent of the current account balance from a BOP perspective.[26] When a country's imports exceed its exports, the current account is in deficit. When a country's exports exceed its imports, the current account is in surplus. As the right side of Equation 2 shows, a current account surplus or deficit can affect GDP (and also employment). The balance of the current account is also important because it measures the size and direction of international borrowing.

In order for the balance of payments to balance, a deficit or surplus in the current account must be offset by an opposite balance in the sum of the capital and financial accounts. This requirement means that a country with a current account deficit has to increase its net foreign debts by the amount of the current account deficit. For example, the United States has run current account deficits for many years while accumulating net foreign liabilities: The current account deficit was financed by net capital imports (i.e., direct investments by foreigners), loans by foreign banks, and the sale of US equities and fixed-income securities to foreign investors. By the same token, an economy with a current account surplus is earning more for its exports than it spends for its imports. Japan, Germany, and China are traditional current account surplus countries accumulating substantial net foreign claims, especially against the United States. An economy with a current account surplus finances the current account deficit of its trading partners by lending to them—that is, granting bank loans and investing in financial and real assets. As a result, the foreign wealth of a surplus country rises because foreigners pay for imports by issuing liabilities that they will eventually have to redeem.

By rearranging Equation 2, we can define the current account balance from the perspective of the national income accounts as:

$$CA = X - M = Y - (C + I + G) \tag{3}$$

Only by borrowing money from foreigners can a country have a current account deficit and consume more output than it produces. If it consumes less output than it produces, it has a current account surplus and can (indeed must) lend the surplus to foreigners. International capital flows essentially reflect an *inter-temporal trade*. An economy with a current account deficit is effectively importing present consumption and exporting future consumption.

Let us now turn to the relationship between output Y and disposable income Y^d. We have to recognize that part of income is spent on taxes T, and that the private sector receives net transfers R in addition to (national) income. Disposable income Y^d is thus equal to income plus transfers minus taxes:

$$Y^d = Y + R - T \tag{4}$$

Disposable income, in turn, is allocated to consumption and saving so that we can write

$$Y^d = C + S_p \tag{5}$$

where S_p denotes private sector saving. Combining Equations 4 and 5 allows us to write consumption as income plus transfers minus taxes and saving.

$$C = Y^d - S_p = Y + R - T - S_p \tag{6}$$

---

[26] Strictly speaking, net exports as defined here is the trade balance rather than the current account balance because it excludes income receipts and unilateral transfers. This distinction arises because we have defined income Y as GDP rather than GNP (see section 1.1.2). Because the trade balance is usually the dominant component of the current account, the terms "trade balance" and "current account" are often used interchangeably. We will do so here unless the distinction is important to the discussion.

We can now use the right side of Equation 6 to substitute for C in Equation 3. With some rearrangement we obtain

$$CA = S_p - I + (T - G - R) \tag{7}$$

Because $(T - G - R)$ is taxes minus government spending and transfers, it is the government surplus, or put differently, government savings S_g. Equation 7 can therefore be restated as

$$S_p + S_g = I + CA \tag{8}$$

Equation 8 highlights an essential difference between open and closed economies: An open economy can use its saving for domestic investment or for foreign investment (i.e., by exporting its savings and acquiring foreign assets), while in a closed economy savings can only be used for domestic investment. Put another way, an open economy with promising investment opportunities is not constrained by its domestic savings rate in order to exploit these opportunities. As Equation 8 shows, it can raise investment by increasing foreign borrowing (a reduction in CA) without increasing domestic savings. For example, if India decides to build a network of high-speed trains, it can import all the required materials it needs from France and then borrow the funds, perhaps also from France, to pay for the materials. This transaction increases India's domestic investment because the imported materials contribute to the expansion in the country's capital stock. All else being equal, this transaction will also produce a current account deficit for India by an amount equal to the increase in investment. India's savings does not have to increase, even though investment increases. This example can be interpreted as an inter-temporal trade, in which India imports present consumption (when it borrows to fund current expenditure) and exports future consumption (when it repays the loan).

Rearranging Equation 8, we can write

$$S_p = I + CA - S_g \tag{9}$$

Equation 9 states that an economy's private savings can be used in three ways: (1) investment in domestic capital (I), (2) purchases of assets from foreigners (CA), and (3) net purchases (or redemptions) of government debt ($-S_g$).

Finally, we can rearrange Equation 8 again to illustrate the macroeconomic sources of a current account imbalance:

$$CA = S_p + S_g - I \tag{10}$$

A current account deficit tends to result from low private savings, high private investment, a government deficit ($S_g < 0$), or a combination of the three. Alternatively, a current account surplus reflects high private savings, low private investment, or a government surplus.

As outlined above, trade deficits can result from a lack of private or government savings or booming investments. If trade deficits primarily reflect high private or government consumption (i.e., scarce savings = $S_p + S_g$), the deficit country's capacity to repay its liabilities from future production remains unchanged. If a trade deficit primarily reflects strong investments (I), however, the deficit country can increase its productive resources and its ability to repay its liabilities.

We can also see from Equation 3 that a current account deficit tends to reflect a strong domestic economy (elevated consumer, government, and investment spending), which is usually accompanied by elevated domestic credit demand and high interest rates. In such an environment, widening interest rate differentials vis-à-vis other countries can lead to growing net capital imports and produce an appreciating currency. In the long run, however, a persistent current account deficit leads to a permanent increase in the claims held by other countries against the deficit country. As a result, foreign investors may require rising risk premiums for such claims, a process that appears to lead to a depreciating currency.

EXAMPLE 10

Historical Example: The United Kingdom Budget

A financial newspaper had the following item:

> The UK's budget deficit is the highest in the G–20; in Europe, only Ireland borrows more. These are the stark facts facing Chancellor of the Exchequer George Osborne as he plans his first Budget tomorrow. He intends to tackle the problem even if that involves severe spending cuts and large tax increases.
>
> *Source*: *Financial Times*, 21 June 2010.

1 What are the likely consequences for the UK current account balance from the planned fiscal policy moves mentioned in the above article?

2 Describe the impact spending cuts and tax increases are likely to have on UK imports.

Solution to 1:

The combination of spending cuts and tax increases will, all else the same, lead to an improvement in the UK current account position.

Solution to 2:

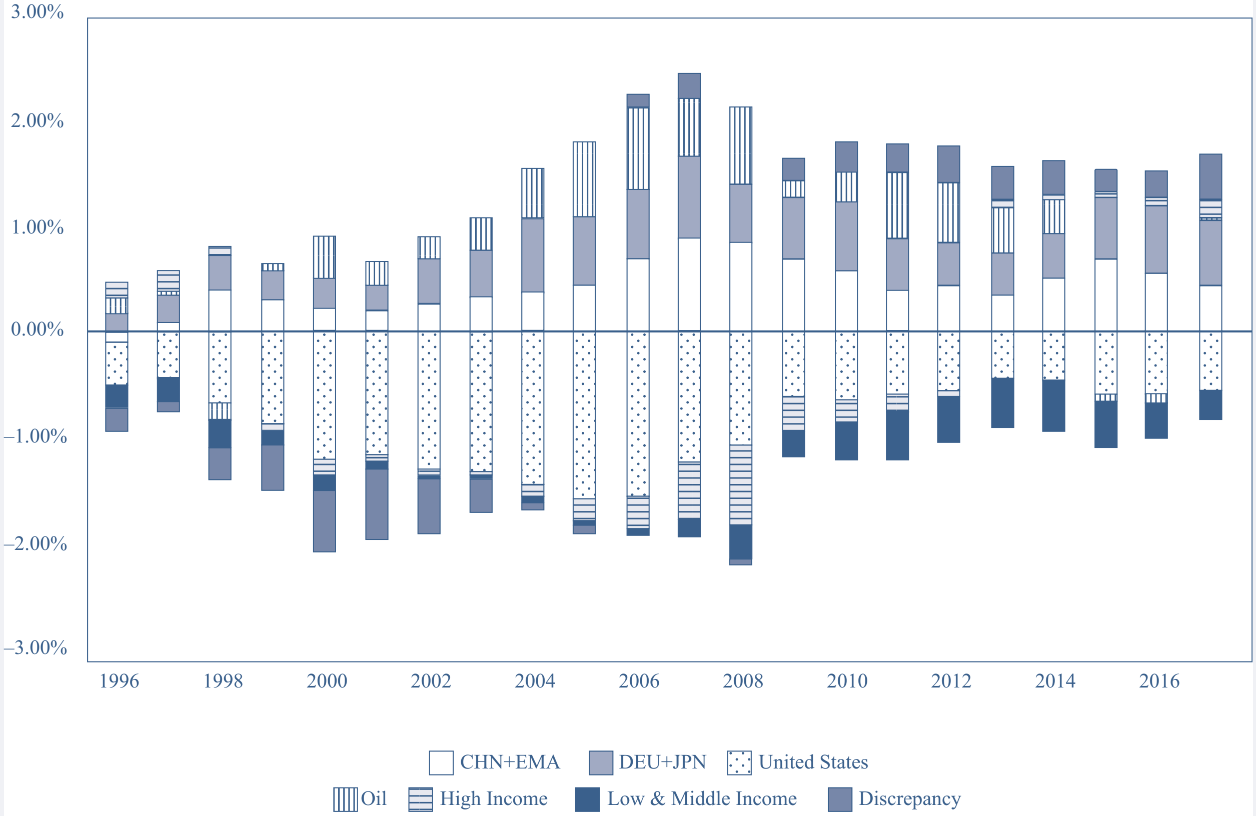
UK imports are likely to be reduced by tax increases and spending cuts because government demand for foreign goods will fall and growth in private household income, which finances private imports, will be restricted as more household income goes to taxes.

EXAMPLE 11

Global Current Account Imbalances since 1996

As a result of growing financial integration and trade liberalization, the world economy has entered a period of rapid growth in cross-border trade since the late 1980s. In synch with surging international trade, current account imbalances widened substantially in the 1990s and the first decade of the new millennium. Exhibit 16 shows current account balances for 1996–2017 for five specific groups—the United States, the top 20 Oil exporting countries as of 2016, Germany and Japan (DEU + JPN), China and emerging Asia (CHN + EMA)—and two broad categories: all other High Income countries and all remaining Low and Middle Income countries. The United States ran a current account deficit in every year, and in every year its deficit represented most of the aggregate value of such deficits worldwide. Only after the 2007–2009 recession has the US deficit declined both in absolute terms and relative to the global aggregate of current account deficits. In the first half of the 1990s, Germany and Japan were the traditional current account surplus countries, providing net exports of goods and services to and accumulating net claims against the United States. Since the late 1990s, among the largest current account surpluses are those of China and emerging Asia. Oil exporting countries, who have traditionally had significant current account surpluses, saw a change to current account deficits in 2015.

Exhibit 16 Global Imbalances (Current Account Balance in Percent of World GDP)



Note: 1. CHN+EMA includes the following economies: Chinese mainland, Hong Kong SAR, Indonesia, Korea, Malaysia, Philippines, Singapore, Thailand.

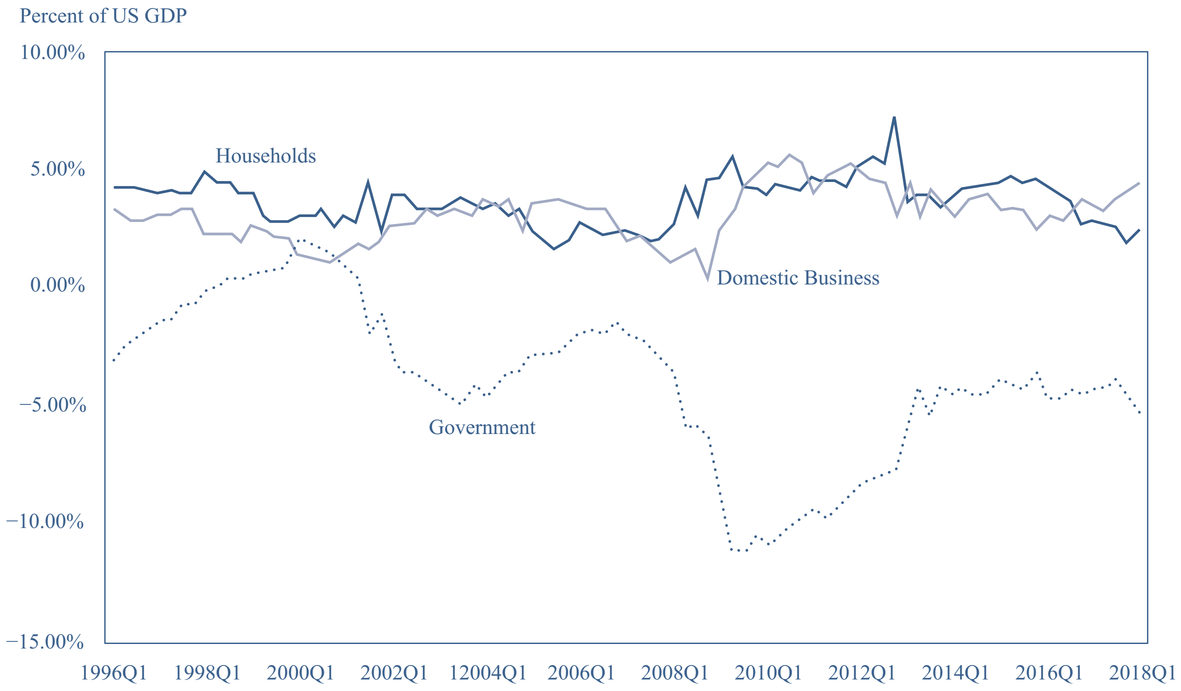
2. The Oil group consists of the 20 largest oil exporting countries in 2016—excluding the United States—as listed in the Economist Intelligence Unit World Factbook 2017.

3. The High Income as well as the Low and Middle Income groups' definitions include all countries that do not fall into any of the other categories.

Source: IMF *Balance of Payments Statistics Yearbook, 2018.*

As illustrated in Equation 10, current account deficits or surpluses reflect imbalances between national savings (including government savings) and investments. Current account deficits are often related to expansionary fiscal policy and government deficits. In the 1980s, for instance, the growing deficit in the US current account was widely seen as the consequence of tax cuts and rising defense spending adopted by the Reagan administration. Since the mid-1990s, however, the current account imbalances depicted in Exhibit 16 appear to reflect other, more complex factors. Exhibit 17 illustrates US net savings (S–I) for private domestic businesses, households and non-profits, and the government (i.e., federal, state, and local) from the first quarter of 1996 to the first quarter of 2018. The exhibit indicates that business sector net savings and government net savings as a percentage of GDP were near mirror-images between 1996 to 2008. During the technology bubble businesses invested heavily and ran progressively larger savings deficits while the government moved to a surplus. After the bubble burst the pattern reversed with businesses moving to net positive savings and the government fiscal balance deteriorating sharply. Meanwhile, the household sector gradually reduced its savings rate. After the global financial crisis in late 2008, households and businesses cut spending and increased savings sharply while the government deficit exploded to more than 12 percent of GDP. As the economy recovered from the financial crisis from 2012 onwards, government borrowing fell to five percent of GDP. Both households and business maintained a higher rate of saving during this period than before the financial crisis.

Exhibit 17 United States: Sectorial Saving–Investment Balance since 1996 (Net Savings in Percent of GDP)



Source: Federal Reserve Board, flow-of-funds data.

12 TRADE ORGANIZATIONS

j describe functions and objectives of the international organizations that facilitate trade, including the World Bank, the International Monetary Fund, and the World Trade Organization.

During the Great Depression in the 1930s, countries attempted to support their failing economies by sharply raising barriers to foreign trade, devaluing their currencies to compete against each other for export markets, and restricting their citizens' freedom to hold foreign exchange. These attempts proved to be self-defeating. World trade declined dramatically and employment and living standards fell sharply in many countries. By the 1940s, it had become a wide-spread conviction that the world economy was in need of organizations that would help promote international economic cooperation. In July 1944, during the United Nations Monetary and Financial Conference in Bretton Woods, New Hampshire, representatives of 45 governments agreed on a framework for international economic cooperation. Two crucial, multinational organizations emanated from this conference—the World Bank, which was founded during the conference, and the International Monetary Fund (IMF), which came into formal existence in December 1945. Although the IMF was founded with the goal to stabilize exchange rates and assist the reconstruction of the world's international payment system, the World Bank was created to facilitate post-war reconstruction and development.

A third institution, the International Trade Organization (ITO), was to be created to handle the trade side of international economic cooperation, joining the other two "Bretton Woods" institutions. The draft ITO charter was ambitious, extending beyond world trade regulations to include rules on employment, commodity agreements, restrictive business practices, international investment, and services. The objective was to create the ITO at a United Nations Conference on Trade and Employment in Havana, Cuba in 1947. Meanwhile, 15 countries had begun negotiations in December 1945 to reduce and regulate customs tariffs. With World War II only barely ended, they wanted to give an early boost to trade liberalization and begin to correct the legacy of protectionist measures that had remained in place since the early 1930s. The group had expanded to 23 nations by the time the deal was signed on 30 October 1947 and the General Agreement on Tariffs and Trade (GATT) was born. The Havana conference began on 21 November 1947, less than a month after GATT was signed. The ITO charter was finally approved in Havana in March 1948, but ratification in some national legislatures proved impossible. The most serious opposition was in the US Congress, even though the US government had been one of the driving forces. In 1950, the United States government announced that it would not seek congressional ratification of the Havana Charter, and the ITO was effectively dead. As a consequence, the GATT became the only multilateral instrument governing international trade from 1948 until the World Trade Organization (WTO) was officially established in 1995.

12.1 International Monetary Fund

As we saw earlier, current account deficits reflect a shortage of net savings in an economy and can be addressed by policies designed to rein in domestic demand. This approach could, however, have adverse consequences for domestic employment. The IMF stands ready to lend foreign currencies to member countries to assist them during periods of significant external deficits. A pool of gold and currencies contributed by members provides the IMF with the resources required for these lending operations. The funds are only lent under strict conditions and borrowing countries' macroeconomic policies are continually monitored. The IMF's main mandate is to ensure the stability of the international monetary system, the system of exchange rates and international payments that enables countries to buy goods and services from each other. More specifically, the IMF:

- provides a forum for cooperation on international monetary problems;
- facilitates the growth of international trade and promotes employment, economic growth, and poverty reduction;
- supports exchange rate stability and an open system of international payments; and
- lends foreign exchange to members when needed, on a temporary basis and under adequate safeguards, to help them address balance of payments problems.

The global financial crisis of 2007–2009 demonstrated that domestic and international financial stability cannot be taken for granted, even in the world's most developed countries. In light of these events, the IMF has redefined and deepened its operations by:[27]

- *enhancing its lending facilities*: The IMF has upgraded its lending facilities to better serve its members. As part of a wide-ranging reform of its lending practices, it has also redefined the way it engages with countries on issues related to structural reform of their economies. In this context, it has doubled member countries' access to fund resources and streamlined its lending approach to reduce the stigma of borrowing for countries in need of financial help.
- *improving the monitoring of global, regional, and country economies*: The IMF has taken several steps to improve economic and financial surveillance, which is its framework for providing advice to member countries on macroeconomic policies and warning member countries of risks and vulnerabilities in their economies.
- *helping resolve global economic imbalances*: The IMF's analysis of global economic developments provides finance ministers and central bank governors with a common framework for discussing the global economy.
- *analyzing capital market developments*: The IMF is devoting more resources to the analysis of global financial markets and their links with macroeconomic policy. It also offers training to country officials on how to manage their financial systems, monetary and exchange regimes, and capital markets.
- *assessing financial sector vulnerabilities*: Resilient, well-regulated financial systems are essential for macroeconomic stability in a world of ever-growing capital flows. The IMF and the World Bank jointly run an assessment program aimed at alerting countries to vulnerabilities and risks in their financial sectors.

From an investment perspective, the IMF helps to keep country-specific market risk and global systemic risk under control. The Greek sovereign debt crisis, which threatened to destabilize the entire European banking system, is a recent example. In early 2010, the Greek sovereign debt rating was downgraded to non-investment grade by leading rating agencies as a result of serious concerns about the sustainability of Greece's public sector debt load. Yields on Greek government bonds rose substantially following the downgrading and the country's ability to refinance its national debt was seriously questioned in international capital markets. Bonds issued by some other European governments fell and equity markets worldwide declined in response to spreading concerns of a Greek debt default. The downgrading of Greek sovereign debt was the ultimate consequence of persistent and growing budget deficits the Greek government had run before and after the country had joined the European Monetary Union (EMU) in 2001. Most of the budget shortfalls reflected elevated outlays for public-sector jobs, pensions, and other social benefits as well as persistent tax evasion. Reports that the Greek government had consistently and deliberately misreported the country's official economic and budget statistics contributed to further erosion of confidence in Greek government bonds in international financial markets. Facing default, the Greek government requested that a joint European Union/IMF bailout package be activated, and a loan agreement was reached between Greece, the other EMU member countries, and the IMF. The deal consisted of an immediate EUR45 billion in loans to be provided in 2010, with more funds available later. A total of EUR110 billion was agreed depending on strict economic policy conditions that included cuts in wages and benefits, an increase in the retirement age for public-sector employees, limits on public pensions, increases in direct and indirect taxes, and a substantial reduction in state-owned companies. By providing conditional emergency lending facilities to the Greek government and designing a joint program with the European Union on how to achieve fiscal consolidation, the IMF prevented a contagious wave of sovereign debt crises in global capital markets.

---

[27] Visit www.imf.org/ for more information.

Another example of IMF activities is the East Asian Financial Crisis in the late 1990s. It began in July 1997, when Thailand was forced to abandon its currency's peg with the US dollar. Currency devaluation subsequently hit other East Asian countries that had similar balance of payment problems, such as South Korea, Malaysia, the Philippines, and Indonesia. They had run persistent and increasing current account deficits, financed mainly with short-term capital imports, in particular, domestic banks borrowing in international financial markets. External financing was popular because of the combination of lower foreign, especially US, interest rates and fixed exchange rates. Easy money obtained from abroad led to imprudent investment, which contributed to overcapacities in several industries and inflated prices on real estate and stock markets. The IMF came to the rescue of the affected countries with considerable loans, accompanied by policies designed to control domestic demand, which included fiscal austerity and tightened monetary reins.

12.2 World Bank Group

The World Bank's main objective is to help developing countries fight poverty and enhance environmentally sound economic growth. For developing countries to grow and attract business, they have to

- strengthen their governments and educate their government officials;
- implement legal and judicial systems that encourage business;
- protect individual and property rights and honour contracts;
- develop financial systems robust enough to support endeavours ranging from micro credit to financing larger corporate ventures; and
- combat corruption.

Given these targets, the World Bank provides funds for a wide range of projects in developing countries worldwide and financial and technical expertise aimed at helping those countries reduce poverty.

The World Bank's two closely affiliated entities—the International Bank for Reconstruction and Development (IBRD) and the International Development Association (IDA)—provide low or no-interest loans and grants to countries that have unfavourable or no access to international credit markets. Unlike private financial institutions, neither the IBRD nor the IDA operates for profit. The IBRD is market-based, and uses its high credit rating to pass the low interest it pays for funds on to its borrowers—developing countries. It pays for its own operating costs because it does not look to outside sources to furnish funds for overhead.

IBRD lending to developing countries is primarily financed by selling AAA-rated bonds in the world's financial markets. Although the IBRD earns a small margin on this lending, the greater proportion of its income comes from lending out its own capital. This capital consists of reserves built up over the years and money paid in from the Bank's 185 member country shareholders. IBRD's income also pays for World Bank operating expenses and has contributed to IDA and debt relief. IDA is the world's largest source of interest-free loans and grant assistance to the poorest countries. IDA's funds are replenished every three years by 40 donor countries. Additional funds are regenerated through repayments of loan principal on 35-to-40-year, no-interest loans, which are then available for re-lending. At the end of September 2010, the IBRD had net loans outstanding of USD125.5 billion, while its borrowings amounted to USD132 billion.

Besides acting as a financier, the World Bank also provides analysis, advice, and information to its member countries to enable them to achieve the lasting economic and social improvements their people need. Another of the Bank's core functions is to increase the capabilities of its partners, people in developing countries, and its own staff. Links to a wide range of knowledge-sharing networks have been set up by the Bank to address the vast need for information and dialogue about development.

From an investment perspective, the World Bank helps to create the basic economic infrastructure that is essential for the creation of domestic financial markets and a well-functioning financial industry in developing countries. Moreover, the IBRD is one of the most important supranational borrowers in the international capital markets. Because of its strong capital position and its very conservative financial, liquidity, and lending policies, it enjoys the top investment-grade rating from the leading agencies and investors have confidence in its ability to withstand adverse events. As a result, IBRD bonds denominated in various major currencies are widely held by institutional and private investors.

12.3 World Trade Organization

The WTO provides the legal and institutional foundation of the multinational trading system. It is the only international organization that regulates cross-border trade relationships among nations on a global scale. It was founded on 1 January 1995, replacing the General Agreement on Tariffs and Trade (GATT) that had come into existence in 1947. The GATT was the only multilateral body governing international trade from 1947 to 1995. It operated for almost half a century as a quasi-institutionalized, provisional system of multilateral treaties. Several rounds of negotiations took place under the GATT, of which the Tokyo round and the Uruguay round may have been the most far reaching. The Tokyo round was the first major effort to address a wide range of non-tariff trade barriers, whereas the Uruguay round focused on the extension of the world trading system into several new areas, particularly trade in services and intellectual property, but also to reform trade in agricultural products and textiles. The GATT still exists in an updated 1994 version and is the WTO's principal treaty for trade in goods. The GATT and the General Agreement on Trade in Services (GATS) are the major agreements within the WTO's body of treaties that encompasses a total of about 60 agreements, annexes, decisions, and understandings.

In November 2001, the most recent and still ongoing round of negotiations was launched by the WTO in Doha, Qatar. The Doha round was an ambitious effort to enhance globalization by slashing barriers and subsidies in agriculture and addressing a wide range of cross-border services. So far, under GATS, which came into force in January 1995, banks, insurance companies, telecommunication firms, tour operators, hotel chains, and transport companies that want to do business abroad can enjoy the same principles of free and fair trade that had previously applied only to international trade in goods. No final agreement has been reached in the Doha round as of mid-2018, however, it marked one of the most crucial events in global trade over the past several decades: China's accession to the WTO in December 2001. The inability to reach agreement in the Doha round has led to an increasing number of bilateral and multi-lateral trade agreements, such as the Trans-Pacific Partnership with Japan, Vietnam and nine other countries.

The WTO's most important functions are the implementation, administration, and operation of individual agreements; acting as a platform for negotiations; and settling disputes. Moreover, the WTO has the mandate to review and propagate its members' trade policies and ensure the coherence and transparency of trade policies through surveillance in a global policy setting. The WTO also provides technical cooperation and training to developing, least-developed, and low-income countries to assist with their adjustment to WTO rules. In addition, the WTO is a major source of economic research and analysis, producing ongoing assessments of global trade in its publications and research reports on special topics. Finally, the WTO is in close cooperation with the other two Bretton Woods institutions, the IMF and the World Bank.

From an investment perspective, the WTO's framework of global trade rules provides the major institutional and regulatory base without which today's global multinational corporations would be hard to conceive. Modern financial markets would look different without the large, multinational companies whose stocks and bonds have become key elements in investment portfolios. In the equity universe, for instance, investment considerations focusing on global sectors rather than national markets would make little sense without a critical mass of multinational firms competing with each other in a globally defined business environment.

EXAMPLE 12

Historical Example: Function and Objective of International Organizations

On 10 May 2010, the Greek government officially applied for emergency lending facilities extended by the International Monetary Fund. It sent the following letter to the IMF's Managing Director:

Request for Stand-By Arrangement
This paper was prepared based on the information available at the time it was completed on Monday, May 10, 2010. The views expressed in this document are those of the staff team and do not necessarily reflect the views of the government of Greece or the Executive Board of the IMF. The policy of publication of staff reports and other documents by the IMF allows for the deletion of market-sensitive information.

May 3, 2010
Managing Director
International Monetary Fund
Washington DC

The attached Memorandum of Economic and Financial Policies (MEFP)[28] outlines the economic and financial policies that the Greek government and the Bank of Greece, respectively, will implement during the remainder of 2010 and in the period 2011–2013 to strengthen market confidence and Greece's fiscal and financial position during a difficult transition period toward a more open and competitive economy. The government is fully committed to the policies stipulated in this document and its attachments, to frame tight budgets in the coming years with the aim to reduce the fiscal deficit to below 3 percent in 2014 and achieve a downward trajectory in the public debt-GDP ratio beginning in 2013, to safeguard the stability of the Greek financial system, and to implement structural reforms to boost competitiveness and the economy's capacity to produce, save, and export. (…) The government is strongly determined to lower the fiscal deficit, (…) by achieving higher and more equitable tax collections,

---

[28] The detailed memorandum is available from www.imf.org/external/pubs/ft/scr/2010/cr10111.pdf.

and constraining spending in the government wage bill and entitlement outlays, among other items. In view of these efforts and to signal the commitment to effective macroeconomic policies, the Greek government requests that the Fund supports this multi-year program under a Stand-By Arrangement (SBA) for a period of 36 months in an amount equivalent to SDR26.4 billion.[29] (…) A parallel request for financial assistance to euro area countries for a total amount of €80 billion has been sent. The implementation of the program will be monitored through quantitative performance criteria and structural benchmarks as described in the attached MEFP and Technical Memorandum of Understanding (TMU). There will be twelve quarterly reviews of the program supported under the SBA by the Fund, (….) to begin with the first review that is expected to be completed in the course of the third calendar quarter of 2010, and then every quarter thereafter until the last quarterly review envisaged to be completed during the second calendar quarter of 2013, to assess progress in implementing the program and reach understandings on any additional measures that may be needed to achieve its objectives. (….) The Greek authorities believe that the policies set forth in the attached memorandum are adequate to achieve the objectives of the economic program, and stand ready to take any further measures that may become appropriate for this purpose. The authorities will consult with the Fund in accordance with its policies on such consultations, (….) and in advance of revisions to the policies contained in the MEFP. All information requested by the Fund (….) to assess implementation of the program will be provided.
(….)
Sincerely,

George Papaconstantinou	George Provopoulos
Minister of Finance	Governor of the Bank of Greece

1. What is the objective of the IMF's emergency lending facilities?
2. What are the macroeconomic policy conditions under which the IMF provides emergency lending to Greece?
3. What is the amount Greece requests from the IMF as emergency funds?

Solution to 1:

The program seeks to safeguard the stability of the Greek financial system and to implement structural reforms to boost competitiveness and the economy's capacity to produce, save and export.

Solution to 2:

The Greek government has to reduce the country's fiscal deficit by achieving higher and more equitable tax collections as well as constrain spending in the government wage bill and entitlement outlays.

Solution to 3:

Greece applied for a standby arrangement in an amount equivalent to SDR26.4 billion (approximately USD39.5 billion, based on the 10 May 2010 exchange rate).

---

[29] A SDR (special drawing right) is a basket of four leading currencies: Japanese yen (JPY), US dollar (USD), British pound (GBP), and euro (EUR). It consists of 18.4 yen, 0.6320 USD, 0.0903 GBP, and 0.41 EUR. One SDR was worth 1.4975 USD or 1.1547 EUR on 10 May 2010.

SUMMARY

This reading provides a framework for analyzing a country's trade and capital flows and their economic implications. It examines basic models that explain trade based on comparative advantage and provides a basis for understanding how international trade can affect the rate and composition of economic growth as well as the attractiveness of investment in various sectors.

- The benefits of trade include
  - gains from exchange and specialization;
  - gains from economies of scale as companies add new markets for their products;
  - greater variety of products available to households and firms; and
  - increased competition and more efficient allocation of resources.
- A country has an absolute advantage in producing a good (or service) if it is able to produce that good at a lower absolute cost or use fewer resources in its production than its trading partner. A country has a comparative advantage in producing a good if its *opportunity cost* of producing that good is less than that of its trading partner.
- Even if a country does not have an absolute advantage in the production of any good, it can gain from trade by producing and exporting the good(s) in which it has a comparative advantage and importing good(s) in which it has a comparative disadvantage.
- In the Ricardian model of trade, comparative advantage and the pattern of trade are determined by differences in technology between countries. In the Heckscher–Ohlin model of trade, comparative advantage and the pattern of trade are determined by differences in factor endowments between countries. In reality, technology and factor endowments are complementary, not mutually exclusive, determinants of trade patterns.
- Trade barriers prevent the free flow of goods and services among countries. Governments impose trade barriers for various reasons including: to promote specific developmental objectives, to counteract certain imperfections in the functioning of markets, or to respond to problems facing their economies.
- For purposes of international trade policy and analysis, a small country is defined as one that cannot affect the world price of traded goods. A large country's production and/or consumption decisions do alter the relative prices of trade goods.
- In a small country, trade barriers generate a net welfare loss arising from distortion of production and consumption decisions and the associated inefficient allocation of resources.
- Trade barriers can generate a net welfare gain in a large country if the gain from improving its terms of trade (higher export prices and lower import prices) more than offsets the loss from the distortion of resource allocations. However, the large country can only gain if it imposes an even larger welfare loss on its trading partner(s).
- An import tariff and an import quota have the same effect on price, production, and trade. With a quota, however, some or all of the revenue that would be raised by the equivalent tariff is instead captured by foreign producers (or the foreign government) as quota rents. Thus, the welfare loss suffered by the importing country is generally greater with a quota.

- A voluntary export restraint is imposed by the exporting country. It has the same impact on the importing country as an import quota from which foreigners capture all of the quota rents.
- An export subsidy encourages firms to export their product rather than sell it in the domestic market. The distortion of production, consumption, and trade decisions generates a welfare loss. The welfare loss is greater for a large country because increased production and export of the subsidized product reduces its global price—that is, worsens the country's terms of trade.
- Capital restrictions are defined as controls placed on foreigners' ability to own domestic assets and/or domestic residents' ability to own foreign assets. In contrast to trade restrictions, which limit the openness of goods markets, capital restrictions limit the openness of financial markets.
- A regional trading bloc is a group of countries who have signed an agreement to reduce and progressively eliminate barriers to trade and movement of factors of production among the members of the bloc.
  - They may or may not have common trade barriers against those countries that are not members of the bloc. In a free trade area all barriers to the flow of goods and services among members are eliminated, but each country maintains its own polices against non-members.
  - A customs union extends the FTA by not only allowing free movement of goods and services among members but also creating a common trade policy against non-members.
  - A common market incorporates all aspects of a customs union and extends it by allowing free movement of factors of production among members.
  - An economic union incorporates all aspects of a common market and requires common economic institutions and coordination of economic policies among members.
  - Members of a monetary union adopt a common currency.
- From an investment perspective, it is important to understand the complex and dynamic nature of trading relationships because they can help identify potential profitable investment opportunities as well as provide some advance warning signals regarding when to disinvest in a market or industry.
- The major components of the balance of payments are the
  - current account balance, which largely reflects trade in goods and services.
  - capital account balance, which mainly consists of capital transfers and net sales of non-produced, non-financial assets.
  - financial account, which measures net capital flows based on sales and purchases of domestic and foreign financial assets.
- Decisions by consumers, firms, and governments influence the balance of payments.
  - Low private savings and/or high investment tend to produce a current account deficit that must be financed by net capital imports; high private savings and/or low investment, however, produce a current account surplus, balanced by net capital exports.
  - All else the same, a government deficit produces a current account deficit and a government surplus leads to a current account surplus.
  - All else the same, a sustained current account deficit contributes to a rise in the risk premium for financial assets of the deficit country. Current account surplus countries tend to enjoy lower risk premiums than current account deficit countries.

- Created after WWII, the International Monetary Fund, the World Bank, and the World Trade Organization are the three major international organizations that provide necessary stability to the international monetary system and facilitate international trade and development.
    - The IMF's mission is to ensure the stability of the international monetary system, the system of exchange rates and international payments that enables countries to buy goods and services from each other. The IMF helps to keep country-specific market risk and global systemic risk under control.
    - The World Bank helps to create the basic economic infrastructure essential for creation and maintenance of domestic financial markets and a well-functioning financial industry in developing countries.
    - The World Trade Organization's mission is to foster free trade by providing a major institutional and regulatory framework of global trade rules without which today's global multinational corporations would be hard to conceive.

REFERENCES

Appleyard, Dennis, Alfred Field, and Steven Cobb. 2010. *International Economics*. 7th edition. Boston: McGraw-Hill/Irwin.

Ariyoshi, Akira, Karl Habermeier, Bernard Laurens, Inci Otker-Robe, Jorge Iván Canales-Kriljenko, and Andrei Kirilenko. 2000. "Capital Controls: Country Experiences with Their Use and Liberalization." IMF Occasional Paper 190, Washington, DC (May 17).

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott. 2010. "Intrafirm Trade and Product Contractibility." *American Economic Review*, vol. 100, no. 2 (May):444–448.

Bureau of Labor Statistics."Textile, Textile Product, and Apparel Manufacturing." In *Career Guide to Industries*: 2010–11 Edition.

Coe, David T., and Elhanan Helpman. 1995. "International R&D Spillovers." *European Economic Review*, vol. 39, no. 5 (May):859–887.

Collier, Paul, and Stephen A. O'Connell. 2007. "Opportunities and Choices." In *The Political Economy of Economic Growth in Africa, 1960–2000*, vol. 1. Edited by Benno J. Ndulu, Stephen A. O'Connell, Robert H. Bates, Paul Collier, and Charles C. Soludo. Cambridge, U.K.: Cambridge University Press.

Feenstra, Robert C., and Alan M. Taylor. 2008. *International Economics*. New York: Worth Publishers.

Gerber, James. 2017. *International Economics*. 7th edition. New York: Prentice Hall.

Hill, Charles W.L., and G. Tomas M. Hult. 2019. *International Business: Competing in the Global Marketplace*. 12th edition. Boston: Irwin/McGraw-Hill.

IMF. 2008. *Globalization: A Brief Overview*. Issues Brief, International Monetary Fund (May).

IMF. 2010a. *Balance of Payments and International Investment Position Manual*. 6th ed. Washington, DC: International Monetary Fund.

IMF. 2010b. *World Economic Outlook*: April 2010. Washington, DC: International Monetary Fund.

IMF. 2011. "The IMF at a Glance." International Monetary Fund (February): www.imf.org/external/np/exr/facts/glance.htm.

Kawai, Masahiro, and Shinji Takagi. 2003. "Rethinking Capital Controls: The Malaysian Experience." PRI Discussion Paper Series No. 03A-05, Policy Research Institute, Ministry of Finance Japan, Tokyo (May).

Meier, Gerald M. 1998. *The International Environment of Business: Competition and Governance in the Global Economy*. New York: Oxford University Press.

Roberts, Mark, and Uwe Deichmann. 2008. "Regional Spillover Estimation." Background paper for the *World Development Report 2009: Reshaping Economic Geography*, World Bank.

Salvatore, Dominick. 2011. *Introduction to International Economics*. 3rd edition. Hoboken, NJ: John Wiley & Sons.

United Nations. 2002. *World Investment Report 2002: Transnational Corporations and Export Competitiveness*. New York: United Nations Conference on Trade and Development (UNCTAD).

World Bank. 2009. *World Development Report 2009: Reshaping Economic Geography*. Washington, DC: World Bank.

World Trade Organization. 2008. *World Trade Report 2008: Trade in a Globalizing World*. Geneva: World Trade Organization.

PRACTICE PROBLEMS

1 Which of the following statements *best* describes the benefits of international trade?

A Countries gain from exchange and specialization.

B Countries receive lower prices for their exports and pay higher prices for imports.

C Absolute advantage is required for a country to benefit from trade in the long term.

2 Which of the following statements *best* describes the costs of international trade?

A Countries without an absolute advantage in producing a good cannot benefit significantly from international trade.

B Resources may need to be allocated into or out of an industry and less-efficient companies may be forced to exit an industry, which in turn may lead to higher unemployment.

C Loss of manufacturing jobs in developed countries as a result of import competition means that developed countries benefit far less than developing countries from trade.

3 Suppose the cost of producing tea relative to copper is lower in Tealand than in Copperland. With trade, the copper industry in Copperland would *most likely*:

A expand.

B contract.

C remain stable.

4 A country has a comparative advantage in producing a good if:

A it is able to produce the good at a lower cost than its trading partner.

B its opportunity cost of producing the good is less than that of its trading partner.

C its opportunity cost of producing the good is more than that of its trading partner.

5 Suppose Mexico exports vegetables to Brazil and imports flashlights used for mining from Brazil. The output per worker per day in each country is as follows:

	Flashlights	**Vegetables**
Mexico	20	60
Brazil	40	80

Which country has a comparative advantage in the production of vegetables and what is the *most* relevant opportunity cost?

A Brazil: 2 vegetables per flashlight.

B Mexico: 1.5 vegetables per flashlight.

C Mexico: ⅓ flashlight per vegetable.

6 Suppose three countries produce bananas and pencils with output per worker per day in each country as follows:

© 2019 CFA Institute. All rights reserved.

	Bananas	Pencils
Mexico	20	40
Brazil	30	90
Canada	40	160

Which country has the greatest comparative advantage in the production of bananas?

A Canada.

B Brazil.

C Mexico.

7 In the Ricardian trade model, a country captures more of the gains from trade if:

A it produces all products while its trade partner specializes in one good.

B the terms of trade are closer to its autarkic prices than to its partner's autarkic prices.

C the terms of trade are closer to its partner's autarkic prices than to its autarkic prices.

8 Germany has much more capital per worker than Portugal. In autarky each country produces and consumes both machine tools and wine. Production of machine tools is relatively capital intensive whereas winemaking is labor intensive. According to the Heckscher–Ohlin model, when trade opens:

A Germany should export machine tools and Portugal should export wine.

B Germany should export wine and Portugal should export machine tools.

C Germany should produce only machine tools and Portugal should produce only wine.

9 According to the Heckscher–Ohlin model, when trade opens:

A the scarce factor gains relative to the abundant factor in each country.

B the abundant factor gains relative to the scarce factor in each country.

C income is redistributed between countries but not within each country.

10 Which type of trade restriction would *most likely* increase domestic government revenue?

A Tariff.

B Import quota.

C Export subsidy.

11 Which of the following trade restrictions is likely to result in the greatest welfare loss for the importing country?

A A tariff.

B An import quota.

C A voluntary export restraint.

12 A large country can:

A benefit by imposing a tariff.

B benefit with an export subsidy.

C not benefit from any trade restriction.

13 If Brazil and South Africa have free trade with each other, a common trade policy against all other countries, but no free movement of factors of production between them, then Brazil and South Africa are part of a:

- **A** customs union.
- **B** common market.
- **C** free trade area (FTA).

14 Which of the following factors *best* explains why regional trading agreements are more popular than larger multilateral trade agreements?

- **A** Minimal displacement costs.
- **B** Trade diversions benefit members.
- **C** Quicker and easier policy coordination.

15 The sale of mineral rights would be captured in which of the following balance of payments components?

- **A** Capital account.
- **B** Current account.
- **C** Financial account.

16 Patent fees and legal services are recorded in which of the following balance of payments components?

- **A** Capital account.
- **B** Current account.
- **C** Financial account.

17 During the most recent quarter, a steel company in South Korea had the following transactions

- Bought iron ore from Australia for AUD50 million.
- Sold finished steel to the United States for USD65 million.
- Borrowed AUD50 million from a bank in Sydney, Australia.
- Received a USD10 million dividend from US subsidiary.
- Paid KRW550 million to a Korean shipping company.

Which of the following would be reflected in South Korea's current account balance for the quarter?

- **A** The loan.
- **B** The shipping.
- **C** The dividend.

18 Which of the following *most likely* contributes to a current account deficit?

- **A** High taxes.
- **B** Low private savings.
- **C** Low private investment.

19 Which of the following chronic deficit conditions is *least* alarming to the deficit country's creditors?

- **A** High consumption.
- **B** High private investment.
- **C** High government spending.

20 Which of the following international trade organizations regulates cross-border exchange among nations on a global scale?

- **A** World Bank Group (World Bank).
- **B** World Trade Organization (WTO).
- **C** International Monetary Fund (IMF).

21 Which of the following international trade organizations has a mission to help developing countries fight poverty and enhance environmentally sound economic growth?

- **A** World Bank Group (World Bank).
- **B** World Trade Organization (WTO).
- **C** International Monetary Fund (IMF).

22 Which of the following organizations helps to keep global systemic risk under control by preventing contagion in scenarios such as the 2010 Greek sovereign debt crisis?

- **A** World Bank Group (World Bank).
- **B** World Trade Organization (WTO).
- **C** International Monetary Fund (IMF).

23 Which of the following international trade bodies was the only multilateral body governing international trade from 1948 to 1995?

- **A** World Trade Organization (WTO).
- **B** International Trade Organization (ITO).
- **C** General Agreement on Tariffs and Trade (GATT).

SOLUTIONS

1 A is correct. Countries gain from exchange when trade enables each country to receive a higher price for exported goods and/or pay a lower price for imported goods. This leads to more efficient resource allocation and allows consumption of a larger variety of goods.

2 B is correct. Resources may need to be reallocated into or out of an industry, depending on whether that industry is an exporting sector or an import-competing sector of that economy. As a result of this adjustment process, less-efficient companies may be forced to exit the industry, which in turn may lead to higher unemployment and the need for retraining in order for displaced workers to find jobs in expanding industries.

3 A is correct. The copper industry in Copperland would benefit from trade. Because the cost of producing copper relative to producing tea is lower in Copperland than in Tealand, Copperland will export copper and the industry will expand.

4 B is correct. Comparative advantage is present when the opportunity cost of producing a good is less than that of a trading partner.

5 C is correct. While Brazil has an absolute advantage in the production of both flashlights and vegetables, Mexico has a comparative advantage in the production of vegetables. The opportunity cost of vegetables in Mexico is ⅓ per flashlight, while the opportunity cost of vegetables in Brazil is ½ per flashlight.

6 C is correct. Mexico has the lowest opportunity cost to produce an extra banana. The opportunity cost is 2 pencils per banana in Mexico, 3 pencils per banana in Brazil, and 4 pencils per banana in Canada.

7 C is correct. A country gains if trade increases the price of its exports relative to its imports as compared to its autarkic prices, i.e. the final terms of trade are more favorable than its autarkic prices. If the relative price of exports and imports remains the same after trade opens, then the country will consume the same basket of goods before and after trade opens, and it gains nothing from the ability to trade. In that case, its trade partner will capture all of the gains. Of course, the opposite is true if the roles are reversed. More generally, a country captures more of the gains from trade the more the final terms of trade differ from its autarkic prices.

8 A is correct. In the Heckscher–Ohlin model a country has a comparative advantage in goods whose production is intensive in the factor with which it is relatively abundantly endowed. In this case, capital is relatively abundant in Germany so Germany has a comparative advantage in producing the capital-intensive product: machine tools. Portugal is relatively labor abundant, hence should produce and export the labor-intensive product: wine.

9 B is correct. As a country opens up to trade, it has a favorable impact on the abundant factor, and a negative impact on the scarce factor. This is because trade causes the output mix to change and therefore changes the relative demand for the factors of production. Increased output of the export product increases demand for the factor that is used intensively in its production, while reduced output of the import product decreases demand for the factor used intensively in its production. Because the export (import) product uses the abundant (scarce) factor intensively, the abundant factor gains relative to the scarce factor in each country.

10 A is correct. The imposition of a tariff will most likely increase domestic government revenue. A tariff is a tax on imports collected by the importing country's government.

11 C is correct. With a voluntary export restraint, the price increase induced by restricting the quantity of imports (= quota rent for equivalent quota = tariff revenue for equivalent tariff) accrues to foreign exporters and/or the foreign government.

12 A is correct. By definition, a large country is big enough to affect the world price of its imports and exports. A large country can benefit by imposing a tariff if its terms of trade improve by enough to outweigh the welfare loss arising from inefficient allocation of resources.

13 A is correct. A customs union extends a free trade area (FTA) by not only allowing free movement of goods and services among members, but also creating common trade policy against non-members. Unlike a more integrated common market, a customs union does not allow free movement of factors of production among members.

14 C is correct. Regional trading agreements are politically less contentious and quicker to establish than multilateral trade negotiations (for example, under the World Trade Organization). Policy coordination and harmonization is easier among a smaller group of countries.

15 A is correct. The capital account measures capital transfers and sale and purchase of non-produced, non-financial assets such as mineral rights and intangible assets.

16 B is correct. The current account measures the flows of goods and services (including income from foreign investments). Patent fees and legal services are both captured in the services sub-account of the current account.

17 C is correct. The current account includes income received on foreign investments. The Korean company effectively "exported" the use of its capital during the quarter to its US subsidiary, and the dividend represents payment for those services.

18 B is correct. A current account deficit tends to result from low private saving, high private investment, a government deficit, or a combination of the three. Of the choices, only low private savings contributes toward a current account deficit.

19 B is correct. A current account deficit tends to result from low private saving, high private investment, low government savings, or a combination of the three. Of these choices, only high investments can increase productive resources and improve future ability to repay creditors.

20 B is correct. The WTO provides the legal and institutional foundation of the multinational trading system and is the only international organization that regulates cross-border trade relations among nations on a global scale. The WTO's mission is to foster free trade by providing a major institutional and regulatory framework of global trade rules. Without such global trading rules, today's global transnational corporations would be hard to conceive.

21 A is correct. The World Bank's mission is to help developing countries fight poverty and enhance environmentally sound economic growth. The World Bank helps to create the basic economic infrastructure essential for creation and maintenance of domestic financial markets and a well-functioning financial industry in developing countries.

22 C is correct. From an investment perspective, the IMF helps to keep country-specific market risk and global systemic risk under control. The Greek sovereign debt crisis on 2010, which threatened to destabilize the entire European banking system, is a recent example. The IMF's mission is to ensure the stability of the international monetary system, the system of exchange rates and international payments which enables countries to buy goods and services from each other.

23 C is correct. The GATT was the only multilateral body governing international trade from 1948 to 1995. It operated for almost half a century as a quasi-institutionalized, provisional system of multilateral treaties and included several rounds of negotiations.

READING

14

Currency Exchange Rates

by William A. Barker, PhD, CFA, Paul D. McNelis, and
Jerry Nickelsburg, PhD

William A. Barker, PhD, CFA (Canada). Paul D. McNelis is at Gabelli School of Business, Fordham University (USA). Jerry Nickelsburg, PhD, is at the Anderson School of Management, University of California, Los Angeles (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** define an exchange rate and distinguish between nominal and real exchange rates and spot and forward exchange rates;
☐	**b.** describe functions of and participants in the foreign exchange market;
☐	**c.** calculate and interpret the percentage change in a currency relative to another currency;
☐	**d.** calculate and interpret currency cross-rates;
☐	**e.** calculate an outright forward quotation from forward quotations expressed on a points basis or in percentage terms;
☐	**f.** explain the arbitrage relationship between spot rates, forward rates, and interest rates;
☐	**g.** calculate and interpret a forward discount or premium;
☐	**h.** calculate and interpret the forward rate consistent with the spot rate and the interest rate in each currency;
☐	**i.** describe exchange rate regimes;
☐	**j.** explain the effects of exchange rates on countries' international trade and capital flows.

© 2019 CFA Institute. All rights reserved.

1

INTRODUCTION & THE FOREIGN EXCHANGE MARKET

a define an exchange rate and distinguish between nominal and real exchange rates and spot and forward exchange rates;

c calculate and interpret the percentage change in a currency relative to another currency;

Measured by daily turnover, the foreign exchange (FX) market—the market in which currencies are traded against each other—is by far the world's largest market. Current estimates put daily turnover at approximately USD5.1 trillion for 2016. This is about 10 to 15 times larger than daily turnover in global fixed-income markets and about 50 times larger than global turnover in equities.

The FX market is also a truly global market that operates 24 hours a day, each business day. It involves market participants from every time zone connected through electronic communications networks that link players as large as multibillion-dollar investment funds and as small as individuals trading for their own account—all brought together in real time. International trade would be impossible without the trade in currencies that facilitates it, and so too would cross-border capital flows that connect all financial markets globally through the FX market.

These factors make foreign exchange a key market for investors and market participants to understand. The world economy is increasingly transnational in nature, with both production processes and trade flows often determined more by global factors than by domestic considerations. Likewise, investment portfolio performance increasingly reflects global determinants because pricing in financial markets responds to the array of investment opportunities available worldwide, not just locally. All of these factors funnel through, and are reflected in, the foreign exchange market. As investors shed their "home bias" and invest in foreign markets, the exchange rate—the price at which foreign-currency-denominated investments are valued in terms of the domestic currency—becomes an increasingly important determinant of portfolio performance.

Even investors adhering to a purely "domestic" portfolio mandate are increasingly affected by what happens in the foreign exchange market. Given the globalization of the world economy, most large companies depend heavily on their foreign operations (for example, by some estimates about 30 percent of S&P 500 Index earnings are from outside the United States). Almost all companies are exposed to some degree of foreign competition, and the pricing for domestic assets—equities, bonds, real estate, and others—will also depend on demand from foreign investors. All of these various influences on investment performance reflect developments in the foreign exchange market.

This reading introduces the foreign exchange market, providing the basic concepts and terminology necessary to understand exchange rates as well as some of the basics of exchange rate economics.

The reading is divided up as follows. Sections 2 and 3 describe the organization of the foreign exchange market and discuss the major players—who they are, how they conduct their business, and how they respond to exchange rate changes. Sections 4–6 take up the mechanics of exchange rates: definitions, quotes, and calculations. These sections show that the reader has to pay close attention to conventions used in various foreign exchange markets around the world because they can vary widely. Sometimes exchange rates are quoted in the number of domestic currency units per unit of foreign currency, and sometimes they are quoted in the opposite way. The exact notation used to represent exchange rates can vary widely as well, and occasionally the same exchange rate notation will be used by different sources to mean completely different things. The notation used here may not be the same as that encountered elsewhere. Therefore, the focus should be on understanding the underlying concepts rather than relying on rote memorization of formulas. We also show how to calculate cross-exchange rates and how to compute the forward exchange rate given either the forward points or the percentage forward premium or discount. In Sections 7 and 8, we discuss alternative exchange rate regimes operating throughout the world. Finally, in Sections 9–11, we discuss how exchange rates affect a country's international trade (exports and imports) and capital flows. A summary and practice problems conclude the reading.

1.1 The Foreign Exchange Market

To understand the FX market, it is necessary to become familiar with some of its basic conventions. Individual currencies are often referred to by standardized three-letter codes that the market has agreed upon through the International Organization for Standardization (ISO). Exhibit 1 lists some of the major global currencies and their identification codes.

Exhibit 1 Standard Currency Codes

Three-Letter Currency Code	Currency
USD	US dollar
EUR	Euro
JPY	Japanese yen
GBP	British pound
CHF	Swiss franc
CAD	Canadian dollar
AUD	Australian dollar
NZD	New Zealand dollar
ZAR	South African rand
SEK	Swedish krona
NOK	Norwegian krone
BRL	Brazilian real
SGD	Singapore dollar
MXN	Mexican peso
CNY	Chinese yuan
HKD	Hong Kong dollar
INR	Indian rupee
KRW	South Korean won
RUB	Russian ruble

It is important to understand that there is a difference between referring to an *individual currency* and an *exchange rate*. One can hold an individual currency (for example, in a EUR100 million deposit), but an exchange rate refers to the price of one currency in terms of another (for example, the exchange rate between the EUR and USD). An individual currency can be singular, but there are always two currencies involved in an exchange rate: the price of one currency relative to another. The exchange rate is the number of units of one currency (called the *price currency*) that one unit of another currency (called the *base currency*) will buy. An equivalent way of describing the exchange rate is as the cost of one unit of the base currency in terms of the price currency.

This distinction between individual currencies and exchange rates is important because, as we will see in a later section, these three-letter currency codes can be used both ways. (For example, when used as an exchange rate in the professional FX market, EUR is understood to be the exchange rate between the euro and US dollar). But be aware of the context (either as a currency or as an exchange rate) in which these three-letter currency codes are being used. To avoid confusion, this reading will identify exchange rates using the convention of "A/B," referring to the number of units of currency A that one unit of currency B will buy. For example, a USD/EUR exchange rate of 1.1700 means that 1 euro will buy 1.1700 US dollars (i.e., 1 euro costs 1.1700 US dollars).[1] In this case, the euro is the base currency and the US dollar is the price currency. A decrease in this exchange rate would mean that the euro costs less or that fewer US dollars are needed to buy one euro. In other words, a decline in this exchange rate indicates that the USD is *appreciating* against the EUR or, equivalently, the EUR is *depreciating* against the USD.

The exchange rates described above are referred to as *nominal* exchange rates. This is to distinguish them from *real* exchange rates, which are indexes often constructed by economists and other market analysts to assess changes in the relative purchasing power of one currency compared with another. Creating these indexes requires adjusting the nominal exchange rate by using the price levels in each country of the currency pair (hence the name "real exchange rates") in order to compare the relative purchasing power between countries.

In a world of homogenous goods and services and with no market frictions or trade barriers, the relative purchasing power across countries would tend to equalize: Why would you pay more, in real terms, domestically for a "widget" if you could import an identical "widget" from overseas at a cheaper price? This basic concept is the intuition behind a theory known as "purchasing power parity" (PPP), which describes the long-term equilibrium of nominal exchange rates. PPP asserts that nominal exchange rates adjust so that identical goods (or baskets of goods) will have the same price in different markets. Or, put differently, the purchasing power of different currencies is equalized for a standardized basket of goods.

In practice, the conditions required to enforce PPP are not satisfied: Goods and services are not identical across countries; countries typically have different baskets of goods and services produced and consumed; many goods and services are not traded internationally; there are trade barriers and transaction costs (e.g., shipping costs and import taxes); and capital flows are at least as important as trade flows in determining nominal exchange rates. As a result, nominal exchange rates exhibit persistent deviations from PPP. Moreover, relative purchasing power among countries displays a weak, if any, tendency toward long-term equalization. A simple example of a cross-country comparison of the purchasing power of a standardized good is the "Big Mac" index produced by the *Economist*, which shows the relative price of this standardized hamburger in different countries. The Big Mac index shows that fast-food hamburger prices can vary widely internationally and that this difference in purchasing power is typical of most goods and services. Hence, movements in real exchange rates provide meaningful information about changes in relative purchasing power among countries.

---

[1] This convention is consistent with the meaning of "/" in mathematics and the straightforward interpretation of "A/B" as "A per B" is helpful in understanding exchange rates as the price of one currency in terms of another. Nevertheless, other notation conventions exist. "B/A" and "B:A" are sometimes used to denote what this reading denotes as "A/B." Careful attention to the context will usually make the convention clear.

Consider the case of an individual who wants to purchase goods from a foreign country. The individual would be able to buy fewer of these goods if the nominal spot exchange rate for the foreign currency appreciated or if the foreign price level increased. Conversely, the individual could buy more foreign goods if the individual's domestic income increased. (For this example, we will assume that changes in the individual's income are proportional to changes in the domestic price level.) Hence, in *real* purchasing power terms, the real exchange rate that an individual faces is an increasing function of the nominal exchange rate (quoted in terms of the number of units of domestic currency per one unit of foreign currency) and the foreign price level and a decreasing function of the domestic price level. The *higher* the real exchange rate that this individual faces, the *fewer* foreign goods, in real terms, the individual can purchase and the *lower* that individual's relative purchasing power compared with the other country.

An equivalent way of viewing the real exchange rate is that it represents the relative price levels in the domestic and foreign countries. Mathematically, we can represent the foreign price level in terms of the domestic currency as:

$$\text{Foreign price level in domestic currency} = S_{d/f} \times P_f$$

where $S_{d/f}$ is the spot exchange rate (quoted in terms of the number of units of domestic currency per one unit of foreign currency) and P_f is foreign price level quoted in terms of the foreign currency. We can define the domestic price level, in terms of the domestic currency, as P_d. Hence, the ratio between the foreign and domestic price levels is:

$$\text{Real exchange rate}_{(d/f)} = (S_{d/f} \times P_f)/P_d = S_{d/f} \times (P_f/P_d)$$

For example, for a British consumer wanting to buy goods made in the Eurozone, the real exchange rate (defined in GBP/EUR terms; note that the domestic currency for the United Kingdom is the price currency, not the base currency) will be an increasing function of the nominal spot exchange rate (GBP/EUR) and the Eurozone price level and a decreasing function of the UK price level. This is written as:

$$\text{Real exchange rate}_{\frac{GBP}{EUR}} = S_{\frac{GBP}{EUR}} \times \left(\frac{CPI_{eur}}{CPI_{UK}}\right)$$

Let's examine the effect of movements in the domestic and foreign price levels, and the nominal spot exchange rate, on the real purchasing power of an individual in the United Kingdom wanting to purchase Eurozone goods. Assume that the nominal spot exchange rate (GBP/EUR) increases by 10 percent, the Eurozone price level by 5 percent, and the UK price level by 2 percent. The change in the real exchange rate is then:

$$\left(1+\frac{\Delta S_{\frac{d}{f}}}{S_{\frac{d}{f}}}\right) \times \frac{\left(1+\frac{\Delta P_f}{P_f}\right)}{\left(1+\frac{\Delta P_d}{P_d}\right)} - 1 = (1+10\%) \times \frac{1+5\%}{1+2\%} - 1 \approx 10\% + 5\% - 2\% \approx 13\%$$

In this case, the real exchange rate for the UK-based individual has *increased* about 13 percent, meaning that it now costs *more*, in real terms, to buy Eurozone goods. Or put differently, the UK individual's real purchasing power relative to Eurozone goods has *declined* by about 13 percent. An easy way to remember this relationship is to consider the real exchange rate (stated with the domestic currency as the price currency) as representing the real price you face in order to purchase foreign goods and services: The *higher* the price (real exchange rate), the *lower* your relative purchasing power.

The real exchange rate for a currency can be constructed for the domestic currency relative to a single foreign currency or relative to a basket of foreign currencies. In either case, these real exchange rate indexes depend on the assumptions made by the analyst creating them. Several investment banks and central banks create proprietary measures of real exchange rates. It is important to note that real exchange rates are *not* quoted or traded in global FX markets: They are only indexes created by analysts to understand the international competitiveness of an economy and the real purchasing power of a currency.

In this context, real exchange rates can be useful for understanding trends in international trade and capital flows and hence can be seen as one of the influences on nominal spot exchange rates. As an example, consider the exchange rate between the Indian rupee and the US dollar. During 2018, the nominal rupee exchange rate against the US dollar (INR/USD) rose by approximately 6.7 percent—meaning that the US dollar appreciated against the rupee. However, the annual inflation rates in the United States and India were different during 2018—approximately 2.5 percent for the United States and 4.7 percent for India. This means that the real exchange rate (in INR/USD terms) was depreciating less rapidly than the nominal INR/USD exchange rate:

$$\left(1+\%\Delta S_{\frac{INR}{USD}}\right)\times\frac{\left(1+\%\Delta P_{US}\right)}{\left(1+\%\Delta P_{India}\right)}-1\approx+6.7\%+2.5\%-4.7\%\approx4.5\%$$

This combination of a much weaker rupee and a higher Indian inflation rate meant that the real exchange rate faced by India was increasing, thus decreasing Indian purchasing power in USD terms.

Movements in real exchange rates can have a similar effect as movements in nominal exchange rates in terms of affecting relative prices and hence trade flows. Even if the nominal spot exchange rate does not move, differences in inflation rates between countries affect their relative competitiveness.

Although real exchange rates can exert some influence on nominal exchange rate movements, they are only one of many factors; it can be difficult to disentangle all of these inter-relationships in a complex and dynamic FX market. As discussed earlier, PPP is a poor guide to predicting future movements in nominal exchange rates because these rates can deviate from PPP equilibrium—and even continue to trend away from their PPP level—for years at a time. Hence, it should not be surprising that real exchange rates, which reflect changes in relative purchasing power, have a poor track record as a predictor of future nominal exchange rate movements.

EXAMPLE 1

Nominal and Real Exchange Rates

An investment adviser located in Sydney, Australia, is meeting with a local client who is looking to diversify her domestic bond portfolio by adding investments in fixed-rate, long-term bonds denominated in HKD. The client frequently visits Hong Kong SAR, and many of her annual expenses are denominated in HKD. The client, however, is concerned about the foreign currency risks of offshore investments and whether the investment return on her HKD-denominated investments will maintain her purchasing power—both domestically (i.e., for her AUD-denominated expenses) and in terms of her foreign trips (i.e., denominated in HKD, for her visits to Hong Kong SAR). The investment adviser explains the effect of changes in nominal and real exchange rates to the client and illustrates this explanation by making the following statements:

Statement 1	All else equal, an increase in the nominal AUD/HKD exchange rate will lead to an increase in the AUD-denominated value of your foreign investment.

Statement 2	All else equal, an increase in the nominal AUD/HKD exchange rate means that your relative purchasing power for your Hong Kong SAR trips will increase (based on paying for your trip with the income from your HKD-denominated bonds).
Statement 3	All else equal, an increase in the Australian inflation rate will lead to an increase in the real exchange rate (AUD/HKD). A higher real exchange rate means that the relative purchasing power of your AUD-denominated income is higher.
Statement 4	All else equal, a decrease in the nominal exchange rate (AUD/HKD) will decrease the real exchange rate (AUD/HKD) and increase the relative purchasing power of your AUD-denominated income.

To demonstrate the effects of the changes in inflation and nominal exchange rates on relative purchasing power, the adviser uses the following scenario: "Suppose that the AUD/HKD exchange rate increases by 5 percent, the price of goods and services in Hong Kong SAR goes up by 5 percent, and the price of Australian goods and services goes up by 2 percent."

1 Statement 1 is:

- **A** correct.
- **B** incorrect, because based on the quote convention the investment's value would be decreasing in AUD terms.
- **C** incorrect, because the nominal AUD value of the foreign investments will depend on movements in the Australian inflation rate.

2 Statement 2 is:

- **A** correct.
- **B** incorrect, because purchasing power is not affected in this case.
- **C** incorrect, because based on the quote convention, the client's relative purchasing power would be decreasing.

3 Statement 3 is:

- **A** correct.
- **B** incorrect with respect to the real exchange rate only.
- **C** incorrect with respect to both the real exchange rate and the purchasing power of AUD-denominated income.

4 Statement 4 is:

- **A** correct.
- **B** incorrect with respect to the real exchange rate.
- **C** incorrect with respect to the purchasing power of AUD-denominated income.

5 Based on the adviser's scenario and assuming that the HKD value of the HKD bonds remained unchanged, the nominal AUD value of the client's HKD investments would:

- **A** decrease by about 5 percent.
- **B** increase by about 5 percent.
- **C** remain approximately the same.

6 Based on the adviser's scenario, the change in the relative purchasing power of the client's AUD-denominated income is *closest* to:

A –8 percent.

B +8 percent.

C +12 percent.

Solution to 1:

A is correct. Given the quoting convention, an increase in the AUD/HKD rate means that the base currency (HKD) is appreciating (one HKD will buy more AUD). This is increasing the nominal value of the HKD-denominated investments when measured in AUD terms.

Solution to 2:

B is correct. When paying for HKD-denominated expenses with HKD-denominated income, the value of the AUD/HKD spot exchange rate (or any other spot exchange rate) would not be relevant. In fact, this is a basic principle of currency risk management: reducing FX risk exposures by denominating assets and liabilities (or income and expenses) in the same currency.

Solution to 3:

C is correct. An increase in the Australian (i.e., domestic) inflation rate means that the real exchange rate (measured in domestic/foreign, or AUD/HKD, terms) would be decreasing, not increasing. Moreover, an increase in the real exchange rate ($R_{AUD/HKD}$) would be equivalent to a reduction of the purchasing power of the Australian client: Goods and services denominated in HKD would cost more.

Solution to 4:

A is correct. As the spot AUD/HKD exchange rate decreases, the HKD is depreciating against the AUD; or equivalently, the AUD is appreciating against the HKD. This is reducing the real exchange rate ($R_{AUD/HKD}$) and increasing the Australian client's purchasing power.

Solution to 5:

B is correct. As the AUD/HKD spot exchange rate increases by 5 percent, the HKD is appreciating against the AUD by 5 percent and, all else equal, the value of the HKD-denominated investment is increasing by 5 percent in AUD terms.

Solution to 6:

A is correct. The real exchange rate ($R_{AUD/HKD}$) is expressed as:

$$R_{\frac{AUD}{HKD}} = S_{\frac{AUD}{HKD}} \times \frac{P_{HKD}}{P_{AUD}}$$

The information in the adviser's scenario can be expressed as:

$$\%\Delta R_{\frac{AUD}{HKD}} \approx \%\Delta S_{\frac{AUD}{HKD}} + \%\Delta P_{HKD} - \%\Delta P_{AUD} \approx +5\% + 5\% - 2\% \approx +8\%$$

Because the real exchange rate (expressed in AUD/HKD terms) has gone up by about 8 percent, the real purchasing power of the investor based in Australia has declined by about 8 percent. This can be seen from the fact that HKD has appreciated against the AUD in nominal terms, and the Hong Kong SAR price level has also increased. This increase in the cost of Hong Kong SAR goods and services (measured in AUD) is only partially offset by the small (2 percent) increase in the investor's income (assumed equal to the change in the Australian price level).

MARKET FUNCTIONS

2

a define an exchange rate and distinguish between nominal and real exchange rates and spot and forward exchange rates;

b describe functions of and participants in the foreign exchange market;

FX markets facilitate international trade in goods and services, where companies and individuals need to make transactions in foreign currencies. This would cover everything from companies and governments buying and selling products in other countries, to tourists engaged in cross-border travel (for example, a German tourist selling euros and buying sterling for a visit to London). Although this is an important dimension of FX markets, and despite the growth of global trade in recent years, an even larger proportion of the daily turnover in FX markets is accounted for by capital market transactions, where investors convert between currencies for the purpose of moving funds into (or out of) foreign assets. These types of transactions cover the range from direct investments (for example, companies buying such fixed assets as factories) in other countries to portfolio investments (the purchase of stocks, bonds, and other financial assets denominated in foreign currencies). Because capital is extremely mobile in modern financial markets, this ebb and flow of money across international borders and currencies generates a huge and growing volume of FX transactions.

Regardless of the underlying motivation for the FX transaction, it will eventually require that one currency be exchanged for another in the FX market. In advance of that required transaction, market participants are exposed to the risk that the exchange rate will move against them. Often they will try to reduce (hedge) this risk through a variety of FX instruments (described in more detail later). Conversely, market participants may form opinions about future FX movements and undertake speculative FX risk exposures through a variety of FX instruments in order to profit from their views.

The distinction between hedging and speculative positions is not always clear cut. For example, consider the case of a corporation selling its products overseas. This creates an FX risk exposure because the revenue from foreign sales will ultimately need to be converted into the corporation's home currency. This risk exposure is typically hedged, and corporate hedging often accounts for large FX flows passing through the market. The amount and timing of foreign revenue, however, are generally hard to predict with precision: They will depend on the pace of foreign sales, the sales prices realized, the pace at which foreign clients pay for their purchases, and so forth. In the face of this uncertainty, the corporate treasury will estimate the timing and amount of foreign revenue and will then hedge a portion of this estimated amount. Many corporate treasuries have hedging targets based on this estimate, but they also have the flexibility to under-hedge or over-hedge based on their opinions about future FX rate movements. In order to judge the effectiveness of these discretionary trades, the performance of the corporate treasury is compared with a benchmark, usually stated in terms of a fixed amount hedged relative to total sales. (For example, the benchmark may be a 100 percent fully hedged position. The profitability of the hedge actually implemented—which, based on the treasury's discretion, can vary above or below 100 percent—is then compared with what would have been achieved with a passive, 100 percent fully hedged position.) Treasury managers' performance is judged based on gains or losses relative to the benchmark, just as an investment fund manager's performance is benchmarked against performance targets.

At the other end of the spectrum between hedging and speculation, consider the archetypical speculative account: a hedge fund. Although it is true that hedge funds will seek out, accept, and manage risk for profit, a hedge fund is, after all, a hedge fund: Strict risk control procedures are critical to the fund's success, especially when leverage is involved. This mixture of speculative and hedging motives is common throughout the FX space as market participants shape their FX exposures to suit their market forecasts, operational mandates, and appetites for risk.

The FX market provides a variety of products that provide the flexibility to meet this varied and complex set of financial goals. *Spot* transactions involve the exchange of currencies for immediate delivery. For most currencies, this corresponds to "T + 2" delivery, meaning that the exchange of currencies is settled two business days after the trade is agreed to by the two sides of the deal. (One exception is the Canadian dollar, for which spot settlement against the US dollar is on a T + 1 basis.) The exchange rate used for these spot transactions is referred to as the spot exchange rate, and it is the exchange rate that most people refer to in their daily lives (for example, this is the exchange rate usually quoted by the financial press, on the evening news, and so forth).

It is important to realize, however, that spot transactions make up only a minority of total daily turnover in the global FX market: The rest is accounted for by trade in outright forward contracts, FX swaps, and FX options. Although these products will be covered in more depth in a subsequent section, and at Level II of the CFA curriculum, we will provide a brief introduction to these products here.

Outright *forward contracts* (often referred to simply as forwards) are agreements to deliver foreign exchange at a future date at an exchange rate agreed upon today. For example, suppose that a UK-based company expects to receive a payment of 100 million euros in 85 days. Although it could convert these euros to British pounds with a spot transaction (the spot rate would be the GBP/EUR rate in 83 days, because of T + 2 settlement), this future spot rate is currently unknown and represents a foreign exchange risk to the company. The company can avoid this risk by entering into a transaction with a foreign exchange dealer to sell 100 million euros against the British pound for settlement 85 days from today at a rate—the forward exchange rate—agreed upon today.

As such, forward contracts are any exchange rate transactions that occur with currency settlement longer than the usual T + 2 settlement for spot delivery. Each of these contracts requires two specifications: the date at which the currencies are to be exchanged and the exchange rate to be applied on the settlement date. Accordingly, exchange rates for these transactions are called *forward exchange rates* to distinguish them from spot rates.

Dealers will typically quote forward rates for a variety of standard forward settlement dates (for example, one week, one month, or 90 days) on their dealing screens. In an over-the-counter (OTC) market, however, traders can arrange forward settlement at *any* future date they agree upon, with the forward exchange rate scaled appropriately for the specific term to settlement. Standard forward settlement dates (such as three months) are defined in terms of the spot settlement date, which is generally T + 2. For example, if today is 18 October and spot settlement is for 20 October, then a three-month forward settlement would be defined as 20 January of the following year. Note as well that these standard forward settlement dates may not always be good business days: 20 January could be a weekend or a holiday. In that case, the forward settlement date is set to the closest good business day. Traders always confirm the exact forward settlement date when making these types of trades, and the forward rate is scaled by the exact number of days to settlement.

In an OTC market, the size of the forward contracts can also be any size that the two counterparties agree upon. In general, however, liquidity in forward markets declines the longer the term to maturity and the larger the trade size. The concept of the forward exchange rate and exchange hedging is developed further in Sections 4–6.

Although the OTC market accounts for the majority of foreign exchange trades with future (i.e., greater than T + 2) settlement dates, there is also a deep, liquid market in exchange-traded *futures* contracts for currencies. Although there are technical differences between futures and forward contracts, the basic concept is the same: The price is set today for settlement on a specified future date. Futures contracts on currencies trade on several exchanges globally, but the majority of volume in exchange-traded currency futures contracts is found on the International Monetary Market (IMM) division of the Chicago Mercantile Exchange (CME). Futures contracts differ from OTC forward contracts in several important ways: They trade on exchanges (such as the CME) rather than OTC; they are only available for fixed contract amounts and fixed settlement dates; the exchanges demand that a fixed amount of collateral be posted against the futures contract trade; and this collateral is marked-to-market daily, with counterparties asked to post further collateral if their positions generate losses. On balance, futures contracts are somewhat less flexible than forward contracts. Nonetheless, they provide deep, liquid markets for deferred delivery with a minimum of counterparty (i.e., default) risk—a proposition that many FX traders find attractive. Accordingly, daily turnover in FX futures contracts is huge. As of 2010, the average daily trading volume of FX futures on the CME alone was estimated to be about USD140 billion, which is almost comparable in size to the interbank volume of spot transactions.

Because forward contracts eventually expire, existing speculative positions or FX hedges that need to be extended must be rolled prior to their settlement dates. This typically involves a spot transaction to offset (settle) the expiring forward contract and a new forward contract to be set at a new, more distant settlement date. The combination of an offsetting spot transaction and a new forward contract is referred to as an **FX swap**.[2]

An FX swap is best illustrated by an example. Suppose that a trader sells 100 million euros with settlement 95 days from today at a forward exchange rate (USD/EUR) of 1.2000. In 93 days, the forward contract is two days from settlement, specifically the T + 2 days to spot settlement. To roll the forward contract, the trader will engage in the following FX swap. First, the trader will need to buy 100 million euros spot, for which T + 2 settlement will fall on day 95, the same day as the settlement of the expiring forward contract. The purchase of the 100 million euros spot will be used to satisfy the delivery of the 100 million euros sold in the expiring forward contract. Because 100 million euros are being both bought and sold on day 95, there is no exchange of euros between counterparties on that day: The amounts net to zero. However, there will be an exchange of US dollars, reflecting the movement in exchange rates between the date the forward contract was agreed to (day 0) and day 93. Suppose that on day 93 the spot exchange rate for USD/EUR is 1.1900. This means that the trader will see a cash flow on day 95 of USD1,000,000. This is calculated as follows:

$$\text{EUR}100{,}000{,}000 \times (1.2000 - 1.1900) = \text{USD}1{,}000{,}000$$

The trader receives USD1,000,000 from the counterparty because the euro was *sold* forward to day 95 at a price of 1.2000; it was *bought* (on day 93) for spot settlement on day 95 at a price of 1.1900. This *price* movement in the euro indicates a profit to the trader, but because the euro *quantities* exchanged on day 95 net to zero (100,000,000 euros both bought and sold), this cash flow is realized in US dollars. The second leg of the FX swap is then to initiate a new forward sale of 100 million euros at the USD/EUR forward exchange rate being quoted on day 93. This renews the forward position (a forward sale of the euro) to a new date.

FX swaps will be dealt with in more detail at Level II in the curriculum. For the purposes of this reading, it is only necessary to understand that (1) an FX swap consists of a simultaneous spot and forward transaction; (2) these swap transactions can extend (roll) an existing forward position to a new future date; and (3) rolling the position forward leads to a cash flow on settlement day. This cash flow can be thought of as a mark-to-market on the forward position. FX swaps are a large component of daily FX market turnover because market participants have to roll over existing speculative or hedging positions as the underlying forward contracts mature in order to extend the hedge or speculative position (otherwise, the position is closed out on the forward settlement date).

---

[2] Note that an "FX swap" is not the same as a "currency swap." An FX swap is simply the combination of a spot and a forward FX transaction (i.e., only two settlement dates—spot and forward—are involved). A currency swap is generally used for multiple periods and payments.

One other area where FX swaps are used in FX markets also bears mentioning: They are often used by market participants as a funding source (called swap funding). Consider the case of a UK-based firm that needs to borrow GBP100 million for 90 days, starting 2 days from today. One way to do this is simply to borrow 90-day money in GBP-denominated funds starting at T + 2. An alternative is to borrow in US dollars and exchange these for British pounds in the spot FX market (both with T + 2 settlement) and then sell British pounds 90 days forward against the US dollar. (Recall that the maturity of a forward rate contract is defined in terms of the spot settlement date, so the 90-day forward rate would be for settlement in 92 days from today.) The company has the use of GBP100 million for 90 days, starting on T + 2, and at the end of this period can pay off the US dollar loan at a known, pre-determined exchange rate (the 90-day forward rate). By engaging in simultaneous spot and forward transactions (i.e., an FX swap), the company has eliminated any FX risk from the foreign borrowing. The all-in financing rate using an FX swap will typically be close to that of domestic borrowing, usually within a few basis points. This near equivalence is enforced by an arbitrage relationship that will be described in Section 6. On large borrowing amounts, however, even a small differential can add up to substantial cost savings.

Another way to hedge FX exposures, or implement speculative FX positions, is to use options on currencies. FX options are contracts that, for an upfront premium or fee, give the purchaser the right, but not the obligation, to make an FX transaction at some future date at an exchange rate agreed upon today (when the contract is agreed to). The holder of an FX option will exercise the option only if it is advantageous to do so—that is, if the agreed upon exchange rate for the FX option contract is better than the FX rate available in the market at option expiry. As such, options are extremely flexible tools for managing FX exposures and account for a large percentage of daily turnover in the FX market.

Another concept to bear in mind is that spot, forward, swap, and option products are typically not used in isolation. Most major market participants manage their FX transactions and FX risk exposures through concurrent spot, forward, swap, and option positions. Taken together, these instruments (the building blocks of the FX market) provide an extremely flexible way for market participants to shape their FX risk exposures to match their operational mandate, risk tolerance, and market opinion. Moreover, FX transactions are often made in conjunction with transactions in other financial markets—such as equities, fixed income, and commodities. These markets have a variety of instruments as well, and market participants jointly tailor their *overall* position simultaneously using the building blocks of the FX market and these other markets.

EXAMPLE 2

Spot and Forward Exchange Rates

The investment adviser based in Sydney, Australia, continues her meeting with the local client who has diversified her domestic bond portfolio by adding investments in fixed-rate, long-term bonds denominated in HKD. Given that the client spends most of the year in Australia, she remains concerned about the foreign exchange risk of her foreign investments and asks the adviser how these might be managed. The investment adviser explains the difference between spot and forward exchange rates and their role in determining foreign exchange risk exposures. The investment adviser suggests the following investment strategy to the client: "You can exchange AUD for HKD in the spot exchange market, invest in a risk-free, one-year HKD-denominated zero coupon bond, and use a one-year forward contract for converting the proceeds back into AUD."

Spot exchange rate (AUD/HKD)	0.1714
One-year HKD interest rate	2.20%
One-year forward exchange rate (AUD/HKD)	0.1724

1 Which of the following statements is *most* correct? Over a one-year horizon, the exchange rate risk of the client's investment in HKD-denominated bonds is determined by uncertainty over:

A today's AUD/HKD forward rate.

B the AUD/HKD spot rate one year from now.

C the AUD/HKD forward rate one year from now.

2 To reduce the exchange rate risk of the Hong Kong SAR investment, the client should:

A sell AUD spot.

B sell AUD forward.

C sell HKD forward.

3 Over a one-year horizon, the investment proposed by the investment adviser is *most* likely:

A risk free.

B exposed to interest rate risk.

C exposed to exchange rate risk.

4 To set up the investment proposed by the adviser, the client would need to:

A sell AUD spot; sell a one-year, HKD-denominated bond; and buy AUD forward.

B buy AUD spot; buy a one-year, HKD-denominated bond; and sell AUD forward.

C sell AUD spot; buy a one-year, HKD-denominated bond; and buy AUD forward.

5 The return (in AUD) on the investment proposed by the investment adviser is *closest* to:

A 2.00 percent.

B 3.00 percent.

C 5.00 percent.

Solution to 1:

B is correct. The exchange rate risk (for an unhedged investment) is defined by the uncertainty over future spot rates. In this case, the relevant spot rate is that which would prevail one year from now. Forward rates that would be in effect one year from now would be irrelevant, and the current forward rate is known with certainty.

Solution to 2:

C is correct. The Australian-based investor owns HKD-denominated bonds, meaning that she is long HKD exposure. To hedge this exposure, she could enter into a forward contract to sell the HKD against the AUD for future delivery (that is, match a long HKD exposure in the cash market with a short HKD exposure in the derivatives market). The forward rate is established at the time the forward contract is entered into, eliminating any uncertainty about what exchange rate would be used to convert HKD-denominated cash flows back into AUD.

Solution to 3:

A is correct. The investment is risk free because the investment is based on a risk-free, one-year, zero coupon, HKD-denominated bond—meaning there is no default or reinvestment risk. The investment will mature in one-year at par; there is no interest rate risk. The use of a forward contract to convert the HKD-denominated proceeds back to AUD eliminates any exchange rate risk.

Solution to 4:

C is correct. To create the investment, the client needs to convert AUD to HKD in the spot exchange market, invest in (buy) the one-year HKD bond, and sell the HKD forward/buy the AUD forward. Note that this process is directly comparable to the swap financing approach described in this section of the reading.

Solution to 5:

B is correct. Converting one AUD to HKD in the spot market gives the client (1/0.1714) = HKD5.83. Investing this for one year leads to 5.83 × (1.022) = HKD5.96. Selling this amount of HKD at the forward rate gives 5.96 × 0.1724= AUD1.028 (rounding to three decimal places). This implies an AUD-denominated return of 2.8 percent which rounds up to 3 percent.

3 MARKET PARTICIPANTS, SIZE AND COMPOSITION

b describe functions of and participants in the foreign exchange market;

We now turn to the counterparties that participate in FX markets. As mentioned previously, there is an extremely diverse range of market participants, ranging in size from multi-billion-dollar investment funds down to individuals trading for their own account (including foreign tourists exchanging currencies at airport kiosks).

To understand the various market participants, it is useful to separate them into broad categories. One broad distinction is between what the market refers to as the *buy side* and the *sell side*. The sell side generally consists of large FX trading banks (such as Citigroup, UBS, and Deutsche Bank); the buy side consists of clients who use these banks to undertake FX transactions (i.e., buy FX products) from the sell-side banks.

The buy side can be further broken down into several categories:

- *Corporate accounts*: Corporations of all sizes undertake FX transactions during cross-border purchases and sales of goods and services. Many of their FX flows can also be related to cross-border investment flows—such as international mergers and acquisitions (M&A) transactions, investment of corporate funds in foreign assets, and foreign currency borrowing.
- *Real money accounts*: These are investment funds managed by insurance companies, mutual funds, pension funds, endowments, exchange-traded funds (ETFs), and other institutional investors. These accounts are referred to as real

money because they are usually restricted in their use of leverage or financial derivatives. This distinguishes them from leveraged accounts (discussed next); although, many institutional investors often engage in some form of leverage, either directly through some use of borrowed funds or indirectly using financial derivatives.

- *Leveraged accounts*: This category, often referred to as the professional trading community, consists of hedge funds, proprietary trading shops, commodity trading advisers (CTAs), high-frequency algorithmic traders, and the proprietary trading desks at banks—and indeed, almost any active trading account that accepts and manages FX risk for profit. The professional trading community accounts for a large and growing proportion of daily FX market turnover. These active trading accounts also have a wide diversity of trading styles. Some are macro-hedge funds that take longer term FX positions based on their views of the underlying economic fundamentals of a currency. Others are high-frequency algorithmic traders that use technical trading strategies (such as those based on moving averages or Fibonacci levels) and whose trading cycles and investment horizons are sometimes measured in milliseconds.
- *Retail accounts*: The simplest example of a retail account is the archetypical foreign tourist exchanging currency at an airport kiosk. However, it is important to realize that as electronic trading technology has reduced the barriers to entry into FX markets and the costs of FX trading, there has been a huge surge in speculative trading activity by retail accounts—consisting of individuals trading for their own accounts as well as smaller hedge funds and other active traders. This also includes households using electronic trading technology to move their savings into foreign currencies (this is relatively widespread among households in Japan, for example). It is estimated that retail trading accounts for as much as 10 percent of all spot transactions in some currency pairs and that this proportion is growing.
- *Governments*: Public entities of all types often have FX needs, ranging from relatively small (e.g., maintaining consulates in foreign countries) to large (e.g., military equipment purchases or maintaining overseas military bases). Sometimes these flows are purely transactional—the business simply needs to be done—and sometimes government FX flows reflect, at least in part, the public policy goals of the government. Some government FX business resembles that of investment funds, although sometimes with a public policy mandate as well. In some countries, public sector pension plans and public insurance schemes are run by a branch of the government. One example is the Caisse de dépôt et placement du Québec, which was created by the Québec provincial government in Canada to manage that province's public sector pension plans. The Caisse, as it is called, is a relatively large player in financial markets, with about CAD308 billion of assets under management as of mid-2018. Although it has a mandate to invest these assets for optimal return, it is also called upon to help promote the economic development of Québec. It should be noted that many governments—both at the federal and provincial/state levels—issue debt in foreign currencies; this, too, creates FX flows. Such supranational agencies as the World Bank and the African Development Bank issue debt in a variety of currencies as well.
- *Central banks*: These entities sometimes intervene in FX markets in order to influence either the level or trend in the domestic exchange rate. This often occurs when the central banks judge their domestic currency to be too weak and when the exchange rate has overshot any concept of equilibrium level (e.g., because of a speculative attack) to the degree that the exchange rate no longer reflects underlying economic fundamentals. Alternatively, central banks also

intervene when the FX market has become so erratic and dysfunctional that end-users such as corporations can no longer transact necessary FX business. Conversely, sometimes central banks intervene when they believe that their domestic currency has become too strong, to the point that it undercuts that country's export competitiveness. The Bank of Japan intervened against yen strength versus the US dollar in 2004 and again in March 2011 after the massive earthquake and nuclear disaster. Similarly, in 2010, 2013, and again in 2015, the Swiss National Bank intervened against strength in the Swiss franc versus the euro by selling the Swiss franc on the euro–Swiss (CHF/EUR) cross-rate. Central bank reserve managers are also frequent participants in FX markets in order to manage their country's FX reserves. In this context, they act much like real money investment funds—although generally with a cautious, conservative mandate to safeguard the value of their country's foreign exchange reserves. The foreign exchange reserves of some countries are enormous, and central bank participation in FX markets can sometimes have a material impact on exchange rates even when these reserve managers are not intervening for public policy purposes. Exhibit 2 provides information on central bank reserve holdings as of the second quarter of 2015.[3] Total central bank reserve holdings have held steady for several years and as of the first quarter of 2018 were $11,594 billion.

Exhibit 2 Currency Composition of Official Foreign Exchange Reserves, as of 1st Quarter 2015 (USD billion)

Total foreign exchange holdings globally	11,433
Held by advanced economies	3,946
Held by emerging and developing economies	7,487
Percent of global holdings held in the US dollar[a]	66%

[a] This percentage is calculated using that amount of global currency reserves for which the currency composition can be identified.

Note that the amount of foreign exchange reserves now held by emerging economies comfortably exceeds those held by developed economies. This largely reflects the rapid growth in foreign reserves held by Asian central banks, because these countries typically run large current account surpluses with the United States and other developed economies. Reserve accumulation by energy exporting countries in the Middle East and elsewhere is also a factor. Most of the global currency reserves are held in US dollars; the percentage held in USD is more than twice the portion held in the euro, the second most widely held currency in central bank foreign exchange reserves.

- *Sovereign wealth funds (SWFs)*: Many countries with large current account surpluses have diverted some of the resultant international capital flows into SWFs rather than into foreign exchange reserves managed by central banks. Although SWFs are government entities, their mandate is usually more oriented to purely investment purposes rather than public policy purposes. As such, SWFs can be thought of as akin to real money accounts, although some SWFs can employ derivatives or engage in aggressive trading strategies. It is generally understood that SWFs use their resources to help fulfill the public policy mandate of their government owners. The SWFs of many current account surplus countries

---

[3] See International Monetary Fund (2018) Currency Composition of Foreign Exchange Reserves (COFER) Tables 1-3.

(such as exporting countries in East Asia or oil-exporting countries) are enormous, and their FX flows can be an important determinant of exchange rate movements in almost all of the major currency pairs.

As mentioned, the sell side generally consists of the FX dealing banks that sell FX products to the buy side. Even here, however, distinctions can be made.

- A large and growing proportion of the daily FX turnover is accounted for by the very largest dealing banks, such as Deutsche Bank, Citigroup, UBS, HSBC, and a few other multinational banking behemoths. Maintaining a competitive advantage in FX requires huge fixed-cost investments in the electronic technology that connects the FX market, and it also requires a broad, global client base. As a result, only the largest banks are able to compete successfully in providing competitive price quotes to clients across the broad range of FX products. In fact, among the largest FX dealing banks, a large proportion of their business is crossed internally, meaning that these banks are able to connect buyers and sellers within their own extremely diverse client base and have no need to show these FX flows outside of the bank.
- All other banks fall into the second and third tier of the FX market sell side. Many of these financial institutions are regional or local banks with well-developed business relationships, but they lack the economies of scale, broad global client base, or information technology (IT) expertise required to offer competitive pricing across a wide range of currencies and FX products. In many cases, these are banks in emerging markets that don't have the business connections or credit lines required to access the FX market on a cost-effective basis on their own. As a result, these banks often outsource FX services by forming business relationships with the larger tier-one banks; otherwise, they depend on the deep, competitive liquidity provided by the largest FX market participants.

The categories presented are based on functions that are closely associated with the named groups. However, in some cases, functions typifying a group may also be assumed by or shared with another group. For example, sell-side banks provide FX price quotes. However, hedge funds and other large players may access the professional FX market on equal terms with the dealing banks and effectively act as market makers.

One of the most important ideas to draw from this categorization of market participants is that there is an extremely wide variety of FX market participants, reflecting a complex mix of trading motives and strategies that can vary with time. Most market participants reflect a combination of hedging and speculative motives in tailoring their FX risk exposures. Among public sector market participants, public policy motives may also be a factor. The dynamic, complex interaction of FX market participants and their trading objectives makes it difficult to analyze or predict movements in FX rates with any precision, or to describe the FX market adequately with simple characterizations.

3.1 Market Size and Composition

In this section, we present a descriptive overview of the global FX market drawn from the 2016 Triennial Survey undertaken by the Bank for International Settlements (BIS). The BIS is an umbrella organization for the world's central banks. Every three years, participating central banks undertake a survey of the FX market in their jurisdictions, the results of which are aggregated and compiled at the BIS. The most recent survey, taken in April 2016, gives a broad indication of the current size and distribution of global FX market flows.

As of April 2016, the BIS estimates that average daily turnover in the traditional FX market (comprised of spot, outright forward, and FX swap transactions) totaled approximately USD5.1 trillion. Exhibit 3 shows the approximate percentage allocation among FX product types, including both traditional FX products and exchange-traded FX derivatives. Note that this table of percentage allocations adds exchange-traded derivatives to the BIS estimate of average daily turnover of USD5.1 trillion; the "Spot" and "Outright forwards" categories include only transactions that are not executed as part of a swap transaction.

Exhibit 3 FX Turnover by Instrument

Spot	33%
Outright forwards	14
Swaps[a]	49
FX options	5
Total	**100%**

[a] Includes both FX and currency swaps.

The survey also provides a percentage breakdown of the average daily flows between sell-side banks (called the interbank market), between banks and financial customers (all non-bank financial entities, such as real money and leveraged accounts, SWFs, and central banks), and between banks and non-financial customers (such as corporations, retail accounts, and governments). The breakdown is provided in Exhibit 4. It bears noting that the proportion of average daily FX flow accounted for by financial clients is much larger than that for non-financial clients. The BIS also reports that the proportion of financial client flows has been growing rapidly, and in 2010 it exceeded interbank trading volume for the first time. This underscores the fact that only a minority of the daily FX flow is accounted for by corporations and individuals buying and selling foreign goods and services. Huge investment pools and professional traders are accounting for a large and growing proportion of the FX business.

Exhibit 4 FX Flows by Counterparty

Interbank	42%
Financial clients	51
Non-financial clients	8

The 2016 BIS survey also identifies the top five currency pairs in terms of their percentage share of average daily global FX turnover. These are shown in Exhibit 5. Note that each of these most active pairs includes the US dollar (USD).

Exhibit 5 FX Turnover by Currency Pair

Currency Pair	% of Market
USD/EUR	23.1%
JPY/USD	17.8%
USD/GBP	9.3%

Exhibit 5 (Continued)

Currency Pair	% of Market
USD/AUD	5.2%
CAD/USD	4.3%

The largest proportion of global FX trading occurs in London, followed by New York. This means that FX markets are most active between approximately 8:00 A.M. and 11:30 A.M. New York time, when banks in both cities are open. (The official London close is at 11:00 A.M. New York time, but London markets remain relatively active for a period after that.) Tokyo is the third-largest FX trading hub.

EXAMPLE 3

Market Participants and Composition of Trades

The investment adviser based in Sydney, Australia, makes the following statements to her client when describing some of the basic characteristics of the foreign exchange market:

Statement 1	"Foreign exchange transactions for spot settlement see the most trade volume in terms of average daily turnover because the FX market is primarily focused on settling international trade flows."
Statement 2	"The most important foreign exchange market participants on the buy side are corporations engaged in international trade; on the sell side they are the local banks that service their FX needs."

1 Statement 1 is:

- **A** correct.
- **B** incorrect with respect to the importance of spot settlements.
- **C** incorrect both with respect to the importance of spot settlements and international trade flows.

2 Statement 2 is:

- **A** correct.
- **B** incorrect with respect to corporations engaged in international trade.
- **C** incorrect with respect to both corporations and the local banks that service their trade needs.

Solution to 1:

C is correct. Although the media generally focus on the spot market when discussing foreign exchange, the majority of average daily trade volume involves the FX swap market as market participants either roll over or modify their existing hedging and speculative positions (or engage in FX swap financing). Although it is true that all international trade transactions eventually result in some form of spot settlement, this typically generates a great deal of hedging (and speculative) activity in advance of spot settlement. Moreover, an important group of FX market participants engages in purely speculative positioning with no intention of ever delivering/receiving the principal amount of the trades. Most FX trading volume is not related to international trade: Portfolio flows (cross-border capital movements) and speculative activities dominate.

Solution to 2:

C is correct. As of 2016, the most important foreign exchange market participants in terms of average daily turnover are found not among corporations engaged in international trade but among huge investment managers, both private (e.g., pension funds) and public (e.g., central bank reserve managers or sovereign wealth funds). A large and growing amount of daily turnover is also being generated by high-frequency traders who use computer algorithms to automatically execute extremely high numbers of speculative trades (although their individual ticket sizes are generally small, they add up to large aggregate flows). On the sell side, the largest money center banks (e.g., Deutsche Bank, Citigroup, HSBC, UBS) are increasingly dominating the amount of trading activity routed through dealers. Regional and local banks are increasingly being marginalized in terms of their share of average daily turnover in FX markets.

4 EXCHANGE RATE QUOTATIONS

c calculate and interpret the percentage change in a currency relative to another currency;

4.1 Exchange Rate Quotations

Exchange rates represent the relative price of one currency in terms of another. This price can be represented in two ways: 1) currency A buys how many units of currency B; or 2) currency B buys how many units of currency A. Of course, these two prices are simply the inverse of each other.

To distinguish between these two prices, market participants sometimes distinguish between *direct* and *indirect* exchange rates. In the quoting convention A/B (where there is a certain number of units of currency A per one unit of currency B), we refer to currency A as the *price currency* (or quote currency); currency B is referred to as the *base currency*. (The reason for this choice of names will become clearer below.) The base currency is always set at a quantity of one. A *direct* currency quote takes the domestic country as the price currency and the foreign country as the base currency. For example, for a Paris-based trader, the domestic currency would be the euro (EUR) and a foreign currency would be the UK pound (GBP). For this Paris-based trader, a *direct* quote would be EUR/GBP. An exchange rate quote of EUR/GBP = 1.1211 means that 1 GBP costs 1.1211 EUR. For this Paris-based trader, an *indirect* quote has the domestic currency—the euro—as the base currency. An indirect quote of GBP/EUR = 0.8920 means that 1 EUR costs 0.8920 GBP. *Direct and indirect quotes are just the inverse (reciprocal) of each other.*

It can be confusing to describe exchange rates as either being direct or indirect because determining the domestic currency and the foreign currency depends on where one is located. For a London-based market participant, the UK pound (GBP) is the domestic currency and the euro (EUR) is a foreign currency. For a Paris-based market participant, it would be the other way around.

To avoid confusion, the professional FX market has developed a set of market conventions that all market participants typically adhere to when making and asking for FX quotes. Exhibit 6 displays some of these for the major currencies: the currency code used for obtaining exchange rate quotes, how the market lingo refers to this currency pair, and the actual ratio—price currency per unit of base currency—represented by the quote.

Exhibit 6 Exchange Rate Quote Conventions

FX Rate Quote Convention	Name Convention	Actual Ratio (Price currency/Base currency)
EUR	Euro	USD/EUR
JPY	Dollar–yen	JPY/USD
GBP	Sterling	USD/GBP
CAD	Dollar–Canada	CAD/USD
AUD	Aussie	USD/AUD
NZD	Kiwi	USD/NZD
CHF	Swiss franc	CHF/USD
EURJPY	Euro–yen	JPY/EUR
EURGBP	Euro–sterling	GBP/EUR
EURCHF	Euro–Swiss	CHF/EUR
GBPJPY	Sterling–yen	JPY/GBP
EURCAD	Euro–Canada	CAD/EUR
CADJPY	Canada–yen	JPY/CAD

Several things should be noted in this exhibit. First, the three-letter currency codes in the first column (for FX rate quotes) refer to what are considered the major exchange rates. Remember that an exchange rate is the price of one currency in terms of another: There are always two currencies involved in the price. This is different from referring to a single currency in its own right. For example, one can refer to the euro (EUR) as a *currency*; but if we refer to a euro *exchange rate* (EUR), it is always the price of the euro in terms of another currency, in this case the US dollar. This is because in the professional FX market, the three-letter code EUR is always taken to refer to the euro–US dollar exchange rate, which is quoted in terms of the number of US dollars per euro (USD/EUR). Second, where there are six-letter currency codes in the first column, these refer to some of the major *cross-rates*. This topic will be covered in the next section, but generally these are secondary exchange rates and they are not as common as the main exchange rates. (It can be noted that three-letter codes are always in terms of an exchange rate involving the US dollar, while the six-letter codes are not.) Third, when both currencies are mentioned in the code or the name convention, *the base currency is always mentioned first, the opposite order of the actual ratio (price currency/base currency).* Thus, the code for "Sterling–yen" is "GBPJPY," but the actual number quoted is the number of yen per sterling (JPY/GBP). It should also be noted that *the codes may appear in a variety of formats that all mean the same thing.* For example, GBPJPY might instead appear as GBP:JPY or GBP–JPY. Fourth, regardless of where a market participant is located, there is always a mix of direct and indirect quotes in common market usage. For example, a trader based in Toronto will typically refer to the euro–Canada and Canada–yen exchange rates—a mixture of direct (CAD/EUR) and indirect (JPY/CAD) quotes for a Canadian-based trader. There is no overall consistency in this mixture of direct and indirect quoting conventions in the professional FX market; a market participant just has to get familiar with how the conventions are used.[4]

Another concept involving exchange rate quotes in professional FX markets is that of a *two-sided price*. When a client asks a bank for an exchange rate quote, the bank will provide a "*bid*" (the price at which the bank is willing to buy the currency) and an "*offer*" (the price at which the bank is willing to sell the currency). But there are *two* currencies involved in an exchange rate quote, which is always the price of one currency relative to the other. So, which one is being bought and sold in this two-sided price quote? This is where the lingo involving the price currency (or quote currency) and the base currency, explained above, becomes useful. *The two-sided price quoted by the dealer is in terms of buying/selling the base currency*. It shows the number of units of the *price* currency that the client will receive from the dealer for one unit of the base currency (the bid) and the number of units of the price currency that the client must sell to the dealer to obtain one unit of the base currency (the offer). Consider the case of a client that is interested in a transaction involving the Swiss franc (CHF) and the euro (EUR). As we have seen above, the market convention is to quote this as euro–Swiss (CHF/EUR). The EUR is the base currency, and the two-sided quote (price) shows the number of units of the price currency (CHF) that must be paid or will be received for 1 euro. For example, a two-sided price in euro–Swiss (CHF/EUR) might look like: 1.1583–1.1585. The client will receive CHF1.1583 for selling EUR1 to the dealer and must pay CHF1.1585 to the dealer to buy EUR1. Note that *the price is shown in terms of the price currency* and that *the bid is always less than the offer*: The bank buys the base currency (EUR, in this case) at the low price and sells the base currency at the high price. Buying low and selling high is profitable for banks, and spreading clients—trying to widen the bid/offer spread—is how dealers try to increase their profit margins. However, it should be noted that the electronic dealing systems currently used in professional FX markets are extremely efficient in connecting buyers and sellers globally. Moreover, this worldwide competition for business has compressed most bid/offer spreads to very tight levels. For simplicity, in the remainder of this reading we will focus on exchange rates as a single number (with no bid/offer spread).

One last thing that can be pointed out in exchange rate quoting conventions is that most major spot exchange rates are typically quoted to four decimal places. One exception among the major currencies involves the yen, for which spot exchange rates are usually quoted to two decimal places. (For example, using spot exchange rates from the middle of 2018, a USD/EUR quote would be expressed as 1.1701, while a JPY/EUR quote would be expressed as 130.9761.) This difference involving the yen comes from the fact that the units of yen per unit of other currencies is typically relatively large already, and hence extending the exchange rate quote to four decimal places is viewed as unnecessary.

Regardless of what quoting convention is used, changes in an exchange rate can be expressed as a percentage appreciation of one currency against the other: One simply has to be careful in identifying which currency is the price currency and which is the base currency. For example, let's suppose the exchange rate for the euro (USD/EUR) increases from 1.1500 to 1.2000. This represents an (un-annualized) percentage change of:

---

[4] In general, however, there is a hierarchy for quoting conventions. For quotes involving the EUR, it serves as the base currency (e.g., GBP/EUR). Next in the priority sequence, for quotes involving the GBP (but not the EUR) it serves as the base currency (e.g., USD/GBP). Finally, for quotes involving the USD (but not the GBP or EUR) it serves as the base currency (e.g., CAD/USD). Exceptions among the major currencies are the AUD and NZD: they serve as the base currency when quoted against the USD (i.e., USD/AUD, USD/NZD).

$$\frac{1.2000}{1.1500} - 1 = 4.35\%$$

This represents a 4.35 percent appreciation in the euro against the US dollar (and not an appreciation of the US dollar against the euro) because the USD/EUR exchange rate is expressed with the dollar as the price currency.

Note that this appreciation of the euro against the US dollar can also be expressed as a depreciation of the US dollar against the euro; but in this case, the depreciation is not equal to 4.35 percent. Inverting the exchange rate quote from USD/EUR to EUR/USD, so that the euro is now the price currency, leads to:

$$\frac{\left(\frac{1}{1.2000}\right)}{\left(\frac{1}{1.1500}\right)} - 1 = \frac{1.1500}{1.2000} - 1 = -4.17\%$$

Note that the US dollar depreciation is not the same, in percentage terms, as the euro appreciation. This will always be true; it is simply a matter of arithmetic.

EXAMPLE 4

Exchange Rate Conventions

A dealer based in New York City provides a spot exchange rate quote of 18.8590 MXN/USD to a client in Mexico City. The inverse of 18.8590 is 0.0530.

1 From the perspective of the Mexican client, the *most* accurate statement is that the:

  A direct exchange rate quotation is equal to 0.0530.

  B direct exchange rate quotation is equal to 18.8590.

  C indirect exchange rate quotation is equal to 18.8590.

2 If the bid/offer quote from the dealer was 18.8580 ~ 18.8600 MXN/USD, then the bid/offer quote in USD/MXN terms would be *closest* to:

  A 0.05302 ~ 0.05303.

  B 0.05303 ~ 0.05302.

  C 0.053025 ~ 0.053025.

Solution to 1:

B is correct. A direct exchange rate uses the domestic currency as the price currency and the foreign currency as the base currency. For an MXN/USD quote, the MXN is the price currency; therefore, the direct quote for the Mexican client is 18.8590 (it costs 18.8590 pesos to purchase 1 US dollar). Another way of understanding a *direct* exchange rate quote is that it is the price of one unit of foreign currency in terms of your own currency. This purchase of a unit of foreign currency can be thought of as a purchase much like any other you might make; think of the unit of foreign currency as just another item that you might be purchasing with your domestic currency. For example, for someone based in Canada, a liter of milk currently costs about CAD1.25 and USD1 costs about CAD1.30. This *direct* currency quote uses the *domestic* currency (the Canadian dollar, in this case) as the *price* currency and simply gives the price of a unit of foreign currency that is being purchased.

Solution to 2:

A is correct. An MXN/USD quote means the amount of MXN the dealer is bidding (offering) to buy (sell) USD1. The dealer's bid to buy USD1 at MXN18.8580 is equivalent to the dealer paying MXN18.8580 to buy USD1. Dividing both terms by 18.8580 means the dealer is paying (i.e., selling) MXN1 to buy USD0.05303. This is the offer in USD/MXN terms: The dealer offers to sell MXN1 at a price of USD0.05303. In USD/MXN terms, the dealer's bid for MXN1 is 0.05302, calculated by inverting the offer of 18.8600 in MXN/USD terms (1/18.8600 = 0.05302). Note that in any bid/offer quote, no matter what the base or price currencies, the bid is always lower than the offer.

5 CROSS-RATE CALCULATIONS

d calculate and interpret currency cross-rates;

Given two exchange rates involving three currencies, it is possible to back out what the cross-rate must be. For example, as we have seen, the FX market convention is to quote the exchange rate between the US dollar and the euro as euro–dollar (USD/EUR). The FX market also quotes the exchange rate between the Canadian dollar and US dollar as dollar–Canada (CAD/USD). Given these two exchange rates, it is possible to back out the cross-rate between the euro and the Canadian dollar, which according to market convention is quoted as euro–Canada (CAD/EUR). This calculation is shown as:

$$\frac{\text{CAD}}{\text{USD}} \times \frac{\text{USD}}{\text{EUR}} = \frac{\text{CAD}}{\cancel{\text{USD}}} \times \frac{\cancel{\text{USD}}}{\text{EUR}} = \frac{\text{CAD}}{\text{EUR}}$$

Hence, to get a euro–Canada (CAD/EUR) quote, we must multiply the dollar–Canada (CAD/USD) quote by the euro–dollar (USD/EUR) quote. For example, assume the exchange rate for dollar–Canada is 1.3020 and the exchange rate for euro–dollar is 1.1701. Using these sample spot exchange rates, calculating the euro–Canada cross-rate equals:

$$1.3020 \times 1.1701 = 1.5235 \text{ CAD per EUR}$$

It is best to avoid talking in terms of direct or indirect quotes because, as noted above, these conventions depend on where one is located and hence what the domestic and foreign currencies are. Instead, focus on how the math works: Sometimes it is necessary to invert one of the quotes in order to get the intermediary currency to cancel out in the equation to get the cross-rate. For example, to get a Canada–yen (JPY/CAD) quote, one is typically using the dollar–Canada (CAD/USD) rate and dollar–yen (JPY/USD) rate, which are the market conventions. This Canada–yen calculation requires that the dollar–Canada rate (CAD/USD) be inverted to a USD/CAD quote for the calculations to work, as shown below:

$$\left(\frac{\text{CAD}}{\text{USD}}\right)^{-1} \times \frac{\text{JPY}}{\text{USD}} = \frac{\text{USD}}{\text{CAD}} \times \frac{\text{JPY}}{\text{USD}} = \frac{\cancel{\text{USD}}}{\text{CAD}} \times \frac{\text{JPY}}{\cancel{\text{USD}}} = \frac{\text{JPY}}{\text{CAD}}$$

Hence, to get a Canada–yen (JPY/CAD) quote, we must first invert the dollar–Canada (CAD/USD) quote before multiplying by the dollar–yen (JPY/USD) quote. As an example, let's assume that we have spot exchange rates of 1.3020 for dollar–Canada

(CAD/USD) and 111.94 for dollar–yen (JPY/USD). The dollar–Canada rate of 1.3020 inverts to 0.7680; multiplying this value by the dollar–yen quote of 111.94 gives a Canada–yen quote of:

$$0.7680 \times 111.94 = 85.97 \text{ JPY per CAD}$$

Market participants asking for a quote in a cross-rate currency pair typically will not have to do this calculation themselves: Either the dealer or the electronic trading platform will provide a quote in the specified currency pair. (For example, a client asking for a quote in Canada–yen will receive that quote from the dealer; he will not be given separate dollar–Canada and dollar–yen quotes in order to do the math.) But be aware that dealers providing the quotes often have to do this calculation themselves if only because the dollar–Canada and dollar–yen currency pairs often trade on different trading desks and involve different traders. Electronic dealing machines used in both the interbank market and bank-to-client markets often provide this mathematical operation to calculate cross-rates automatically.

Because market participants can receive both a cross-rate quote (for example, Canada–yen) as well as the component underlying exchange rate quotes (for example, dollar–Canada and dollar–yen), these cross-rate quotes must be consistent with the above equation; otherwise, the market will arbitrage the mispricing. Extending our example above, we calculate a Canada–yen (JPY/CAD) rate of 85.97 based on underlying dollar–Canada (CAD/USD) and dollar–yen (JPY/USD) rates of 1.3020 and 111.94, respectively. Now suppose that at the same time a misguided dealer quotes a Canada–yen rate of 86.20. This is a different price in JPY/CAD for an identical service: converting yen into Canadian dollars. Hence, any trader could buy CAD1 at the lower price of JPY85.97and then turn around and sell CAD1 at JPY86.20(recall our earlier discussion of how price and base currencies are defined). The riskless arbitrage profit is JPY0.23 per CAD1. The arbitrage—called *triangular arbitrage*, "tri-," because it involves three currencies—would continue until the price discrepancy was removed.

In reality, however, these discrepancies in cross-rates almost never occur because both human traders and automatic trading algorithms are constantly on alert for any pricing inefficiencies. In practice, and for the purposes of this reading, we can consider cross-rates as being consistent with their underlying exchange rate quotes and that given any two exchange rates involving three currencies, we can back out the third cross-rate.

EXAMPLE 5

Cross Exchange Rates and Percentage Changes

A research report produced by a dealer includes the following exhibit:

	Spot Rate	Expected Spot Rate in One Year
USD/EUR	1.1701	1.1619
CHF/USD	0.9900	0.9866
USD/GBP	1.3118	1.3066

1 The spot CHF/EUR cross-rate is *closest* to:

A 0.8461.

B 0.8546.

C 1.1584.

2 The spot GBP/EUR cross-rate is *closest* to:

A 0.8920.

B 1.1211.

C 1.4653.

3 Based on the exhibit, the euro is expected to appreciate by how much against the US dollar over the next year?

A –0.7 percent

B +0.7 percent

C +1.0 percent

4 Based on the exhibit, the US dollar is expected to appreciate by how much against the British pound over the next year?

A +0.6 percent

B –0.4 percent

C +0.4 percent

5 Over the next year, the Swiss franc is expected to:

A depreciate against the GBP.

B depreciate against the EUR.

C appreciate against the GBP, EUR, and USD.

6 Based on the exhibit, which of the following lists the three currencies from strongest to weakest over the next year?

A USD, GBP, EUR

B USD, EUR, GBP

C EUR, USD, GBP

7 Based on the exhibit, which of the following lists the three currencies in order of appreciating the most to appreciating the least (in percentage terms) against the USD over the next year?

A GBP, CHF, EUR

B CHF, GBP, EUR

C EUR, CHF, GBP

Solution to 1:

C is correct:

$$\frac{\text{CHF}}{\text{EUR}} = \frac{\text{USD}}{\text{EUR}} \times \frac{\text{CHF}}{\text{USD}} = 1.1701 \times 0.9900 = 1.1584$$

Solution to 2:

A is correct:

$$\frac{\text{GBP}}{\text{EUR}} = \frac{\text{USD}}{\text{EUR}} \times \left(\frac{\text{USD}}{\text{GBP}}\right)^{-1} = \frac{\text{USD}}{\text{EUR}} \times \frac{\text{GBP}}{\text{USD}} = \frac{1.1701}{1.3118} = 0.8920$$

Solution to 3:

A is correct. The euro is the base currency in the USD/EUR quote, and the expected decrease in the USD/EUR rate indicates that the EUR is depreciating (in one year it will cost less USD to buy one EUR). Mathematically:

$$\frac{1.1619}{1.1701} - 1 = -0.7\%$$

Solution to 4:

C is correct. The GBP is the base currency in the USD/GBP quote, and the expected decrease in the USD/GBP rate means that the GBP is expected to depreciate against the USD. Or equivalently, the USD is expected to appreciate against the GBP. Mathematically:

$$\left(\frac{1.3066}{1.3118}\right)^{-1} - 1 = \frac{1.3118}{1.3066} - 1 = +0.4\%$$

Solution to 5:

C is correct: Because the question does not require calculating the magnitude of the appreciation or depreciation, we can work with CHF as either the price currency or the base currency. In this case, it is easiest to use it as the price currency. According to the table, CHF/USD is expected to decline from 0.9900 to 0.9866, so CHF is expected to be stronger (i.e., it should appreciate against the USD). CHF/EUR is currently 1.1584 (see the solution to problem 1) and is expected to be 1.1463 (= 0.9866 × 1.1619), so CHF is expected to appreciate against the EUR. CHF/GBP is currently 1.2987 (= 0.9900 × 1.3118) and is expected to be 1.2891 (= 0.9866 × 1.3066), so CHF is also expected to appreciate against the GBP.

Alternatively, we can derive this answer intuitively. The table shows that the CHF/USD rate is expected to decline: That is, the USD is expected to depreciate against the CHF, or alternatively, the CHF is expected to appreciate against the USD. The table also shows that the USD/EUR and USD/GBP rates are also decreasing, meaning that the EUR and GBP are expected to depreciate against the USD, or alternatively, the USD is expected to appreciate against the EUR and GBP. If the CHF is expected to appreciate against the USD and the USD is expected to appreciate against both the EUR and GBP, it follows that the CHF is expected to appreciate against both the EUR and GBP.

Solution to 6:

A is correct. According to the table, USD/EUR is expected to decline from 1.1701 to 1.1619, while USD/GBP is expected to decline from 1.3118 to 1.3066. So, the USD is expected to be stronger than both the EUR and GBP. GBP/EUR is currently 0.8920 [= $(1.3118)^{-1}$ × 1.1701] and is expected to be 0.8893 [= $(1.3066)^{-1}$ × 1.1619], so the GBP is expected to be stronger than the EUR.

Solution to 7:

B is correct. The USD/EUR rate depreciates by –0.7 percent (= [1.1619/1.1701] – 1), which is the depreciation of the base currency EUR against the USD. The USD/GBP rate declines –0.4 percent (= [1.3066/1.3118] – 1), which is the depreciation of the GBP against the USD. Inverting the CHF/USD rate to a USD/CHF convention shows that the base currency CHF appreciates by +0.35 percent against the USD (= [1.0136/1.0101] – 1).

FORWARD CALCULATIONS

6

e calculate an outright forward quotation from forward quotations expressed on a points basis or in percentage terms;

f explain the arbitrage relationship between spot rates, forward rates, and interest rates;

g calculate and interpret a forward discount or premium;

In professional FX markets, forward exchange rates are typically quoted in terms of points (also sometimes referred to as "pips"). The points on a forward rate quote are simply the difference between the forward exchange rate quote and the spot exchange rate quote, with the points scaled so that they can be related to the last decimal in the spot quote. When the forward rate is higher than the spot rate, the points are positive and the base currency is said to be trading at a *forward premium*. Conversely, if the forward rate is less than the spot rate, the points (forward rate minus spot rate) are negative and the base currency is said to be trading at a *forward discount*. Of course, if the base currency is trading at a forward premium, then the price currency is trading at a forward discount, and vice versa.

This can best be explained by means of an example. Mid-2018, the spot euro–dollar exchange rate (USD/EUR) was 1.15885 and the one-year forward rate was 1.19532. Hence, the forward rate was trading at a premium to the spot rate (the forward rate was larger than the spot rate) and the one-year forward points were quoted as +364.7. This +364.7 comes from:

$$1.19532 - 1.15885 = +0.03647$$

Recall that most non-yen exchange rates are quoted to four decimal places, so in this case we would scale up by four decimal places (multiply by 10,000) so that this +0.03647 would be represented as +364.7 points. Notice that the points are scaled to the size of the last digit in the spot exchange rate quote—usually the fourth decimal place. Notice as well that points are typically quoted to one (or more) decimal places, meaning that the forward rate will typically be quoted to five or more decimal places. The exception among the major currencies is the yen, which is typically quoted to two decimal places for spot rates. Here, forward points are scaled up by two decimal places—the last digit in the spot rate quote—by multiplying the difference between forward and spot rates by 100.

Typically, quotes for forward rates are shown as the number of forward points at each maturity.[5] These forward points are also called *swap points* because an FX swap consists of simultaneous spot and forward transactions. In the middle of 2018, a trader would have faced a spot rate and forward points in the euro–dollar (USD/EUR) currency pair similar to those in Exhibit 7:

Exhibit 7 Sample Spot and Forward Quotes

Maturity	Spot Rate or Forward Points
Spot	1.15885
One week	+5.6
One month	+27.1
Three months	+80.9
Six months	+175.6
Twelve months	+364.7

---

[5] As mentioned earlier, "maturity" is defined in terms of the time between spot settlement (usually T + 2) and the settlement of the forward contract.

Notice that the absolute number of points generally increases with maturity. This is because the number of points is proportional to the yield differential between the two countries (the Eurozone and the United States, in this case) scaled by the term to maturity. Given the interest rate differential, the longer the term to maturity, the greater the absolute number of forward points. Similarly, given the term to maturity, a wider interest rate differential implies a greater absolute number of forward points. (This will be explained and demonstrated in more detail later in this section.)

To convert any of these quoted forward points into a forward rate, one would divide the number of points by 10,000 (to scale down to the fourth decimal place, the last decimal place in the spot quote) and then add the result to the spot exchange rate quote.[6] For example, using the data in Exhibit 7, the three-month forward rate in this case would be:

$$1.15885 + \left(\frac{+80.9}{10{,}000}\right) = 1.15885 + 0.00809 = 1.16694$$

Occasionally, one will see the forward rate or forward points represented as a percentage of the spot rate rather than as an absolute number of points. Continuing our example from above, the three-month forward rate for USD/EUR can be represented as:

$$\frac{1.15885 + 0.00809}{1.15885} - 1 = \left(\frac{1.16694}{1.15885}\right) - 1 = +0.698\%$$

This shows that either the forward rate or the forward points can be used to calculate the percentage discount (or premium) in the forward market—in this case, +0.698 percent rounding to three decimal places. To convert a spot quote into a forward quote when the points are shown as a percentage, one simply multiples the spot rate by one plus the percentage premium or discount:

$$1.15885 \times (1 + 0.698\%) = 1.15885 \times (1.0000 + 0.00698) \approx 1.16694$$

Note that, rounded to the fifth decimal place, this is equal to our previous calculation. However, it is typically the case in professional FX markets that forward rates will be quoted in terms of pips rather than percentages.

We now turn to the relationship between spot rates, forward rates, and interest rates and how their relationship is derived. Forward exchange rates are based on an arbitrage relationship that equates the investment return on two alternative but equivalent investments. Consider the case of an investor with funds to invest. For simplicity, we will assume that there is one unit of the investor's domestic currency to be invested for one period. One alternative is to invest for one period at the domestic risk-free rate (i_d); at the end of the period, the amount of funds held is equal to $(1 + i_d)$. An alternative investment is to convert this one unit of domestic currency to foreign currency using the spot rate of $S_{f/d}$ (number of units of foreign currency per one unit of domestic currency). This can be invested for one period at the foreign risk-free rate; at the end of the period, the investor would have $S_{f/d}(1 + i_f)$ units of foreign currency. These funds must then be converted back to the investor's domestic currency. If the exchange rate to be used for this end-of-period conversion was precontracted at the start of the period (i.e., a forward rate was used), it would eliminate any foreign exchange risk from converting at a future, unknown spot rate. Given the assumed exchange rate convention here (foreign/domestic), the investor would obtain $(1/F_{f/d})$ units of the domestic currency for each unit of foreign currency sold forward. Note that this process of converting domestic funds in the spot FX market, investing at the foreign risk-free rate, and then converting back to the domestic currency with a forward rate is identical to the concept of swap financing described in an earlier section of this reading.

---

[6] Because the JPY/USD exchange rate is only quoted to two decimal places, forward points for the dollar–yen currency pair are divided by 100.

Hence, we have two alternative investments—both risk-free because both are invested at risk-free interest rates and because any foreign exchange risk was eliminated (hedged) by using a forward rate. Because these two investments are equal in risk characteristics, they must have the same return. Bearing in mind that the currency quoting convention is the number of foreign currency units per single domestic unit (*f/d*), this relationship can be stated as:

$$\left(1+i_d\right) = S_{f/d}\left(1+i_f\right)\left(\frac{1}{F_{f/d}}\right)$$

This is an arbitrage relationship because it describes two alternative investments (one on either side of the equal sign) that should have equal returns. If they do not, a riskless arbitrage opportunity exists because an investor can sell short the investment with the lower return and invest the funds in the investment with the higher return; the difference between the two returns is pure profit.[7]

This formula is perhaps the easiest and most intuitive way to remember the formula for the forward rate because it is based directly on the underlying intuition (the arbitrage relationship of two alternative but equivalent investments, one on either side of the equal sign). Also, the right-hand side of the equation, for the hedged foreign investment alternative, is arranged in proper time sequence: a) convert domestic to foreign currency; then b) invest the foreign currency at the foreign interest rate; and finally c) convert the foreign currency back to the domestic currency.[8]

This arbitrage equation can be re-arranged as needs require. For example, to get the formula for the forward rate, the above equation can be restated as:

$$F_{f/d} = S_{f/d}\left(\frac{1+i_f}{1+i_d}\right)$$

Another way of looking at this is, given the spot exchange rate and the domestic and foreign risk-free interest rates, the forward rate is whatever value completes this equation and eliminates any arbitrage opportunity. For example, let's assume that the spot exchange rate ($S_{f/d}$) is 1.6535, the domestic 12-month risk-free rate is 3.50 percent, and the foreign 12-month risk-free rate is 5.00 percent. The 12-month forward rate ($F_{f/d}$) must then be equal to:

$$1.6535\left(\frac{1.0500}{1.0350}\right) = 1.6775$$

Suppose instead that, with the spot exchange rate and interest rates unchanged, you were given a quote on the 12-month forward rate ($F_{f/d}$) of 1.6900. Because this misquoted forward rate does not agree with the arbitrage equation, it would present a riskless arbitrage opportunity. This can be seen by using the arbitrage equation to compute the return on the two alternative investment strategies. The return on the domestic-only investment approach is the domestic risk-free rate (3.50 percent). In contrast, the return on the hedged foreign investment when this misquoted forward rate is put into the arbitrage equation equals:

---

[7] It is because of this arbitrage relationship that the all-in financing rate using swap financing is close to the domestic interest rate.

[8] Recall that this equation is based on an *f*/*d* exchange rate quoting convention. If the exchange rate data were presented in *d*/*f* form, one could either invert these quotes back to *f*/*d* form and use the above equation or use the following equivalent equation: $(1 + i_d) = (1/S_{d/f})(1 + i_f)F_{d/f}$. If this latter equation were used, remember that forward and spot exchange rates are now being quoted on a *d*/*f* convention.

$$S_{f/d}\left(1+i_f\right)\left(\frac{1}{F_{f/d}}\right) = 1.6535(1.05)\left(\frac{1}{1.6900}\right) = 1.0273$$

This defines a return of 2.73 percent. Hence, the investor could make riskless arbitrage profits by borrowing at the higher foreign risk-free rate, selling the foreign currency at the spot exchange rate, hedging the currency exposure (buying the foreign currency back) at the misquoted forward rate, investing the funds at the lower domestic risk-free rate, and thereby getting a profit of 77 basis points (3.50% – 2.73%) for each unit of domestic currency involved—all with no upfront commitment of the investor's own capital. Any such opportunity in real-world financial markets would be quickly "arbed" away. It is interesting to note that in this example, the investor actually borrows at the higher of the two interest rates but makes a profit because the foreign currency is underpriced in the forward market.

The underlying arbitrage equation can also be re-arranged to show the forward rate as a percentage of the spot rate:

$$\frac{F_{f/d}}{S_{f/d}} = \left(\frac{1+i_f}{1+i_d}\right)$$

This shows that, given an *f/d* quoting convention, the forward rate will be higher than (be at a premium to) the spot rate if foreign interest rates are higher than domestic interest rates. More generally, and regardless of the quoting convention, *the currency with the higher (lower) interest rate will always trade at a discount (premium) in the forward market*.

One context in which forward rates are quoted as a percentage of spot rates occurs when forward rates are interpreted as expected future spot rates, or:

$$F_t = \hat{S}_{t+1}$$

Substituting this expression into the previous equation and doing some re-arranging leads to:

$$\frac{\hat{S}_{t+1}}{S_t} - 1 = \%\Delta\hat{S}_{t+1} = \left(\frac{i_f - i_d}{1+i_d}\right)$$

This shows that if forward rates are interpreted as expected future spot rates, the expected percentage change in the spot rate is proportional to the interest rate differential ($i_f - i_d$).

It is intuitively appealing to see forward rates as expected future spot rates. However, this interpretation of forward rates should be used cautiously. First, the direction of the expected change in spot rates is somewhat counter-intuitive. All else being equal, an increase in domestic interest rates (for example, the central bank tightens monetary policy) would typically be expected to lead to an increase in the value of the domestic currency. In contrast, the equation above indicates that, all else equal, a higher domestic interest rate implies slower expected appreciation (or greater expected depreciation) of the domestic currency (recall that this equation is based on an *f/d* quoting convention).

More important, historical data show that forward rates are poor predictors of future spot rates. Although various econometric studies suggest that forward rates may be unbiased predictors of future spot rates (i.e., they do not systematically over- or under-estimate future spot rates), this is not particularly useful information because the margin of error for these forecasts is so large. As we have seen in our introductory section, the FX market is far too complex and dynamic to be captured by a single variable, such as the level of the yield differential between countries. Moreover, as can be seen in the formula above for the forward rate, forward rates are based on domestic and foreign interest rates. This means that anything that affects the level and shape of the yield curve in either the domestic or foreign market will also affect the relationship between spot and forward exchange rates. In other words, FX markets do not operate in isolation but will reflect almost all factors affecting other markets globally; anything that affects expectations or risk premia in these other markets will reverberate in forward exchange rates as well. Although the level of the yield differential is one factor that the market may look at in forming spot exchange rate expectations, it is only one of many factors. (Many traders look to the trend in the yield differential rather than the level of the differential.) Moreover, there is a lot of noise in FX markets that makes almost any model—no matter how complex—a relatively poor predictor of spot rates at any given point in the future. In practice, FX traders and market strategists do *not* base either their currency expectations or trading strategies solely on forward rates.

For the purposes of this reading, *it is best to understand forward exchange rates simply as a product of the arbitrage equation outlined earlier and forward points as being related to the (time-scaled) interest rate differential between the two countries.* Reading any more than that into forward rates or interpreting them as the “market forecast” can be potentially misleading.

To understand the relationship between maturity and forward points, we need to generalize our arbitrage formula slightly. Suppose the investment horizon is a fraction, τ, of the period for which the interest rates are quoted. Then the interest earned in the domestic and foreign markets would be (i_d τ) and (i_f τ), respectively. Substituting this into our arbitrage relationship and solving for the difference between the forward and spot exchange rates gives:

$$F_{f/d} - S_{f/d} = S_{f/d}\left(\frac{i_f - i_d}{1 + i_d\tau}\right)\tau$$

This equation shows that forward points (appropriately scaled) are proportional to the spot exchange rate and to the interest rate differential and approximately (but not exactly) proportional to the horizon of the forward contract.

Let’s demonstrate this using an example. Suppose that we wanted to determine the 30-day forward exchange rate given a 30-day domestic risk-free interest rate of 2.00 percent per year, a 30-day foreign risk-free interest rate of 3.00 percent per year, and a spot exchange rate ($S_{f/d}$) of 1.6555. The risk-free assets used in this arbitrage relationship are typically bank deposits quoted using the London Interbank Offered Rate (Libor) for the currencies involved. The day count convention for Libor deposits is Actual/360.[9] Incorporating the fractional period (τ) as above and inserting the data into the forward rate equation leads to a 30-day forward rate of:

$$F_{f/d} = S_{f/d}\left(\frac{1 + i_f\tau}{1 + i_d\tau}\right) = 1.6555\left(\frac{1 + 0.0300\left[\frac{30}{360}\right]}{1 + 0.0200\left[\frac{30}{360}\right]}\right) = 1.6569$$

---

[9] This means that for interest calculation purposes, it is assumed that there are 360 days in the year. However, the actual number of days the funds are on deposit is used to calculate the interest payable.

This means that, for a 30-day term, forward rates are trading at a premium of 14 pips (1.6569 – 1.6555). This can also be calculated using the above formula for swap points:

$$F_{f/d} - S_{f/d} = S_{f/d}\left(\frac{i_f - i_d}{1 + i_d\tau}\right)\tau = 1.6555\left(\frac{0.0300 - 0.0200}{1 + 0.0200\left[\frac{30}{360}\right]}\right)\left[\frac{30}{360}\right] = 0.0014$$

As should be clear from this expression, the absolute number of swap points will be closely related to the term of the forward contract (i.e., approximately proportional to τ = Actual/360). For example, leaving the spot exchange rate and interest rates unchanged, let's set the term of the forward contract to 180 days:

$$F_{f/d} - S_{f/d} = 1.6555\left(\frac{0.0300 - 0.0200}{1 + 0.0200\left[\frac{180}{360}\right]}\right)\left[\frac{180}{360}\right] = 0.0082$$

This leads to the forward rate trading at a premium of 82 pips. The increase in the number of forward points is approximately proportional to the increase in the term of the contract (from 30 days to 180 days). Note that although the term of the 180-day forward contract is six times longer than that of a 30-day contract, the number of forward points is not exactly six times larger: 6 × 14 = 84.

Similarly, the number of forward points is proportional to the spread between foreign and domestic interest rates ($i_f - i_d$). For example, with reference to the original 30-day forward contract, let's set the foreign interest rate to 4.00 percent leaving the domestic interest rate and spot exchange rate unchanged. This doubles the interest rate differential ($i_f - i_d$) from 1.00 percent to 2.00 percent; it also doubles the forward points (rounding to four decimal places):

$$F_{f/d} - S_{f/d} = 1.6555\left(\frac{0.0400 - 0.0200}{1 + 0.0200\left[\frac{30}{360}\right]}\right)\left[\frac{30}{360}\right] = 0.0028$$

EXAMPLE 6

Forward Rates

A French company has recently finalized a sale of goods to a UK-based client and expects to receive a payment of GBP50 million in 32 days. The corporate treasurer at the French company wants to hedge the foreign exchange risk of this transaction and receives the following exchange rate information from a dealer:

GBP/EUR spot rate	0.8752
One-month forward points	–1.4

1 Given the above data, the treasurer could hedge the foreign exchange risk by:

A buying EUR (selling GBP) at a forward rate of 0.87380.

B buying EUR (selling GBP) at a forward rate of 0.87506.

C selling EUR (buying GBP) at a forward rate of 0.87506.

2 The *best* interpretation of the forward discount shown is that:

A the euro is expected to depreciate over the next 30 days.

- **B** one-month UK interest rates are higher than those in the Eurozone.
- **C** one-month Eurozone interest rates are higher than those in the United Kingdom.

3 If the 12-month forward rate is 0.87295 GBP/EUR, then based on the data the 12-month forward points are *closest* to:

- **A** –22.5.
- **B** –2.25.
- **C** –0.00225.

4 If a second dealer quotes GBP/EUR at a 12-month forward discount of 0.30 percent on the same spot rate, the French company could:

- **A** trade with either dealer because the 12-month forward quotes are equivalent.
- **B** lock in a profit in 12 months by buying EUR from the second dealer and selling it to the original dealer.
- **C** lock in a profit in 12 months by buying EUR from the original dealer and selling it to the second dealer.

5 If the 270-day Libor rates (annualized) for the EUR and GBP are 1.370% and 1.325%, respectively, and the spot GBP/EUR exchange rate is 0.8489, then the number of forward points for a 270-day forward rate ($F_{GBP/EUR}$) is *closest* to:

- **A** –22.8.
- **B** –3.8.
- **C** –2.8.

Solution to 1:

B is correct. The French company would want to convert the GBP to its domestic currency, the EUR (it wants to sell GBP, buy EUR). The forward rate would be equal to: 0.8752 + (–1.4/10,000) = 0.87506.

Solution to 2:

C is correct. A forward discount indicates that interest rates in the base currency country (France in this case, which uses the euro) are higher than those in the price currency country (the United Kingdom).

Solution to 3:

A is correct. The number of forward points is equal to the scaled difference between the forward rate and the spot rate. In this case: 0.87295 – 0.87520 = –0.00225. This is then multiplied by 10,000 to convert to the number of forward points.

Solution to 4:

B is correct. A 0.30 percent discount means that the second dealer will sell euros 12 months forward at 0.8752 × (1 – 0.0030) = 0.87257, a lower price per euro than the original dealer's quote of 0.87295. Buying euros at the cheaper 12-month forward rate (0.87257) and selling the same amount of euros 12 months forward at the higher 12-month forward rate (0.87295) means a profit of (0.87295 – 0.87257 = GBP 0.00038) per euro transacted, receivable when both forward contracts settle in 12 months.

Solution to 5:

C is correct, because the forward rate is calculated as:

$$F_{\frac{GBP}{EUR}} = S_{\frac{GBP}{EUR}}\left(\frac{1+i_{GBP}\left[\frac{Actual}{360}\right]}{1+i_{EUR}\left[\frac{Actual}{360}\right]}\right) = 0.8489\left(\frac{1+0.01325\left[\frac{270}{360}\right]}{1+0.01370\left[\frac{270}{360}\right]}\right) = 0.84862$$

This shows that the forward points are at a discount of: 0.84862 – 0.84890 = –0.00028, or –2.8 points. This can also be seen using the swap points formula:

$$F_{\frac{GBP}{EUR}} - S_{\frac{GBP}{EUR}} = 0.8489\left(\frac{0.01325-0.01370}{1+0.01370\left[\frac{270}{360}\right]}\right)\left[\frac{270}{360}\right] = -0.00028$$

The calculation of –3.8 points omits the day count (270/360), and –22.8 points gets the scaling wrong.

EXCHANGE RATE REGIMES- IDEALS AND HISTORICAL PERSPECTIVE

7

i. describe exchange rate regimes;

Highly volatile exchange rates create uncertainty that undermines the efficiency of real economic activity and the financial transactions required to facilitate that activity. Exchange rate volatility also has a direct impact on investment decisions because it is a key component of the risk inherent in foreign (i.e., foreign-currency-denominated) assets. Exchange rate volatility is also a critical factor in selecting hedging strategies for foreign currency exposures.

The amount of foreign exchange rate volatility will depend, at least in part, on the institutional and policy arrangements associated with trade in any given currency. Virtually every exchange rate is managed to some degree by central banks. The policy framework that each central bank adopts is called an *exchange rate regime*. Although there are many potential variations, these regimes fall into a few general categories. Before describing each of these types, we consider the possibility of an ideal regime and provide some historical perspective on the evolution of currency arrangements.

7.1 The Ideal Currency Regime

The ideal currency regime would have three properties. First, the exchange rate between any two currencies would be credibly fixed. This would eliminate currency-related uncertainty with respect to the prices of goods and services as well as real and financial assets. Second, all currencies would be fully convertible (i.e., currencies could be freely exchanged for any purpose and in any amount). This condition ensures unrestricted flow of capital. Third, each country would be able to undertake fully independent monetary policy in pursuit of domestic objectives, such as growth and inflation targets.

Unfortunately, these three conditions are not consistent. If the first two conditions were satisfied—credibly fixed exchange rates and full convertibility—then there would really be only one currency in the world. Converting from one national currency to another would have no more significance (indeed less) than deciding whether to carry coins or paper currency in your wallet. Any attempt to influence interest rates, asset prices, or inflation by adjusting the supply of one currency versus another would be futile. Thus, it should be clear that independent monetary policy is not possible if exchange rates are credibly fixed and currencies are fully convertible. *There can be no ideal currency regime.*

The impact of the currency regime on a country's ability to exercise independent monetary policy is a recurring theme in open-economy macroeconomics. It will be covered in more detail in other readings; however, it is worthwhile to emphasize the basic point by considering what would happen in an idealized world of perfect capital mobility. If the exchange rate were credibly fixed, then any attempt to decrease default-free interest rates in one country below those in another—that is, to undertake independent, expansionary monetary policy—would result in a potentially unlimited outflow of capital because funds would seek the higher return. The central bank would be forced to sell foreign currency and buy domestic currency to maintain the fixed exchange rate. The loss of reserves and reduction in the domestic money supply would put upward pressure on domestic interest rates until rates were forced back to equality, negating the initial expansionary policy. Similarly, contractionary monetary policy (higher interest rates) would be thwarted by an inflow of capital.

The situation is quite different, however, with a floating exchange rate. A decrease in the domestic interest rate would make the domestic currency less attractive. The resulting depreciation of the domestic currency would shift demand toward domestically produced goods (i.e., exports rise and imports fall), reinforcing the expansionary impact of the initial decline in the interest rate. Similarly, a contractionary increase in the interest rate would be reinforced by appreciation of the domestic currency.

In practice, of course, capital is not perfectly mobile and the impact on monetary policy is not so stark. The fact remains, however, that fixed exchange rates limit the scope for independent monetary policy and that national monetary policy regains potency and independence, at least to some degree, if the exchange rate is allowed to fluctuate and/or restrictions are placed on convertibility. In general, the more freely the exchange rate is allowed to float and the more tightly convertibility is controlled, the more effective the central bank can be in addressing domestic macroeconomic objectives. The downside, of course, is the potential distortion of economic activity caused by exchange rate risk and inefficient allocation of financial capital.

7.2 Historical Perspective on Currency Regimes

How currencies exchange for one another has evolved over the centuries. At any point in time, different exchange rate systems may coexist; still, there tends to be one dominant system in the world economy. Throughout most of the 19th century and the early 20th century until the start of World War I, the US dollar and the UK pound sterling operated on the "classical gold standard." The price of each currency was fixed in terms of gold. Gold was the numeraire[10] for each currency; therefore, it was indirectly the numeraire for all other prices in the economy. Many countries (e.g., the colonies of the United Kingdom) fixed their currencies relative to sterling and were therefore implicitly also operating on the classical gold standard.

The classical gold standard operated by what is called the *price-specie-flow mechanism*. This mechanism operated through the impact of trade imbalances on capital flows, namely gold. As countries experienced a trade surplus, they accumulated gold as payment, their domestic money supply expanded by the amount dictated by the fixed parity, prices rose, and exports fell. Similarly, when a country ran a trade deficit, there was an automatic outflow of gold, a contraction of the domestic money supply, and a fall in prices leading to increased exports.

---

[10] Economists refer to the unit of account in terms of which other goods, services, and assets are priced as the *numeraire*. Under the classical gold standard, the official value of each currency was expressed in ounces of gold.

In this system, national currencies were backed by gold. A country could only print as much money as its gold reserve warranted. The system was limited by the amount of gold, but it was self-adjusting and inspired confidence. With a fixed stock of gold, the price-specie-flow mechanism would work well. Still, new gold discoveries as well as more efficient methods of refining gold would enable a country to increase its gold reserves and increase its money supply apart from the effect of trade flows. In general, however, trade flows drove changes in national money supplies.[11]

There is much disagreement among economic historians about the effect of the classical gold standard on overall macroeconomic stability. Was it destabilizing? On the one hand, monetary policy was tied to trade flows, so a country could not engage in expansionary policies when there was a downturn in the non-traded sector. On the other hand, it has been argued that tying monetary policy to trade flows kept inflation in check.

During the 1930s, the use of gold as a clearing device for settlement of trade imbalances, combined with increasing protectionism on the part of economies struggling with depression as well as episodes of deflation and hyperinflation, created a chaotic environment for world trade. As a consequence of these factors, world trade dropped by over 50 percent and the gold standard was abandoned.

In the later stages of World War II, a new system of fixed exchange rates with periodic realignments was devised by John Maynard Keynes and Harry Dexter White, representing the UK and US Treasuries, respectively. The Bretton Woods system, named after the town where it was negotiated, was adopted by 44 countries in 1944. From the end of the war until the collapse of the system in the early 1970s, the United States, Japan, and most of the industrialized countries of Europe maintained a system of fixed parities for exchange rates between currencies. When the parities were significantly and persistently out of line with the balancing of supply and demand, there would be a realignment of currencies with some appreciating in value and others depreciating in value. These periodic realignments were viewed as a part of standard monetary policy.

By 1973, with chronic inflation taking hold throughout the world, most nations abandoned the Bretton Woods system in favor of a flexible exchange rate system under what are known as the Smithsonian Agreements. Milton Friedman had called for such a system as far back as the 1950s.[12] His argument was that the fixed parity system with periodic realignments would become unsustainable. When the inevitable realignments were imminent, large speculative profit opportunities would appear. Speculators would force the hand of monetary policy authorities, and their actions would distort the data needed to ascertain appropriate trade-related parities. It is better, he argued, to let the market, rather than central bank governors and treasury ministers, determine the exchange rate.

After 1973, most of the industrialized world changed to a system of flexible exchange rates. The original thinking was that the forces that caused exchange rate chaos in the 1930s—poor domestic monetary policy and trade barriers—would not be present in a flexible exchange rate regime, and therefore exchange rates would move in response to the exchange of goods and services among countries. As it turned out, however, exchange rates moved around much more than anyone expected. Academic economists and financial analysts alike soon realized that the high degree of exchange rate volatility was the manifestation of a highly liquid, forward-looking asset market.[13] Investment-driven FX transactions—for both long-term investment and short-term speculation—mattered much more in setting the spot exchange rate than anyone had previously imagined.

---

[11] The European inflation of the 17th century was an important exception. Discoveries of gold in South America led to an increase in the world gold stock and in prices throughout Europe. The impact was especially pronounced in Imperial Spain, the primary importing country. Historians have attributed the decline of the Spanish Empire, in part, to the loss of control of domestic prices.

[12] Friedman (1953).

There are costs, of course, to a high degree of exchange rate volatility. These include difficulty planning without hedging exchange rate risks—a form of insurance cost, domestic price fluctuations, uncertain costs of raw materials, and short-term interruptions in financing transactions. For these reasons, in 1979 the European Economic Community opted for a system of limited flexibility, the European Exchange Rate Mechanism (ERM).

Initially, the system called for European currency values to fluctuate within a narrow band called "the snake." This did not last long. The end of the Cold War and the re-unification of Germany created conditions ripe for speculative attack. In the early 1990s, the United Kingdom was in a recession and the government's monetary policy leaned toward low interest rates to stimulate economic recovery. Germany was issuing large amounts of debt to pay for re-unification, and the German central bank (the Deutsche Bundesbank) opted for high interest rates to ensure price stability. Capital began to flow from sterling to Deutsche marks to obtain the higher interest rate. The Bank of England tried to lean against these flows and maintain the exchange rate within the Exchange Rate Mechanism, but eventually it began to run out of marks to sell. Because it was almost certain that devaluation would be required, holders of sterling rushed to purchase marks at the old rate and the speculative attack forced the United Kingdom out of the ERM in September 1992, only two years after it finally joined the system.

Despite these difficulties, 1999 saw the creation of a common currency for most Western European countries, without Switzerland or the United Kingdom, called the euro.[14] The hope was that the common currency would increase transparency of prices across borders in Europe, enhance market competition, and facilitate more efficient allocation of resources. The drawback, of course, is that each member country lost the ability to manage its exchange rate and therefore to engage in independent monetary policy.

8 A TAXONOMY OF CURRENCY REGIMES

i. describe exchange rate regimes;

Although the pros and cons of fixed and flexible exchange rate regimes continue to be debated, regimes have been adopted that lie somewhere between these polar cases. In some cases, the driving force is the lack of credibility with respect to sound monetary policy. An economy with a history of hyperinflation may be forced to adopt a form of fixed-rate regime because its promise to maintain a sound currency with a floating rate regime would not be credible. This has been a persistent issue in Latin America. In other cases, the driving force is as much political as economic. The decision to create the euro was strongly influenced by the desire to enhance political union within the European Community, whose members had been at war with each other twice in the 20th century.

---

[13] Whether or not FX markets satisfy recognized definitions of market efficiency—correctly reflecting all available information—is debatable (e.g., some point to evidence of trending as a clear violation of efficiency). However, there is no doubt that FX market participants attempt to incorporate new information, which is often lumpy and difficult to decipher, into their expectations about the future. Changing expectations—accurate or otherwise—affect the value that investors place on holding different currencies and, in a highly liquid market, lead to rapid and sometimes violent exchange rate movements.

[14] The number of European countries adopting the euro has continued to expand since its inception; the most recent country to join the euro was Lithuania, on 1 January 2015.

As of April 2008, the International Monetary Fund (IMF) classified exchange rate regimes into the eight categories shown in Exhibit 8.

Exhibit 8 Exchange Rate Regimes for Selected Economies[15] As of 30 April 2008

	Currency Anchor		
Type of Regime	**USD**	**EUR**	**Basket/None**
No separate legal tender			
Dollarized	Ecuador, El Salvador, Marshall Islands, Micronesia, Palau, Panama, Timor-Leste, Zimbabwe	Kosovo, Montenegro, San Marino	Kiribati, Tuvalu
Monetary union		*EMU*: Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Luxembourg, Lithuania, Malta, Netherlands, Portugal, Slovak Rep., Slovenia, Spain	
Currency board	Djibouti, Hong Kong SAR, Antigua and Barbuda	Bosnia and Herzegovina, Bulgaria	Brunei Darussalam
Fixed parity	Aruba, The Bahamas, Bahrain, Barbados, Belize, Curaçao and Saint Maarten, Eritrea, Jordan, Oman, Qatar, Saudi Arabia, South Sudan, Turkmenistan, UAE, Venezuela	Cabo Verde, Comoros, Denmark, São Tomé and Príncipe WAEMU: Benin, Burkina Faso, Côte d'Ivoire, Guinea-Bissau, Mali, Niger, Senegal, Togo CEMAC: Cameroon, Central African Rep., Chad, Rep. of Congo, Equatorial Guinea, Gabon	Fiji, Kuwait, Libya, Morocco, Samoa, Bhutan, Lesotho, Namibia, Nepal, Swaziland
Target zone		Slovak Republic	Syria
Crawling peg	Nicaragua		Botswana
Crawling band	Honduras, Jamaica	Croatia	China, Ethiopia, Uzbekistan, Armenia, Dominican Republic, Guatemala, Argentina, Belarus, Haiti, Switzerland, Tunisia

(continued)

---

[15] The classifications are described in the IMF's Annual Report on Exchange Arrangements and Exchange Restrictions 2014. In some cases, the labels used by the IMF do not clearly distinguish among the regimes. Hence, the names applied here to the regimes differ somewhat from the IMF's original taxonomy.

Exhibit 8 (Continued)

	Currency Anchor		
Type of Regime	**USD**	**EUR**	**Basket/None**
Managed float	Cambodia, Liberia		Algeria, Iran, Syria, The Gambia, Myanmar, Nigeria, Rwanda, Czech Rep., Costa Rica, Malaysia, Mauritania Pakistan, Russia, Sudan, Vanuatu
Independent float *Currency columns due to floating*	Australia, Canada, Chile, Japan, Mexico, Norway, Poland, Sweden, United Kingdom, Somalia, United States	Albania, Brazil, Colombia, Georgia, Ghana, Hungary, Iceland, Indonesia, Israel, Korea, Moldova, New Zealand, Paraguay, Peru	Philippines, Romania, Serbia, South Africa, Thailand, Turkey, Uganda

It should be noted that global financial markets are too complex and diverse to be fully captured by this (or any other) classification system. A government's control over the domestic currency's exchange rate will depend on many factors; for example, the degree of capital controls used to prevent the free flow of funds in and out of the economy. Also, even under an "independent float" regime monetary authorities will occasionally intervene in foreign exchange markets in order to influence the value of their domestic currency. Additionally, the specifics of exchange rate policy implementation are subject to change.

This means that the classifications in Exhibit 8 are somewhat arbitrary and subject to interpretation, as well as change, over time. The important point to be drawn from this discussion is that the prices and flows in foreign exchange markets will, to varying degrees, reflect the legal and regulatory framework imposed by governments, not just "pure" market forces. Governments have a variety of motives and tools for attempting to manage exchange rates. The taxonomy in Exhibit 8 can be used to help understand the main distinctions among currency regimes and the rationales for adopting them, but the specific definitions should not be interpreted too rigidly. Instead, the focus should be on the diversity of foreign exchange markets globally as well as the implications of these various currency regimes for market pricing.

8.1 Arrangements with No Separate Legal Tender

The IMF identifies two types of arrangements in which a country does not have its own legal tender. In the first, known as *dollarization*, the country uses the currency of another nation as its medium of exchange and unit of account. In the second, the country participates in a monetary union whose members share the same legal tender. In either case, the country gives up the ability to conduct its own monetary policy.

In principle, a country could adopt any currency as its medium of exchange and unit of account, but the main reserve currency, the US dollar, is an obvious choice—hence the name dollarization. Many countries are dollarized: East Timor, El Salvador, Ecuador, and Panama, for example. By adopting another country's currency as legal tender, a dollarized country inherits that country's currency credibility, but not its credit-worthiness. For example, although local banks may borrow, lend, and accept deposits in US dollars, they are not members of the US Federal Reserve System nor are they backed by deposit insurance from the Federal Deposit Insurance Corporation. Thus, interest rates on US dollars in a dollarized economy need not be, and generally are not, the same as on dollar deposits in the United States.

Dollarization imposes fiscal discipline by eliminating the possibility that the central bank will be induced to monetize government debt (i.e., to persistently purchase government debt with newly created local currency). For countries with a history of fiscal excess or lack of monetary discipline, dollarizing the economy can facilitate growth of international trade and capital flows if it creates an expectation of economic and financial stability. In the process, however, it removes another potential source of stabilization—domestic monetary policy.

The European Economic and Monetary Union (EMU) is the most prominent example of the second type of arrangement lacking separate legal tender. Each EMU member uses the euro as its currency. Although member countries cannot have their own monetary policies, they jointly determine monetary policy through their representation at the European Central Bank (ECB). As with dollarization, a monetary union confers currency credibility on members with a history of fiscal excess and/or a lack of monetary discipline. However, as shown by the 2010 EMU sovereign debt crisis, monetary union alone cannot confer credit-worthiness.

8.2 Currency Board System

The IMF defines a *currency board system* (CBS) as:

> A monetary regime based on an explicit legislative commitment to exchange domestic currency for a specified foreign currency at a fixed exchange rate, combined with restrictions on the issuing authority to ensure fulfillment of its legal obligation. This implies that domestic currency will be issued only against foreign exchange and it remains fully backed by foreign assets….[16]

Hong Kong SAR has the leading example of a long-standing (since 1983) currency board. US dollar reserves are held to cover, at the fixed parity, the entire *monetary base*—essentially bank reserves plus all HKD notes and coins in circulation.[17] Note that HKD-denominated bank deposits are not fully collateralized by US dollar reserves; to do so would mean that banks could not lend against their deposits. The Hong Kong Monetary Authority (HKMA) does not function as a traditional central bank under this system because the obligation to maintain 100 percent foreign currency reserves against the monetary base prevents it from acting as a lender-of-last-resort for troubled financial institutions. However, it can provide short-term liquidity by lending against foreign currency collateral.

A CBS works much like the classical gold standard in that expansion and contraction of the monetary base are directly linked to trade and capital flows. As with the gold standard, a CBS works best if domestic prices and wages are very flexible, non-traded sectors of the domestic economy are relatively small, and the global supply of the reserve asset grows at a slow, steady rate consistent with long-run real growth with stable prices. The first two of these conditions are satisfied in Hong Kong SAR. Until and unless Hong Kong SAR selects a new reserve asset, however, the third condition depends on US monetary policy.

In practice, the HKD exhibits modest fluctuations around the official parity of HKD/USD = 7.80 because the HKMA buys (sells) USD at a pre-announced level slightly below (above) the parity. Persistent flows on one side of this convertibility zone or the other result in interest rate adjustments rather than exchange rate adjustments. Inside the zone, however, the exchange rate is determined by the market and the HKMA is free to conduct limited monetary operations aimed at dampening transitory interest rate movements.

---

[16] International Monetary Fund (2006).

[17] For a description of Hong Kong's currency board system, see Hong Kong Monetary Authority (2005).

One of the advantages of a CBS as opposed to dollarization is that the monetary authority can earn a profit by paying little or no interest on its liability—the monetary base—and can earn a market rate on its asset—the foreign currency reserves. This profit is called *seigniorage*.[18] Under dollarization, the seigniorage goes to the monetary authority whose currency is used.

8.3 Fixed Parity

A simple fixed-rate system differs from a CBS in two important respects. First, there is no legislative commitment to maintaining the specified parity. Thus, market participants know that the country may choose to adjust or abandon the parity rather than endure other, potentially more painful, adjustments. Second, the target level of foreign exchange reserves is discretionary; it bears no particular relationship to domestic monetary aggregates. Thus, although monetary independence is ultimately limited as long as the exchange peg is maintained, the central bank can carry out traditional functions, such as serving as lender of last resort.

In the conventional fixed-rate system, the exchange rate may be pegged to a single currency—for example, the US dollar—or to a basket index of the currencies of major trading partners. There is a band of up to ±1 percent around the parity level within which private flows are allowed to determine the exchange rate. The monetary authority stands ready to spend its foreign currency reserves, or buy foreign currency, in order to maintain the rate within these bands.

The credibility of the fixed parity depends on the country's willingness and ability to offset imbalances in private sector demand for its currency. Both excess and deficient private demand for the currency can exert pressure to adjust or abandon the parity. Excess private demand for the domestic currency implies a rapidly growing stock of foreign exchange reserves, expansion of the domestic money supply, and potentially accelerating inflation. Deficient demand for the currency depletes foreign exchange reserves and exerts deflationary pressure on the economy. If market participants believe the foreign exchange reserves are insufficient to sustain the parity, then that belief may be self-fulfilling because the resulting speculative attack will drain reserves and may force an immediate devaluation. Thus, the level of reserves required to maintain credibility is a key issue for a simple fixed exchange rate regime.

8.4 Target Zone

A target zone regime has a fixed parity with fixed horizontal intervention bands that are somewhat wider—up to ±2 percent around the parity—than in the simple fixed parity regime. The wider bands provide the monetary authority with greater scope for discretionary policy.

8.5 Active and Passive Crawling Pegs

Crawling pegs for the exchange rate—usually against a single currency, such as the US dollar—were common in the 1980s in Latin America, particularly Brazil, during the high inflation periods. To prevent a run on the US dollar reserves, the exchange rate was adjusted frequently (weekly or daily) to keep pace with the inflation rate. Such a system was called a passive crawl. An adaptation used in Argentina, Chile, and Uruguay was the active crawl: The exchange rate was pre-announced for the coming weeks with changes taking place in small steps. The aim of the active crawl was to manipulate expectations of inflation. Because the domestic prices of many goods were directly tied to import prices, announced changes in the exchange rate would effectively signal future changes in the inflation rate of these goods.

---

[18] More generally, seigniorage is the profit earned when the value of money issued exceeds the cost of producing it. For physical currency, seigniorage arises when a coin is minted for a fraction of its face value and then issued (sold) at its face value.

8.6 Fixed Parity with Crawling Bands

A country can also have a fixed central parity with crawling bands. Initially, a country may fix its rates to a foreign currency to anchor expectations about future inflation but then gradually permit more and more flexibility in the form of a pre-announced widening band around the central parity. Such a system has the desirable property of allowing a gradual exit strategy from the fixed parity. A country might want to introduce greater flexibility and greater scope for monetary policy, but it may not yet have the credibility or financial infrastructure for full flexibility. So it maintains a fixed parity with slowly widening bands.

8.7 Managed Float

A country may simply follow an exchange rate policy based on either internal or external policy targets—intervening or not to achieve trade balance, price stability, or employment objectives. Such a policy, often called *dirty floating,* invites trading partners to respond likewise with their exchange rate policy and potentially decreases stability in foreign exchange markets as a whole. The exchange rate target, in terms of either a level or a rate of change, is typically not explicit.

8.8 Independently Floating Rates

In this case, the exchange rate is left to market determination and the monetary authority is able to exercise independent monetary policy aimed at achieving such objectives as price stability and full employment. The central bank also has latitude to act as a lender of last resort to troubled financial institutions, if necessary.

It should be clear from recent experience that the concepts of float, managed float, crawl, and target zone are not hard and fast rules. Central banks do occasionally engage, implicitly or explicitly, in regime switches—even in countries nominally following an independently floating exchange rate regime. For example, when the US dollar appreciated in the mid-1980s with record US trade deficits, then-US Treasury Secretary James Baker engineered the Plaza Accord, in which Japan and Germany engineered an appreciation of their currencies against the US dollar. (The "Plaza Accord" is so named because it was negotiated at the Plaza Hotel in New York City.) This 1985 policy agreement involved a combination of fiscal and monetary policy measures by the countries involved as well as direct intervention in foreign exchange markets. The Plaza Accord was a clear departure from a pure independently floating exchange rate system.

There are more recent examples of government intervention in foreign exchange markets. In September 2000, the European Central Bank, the Federal Reserve Board, the Bank of Japan, the Bank of England, and the Bank of Canada engaged in "concerted" intervention in order to support the value of the euro, a "freely floating" currency which was then under pressure within foreign exchange markets. (This intervention was described as "concerted" because it was pre-arranged and coordinated among the central banks involved.) During 2010, many countries engaged in unilateral intervention to prevent the rapid appreciation of their currencies against the US dollar. Several of these countries also employed various fiscal and regulatory measures (for example, taxes on capital inflows) in order to further affect exchange rate movements.

The important point to draw from this discussion is that exchange rates do not only reflect private sector market forces but will also, to varying degrees, be influenced by the legal and regulatory framework (currency regimes) within which foreign exchange markets operate. Moreover, they will occasionally be influenced by government polices (fiscal, monetary, and intervention) intended to manage exchange rates. All of these can vary widely among countries and are subject to change with time.

Nonetheless, the most widely traded currencies in foreign exchange markets (the US dollar, yen, euro, UK pound, Swiss franc, and the Canadian and Australian dollars) are typically considered to be free floating, although subject to relatively infrequent intervention.

EXAMPLE 7

Currency Regimes

An investment adviser in Los Angeles, USA, is meeting with a client who wishes to diversify her portfolio by including more international investments. In order to evaluate the suitability of international diversification for the client, the adviser attempts to explain some of the characteristics of foreign exchange markets. The adviser points out that exchange rate regimes affect the performance of domestic economies as well as the amount of foreign exchange risk posed by international investments.

The client and her adviser discuss potential investments in Hong Kong SAR, Panama, and Canada. The adviser notes that the currency regimes of Hong Kong SAR, Panama, and Canada are a currency board, dollarization, and a free float, respectively. The adviser tells his client that these regimes imply different degrees of foreign exchange risk for her portfolio.

The discussion between the investment adviser and his client then turns to potential investments in other markets with different currency regimes. The adviser notes that some markets are subject to fixed parity regimes against the US dollar. The client asks whether a fixed parity regime would imply less foreign currency risk for her portfolio than would a currency board. The adviser replies: "Yes, a fixed parity regime means a constant exchange rate and is more credible than a currency board."

The adviser goes on to explain that in some markets exchange rates are allowed to vary, although with different degrees of foreign exchange market intervention to limit exchange rate volatility. Citing examples, he notes that mainland China has a crawling peg regime with reference to the US dollar, but the average daily percentage changes in mainland China/US exchange rate are very small compared with the average daily volatility for a freely floating currency. The adviser also indicates that Denmark has a target zone regime with reference to the euro, and South Korea usually follows a freely floating currency regime but sometimes switches to a managed float regime. The currencies of mainland China, Denmark, and South Korea are the yuan renminbi (CNY), krone (DKK), and won (KRW), respectively.

1 Based solely on the exchange rate risk the client would face, what is the correct ranking (from most to least risky) of the following investment locations?

A Panama, Canada, Hong Kong SAR.

B Canada, Hong Kong SAR, Panama.

C Hong Kong SAR, Panama, Canada.

2 Based solely on their foreign exchange regimes, which investment location is least likely to import inflation or deflation from the United States?

  A Canada.

  B Panama.

  C Hong Kong SAR.

3 The adviser's statement about fixed parity regimes is incorrect with regard to:

  A credibility.

  B a constant exchange rate.

  C both a constant exchange rate and credibility.

4 Based on the adviser's categorization of mainland China's currency regime, if the USD is depreciating against the KRW, then it is *most* likely correct that the CNY is:

  A fixed against the KRW.

  B appreciating against the KRW.

  C depreciating against the KRW.

5 Based on the adviser's categorization of Denmark's currency regime, it would be *most* correct to infer that the:

  A krone is allowed to float against the euro within fixed bands.

  B Danish central bank will intervene if the exchange rate strays from its target level.

  C target zone will be adjusted periodically in order to manage inflation expectations.

6 Based on the adviser's categorization of South Korea's currency policy, it would be *most* correct to infer that the:

  A Korean central bank is engineering a gradual exit from a fixed-rate regime.

  B government is attempting to peg the exchange rate within a predefined zone.

  C won is allowed to float, but with occasional intervention by the Korean central bank.

Solution to 1:

B is correct. The CAD/USD exchange rate is a floating exchange rate, and Canadian investments would therefore carry exchange rate risk for a US-based investor. Although Hong Kong SAR follows a currency board system, the HKD/USD exchange rate nonetheless does display some variation, albeit much less than in a floating exchange rate regime. In contrast, Panama has a dollarized economy (i.e., it uses the US dollar as the domestic currency); therefore, there is no foreign exchange risk for a US investor.

Solution to 2:

A is correct. The Canadian dollar floats independently against the US dollar leaving the Bank of Canada able to adjust monetary policy to maintain price stability. Neither Hong Kong SAR (currency board) nor Panama (dollarized) can exercise independent monetary policy to buffer its economy from the inflationary/deflationary consequences of US monetary policy.

Solution to 3:

C is correct. A fixed exchange rate regime does not mean that the exchange rate is rigidly fixed at a constant level. In practice, both a fixed-rate regime and a currency board allow the exchange rate to vary within a band around the stated parity level. Thus, both regimes involve at least a modest amount of exchange rate risk. The fixed parity regime exposes the investor to the additional risk that the parity may not be maintained. In a fixed parity regime, the level of foreign currency reserves is discretionary and typically only a small fraction of the domestic money supply. With no legal obligation to maintain the parity, the monetary authority may adjust the parity (devalue or revalue its currency) or allow its currency to float if doing so is deemed to be less painful than other adjustment mechanisms (e.g., fiscal restraint). In contrast, a currency board entails a legal commitment to maintain the parity and to fully back the domestic currency with reserve currency assets. Hence, there is little risk that the parity will be abandoned.

Solution to 4:

C is correct. If the CNY is subject to a crawling peg with very small daily adjustments versus the USD and the USD is depreciating against the KRW, then the CNY would *most* likely be depreciating against the KRW as well. In fact, this was an important issue in foreign exchange markets through the latter part of 2010: As the USD depreciated against most Asian currencies (and less so against the CNY), many Asian countries felt that they were losing their competitive export advantage because the CNY was so closely tied to the USD. This led many Asian countries to intervene in FX markets against the strength of their domestic currencies in order not to lose an export pricing advantage against the Chinese mainland.

Solution to 5:

A is correct. A target zone means that the exchange rate between the euro and Danish krone (DKK) will be allowed to vary within a fixed band (as of 2010, the target zone for the DKK/EUR is a ±2.5 percent band). This does not mean that the DKK/EUR rate is fixed at a certain level (B is incorrect) or that the target zone will vary in order to manage inflation expectations (this is a description of a crawling peg, which makes C incorrect).

Solution to 6:

C is correct. Similar to the monetary authorities responsible for many of the world's major currencies, the South Korean policy typically involves letting market forces determine the exchange rate (an independent floating rate regime). But this approach does not mean that market forces are the sole determinant of the won exchange rate. As with most governments, the South Korean policy is to intervene in foreign exchange markets when movements in the exchange rate are viewed as undesirable (a managed float). For example, during the later part of 2010, South Korea and many other countries intervened in foreign exchange markets to moderate the appreciation of their currencies against the US dollar. Answer A describes a fixed parity with a crawling bands regime, and B describes a target zone regime: Both answers are incorrect.

EXCHANGE RATES AND THE TRADE BALANCE: INTRODUCTION

9

j explain the effects of exchange rates on countries' international trade and capital flows.

Just as a family that spends more than it earns must borrow or sell assets to finance the excess, a country that imports more goods and services than it exports must borrow from foreigners or sell assets to foreigners to finance the trade deficit. Conversely, a country that exports more goods and services than it imports must invest the excess either by lending to foreigners or by buying assets from foreigners. Thus, a trade deficit (surplus) must be exactly matched by an offsetting *capital account* surplus (deficit).[19] This implies that any factor that affects the trade balance must have an equal and opposite impact on the capital account, and vice versa. To put this somewhat differently, the *impact of exchange rates and other factors on the trade balance must be mirrored by their impact on capital flows*: They cannot affect one without affecting the other.

Using a fundamental identity from macroeconomics, the relationship between the trade balance and expenditure/saving decisions can be expressed as:[20]

$$X - M = (S - I) + (T - G)$$

where X represents exports, M is imports, S is private savings, I is investment in plant and equipment, T is taxes net of transfers, and G is government expenditure. From this relationship, we can see that a trade surplus ($X > M$) must be reflected in a fiscal surplus ($T > G$), an excess of private saving over investment ($S > I$), or both. Because a fiscal surplus can be viewed as government saving, we can summarize this relationship more simply by saying that a trade surplus means the country saves more than enough to fund its investment (I) in plant and equipment. The excess saving is used to accumulate financial claims on the rest of the world. Conversely, a trade deficit means the country does not save enough to fund its investment spending (I) and must reduce its net financial claims on the rest of the world.

Although this identity provides a key link between real expenditure/saving decisions and the aggregate flow of financial assets into or out of a country, it does not tell us what type of financial assets will be exchanged or in what currency they will be denominated. All that can be said is that asset prices and exchange rates at home and abroad must adjust so that all financial assets are willingly held by investors.

If investors anticipate a significant change in an exchange rate, they will try to sell the currency that is expected to depreciate and buy the currency that is expected to appreciate. This implies an incipient (i.e., potential) flow of capital from one country to the other, which must either be accompanied by a simultaneous shift in the trade balance or be discouraged by changes in asset prices and exchange rates. Because expenditure/saving decisions and prices of goods change much more slowly than financial investment decisions and asset prices, most of the adjustment usually occurs within the financial markets. That is, *asset prices and exchange rates adjust so that the potential flow of financial capital is mitigated and actual capital flows remain consistent with trade flows.* In a fixed exchange rate regime, the central bank offsets the private capital flows in the process of maintaining the exchange rate peg and the adjustment occurs in other asset prices, typically interest rates, until and unless the central bank is forced to allow the exchange rate to adjust.[21] In a floating exchange rate regime, the main adjustment is often a rapid change in the exchange rate that dampens an investor's conviction that further movement will be forthcoming. Thus, *capital flows—potential and actual—are the primary determinant of exchange rate movements in the short-to-intermediate term*. Trade flows become increasingly important in the longer term as expenditure/saving decisions and the prices of goods and services adjust.

---

[19] In official balance of payments accounts, investment/financing flows are separated into two categories: the capital account and the financial account. Because the technical distinction is immaterial for present purposes, we will simply refer to the balance of investment/financing flows as the capital account. Similarly, we ignore the technical distinction between the trade balance and the *current account* balance. The details of balance of payments accounting are presented in the Level I curriculum reading on International Trade and Capital Flows.

[20] This relationship is developed in the Level I curriculum reading on Aggregate Output, Prices, and Economic Growth.

With the correspondence between the trade balance and capital flows firmly established, we can now examine the impact of exchange rate changes on the trade balance from two perspectives. The first approach focuses on the effect of changing the relative price of domestic and foreign goods. This approach, which is called the *elasticities approach*, highlights changes in the composition of spending. The second approach, called the *absorption approach*, focuses on the impact of exchange rates on aggregate expenditure/saving decisions.

10 EXCHANGE RATES AND THE TRADE BALANCE: THE ELASTICITIES APPROACH

j explain the effects of exchange rates on countries' international trade and capital flows.

The effectiveness of devaluation (in a fixed system) or depreciation (in a flexible system) of the currency for reducing a trade deficit depends on well-behaved demand and supply curves for goods and services. The condition that guarantees that devaluations improve the trade balance is called the Marshall–Lerner condition. The usual statement of this condition assumes that trade is initially balanced. We will present a generalization of the condition that allows for an initial trade imbalance and hence is more useful in addressing whether exchange rate movements will correct such imbalances.

Recall from microeconomics that the price elasticity of demand is given by:[22]

$$\varepsilon = -\frac{\%\ change\ in\ quantity}{\%\ change\ in\ price} = -\frac{\%\Delta Q}{\%\Delta P}$$

For example, a demand elasticity of 0.6 means that quantity demanded increases by 6 percent if price declines by 10 percent. Note that the elasticity of demand is defined so that it is a positive number. Because expenditure (R) equals price multiplied by quantity ($P \times Q$), by re-arranging the above expression to solve and substitute for $\%\Delta Q$, we can see that:

$$\%\ change\ in\ expenditure = \%\Delta R = \%\Delta P + \%\Delta Q = (1 - \varepsilon)\%\Delta P$$

From this we can see that an increase in price decreases expenditure if $\varepsilon > 1$, but it increases expenditure if $\varepsilon < 1$. By convention, if $\varepsilon > 1$ demand is described as being "elastic," while if $\varepsilon < 1$ demand is described as "inelastic."

---

[21] A classic example of this occurred in September 1992, when the United Kingdom was forced to withdraw from the European Exchange Rate Mechanism, the forerunner of the current European Economic and Monetary Union (EMU).

[22] See the Level I curriculum reading Topics in Demand and Supply Analysis.

The basic idea behind the Marshall–Lerner condition is that demand for imports and exports must be sufficiently price-sensitive so that increasing the relative price of imports increases the difference between export receipts and import expenditures. The generalized Marshall–Lerner condition is:

$$\omega_X \varepsilon_X + \omega_M(\varepsilon_M - 1) > 0$$

where ω_X and ω_M are the shares of exports and imports, respectively, in total trade (i.e., imports + exports) and ε_X and ε_M are the price elasticities of foreign demand for domestic country exports and domestic country demand for imports, respectively. Note that $(\omega_X + \omega_M) = 1$ and that an initial trade deficit implies $\omega_M > \omega_X$. If this condition is satisfied, a devaluation/depreciation of the domestic currency will move the trade balance toward surplus.

The first term in the generalized Marshall–Lerner condition reflects the change in export receipts assuming the domestic currency price of exports is unchanged (i.e., foreigners are billed in the domestic currency). It will be positive as long as export demand is not totally insensitive to price. Depreciation of the domestic currency makes exports cheaper in foreign currency and induces an increase in the quantity demanded by foreigners. This is reflected by the elasticity ε_x. There is no direct price impact on domestic currency export revenue because the domestic currency price is assumed to be unchanged. Hence, the percentage change in export revenue corresponding to a 1 percent depreciation of the currency is simply $\varepsilon_{x.}$ The second term in the generalized Marshall–Lerner condition reflects the impact on import expenditures. Assuming that imports are billed in a foreign currency, the domestic currency price of imports rises as the domestic currency depreciates. The direct price effect increases import expenditures, while the induced reduction in the quantity of imports decreases import expenditures. The net effect depends on the elasticity of import demand, ε_M. Import expenditure declines only if import demand is elastic (i.e., $\varepsilon_M > 1$).

Examination of the generalized Marshall–Lerner condition indicates that more elastic demand—for either imports or exports—makes it more likely that the trade balance will improve. Indeed, if the demand for imports is elastic, $\varepsilon_M > 1$, then the trade balance will definitely improve. It should also be clear that the elasticity of import demand becomes increasingly important, and the export elasticity less important, as the initial trade deficit gets larger—that is, as ω_M increases. In the special case of initially balanced trade, $\omega_X = \omega_M$, the condition reduces to $(\varepsilon_X + \varepsilon_M > 1)$, which is the classic Marshall–Lerner condition.

Exhibit 9 illustrates the impact of depreciation on the trade balance. For ease of reference, we assume the domestic currency is the euro. A 10 percent depreciation of the euro makes imports 10 percent more expensive in euro terms. With an import elasticity of 0.65, this induces a 6.5 percent reduction in the quantity of imports. But import expenditures increase by 3.5 percent [10% × (1 – 0.65)] or €21,000,000 because the drop in quantity is not sufficient to offset the increase in price. On the export side, the euro price of exports does not change but the foreign currency price of exports declines by 10 percent. This induces a 7.5 percent increase in the quantity of exports given an elasticity of 0.75. The euro value of exports therefore increases by 7.5 percent or €30,000,000. The net effect is a €9,000,000 improvement in the trade balance and a €51,000,000 increase in total trade.

Exhibit 9 Marshall–Lerner Condition with a 10 Percent Depreciation of Domestic Currency (€)

Assumptions	Exports	Imports
Demand elasticity	0.75	0.65
Percent price change		

(continued)

Exhibit 9 (Continued)

Assumptions	Exports	Imports
In domestic currency (€)	0	10%
In foreign currency	–10%	0

Results	Initial value(€)	Change(€)
Exports	400,000,000	30,000,000
Imports	600,000,000	21,000,000
Trade balance	–200,000,000	9,000,000
Total trade	1,000,000,000	51,000,000

The balance of trade improves after the depreciation of the euro because the Marshall–Lerner condition is satisfied: The increase in the euro-value of exports exceeds the increase in the value of imports. Based on the data in Exhibit 9, $\omega_M = 0.6$ (i.e., 600,000,000/1,000,000,000) and $\omega_X = 0.4$ (i.e., 1 – 0.6). Thus, the Marshall–Lerner equation is greater than zero:

$$\omega_X\varepsilon_X + \omega_M(\varepsilon_M - 1) = 0.4 \times 0.75 + 0.6(0.65 - 1) = 0.09$$

The elasticity of demand for any good or service depends on at least four factors: 1) the existence or absence of close substitutes, 2) the structure of the market for that product (e.g., a monopoly or perfect competition), 3) its share in people's budgets, and 4) the nature of the product and its role in the economy. Demand for a product with close substitutes is highly price-sensitive, whereas demand for a unique product tends to be much less elastic. The demand curve faced by any producer also depends on the nature and level of competition among producers of that product. If there are many sellers of identical products, then each producer faces highly elastic demand for its output even if global demand for that product is insensitive to price. Producers who are able to differentiate their product, perhaps through branding, face somewhat less elastic demand. In markets with only a few sellers, each producer faces demand that is highly dependent upon strategic maneuvers by its competitors. If competitors match price decreases but not increases, then the producer loses market share by raising his price but fails to gain market share by reducing his price.

Price changes have two effects on demand. The *substitution effect* refers to changes in the composition of spending across different products. As a product gets more expensive (cheaper) relative to other products, customers demand less (more) of it. This is what people usually think of first when they consider the effect of a price change. The *income effect* refers to the fact that price changes affect real purchasing power. When the price of a good rises (falls), people's purchasing power is reduced (increased). The strength of this effect depends on the product's share in people's budgets—the more important the product, the stronger the income effect. The income effect also depends on the nature of the product. The demand for luxuries is highly sensitive to income, whereas the demand for necessities is fairly insensitive to income.

To illustrate the differential impact of the two drivers of the income effect—share of expenditure and nature of the product—consider the demand for food. Clearly, food is a necessity. Based on this fact, we would expect demand to be inelastic. However, the share of expenditures that go to food varies across countries. In poor countries, food represents a much larger share of expenditure than in rich countries. Hence, all else being equal, we would expect the demand for food to be more price elastic in poorer countries. Of course, even in rich countries, the composition of spending on food may change considerably even if overall demand for food does not.

A significant portion of international trade occurs in intermediate products—products that are used as inputs into the production of other goods. Demand for these products derives from supply and demand decisions for the final products. However, the same basic considerations apply for intermediate products as for final products. Are there close substitutes for it in the production process? If not, its demand will tend to be less elastic than would be the case if there were readily available substitutes. How important is it to the overall economy? All else equal, the larger its share in overall production costs for the economy, the bigger its impact on production decisions and therefore the more price-elastic its derived demand. Oil is a classic example of a widely used input with few readily adoptable substitutes, at least in the short run. Lack of substitutes tends to make oil demand price-inelastic. However, it is so important in modern industrial economies that changes in its price can induce expansion or contraction of aggregate output. This makes short-run oil demand somewhat more elastic—at least for significant price changes. In the longer run, the feasibility of substitution among energy sources enhances the price sensitivity of oil demand.

Exhibit 10 shows estimates of demand elasticity for various products. The estimates range from essentially zero for pediatric doctor visits—a necessity for which there is virtually no substitute—to 3.8 for Coca-Cola, a specific brand for which there are many substitutes. Note that the elasticity of demand for soft drinks in general is much lower than for Coca-Cola, roughly 0.9. The elasticity of demand for rice in Japan versus in Bangladesh clearly illustrates the impact of expenditure share on price sensitivity. Similarly, although air travel for pleasure (a luxury) is quite price elastic, demand for first-class air travel is fairly insensitive to price. This is most likely because many first-class passengers are either traveling on business (presumably deemed to have high value added) or wealthy enough that the cost of first-class airfare is inconsequential.

Exhibit 10 Estimates of Demand Elasticities

Product Description	Elasticity	Rationale/Comment
Travel and transport		
Airline travel (US)		
For pleasure	1.5	Luxury
1st class	0.3	Business and wealthy travelers
Car fuel (US, long term)	0.3	
Bus travel (US)	0.2	
Ford compact car	2.8	Large purchase; specific brand
Food and beverages		
Rice		Necessity; staple food
Bangladesh	0.8	Poor country
Japan	0.3	Wealthy country
Soft drinks		
All	0.8–1.0	
Coca-Cola	3.8	Specific brand; competitive market
Medical care (US)		
Health insurance	0.3	
Pediatric doctor visit	0.0–0.1	No good substitute
Materials and energy		Necessary inputs

(continued)

Exhibit 10 (Continued)

Product Description	Elasticity	Rationale/Comment
Steel	0.2–0.3	
Oil	0.4	

Sources: Various studies cited in Wikipedia, "Price Elasticity of Demand," accessed August 2018 (http://en.wikipedia.org/wiki/Price_elasticity_of_demand).

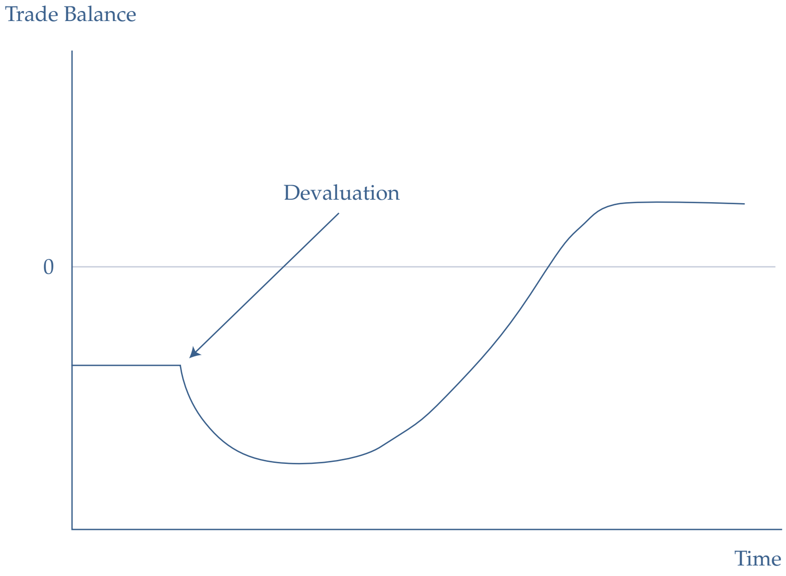
In practice, most countries import and export a variety of products. Hence, the overall price elasticities of their imports and exports reflect a composite of the products they trade. In conjunction with the Marshall–Lerner condition, our review of the factors that determine price elasticities suggests that exchange rate changes will be a more-effective mechanism for trade balance adjustment if a country imports and exports the following:

- Goods for which there are good substitutes
- Goods that trade in competitive markets
- Luxury goods, rather than necessities
- Goods that represent a large portion of consumer expenditures or a large portion of input costs for final producers

Note that each of these conditions is associated with higher demand elasticities (ε_X and ε_M).

Even when the Marshall–Lerner condition is satisfied, it is still possible that devaluation (in a fixed parity regime) or depreciation (in a floating regime) of the currency will initially make the trade balance worse before making it better. This effect, called the *J*-curve effect, is illustrated in Exhibit 11.

Exhibit 11 Trade Balance Dynamics: The *J*-Curve



In the very short run, the *J*-curve reflects the order delivery lags that take place in import and export transactions. Imagine a clothing importer in Washington. Orders are placed in January for French spring fashions. Market forces cause the dollar to depreciate in February, but contracts were already signed for payment in euros. When the fashions arrive in March, more dollars have to go out to pay for the order signed in euros. Thus, the trade balance gets worse. However, after the depreciation, the clothing importer has to put in new orders for summer fashions. As a result of the currency depreciation, the French summer fashions are now more expensive, so the clothing store cuts the demand for imported clothes from France. The depreciation eventually improves the trade balance, even though it initially made it worse.

A *J*-curve pattern may also arise if short-term price elasticities do not satisfy the Marshall–Lerner condition but long-term elasticities do satisfy it. As noted above in the case of oil, significant changes in spending patterns often take time. Thus, the trade balance may worsen initially and then gradually improve following a depreciation of the currency as firms and consumers adapt.

EXCHANGE RATES AND THE TRADE BALANCE: THE ABSORPTION APPROACH

11

j explain the effects of exchange rates on countries' international trade and capital flows.

The elasticities approach focuses on the expenditure-switching effect of changing the relative prices of imports and exports. It is essentially a microeconomic view of the relationship between exchange rates and the trade balance. The absorption approach adopts an explicitly macroeconomic view of this relationship.

Recall from above that the trade balance is equal to the country's saving, including the government fiscal balance, minus its investment in new plants and equipment. Equivalently, it is equal to the difference between income (GDP) and domestic expenditure, or absorption. Thus, in order to move the trade balance toward surplus, a devaluation/depreciation of the domestic currency must increase income relative to expenditure or, equivalently, increase national saving relative to investment in physical capital.

If there is excess capacity in the economy, then by switching demand toward domestically produced goods and services, depreciation of the currency can increase output/income. Because some of the additional income will be saved, income rises relative to expenditure and the trade balance improves. If the economy is at full employment, however, the trade balance cannot improve unless domestic expenditure declines. If expenditure does not decline, then the depreciation will put upward pressure on domestic prices until the stimulative effect of the exchange rate change is negated by the higher price level and the trade balance reverts to its original level.

How might depreciation of the currency reduce domestic expenditure relative to income? The main mechanism is a wealth effect. A weaker currency reduces the purchasing power of domestic-currency-denominated assets (including the present value of current and future earned income). Households respond by reducing expenditure and increasing saving in order to rebuild their wealth. Of course, as real wealth is rebuilt, the effect on saving is likely to be reversed—resulting in only a temporary improvement in the trade balance. Thus, in the absence of excess capacity in the economy, currency depreciation is likely to provide only a temporary solution for a chronic trade imbalance. Lasting correction of the imbalance requires more fundamental changes in expenditure/saving behavior (e.g., a policy shift that improves the fiscal balance or an increase in saving relative to capital investment induced by an increase in real interest rates).

The absorption approach also reminds us that currency depreciation cannot improve the trade balance unless it also induces a corresponding change in the capital account. Not only must domestic saving increase, but that saving must also be willingly channeled into buying financial assets from foreigners. All else equal, this implies that foreign and domestic asset prices must change such that foreign assets become relatively more attractive and domestic assets relatively less attractive to both foreign and domestic investors.

EXAMPLE 8

Exchange Rates and the Trade Balance

An analyst at a foreign exchange dealing bank is examining the exchange rate for the Australian dollar (AUD), which is a freely floating currency. Currently, Australia is running a trade surplus with the rest of the world, primarily reflecting strong demand for Australian resource exports generated by rapid growth in emerging market economies in the Western Pacific region. In turn, Australia imports food and energy from a variety of foreign countries that compete with each other as well as with Australian producers of these products. The analyst uses data in the following table to estimate the effect of changes in the AUD exchange rate on Australia's balance of trade.

	Volume (AUD billions)	Demand Elasticity
Exports	200	0.3
Imports	180	0.6

The analyst's research report on this topic notes that the mix of products that Australia imports and exports seems to be changing and that this will affect the relation between the exchange rates and the trade surplus. The proportion of Australian exports accounted for by fine wines is increasing. These are considered a luxury good and must compete with increased wine exports from comparable-producing regions (such as Chile and New Zealand). At the same time, rising income levels in Australia are allowing the country to increase the proportion of its imports accounted for by luxury goods, and these represent a rising proportion of consumer expenditures. The analyst's report states: "Given the changing export mix, an appreciation of the currency will be more likely to reduce Australia's trade surplus. In contrast, the changing import mix will have the opposite effect."

1 Given the data in the table, an appreciation in the AUD will:

A cause the trade balance to increase.

B cause the trade balance to decrease.

C have no effect on the trade balance.

2 All else equal, an appreciation in the AUD will be *more* likely to reduce the trade surplus if the demand:

A elasticities for imports and exports increase.

B elasticity for exports and the export share in total trade decrease.

C elasticity for imports decreases and the import share in total trade increases.

3 All else equal, an appreciation in the AUD will be *more* likely to reduce the trade surplus if it leads to an increase in Australian:

- **A** tax receipts.
- **B** private sector investment.
- **C** government budget surpluses.

4 The report's statement about the effect of changing import and export mixes is *most* likely:

- **A** correct.
- **B** incorrect with respect to the import effect.
- **C** incorrect with respect to the export effect.

5 Suppose the Australian government imposed capital controls that prohibited the flow of financial capital into or out of the country. What impact would this have on the Australian trade balance?

- **A** The trade surplus would increase.
- **B** The trade balance would go to zero.
- **C** The trade balance would not necessarily be affected.

6 Suppose the Australian government imposed capital controls that prohibited the flow of financial capital into or out of the country. The impact on the trade balance, if any, would most likely take the form of:

- **A** a decrease in private saving.
- **B** a decrease in private investment.
- **C** an increase in the government fiscal balance.

Solution to 1:

A is correct. As the AUD appreciates, the price of exports to *offshore buyers* goes up and they demand fewer of them; hence, the AUD-denominated revenue from exports decreases. (Although export demand is inelastic, or $\varepsilon_X < 1$, recall that the *Australian* price of these exports is assumed not to have changed, so the amount of export revenue received by Australia, in AUD-terms, unambiguously declines as the quantity of exports declines.) Australian expenditure for imports also declines. Although the price of imports declines as the AUD appreciates, the Australians do not increase their import purchases enough to lead to higher expenditures. This is because import demand is also inelastic ($\varepsilon_M < 1$). This effect on import expenditure can be seen from: $\%\Delta R_M = (1 - \varepsilon_M)\%\Delta P_M$, where $\%\Delta P_M$ is negative (import prices are declining) and import demand is inelastic (so $(1 - \varepsilon_M) > 0$). With both import expenditures and export revenues declining, the net effect on the trade balance comes down to the relative size of the import and export weights (ω_M and ω_X, respectively). In this case, $\omega_X = 0.53$ (i.e., 200/380) and $\omega_M = 0.47$ (i.e., 180/380). Putting this into the Marshall–Lerner equation leads to:

$$\omega_X\varepsilon_X + \omega_M(\varepsilon_M - 1) = 0.53 \times 0.3 + 0.47(0.6 - 1) = -0.03$$

Because the Marshall–Lerner condition is not satisfied, exchange rate movements do not move the trade balance in the expected direction [i.e., appreciation (depreciation) of the currency does not decrease (increase) the trade balance]. However, note that with different import/export weights and the same elasticities, the Marshall–Lerner condition would be met. In particular, the condition would be met for any value of ω_X greater than 4/7 (≈ 0.571).

Solution to 2:

A is correct. The basic intuition of the Marshall–Lerner condition is that in order for an exchange rate movement to rebalance trade, the demands for imports and exports must be sufficiently price-sensitive (i.e., they must have sufficiently high elasticities). However, the relative share of imports and exports in total trade must also be considered. The generalized Marshall–Lerner condition requires:

$$\omega_X \varepsilon_X + \omega_M(\varepsilon_M - 1) > 0$$

An increase in both ε_X and ε_M will clearly make this expression increase (A is correct). In contrast, a decrease in both ω_X and ε_X tends to make the expression smaller (B is incorrect).[23] If ε_M decreases and ω_M increases, import demand will respond less to an exchange rate movement and will have a larger role in determining the trade balance (C is incorrect).

Solution to 3:

B is correct. An Australian trade surplus means that Australia is spending less than it earns and is accumulating claims on foreigners. Equivalently, Australian saving, inclusive of both private saving and the government fiscal balance, is more than sufficient to fund Australian private sector investment. The relationship between the trade balance and expenditure/saving decisions is given by:

$$X - M = (S - I) + (T - G) > 0$$

For Australia's trade balance to decline, it must save less (S down), invest more (I up), decrease its fiscal balance ($T - G$ down), or some combination of these. Increasing tax receipts (T up) increases rather than decreases the fiscal balance, so answer A is incorrect. Similarly, answer C, increasing the government budget surplus, is incorrect. Increasing private investment (I up) does decrease the trade balance, so answer B is correct.

Solution to 4:

B is correct. As Australian exports become more dominated by luxury goods that face highly competitive market conditions, the elasticity of export demand (ε_X) is likely to be increasing. Increasing export elasticity makes the trade surplus more responsive to an AUD appreciation (the increase in ε_X will tend to increase the computed value for the Marshall–Lerner equation). Similarly, as Australian imports become more dominated by luxury goods that are an increasing proportion of household expenditure, import elasticity (ε_M) will most likely increase. This will also tend to increase the computed value for the Marshall–Lerner equation.

Solution to 5:

B is correct. A trade deficit (surplus) must be exactly matched by an offsetting capital account surplus (deficit). Anything that impacts the trade balance must impact the capital account, and vice versa. If capital flows are prohibited, then both the capital account and the trade balance must be zero.

Solution to 6:

A is correct. The trade balance must go to zero. An increase in the fiscal balance implies an increase in the existing trade surplus, so answer C is incorrect. A decrease in private investment will also cause an increase in the trade surplus, so answer B is incorrect. A decrease in private saving will decrease the trade surplus as required, so answer A is correct: A decrease in saving will most likely reflect a decline in national income, especially the profit component, as export demand is choked off by the inability to extend credit to foreigners.

---

[23] Because $\omega_M = 1 - \omega_X$ and $\varepsilon_M < 1$ in this example, a decrease in ω_X also decreases the second terms, $\omega_M(\varepsilon_M - 1)$, in the Marshall–Lerner condition.

SUMMARY

Foreign exchange markets are crucial for understanding both the functioning of the global economy as well as the performance of investment portfolios. In this reading, we have described the diverse array of FX market participants and have introduced some of the basic concepts necessary to understand the structure and functions of these markets. The reader should be able to understand how exchange rates—both spot and forward—are quoted and be able to calculate cross exchange rates and forward rates. We also have described the array of exchange rate regimes that characterize foreign exchange markets globally and how these regimes determine the flexibility of exchange rates, and hence, the degree of foreign exchange rate risk that international investments are exposed to. Finally, we have discussed how movements in exchange rates affect international trade flows (imports and exports) and capital flows.

The following points, among others, are made in this reading:

- Measured by average daily turnover, the foreign exchange market is by far the largest financial market in the world. It has important effects, either directly or indirectly, on the pricing and flows in all other financial markets.
- There is a wide diversity of global FX market participants that have a wide variety of motives for entering into foreign exchange transactions.
- Individual currencies are usually referred to by standardized three-character codes. These currency codes can also be used to define exchange rates (the price of one currency in terms of another). There are a variety of exchange rate quoting conventions commonly used.
- A direct currency quote takes the domestic currency as the price currency and the foreign currency as the base currency (i.e., $S_{d/f}$). An indirect quote uses the domestic currency as the base currency (i.e., $S_{f/d}$). To convert between direct and indirect quotes, the inverse (reciprocal) is used. Professional FX markets use standardized conventions for how the exchange rate for specific currency pairs will be quoted.
- Currencies trade in foreign exchange markets based on nominal exchange rates. An increase (decrease) in the exchange rate, quoted in indirect terms, means that the domestic currency is appreciating (depreciating) versus the foreign currency.
- The real exchange rate, defined as the nominal exchange rate multiplied by the ratio of price levels, measures the relative purchasing power of the currencies. An increase in the real exchange rate ($R_{d/f}$) implies a reduction in the relative purchasing power of the domestic currency.
- Given exchange rates for two currency pairs—A/B and A/C—we can compute the cross-rate (B/C) between currencies B and C. Depending on how the rates are quoted, this may require inversion of one of the quoted rates.

- Spot exchange rates are for immediate settlement (typically, T + 2), while forward exchange rates are for settlement at agreed-upon future dates. Forward rates can be used to manage foreign exchange risk exposures or can be combined with spot transactions to create FX swaps.
- The spot exchange rate, the forward exchange rate, and the domestic and foreign interest rates must jointly satisfy an arbitrage relationship that equates the investment return on two alternative but equivalent investments. Given the spot exchange rate and the foreign and domestic interest rates, the forward exchange rate must take the value that prevents riskless arbitrage.
- Forward rates are typically quoted in terms of forward (or swap) points. The swap points are added to the spot exchange rate in order to calculate the forward rate. Occasionally, forward rates are presented in terms of percentages relative to the spot rate.
- The base currency is said to be trading at a forward premium if the forward rate is above the spot rate (forward points are positive). Conversely, the base currency is said to be trading at a forward discount if the forward rate is below the spot rate (forward points are negative).
- The currency with the higher (lower) interest rate will trade at a forward discount (premium).
- Swap points are proportional to the spot exchange rate and to the interest rate differential and approximately proportional to the term of the forward contract.
- Empirical studies suggest that forward exchange rates may be unbiased predictors of future spot rates, but the margin of error on such forecasts is too large for them to be used in practice as a guide to managing exchange rate exposures. FX markets are too complex and too intertwined with other global financial markets to be adequately characterized by a single variable, such as the interest rate differential.
- Virtually every exchange rate is managed to some degree by central banks. The policy framework that each central bank adopts is called an exchange rate regime. These regimes range from using another country's currency (dollarization), to letting the market determine the exchange rate (independent float). In practice, most regimes fall in between these extremes. The type of exchange rate regime used varies widely among countries and over time.
- An ideal currency regime would have three properties: (1) the exchange rate between any two currencies would be credibly fixed; (2) all currencies would be fully convertible; and (3) each country would be able to undertake fully independent monetary policy in pursuit of domestic objectives, such as growth and inflation targets. However, these conditions are inconsistent. In particular, a fixed exchange rate and unfettered capital flows severely limit a country's ability to undertake independent monetary policy. Hence, there cannot be an ideal currency regime.
- The IMF identifies the following types of regimes: arrangements with no separate legal tender (dollarization, monetary union), currency board, fixed parity, target zone, crawling peg, crawling band, managed float, and independent float. Most major currencies traded in FX markets are freely floating, albeit subject to occasional central bank intervention.
- A trade surplus (deficit) must be matched by a corresponding deficit (surplus) in the capital account. Any factor that affects the trade balance must have an equal and opposite impact on the capital account, and vice versa.

- A trade surplus reflects an excess of domestic saving (including the government fiscal balance) over investment spending. A trade deficit indicates that the country invests more than it saves and must finance the excess by borrowing from foreigners or selling assets to foreigners.
- The impact of the exchange rate on trade and capital flows can be analyzed from two perspectives. The elasticities approach focuses on the effect of changing the relative price of domestic and foreign goods. This approach highlights changes in the composition of spending. The absorption approach focuses on the impact of exchange rates on aggregate expenditure/saving decisions.
- The elasticities approach leads to the Marshall–Lerner condition, which describes combinations of export and import demand elasticities such that depreciation (appreciation) of the domestic currency will move the trade balance toward surplus (deficit).
- The idea underlying the Marshall–Lerner condition is that demand for imports and exports must be sufficiently price-sensitive so that an increase in the relative price of imports increases the difference between export receipts and import expenditures.
- In order to move the trade balance toward surplus (deficit), a change in the exchange rate must decrease (increase) domestic expenditure (also called absorption) relative to income. Equivalently, it must increase (decrease) domestic saving relative to domestic investment.
- If there is excess capacity in the economy, then currency depreciation can increase output/income by switching demand toward domestically produced goods and services. Because some of the additional income will be saved, income rises relative to expenditure and the trade balance improves.
- If the economy is at full employment, then currency depreciation must reduce domestic expenditure in order to improve the trade balance. The main mechanism is a wealth effect: A weaker currency reduces the purchasing power of domestic-currency-denominated assets (including the present value of current and future earned income), and households respond by reducing expenditure and increasing saving.

REFERENCES

Bank for International Settlements (BIS). 2010. "Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity in 2010" (www.bis.org).

Friedman, Milton. 1953. "The Monetarist Theory of Flexible Exchange Rate Systems." In *Essays in Positive Economics*. Chicago: University of Chicago Press.

Hong Kong Monetary Authority (HKMA). 2005. *HKMA Background Brief No. 1: Hong Kong's Linked Exchange Rate System*. 2nd ed. (November): www.info.gov.hk.

International Monetary Fund (IMF). 2006. "De Facto Classification of Exchange Rate Regimes and Monetary Policy Framework" (www.imf.org).

International Monetary Fund (IMF). 2010. "Currency Composition of Official Foreign Exchange Reserves (COFER)" report (www.imf.org).

PRACTICE PROBLEMS

1 An exchange rate:

A is most commonly quoted in real terms.

B is the price of one currency in terms of another.

C between two currencies ensures they are fully convertible.

2 A decrease in the real exchange rate (quoted in terms of domestic currency per unit of foreign currency) is *most likely* to be associated with an increase in which of the following?

A Foreign price level.

B Domestic price level.

C Nominal exchange rate.

3 In order to minimize the foreign exchange exposure on a euro-denominated receivable due from a German company in 100 days, a British company would *most likely* initiate a:

A spot transaction.

B forward contract.

C real exchange rate contract.

4 Which of the following counterparties is *most likely* to be considered a sell-side foreign-exchange market participant?

A A large corporation that borrows in foreign currencies.

B A sovereign wealth fund that influences cross-border capital flows.

C A multinational bank that trades foreign exchange with its diverse client base.

5 What will be the effect on a direct exchange rate quote if the domestic currency appreciates?

A Increase

B Decrease

C No change

6 An executive from Switzerland checked into a hotel room in Spain and was told by the hotel manager that 1 EUR will buy 1.2983 CHF. From the executive's perspective, an indirect exchange rate quote would be:

A 0.7702 EUR per CHF.

B 0.7702 CHF per EUR.

C 1.2983 EUR per CHF.

7 Over the past month, the Swiss Franc (CHF) has depreciated 12 percent against pound sterling (GBP). How much has the pound sterling appreciated against the Swiss Franc?

A 12%

B Less than 12%

C More than 12%

8 An exchange rate between two currencies has increased to 1.4500. If the base currency has appreciated by 8% against the price currency, the initial exchange rate between the two currencies was *closest* to:

© 2011 CFA Institute. All rights reserved.

A 1.3340.
B 1.3426.
C 1.5660.

The following information relates to Questions 9–10

A dealer provides the following quotes:

Ratio	Spot rate
CNY/HKD	0.8422
CNY/ZAR	0.9149
CNY/SEK	1.0218

9 The spot ZAR/HKD cross-rate is *closest* to:
A 0.9205.
B 1.0864.
C 1.2978.

10 Another dealer is quoting the ZAR/SEK cross-rate at 1.1210. The arbitrage profit that can be earned is *closest* to:
A ZAR 3671 per million SEK traded.
B SEK 4200 per million ZAR traded.
C ZAR 4200 per million SEK traded.

---

11 A BRL/MXN spot rate is listed by a dealer at 0.1378. The 6-month forward rate is 0.14193. The 6-month forward points are *closest* to:
A –41.3.
B +41.3.
C +299.7.

12 A three-month forward exchange rate in CAD/USD is listed by a dealer at 1.0123. The dealer also quotes 3-month forward points as a percentage at 6.8%. The CAD/USD spot rate is *closest* to:
A 0.9478.
B 1.0550.
C 1.0862.

13 If the base currency in a forward exchange rate quote is trading at a forward discount, which of the following statements is *most* accurate?
A The forward points will be positive.
B The forward percentage will be negative.
C The base currency is expected to appreciate versus the price currency.

14 A forward premium indicates:
A an expected increase in demand for the base currency.
B the interest rate is higher in the base currency than in the price currency.
C the interest rate is higher in the price currency than in the base currency.

15 The JPY/AUD spot exchange rate is 82.42, the JPY interest rate is 0.15%, and the AUD interest rate is 4.95%. If the interest rates are quoted on the basis of a 360-day year, the 90-day forward points in JPY/AUD would be *closest* to:

A –377.0.

B –97.7.

C 98.9.

16 Which of the following is *not* a condition of an ideal currency regime?

A Fully convertible currencies.

B Fully independent monetary policy.

C Independently floating exchange rates.

17 In practice, both a fixed parity regime and a target zone regime allow the exchange rate to float within a band around the parity level. The *most likely* rationale for the band is that the band allows the monetary authority to:

A be less active in the currency market.

B earn a spread on its currency transactions.

C exercise more discretion in monetary policy.

18 A fixed exchange rate regime in which the monetary authority is legally required to hold foreign exchange reserves backing 100% of its domestic currency issuance is best described as:

A dollarization.

B a currency board.

C a monetary union.

19 A country with a trade deficit will *most likely*:

A have an offsetting capital account surplus.

B save enough to fund its investment spending.

C buy assets from foreigners to fund the imbalance.

20 A large industrialized country has recently devalued its currency in an attempt to correct a persistent trade deficit. Which of the following domestic industries is *most likely* to benefit from the devaluation?

A Luxury cars.

B Branded prescription drugs.

C Restaurants and live entertainment venues.

21 A country with a persistent trade surplus is being pressured to let its currency appreciate. Which of the following *best* describes the adjustment that must occur if currency appreciation is to be effective in reducing the trade surplus?

A Domestic investment must decline relative to saving.

B Foreigners must increase investment relative to saving.

C Global capital flows must shift toward the domestic market.

SOLUTIONS

1 B is correct. The exchange rate is the number of units of the price currency that 1 unit of the base currency will buy. Equivalently, it is the number of units of the price currency required to buy 1 unit of the base currency.

2 B is correct. The real exchange rate (quoted in terms of domestic currency per unit of foreign currency) is given by:

$$\text{Real exchange rate}_{(d/f)} = S_{d/f} \times (P_f/P_d)$$

An increase in the domestic price level (P_d) *decreases* the real exchange rate because it implies an *increase* in the relative purchasing power of the domestic currency.

3 B is correct. The receivable is due in 100 days. To reduce the risk of currency exposure, the British company would initiate a forward contract to sell euros/buy pounds at an exchange rate agreed to today. The agreed-upon rate is called the forward exchange rate.

4 C is correct. The sell side generally consists of large banks that sell foreign exchange and related instruments to buy-side clients. These banks act as market makers, quoting exchange rates at which they will buy (the bid price) or sell (the offer price) the base currency.

5 B is correct. In the case of a direct exchange rate, the domestic currency is the price currency (the numerator) and the foreign currency is the base currency (the denominator). If the domestic currency appreciates, then fewer units of the domestic currency are required to buy 1 unit of the foreign currency and the exchange rate (domestic per foreign) declines. For example, if sterling (GBP) appreciates against the euro (EUR), then euro–sterling (GBP/EUR) might decline from 0.8650 to 0.8590.

6 A is correct. An indirect quote takes the foreign country as the price currency and the domestic country as the base currency. To get CHF—which is the executive's domestic currency—as the base currency, the quote must be stated as EUR/CHF. Using the hotel manager's information, the indirect exchange rate is (1/1.2983) = 0.7702.

7 C is correct. The appreciation of sterling against the Swiss franc is simply the inverse of the 12% depreciation of the Swiss franc against Sterling: $[1/(1 - 0.12)] - 1 = (1/0.88) - 1 = 0.1364$, or 13.64%.

8 B is correct. The percentage appreciation of the base currency can be calculated by dividing the appreciated exchange rate by the initial exchange rate. In this case, the unknown is the initial exchange rate. The initial exchange is the value of X that satisfies the formula:

$$1.4500/X = 1.08$$

Solving for X leads to $1.45/1.08 = 1.3426$.

9 A is correct. To get to the ZAR/HKD cross-rate, it is necessary to take the inverse of the CNY/ZAR spot rate and then multiply by the CNY/HKD exchange rate:

$$\begin{aligned}\text{ZAR/HKD} &= (\text{CNY/ZAR})^{-1} \times (\text{CNY/HKD}) \\ &= (1 / 0.9149) \times 0.8422 = 0.9205\end{aligned}$$

10 C is correct. The ZAR/SEK cross-rate from the original dealer is (1.0218/0.9149) = 1.1168, which is lower than the quote from the second dealer. To earn an arbitrage profit, a currency trader would buy SEK (sell ZAR) from the original dealer and sell SEK (buy ZAR) to the second dealer. On 1 million SEK the profit would be

SEK 1,000,000 × (1.1210 – 1.1168) = ZAR 4200

11 B is correct. The number of forward points equals the forward rate minus the spot rate, or 0.14193 – 0.1378 = 0.00413, multiplied by 10,000: 10,000 × 0.00413= 41.3 points. By convention, forward points are scaled so that ±1 forward point corresponds to a change of ±1 in the last decimal place of the spot exchange rate.

12 A is correct. Given the forward rate and forward points as a percentage, the unknown in the calculation is the spot rate. The calculation is as follows:

Spot rate × (1 + Forward points as a percentage) = Forward rate

Spot rate × (1 + 0.068) = 1.0123

Spot = 1.0123/1.068 =0.9478

13 B is correct. The base currency trading at a forward discount means that 1 unit of the base currency costs less for forward delivery than for spot delivery; i.e., the forward exchange rate is less than the spot exchange rate. The forward points, expressed either as an absolute number of points or as a percentage, are negative.

14 C is correct. To eliminate arbitrage opportunities, the spot exchange rate (S), the forward exchange rate (F), the interest rate in the base currency (i_b), and the interest rate in the price currency (i_p) must satisfy:

$$\frac{F}{S} = \left(\frac{1 + i_p}{1 + i_b}\right)$$

According to this formula, the base currency will trade at forward premium ($F > S$) if, and only if, the interest rate in the price currency is higher than the interest rate in the base currency ($i_p > i_b$).

15 B is correct. The forward exchange rate is given by

$$F_{JPY/AUD} = S_{JPY/AUD}\left(\frac{1 + i_{JPY}\tau}{1 + i_{AUD}\tau}\right) = 82.42\left(\frac{1 + .0015\left(\frac{90}{360}\right)}{1 + .0495\left(\frac{90}{360}\right)}\right)$$

$$= 82.42 \times .98815 = 81.443$$

The forward points are 100 × (F – S) = 100 × (81.443 – 82.42) = 100 × (–0.977) = –97.7. Note that because the spot exchange rate is quoted with two decimal places, the forward points are scaled by 100.

16 C is correct. An ideal currency regime would have credibly fixed exchange rates among all currencies. This would eliminate currency-related uncertainty with respect to the prices of goods and services as well as real and financial assets.

17 C is correct. Fixed exchange rates impose severe limitations on the exercise of independent monetary policy. With a rigidly fixed exchange rate, domestic interest rates, monetary aggregates (e.g., money supply), and credit conditions are dictated by the requirement to buy/sell the currency at the rigid parity. Even a narrow band around the parity level allows the monetary authority to exercise some discretionary control over these conditions. In general, the wider the band, the more independent control the monetary authority can exercise.

18 B is correct. With a currency board, the monetary authority is legally required to exchange domestic currency for a specified foreign currency at a fixed exchange rate. It cannot issue domestic currency without receiving foreign currency in exchange, and it must hold that foreign currency as a 100% reserve against the domestic currency issued. Thus, the country's monetary base (bank reserves plus notes and coins in circulation) is fully backed by foreign exchange reserves.

19 A is correct. A trade deficit must be exactly matched by an offsetting capital account surplus to fund the deficit. A capital account surplus reflects borrowing from foreigners (an increase in domestic liabilities) and/or selling assets to foreigners (a decrease in domestic assets). A capital account surplus is often referred to as a "capital inflow" because the net effect is foreign investment in the domestic economy.

20 A is correct. A devaluation of the domestic currency means domestic producers are cutting the price faced by their foreign customers. The impact on their unit sales and their revenue depends on the elasticity of demand. Expensive luxury goods exhibit high price elasticity. Hence, luxury car producers are likely to experience a sharp increase in sales and revenue due to the devaluation.

21 C is correct. The trade surplus cannot decline unless the capital account deficit also declines. Regardless of the mix of assets bought and sold, foreigners must buy more assets from (or sell fewer assets to) domestic issuers/investors.

Financial Statement Analysis

STUDY SESSIONS

Study Session 5	Financial Statement Analysis (1)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to demonstrate a thorough knowledge of financial reporting procedures and the standards that govern financial reporting disclosure. Emphasis is on basic financial statements and how alternative accounting methods affect those statements and the analysis of them.

Financial statement analysis is critical in assessing a company's overall financial position and associated risks over time. Security and business valuation, credit risk assessment, and acquisition due diligence all require an understanding of the major financial statements including general principles and reporting approaches. Because no set of accounting standards has universal acceptance, companies around the world may differ in reporting treatment based on their jurisdiction.

Financial statement analysis requires the ability to analyze a company's reported results with its economic reality, normalize differences in accounting treatment to make valid cross company comparisons, identify quality issues that may exist in reported financial statements, and discern evidence of financial statement manipulation by management.

Candidates should be familiar with the material covered in the following prerequisite reading available in Candidate Resources on the CFA Institute website:

- **Financial Reporting Mechanics**

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2021 CFA Institute. All rights reserved.

FINANCIAL STATEMENT ANALYSIS

STUDY SESSION

5

Financial Statement Analysis (1)

This study session introduces the principal information sources used to evaluate a company's financial performance. Primary financial statements (income statement, balance sheet, cash flow statement, and statement of changes in equity) in addition to notes to these statements and management reporting are examined. A general framework for conducting financial statement analysis is provided. The session also includes a description of the roles played by financial reporting standard-setting bodies and regulatory authorities.

READING ASSIGNMENTS

Reading 15	Introduction to Financial Statement Analysis by Elaine Henry, PhD, CFA, and Thomas R. Robinson, PhD, CFA, CAIA
Reading 16	Financial Reporting Standards by Elaine Henry, PhD, CFA, J. Hennie van Greuning, DCom, CFA, and Thomas R. Robinson, PhD, CFA, CAIA

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2021 CFA Institute. All rights reserved.

READING

15

Introduction to Financial Statement Analysis

by Elaine Henry, PhD, CFA, and Thomas R. Robinson, PhD, CFA, CAIA

Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). Thomas R. Robinson, PhD, CFA, CAIA, Robinson Global Investment Management LLC, (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the roles of financial reporting and financial statement analysis;
☐	**b.** describe the roles of the statement of financial position, statement of comprehensive income, statement of changes in equity, and statement of cash flows in evaluating a company's performance and financial position;
☐	**c.** describe the importance of financial statement notes and supplementary information—including disclosures of accounting policies, methods, and estimates—and management's commentary;
☐	**d.** describe the objective of audits of financial statements, the types of audit reports, and the importance of effective internal controls;
☐	**e.** identify and describe information sources that analysts use in financial statement analysis besides annual financial statements and supplementary information;
☐	**f.** describe the steps in the financial statement analysis framework.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1

INTRODUCTION

Financial analysis is the process of examining a company's performance in the context of its industry and economic environment in order to arrive at a decision or recommendation. Often, the decisions and recommendations addressed by financial analysts pertain to providing capital to companies—specifically, whether to invest in the company's debt or equity securities and at what price. An investor in debt securities is concerned about the company's ability to pay interest and to repay the principal lent. An investor in equity securities is an owner with a residual interest in the company and is concerned about the company's ability to pay dividends and the likelihood that its share price will increase.

Overall, a central focus of financial analysis is evaluating the company's ability to earn a return on its capital that is at least equal to the cost of that capital, to profitably grow its operations, and to generate enough cash to meet obligations and pursue opportunities.

Fundamental financial analysis starts with the information found in a company's financial reports. These financial reports include audited financial statements, additional disclosures required by regulatory authorities, and any accompanying (unaudited) commentary by management. Basic financial statement analysis—as presented in this reading—provides a foundation that enables the analyst to better understand other information gathered from research beyond the financial reports.

This reading is organized as follows: Section 2 discusses the scope of financial statement analysis. Sections 3–8 describe the sources of information used in financial statement analysis, including the primary financial statements (statement of financial position or balance sheet, statement of comprehensive income, statement of changes in equity, and cash flow statement). Section 9 provides a framework for guiding the financial statement analysis process. A summary of the key points conclude the reading.

2

SCOPE OF FINANCIAL STATEMENT ANALYSIS

a describe the roles of financial reporting and financial statement analysis;

The role of financial statements issued by companies is to provide information about a company's performance, financial position, and changes in financial position that is useful to a wide range of users in making economic decisions. The role of financial statement analysis is to use financial reports prepared by companies, combined with other information, to evaluate the past, current, and potential performance and financial position of a company for the purpose of making investment, credit, and other economic decisions. Managers within a company perform financial analysis to make operating, investing, and financing decisions but do not necessarily rely on analysis of related financial statements. They have access to additional financial information that can be reported in whatever format is most useful to their decision.

In evaluating financial reports, analysts typically have a specific economic decision in mind. Examples of these decisions include the following:

- Evaluating an equity investment for inclusion in a portfolio.
- Evaluating a merger or acquisition candidate.
- Evaluating a subsidiary or operating division of a parent company.
- Deciding whether to make a venture capital or other private equity investment.

- Determining the creditworthiness of a company in order to decide whether to extend a loan to the company and if so, what terms to offer.
- Extending credit to a customer.
- Examining compliance with debt covenants or other contractual arrangements.
- Assigning a debt rating to a company or bond issue.
- Valuing a security for making an investment recommendation to others.
- Forecasting future net income and cash flow.

These decisions demonstrate certain themes in financial analysis. In general, analysts seek to examine the past and current performance and financial position of a company in order to form expectations about its future performance and financial position. Analysts are also concerned about factors that affect risks to a company's future performance and financial position. An examination of performance can include an assessment of a company's profitability (the ability to earn a profit from delivering goods and services) and its ability to generate positive cash flows (cash receipts in excess of cash disbursements). Profit and cash flow are not equivalent. Profit (or loss) represents the difference between the prices at which goods or services are provided to customers and the expenses incurred to provide those goods and services.

In addition, profit (or loss) includes other income (such as investing income or income from the sale of items other than goods and services) minus the expenses incurred to earn that income. Overall, profit (or loss) equals income minus expenses, and its recognition is mostly independent from when cash is received or paid. Example 1 illustrates the distinction between profit and cash flow.

EXAMPLE 1

Profit versus Cash Flow

Sennett Designs (SD) sells furniture on a retail basis. SD began operations during December 2017 and sold furniture for €250,000 in cash. The furniture sold by SD was purchased on credit for €150,000 and delivered by the supplier during December. The credit terms granted by the supplier required SD to pay the €150,000 in January for the furniture it received during December. In addition to the purchase and sale of furniture, in December, SD paid €20,000 in cash for rent and salaries.

1 How much is SD's profit for December 2017 if no other transactions occurred?

2 How much is SD's cash flow for December 2017?

3 If SD purchases and sells exactly the same amount in January 2018 as it did in December and under the same terms (receiving cash for the sales and making purchases on credit that will be due in February), how much will the company's profit and cash flow be for the month of January?

Solution to 1:

SD's profit for December 2017 is the excess of the sales price (€250,000) over the cost of the goods that were sold (€150,000) and rent and salaries (€20,000), or €80,000.

Solution to 2:

The December 2017 cash flow is €230,000, the amount of cash received from the customer (€250,000) less the cash paid for rent and salaries (€20,000).

Solution to 3:

SD's profit for January 2018 will be identical to its profit in December: €80,000, calculated as the sales price (€250,000) minus the cost of the goods that were sold (€150,000) and minus rent and salaries (€20,000). SD's cash flow in January 2018 will also equal €80,000, calculated as the amount of cash received from the customer (€250,000) minus the cash paid for rent and salaries (€20,000) *and* minus the €150,000 that SD owes for the goods it had purchased on credit in the prior month.

Although profitability is important, so is a company's ability to generate positive cash flow. Cash flow is important because, ultimately, the company needs cash to pay employees, suppliers, and others in order to continue as a going concern. A company that generates positive cash flow from operations has more flexibility in funding needed for investments and taking advantage of attractive business opportunities than an otherwise comparable company without positive operating cash flow. Additionally, a company needs cash to pay returns (interest and dividends) to providers of debt and equity capital.

The expected magnitude of future cash flows is important in valuing corporate securities and in determining the company's ability to meet its obligations. The ability to meet short-term obligations is generally referred to as **liquidity**, and the ability to meet long-term obligations is generally referred to as **solvency**. Cash flow in any given period is not, however, a complete measure of performance for that period because, as shown in Example 1, a company may be obligated to make future cash payments as a result of a transaction that generates positive cash flow in the current period.

Profits may provide useful information about cash flows, past and future. If the transaction of Example 1 were repeated month after month, the long-term average monthly cash flow of SD would equal €80,000, its monthly profit. Analysts typically not only evaluate past profitability but also forecast future profitability.

Exhibit 1 shows how news coverage of corporate earnings announcements places corporate results in the context of analysts' expectations. Panel A shows the earnings announcement, and Panel B shows a sample of the news coverage of the announcement. Earnings are also frequently used by analysts in valuation. For example, an analyst may value shares of a company by comparing its price-to-earnings ratio (P/E) to the P/Es of peer companies and/or may use forecasted future earnings as direct or indirect inputs into discounted cash flow models of valuation.

Exhibit 1 An Earnings Release and News Media Comparison with Analysts' Expectations

Panel A: Excerpt from Apple Earnings Release

Apple Reports Second Quarter Results

Revenue Grows 16 Percent and EPS Grows 30 Percent to New March Quarter Records

New $100 Billion Share Repurchase Authorization Announced, Dividend Raised by 16 Percent

> Cupertino, California—May 1, 2018—Apple today announced financial results for its fiscal 2018 second quarter ended March 31, 2018. The Company posted quarterly revenue of $61.1 billion, an increase of 16 percent from the year-ago quarter, and quarterly earnings per diluted share of $2.73, up 30 percent. International sales accounted for 65 percent of the quarter's revenue.

Exhibit 1 (Continued)

"We're thrilled to report our best March quarter ever, with strong revenue growth in iPhone, Services and Wearables," said Tim Cook, Apple's CEO. "Customers chose iPhone X more than any other iPhone each week in the March quarter, just as they did following its launch in the December quarter. We also grew revenue in all of our geographic segments, with over 20% growth in Greater China and Japan."

"Our business performed extremely well during the March quarter, as we grew earnings per share by 30 percent and generated over $15 billion in operating cash flow," said Luca Maestri, Apple's CFO. "With the greater flexibility we now have from access to our global cash, we can more efficiently invest in our US operations and work toward a more optimal capital structure. Given our confidence in Apple's future, we are very happy to announce that our Board has approved a new $100 billion share repurchase authorization and a 16 percent increase in our quarterly dividend."

The Company will complete the execution of the previous $210 billion share repurchase authorization during the third fiscal quarter.

Reflecting the approved increase, the Board has declared a cash dividend of $0.73 per share of Apple's common stock payable on May 17, 2018 to shareholders of record as of the close of business on May 14, 2018.

The Company also expects to continue to net-share-settle vesting restricted stock units.

From the inception of its capital return program in August 2012 through March 2018, Apple has returned $275 billion to shareholders, including $200 billion in share repurchases. The management team and the Board will continue to review each element of the capital return program regularly and plan to provide an update on the program on an annual basis.

Apple is providing the following guidance for its fiscal 2018 third quarter:

- revenue between $51.5 billion and $53.5 billion
- gross margin between 38 percent and 38.5 percent
- operating expenses between $7.7 billion and $7.8 billion
- other income/(expense) of $400 million
- tax rate of approximately 14.5 percent

Apple will provide live streaming of its Q2 2018 financial results conference call beginning at 2:00 p.m. PDT on May 1, 2018 at www.apple.com/investor/earnings-call/. This webcast will also be available for replay for approximately two weeks thereafter.

Source: https://www.apple.com/newsroom/2018/05/apple-reports-second-quarter-results/ (retrieved 3 November 2018)

Panel B: Excerpt from News Article: Apple Second Quarter 2018 Earnings Release

Apple reported quarterly earnings and revenue on Tuesday that beat expectations, but sold fewer iPhones than expected.

(continued)

Exhibit 1 (Continued)

Shares rose as much 5 percent after hours, as investors digested the company's better-than-expected outlook for the current quarter, and a hefty capital return program.

The soft iPhone sales were still up from a year ago, and Apple CEO Tim Cook said in a statement that customers "chose iPhone X more than any other iPhone each week in the March quarter."

- Earnings per share: $2.73 vs. $2.67, adjusted, expected by a Thomson Reuters consensus estimate
- Revenue: $61.1 billion vs. $60.82 billion expected by Thomson Reuters consensus
- iPhone unit sales: 52.2 million vs. 52.54 million expected by a StreetAccount estimate
- Fiscal Q3 revenue guidance: $51.5 billion to $53.5 billion vs. $51.61 billion expected by Thomson Reuters consensus

Net income was $13.82 billion, up from $11.03 billion a year ago. A year ago, Apple earned $2.10 a share on revenue of $52.9 billion.

Source: https://cnbc.com/2018/05/01/apple-earnings-q2-2018.html (retrieved 3 November 2018)

Analysts are also interested in the financial position of a company. The financial position can be measured by comparing the resources controlled by the company (**assets**) in relation to the claims against those resources (**liabilities** and **equity**). The combination of liabilities and equity used to finance its assets represents the capital structure of the company. An example of a resource is cash. In Example 1, if no other transactions occur, the company should have €230,000 more in cash at 31 December 2017 than at the start of the period. The cash can be used by the company to pay its obligation to the supplier (a claim against the company) and may also be used to make distributions to the owner (who has a residual claim against the company's assets, net of liabilities).

Financial position and capital structure are particularly important in credit analysis, as depicted in Exhibit 2. Panel A of the exhibit is an excerpt from the company's annual earnings release highlighting the cumulative profitability, strong cash flow, strong balance sheet, and strong return on invested capital. Panel B of the exhibit is an excerpt from an August 2017 news article about an increase in the credit rating of Southwest Airlines due to a long history of profitability and a conservative capital structure.

Exhibit 2

Panel A: Excerpt from Earnings Announcement by Southwest Airlines

Southwest Airlines Reports Fourth Quarter and Record Annual Profit; 44th Consecutive Year of Profitability

DALLAS, Jan. 26, 2017 /PRNewswire/ -- Southwest Airlines Co. (NYSE:LUV) (the "Company") today reported its fourth quarter and annual 2016 results:

Exhibit 2 (Continued)

Fourth quarter net income of $522 million, or $.84 per diluted share, compared with fourth quarter 2015 net income of $536 million, or $.82 per diluted share.

Excluding special items, fourth quarter net income of $463 million, or $.75 per diluted share, compared with fourth quarter 2015 net income of $591 million, or $.90 per diluted share. This exceeded the First Call fourth quarter 2016 consensus estimate of $.70 per diluted share.

Record annual net income of $2.24 billion, or $3.55 per diluted share, compared with 2015 net income of $2.18 billion, or $3.27 per diluted share.

Excluding special items, record annual net income of $2.37 billion, or $3.75 per diluted share, compared with 2015 net income of $2.36 billion, or $3.52 per diluted share.

Annual operating income of $3.76 billion, resulting in an operating margin of 18.4 percent.

Excluding special items, annual operating income of $3.96 billion, resulting in an operating margin of 19.4 percent.

Record annual operating cash flow of $4.29 billion, and record annual free cash flow1 of $2.25 billion.

Returned $1.97 billion to Shareholders in 2016, through a combination of $222 million in dividends and $1.75 billion in share repurchases.

Annual return on invested capital (ROIC)1 of 30.0 percent.

Gary C. Kelly, Chairman of the Board and Chief Executive Officer, stated, "We are delighted to report record annual profits for 2016, our 44th consecutive year of profitability. Our total operating revenues reached a record $20.4 billion, with sustained demand for our legendary low fares and superior Customer Service. Our profit margins were very strong, and our ROIC was a near-record 30.0 percent. Our record profits and balance sheet discipline generated record free cash flow, allowing us to return significant value to our Shareholders. Operationally, our performance was also very solid. We carried a record number of Customers while improving our ontime performance, baggage delivery rate, and net promoter score. My thanks and congratulations to the superb People of Southwest for these outstanding results, which earned them $586 million in profit sharing during 2016.

"We ended the year with a solid fourth quarter 2016 performance. Total operating revenues grew 2.0 percent, year-over-year, to a fourth quarter record $5.1 billion, exceeding our expectations as of the beginning of the fourth quarter. Travel demand and close-in yields improved post-election. In addition, December business travel was stronger than anticipated leading up to the holiday period. Based on current bookings and revenue trends, we estimate first quarter 2017 operating unit revenues will be flat to down one percent, year-over-year. This represents a continued and sequential improvement from the 2.9 percent operating unit revenue year-over-year decline in fourth quarter 2016, which is an encouraging start to the year.

"As expected, our fourth quarter unit costs increased, year-over-year, due to higher fuel costs, pay increases from amended union contracts, and additional depreciation expense associated with the accelerated retirement of our Boeing 737-300 aircraft. While

(continued)

Exhibit 2 (Continued)

> inflationary cost pressures are expected in 2017 due to the union contract pay increases, we are continuing our efforts to drive offsetting cost efficiencies through fleet modernization and ongoing technology investments in our operations.
>
> …"As we close out a year of record results, we begin 2017 with momentum and enthusiasm. We are on track to open a new international terminal in Fort Lauderdale, along with the launch of new service, this June. We are on track to launch the new Boeing 737-8 in the fall. And, we are encouraged by recent revenue trends, as well as the prospects for continued economic growth and moderate fuel prices. We are excited about our current outlook for another strong year with opportunities to win more Customers and reward our People and our Shareholders."
>
> *Source*: http://www.southwestairlinesinvestorrelations.com/news-and-events/news-releases/2017/01-26-2017-111504198 (retrieved November 3, 2018)

Panel B: Excerpt from News Article About Southwest Airlines

Southwest Wins Another Credit Rating Upgrade

> Citing its "consistent record of profitability," S&P Global Ratings upgraded Southwest Airlines (LUV) on Monday to a triple-B-plus rating from triple-B.
>
> This line from the report sums up the rationale for the upgrade nicely:
>
> Southwest is the only U.S. airline that has recorded 44 consecutive years of profitability and remains the largest low-cost airline in the world.
>
> The credit rating agency commends the airline for its record despite headwinds including, "multiple industry cycles, the evolution of the large U.S. hub-and-spoke airlines into more efficient and financially secure competitors, the emergence of ultra-low-cost airlines, the acquisition and integration of AirTran Holdings Inc. in 2011, and the company's entry into more congested but lucrative metropolitan airports such as New York's LaGuardia Airport."
>
> In June, Moody's upgraded Southwest to A3, one notch higher than S&P's new rating on its credit scale. Moody's Senior Credit Officer Jonathan Root wrote then, "the upgrade to A3 reflects Moody's expectation that Southwest will continue to conservatively manage its capital structure, allowing it to sustain credit metrics supportive of the A3 rating category."
>
> *Source*: https://www.barrons.com/articles/southwest-wins-another-credit-rating-upgrade-1502747699 (retrieved November 3, 2018)

In conducting financial analysis of a company, the analyst will regularly refer to the company's financial statements, financial notes, and supplementary schedules and a variety of other information sources. The next section introduces the primary financial statements and some commonly used information sources.

MAJOR FINANCIAL STATEMENTS - BALANCE SHEET

3

b describe the roles of the statement of financial position, statement of comprehensive income, statement of changes in equity, and statement of cash flows in evaluating a company's performance and financial position;

In order to perform an equity or credit analysis of a company, an analyst collects a great deal of information. The nature of the information collected will vary on the basis of the individual decision to be made (or the specific purpose of the analysis) but will typically include information about the economy, industry, and company, as well as information about comparable peer companies. Information from outside the company will likely include economic statistics, industry reports, trade publications, and databases containing information on competitors. The company itself provides core information for analysis in its financial reports, press releases, investor conference calls, and webcasts.

Companies prepare financial reports at regular intervals (annually, semiannually, and/or quarterly depending on the applicable regulatory requirements). Financial reports include financial statements along with supplemental disclosures necessary to assess the company's financial position and periodic performance.

Financial statements are the result of an accounting process that records a company's economic activities, following the applicable accounting standards and principles. These statements summarize the accounting information, mainly for users outside the company (such as investors, creditors, analysts, and others) because insiders have direct access to the underlying financial data summarized in the financial statements and to other information that is not included in the financial reporting process.

Financial statements are almost always audited by independent accountants, who provide an opinion on whether the financial statements present fairly the company's performance and financial position, in accordance with a specified, applicable set of accounting standards and principles.

3.1 Financial Statements and Supplementary Information

A complete set of financial statements include a statement of financial position (i.e., a balance sheet), a statement of comprehensive income (i.e., a single statement of comprehensive income or an income statement and a statement of comprehensive income), a statement of changes in equity, and a statement of cash flows.[1] The balance sheet portrays the company's financial position at a given point in time. The statement of comprehensive income and statement of cash flows present different aspects of a company's performance over a period of time. The statement of changes in equity provides additional information regarding the changes in a company's financial position. In addition, the accompanying required notes, or footnotes, are considered an integral part of a complete set of financial statements.

Along with the required financial statements, a company typically provides additional information in its financial reports. In many jurisdictions, some or all of this additional information is mandated by regulators or accounting standards boards. The additional information provided may include a letter from the chairman of the company, a report from management discussing the results (typically called management discussion and analysis [MD&A] or management commentary), an external auditor's report providing assurances, a governance report describing the structure of the company's board of directors, and a corporate responsibility report. As part of his or her analysis, the financial analyst should read and assess this additional information along with the financial statements. The following sections describe and illustrate each financial statement and some of the additional information.

---

[1] The names of the financial statements are those in IAS 1. Commonly used terms for these financial statements are indicated in parentheses. Later readings will elaborate on each of these financial statements.

3.1.1 *Balance Sheet (Statement of Financial Position)*

The **balance sheet** (also called the **statement of financial position** or **statement of financial condition**) presents a company's financial position by disclosing the resources the company controls (assets) and its obligations to lenders and other creditors (liabilities) at a specific point in time. **Owners' equity** (sometimes called "net assets") represents the excess of assets over liabilities. This amount is attributable to the company's owners or shareholders. Owners' equity is the owners' residual interest in (i.e., residual claim on) the company's assets after deducting its liabilities.

The relationship among the three parts of the balance sheet (assets, liabilities, and owners' equity) can be expressed in the following equation form: Assets = Liabilities + Owners' equity. This equation (sometimes called the accounting equation or the balance sheet equation) shows that the total amount of assets must equal or *balance* with the combined total amounts of liabilities and owners' equity. Alternatively, the equation may be rearranged as follows: Assets − Liabilities = Owners' equity. This formulation emphasizes the residual claim aspect of owners' equity. Depending on the form of the organization, owners' equity may be referred to as "partners' capital" or "shareholders' equity" or "shareholders' funds."

Exhibit 3 presents the balance sheet of the Volkswagen Group from its Annual Report 2017.

Exhibit 3 Balance Sheet of the Volkswagen Group (Excerpt)

€ million	31 Dec. 2017	31 Dec. 2016
Assets		
Noncurrent assets		
Intangible assets	63,419	62,599
Property, plant and equipment	55,243	54,033
Lease assets	39,254	38,439
Investment property	468	512
Equity-accounted investments	8,205	8,616
Other equity investments	1,318	996
Financial services receivables	73,249	68,402
Other financial assets	8,455	8,256
Other receivables	2,252	2,009
Tax receivables	407	392
Deferred tax assets	9,810	9,756
	262,081	**254,010**
Current assets		
Inventories	40,415	38,978
Trade receivables	13,357	12,187
Financial services receivables	53,145	49,673
Other financial assets	11,998	11,844
Other receivables	5,346	5,130
Tax receivables	1,339	1,126

Exhibit 3 (Continued)

€ million	31 Dec. 2017	31 Dec. 2016
Marketable securities	15,939	17,520
Cash, cash equivalents and time deposits	18,457	19,265
Assets held for sale	115	—
	160,112	**155,722**
Total assets	**422,193**	**409,732**
Equity and Liabilities		
Equity		
Subscribed capital	1,283	1,283
Capital reserves	14,551	14,551
Retained earnings	81,367	70,446
Other reserves	560	(1,158)
Equity attributable to Volkswagen AG hybrid capital investors	11,088	7,567
Equity attributable to Volkswagen AG shareholders and hybrid capital investors	108,849	92,689
Noncontrolling interests	229	221
	109,077	**92,910**
Noncurrent liabilities		
Financial liabilities	81,628	66,358
Other financial liabilities	2,665	4,488
Other liabilities	6,199	5,664
Deferred tax liabilities	5,636	4,745
Provisions for pensions	32,730	33,012
Provisions for taxes	3,030	3,556
Other provisions	20,839	21,482
	152,726	**139,306**
Current liabilities		
Put options and compensation rights granted to noncontrolling interest shareholders	3,795	3,849
Financial liabilities	81,844	88,461
Trade payables	23,046	22,794
Tax payables	430	500
Other financial liabilities	8,570	9,438
Other liabilities	15,961	15,461
Provisions for taxes	1,397	1,301
Other provisions	25,347	35,711
	160,389	**177,515**
Total equity and liabilities	**422,193**	**409,732**

Note: Numbers are as shown in the annual report and may not precisely add because of rounding.
Source: Volkswagen 2017 annual report

In Exhibit 3, the balance sheet is presented with the most recent year in the left column and the earlier year in the right column. Although this is a common presentation, analysts should be careful when reading financial statements. In some cases, the ordering may be reversed, with the earlier year(s) on the left and the most recent year on the far right.

At 31 December 2017, Volkswagen's total resources or assets were €422 billion. This number is the sum of non-current assets of €262 billion and current assets of €160 billion.[2] Total equity was €109 billion. Although Volkswagen does not give a total amount for all the balance sheet liabilities, it can be determined by adding the non-current and current liabilities, €153 billion + €160 billion = €313 billion.[3]

Referring back to the basic accounting equation, Assets = Liabilities + Equity, we have €422billion = €313 billion + €109 billion. In other words, Volkswagen has assets of €422 billion, owes €313 billion, and thus has equity of €109 billion. Using the balance sheet and applying financial statement analysis, the analyst can answer such questions as

- Has the company's liquidity (ability to meet short-term obligations) improved?
- Is the company solvent (does it have sufficient resources to cover its obligations)?
- What is the company's financial position relative to the industry?

Volkswagen, a German-based automobile manufacturer, prepares its financial statements in accordance with International Financial Reporting Standards (IFRS). IFRS require companies to present balance sheets that show current and non-current assets and current and non-current liabilities as separate classifications. However, IFRS do not prescribe a particular ordering or format, and the order in which companies present their balance sheet items is largely a function of tradition.

As shown, Volkswagen presents non-current assets before current assets, owners' equity before liabilities, and non-current liabilities before current liabilities. This method generally reflects a presentation from least liquid to most liquid. In other countries, the typical order of presentation may differ. For example, in the United States, Australia, and Canada, companies usually present their assets and liabilities from most liquid to least liquid. Cash is typically the first asset shown, and equity is presented after liabilities.

As a basis for comparison, Exhibit 4 presents the balance sheet of Walmart, Inc. (Walmart) from its 2018 annual report, with a fiscal year end of 31 January.

Exhibit 4 Walmart Consolidated Balance Sheet

	As of 31 January	
(Amounts in $ millions)	**2018**	**2017**
ASSETS		
Current assets:		
Cash and cash equivalents	6,756	6,867
Receivables, net	5,614	5,835
Inventories	43,783	43,046

---

[2] Current assets are defined, in general, as those assets that are cash or cash equivalents; are held for trading; or are expected to be converted to cash (realized), sold, or consumed within 12 months or the company's normal operating cycle. All other assets are classified as non-current.

[3] Current liabilities are defined, in general, as those that are expected to be settled within 12 months or the company's normal operating cycle. All other liabilities are classified as non-current.

Exhibit 4 (Continued)

(Amounts in $ millions)	As of 31 January 2018	As of 31 January 2017
Prepaid expenses and other	3,511	1,941
Total current assets	59,664	57,689
Property and equipment:		
Property and equipment	185,154	179,492
Less accumulated depreciation	(77,479)	(71,782)
Property and equipment, net	107,675	107,710
Property under capital lease and financing obligations:		
Property under capital lease and financing obligations	12,703	11,637
Less accumulated amortization	(5,560)	(5,169)
Property under capital lease and financing obligations, net	7,143	6,468
Goodwill	18,242	17,037
Other assets and deferred charges	11,798	9,921
Total assets	204,522	198,825
LIABILITIES AND EQUITY		
Current liabilities:		
Short-term borrowings	5,257	1,099
Accounts payable	46,092	41,433
Accrued liabilities	22,122	20,654
Accrued income taxes	645	921
Long-term debt due within one year	3,738	2,256
Capital lease and financing obligations due within one year	667	565
Total current liabilities	78,521	66,928
Long-term debt	30,045	36,015
Long-term capital lease and financing obligations	6,780	6,003
Deferred income taxes and other	8,354	9,344
Commitments and contingencies		
Equity:		
Common stock	295	305
Capital in excess of par value	2,648	2,371
Retained earnings	85,107	89,354
Accumulated other comprehensive loss	(10,181)	(14,232)
Total Walmart shareholders' equity	77,869	77,798
Noncontrolling interest	2,953	2,737

(continued)

Exhibit 4 (Continued)

	As of 31 January	
(Amounts in $ millions)	**2018**	**2017**
Total equity	80,822	80,535
Total liabilities and equity	204,522	198,825

Source: Walmart 2018 annual report.

As of 31 January 2018, Walmart has total assets of $205 billion. Liabilities and other non-equity claims total $124 billion, and equity is $81 billion.

4 STATEMENT OF COMPREHENSIVE INCOME

b describe the roles of the statement of financial position, statement of comprehensive income, statement of changes in equity, and statement of cash flows in evaluating a company's performance and financial position;

The statement of comprehensive income can be presented as a single statement of comprehensive income or as two statements, an income statement and a statement of comprehensive income that begins with profit or loss from the income statement. The Volkswagen Group chose the latter form of presentation.

4.1 Income Statement

The income statement presents information on the financial performance of a company's business activities over a period of time. It communicates how much **revenue** and other income the company generated during a period and the expenses it incurred to generate that revenue and other income. Revenue typically refers to amounts charged for the delivery of goods or services in the ordinary activities of a business. Other income may include gains that may or may not arise in the ordinary activities of the business, such as profit on a business disposal. **Expenses** reflect outflows, depletions of assets, and incurrences of liabilities that decrease equity. Expenses typically include such items as cost of sales (cost of goods sold), administrative expenses, and income tax expenses and may be defined to include losses. Net income (revenue plus other income minus expenses) on the income statement is often referred to as the "bottom line" because of its proximity to the bottom of the income statement. Net income may also be referred to as "net earnings," "net profit," and "profit or loss." In the event that expenses exceed revenues and other income, the result is referred to as "net loss."

Income statements are reported on a consolidated basis, meaning that they include the income and expenses of subsidiary companies under the control of the parent (reporting) company. The income statement is sometimes referred to as a **statement of operations** or **profit and loss (P&L) statement**. The basic equation underlying the income statement is Revenue + Other income – Expenses = Income – Expenses = Net income.

In general terms, when one company (the parent) controls another company (the subsidiary), the parent presents its own financial statement information consolidated with that of the subsidiary. (When a parent company owns more than 50 percent of the voting shares of a subsidiary company, it is presumed to control the subsidiary and thus presents consolidated financial statements.) Each line item of the consolidated income statement includes the entire amount from the relevant line item on the subsidiary's income statement (after removing any intercompany transactions). However, if the parent does not own 100 percent of the subsidiary, it is necessary for the parent to present an allocation of net income to the minority interests. Minority interests, also called non-controlling interests, refer to owners of the remaining shares of the subsidiary that are not owned by the parent. The share of consolidated net income attributable to minority interests is shown at the bottom of the income statement along with the net income attributable to shareholders of the parent company. Exhibit 5 presents the income statements of the Volkswagen Group from its 2017 annual report.

Exhibit 5 Income Statement of the Volkswagen Group (Excerpt)

€ million	2017	2016
Sales revenue	**230,682**	**217,267**
Cost of sales	(188,140)	(176,270)
Gross result	**42,542**	**40,997**
Distribution expenses	(22,710)	(22,700)
Administrative expenses	(8,254)	(7,336)
Other operating income	14,500	13,049
Other operating expenses	(12,259)	(16,907)
Operating result	**13,818**	**7,103**
Share of the result of equity-accounted investments	3,482	3,497
Interest income	951	1,285
Interest expenses	(2,317)	(2,955)
Other financial result	(2,022)	(1,638)
Financial result	**94**	**189**
Earnings before tax	**13,913**	**7,292**
Income tax income/expense	(2,275)	(1,912)
Current	(3,205)	(3,273)
Deferred	930	1,361
Earnings after tax	**11,638**	**5,379**
of which attributable to		
Noncontrolling interests	10	10
Volkswagen AG hybrid capital investors	274	225
Volkswagen AG shareholders	11,354	5,144
Basic earnings per ordinary share in €	**22.63**	**10.24**
Diluted earnings per ordinary share in €	**22.63**	**10.24**
Basic earnings per preferred share in €	**22.69**	**10.30**
Diluted earnings per preferred share in €	**22.69**	**10.30**

Note: The numbers are as shown in the annual report and may not add because of rounding.
Source: 2017 Volkswagen annual report.

Exhibit 5 shows that Volkswagen's sales revenue for the year ended 31 December 2017 was €231 billion. Subtracting cost of sales from revenue gives gross profit (called "gross results" by Volkswagen) of €43 billion. After subtracting operating costs and expenses and adding other operating income, the company's operating profit (called "operating result" by Volkswaen) totals €14 billion. Operating profit represents the results of the company's usual business activities before deducting interest expense or taxes. Operating profit (also called operating income in addition to operating result) is thus often referred to as earnings before interest and taxes (EBIT).

Next, operating profit is increased by Volkswagen's share of the profits generated by certain of its investments (€3.5 billion) plus interest income of €1.0 billion, and decreased by losses from its other financial activities (€2.0 billion) and by interest expense of €2.3 billion, resulting in profit (earnings) before tax of €13.9 billion. Total income tax expense for 2017 was €2.3 billion, resulting in profit after tax (net income) of €11.6 billion. After allocating the profits attributable to minority interest ownership in Volkswagen subsidiary companies, the profit attributable to shareholders of Volkswagen for 2017 was €11.4 billion.

Companies present both basic and diluted earnings per share on the face of the income statement. Earnings per share numbers represent net income attributable to the class of shareholders divided by the relevant number of shares outstanding during the period. Basic earnings per share is calculated using the weighted-average number of common (ordinary) shares that were actually outstanding during the period and the profit or loss attributable to the common shareowners. Diluted earnings per share uses **diluted shares**—the number of shares that would hypothetically be outstanding if potentially dilutive claims on common shares (e.g., stock options or convertible bonds) were exercised or converted by their holders—and an appropriately adjusted profit or loss attributable to the common shareowners.

Volkswagen has two types of shareholders, ordinary and preferred, and presents earnings per share information for both, although there is no requirement to present earnings per share information for preferred shareowners. Volkswagen's basic earnings per ordinary share was €22.63. A note to the company's financial statements explains that this number was calculated as follows: €11.4 billion profit attributable to shareholders of Volkswagen, of which €6.8 billion is attributable to ordinary shareholders and the balance is attributable to preferred shareholders. The €6.8 billion attributable to ordinary shareholders divided by the weighted-average number of ordinary shares of 0.295 billion shares equals basic earnings per share of €22.63. Similar detail is provided in the notes for each of the earnings per share numbers.

An analyst examining the income statement might note that Volkswagen was profitable in both years. The company's profits increased substantially in 2017, primarily because of higher sales and lower other operating expenses (footnotes reveal this was largely due to $4 billion less in litigation expenses in 2017). The analyst might formulate questions related to profitability, such as the following:

- Is the change in revenue related to an increase in units sold, an increase in prices, or some combination?
- If the company has multiple business segments (for example, Volkswagen's segments include passenger cars, light commercial vehicles, and financial services, among others), how are the segments' revenue and profits changing?
- How does the company compare with other companies in the industry?

Answering such questions requires the analyst to gather, analyze, and interpret information from a number of sources, including, but not limited to, the income statement.

4.2 Other Comprehensive Income

Comprehensive income includes all items that impact owners' equity but are not the result of transactions with shareowners. Some of these items are included in the calculation of net income, and some are included in other comprehensive income (OCI). When comprehensive income is presented in two statements, the statement of comprehensive income begins with the profit or loss from the income statement and then presents the components of OCI.

Exhibit 6 presents an excerpt from the statement of comprehensive income of the Volkswagen Group from its 2017 annual report.

Exhibit 6 Statement of Comprehensive Income of the Volkswagen Group (Excerpt)

€ million	Fiscal Year Ended 31 December 2017
Earnings after tax	**11,638**
Pension plan remeasurements recognized in other comprehensive income	
Pension plan remeasurements recognized in other comprehensive income, before tax	785
Deferred taxes relating to pension plan remeasurements recognized in other comprehensive income	(198)
Pension plan remeasurements recognized in other comprehensive income, net of tax	588
Share of other comprehensive income of equity-accounted investments that will not be reclassified to profit or loss, net of tax	96
Items that will not be reclassified to profit or loss	**683**
Exchange differences on translating foreign operations	
Unrealized currency translation gains/losses	(2,095)
Transferred to profit or loss	(4)
Exchange differences on translating foreign operations, before tax	(2,099)
Deferred taxes relating to exchange differences on translating foreign operations	(8)
Exchange differences on translating foreign operations, net of tax	(2,107)
Cash flow hedges	
Fair value changes recognized in other comprehensive income	6,137
Transferred to profit or loss	(558)
Cash flow hedges, before tax	5,579
Deferred taxes relating to cash flow hedges	(1,597)
Cash flow hedges, net of tax	3,982
Available-for-sale financial assets	
Fair value changes recognized in other comprehensive income	56
Transferred to profit or loss	62
Available-for-sale financial assets, before tax	118
Deferred taxes relating to available-for-sale financial assets	(25)
Available-for-sale financial assets, net of tax	93
Share of other comprehensive income of equity-accounted investments that may be reclassified subsequently to profit or loss, net of tax	(346)
Items that may be reclassified subsequently to profit or loss	**1,622**
Other comprehensive income, before tax	4,133
Deferred taxes relating to other comprehensive income	(1,828)

(continued)

Exhibit 6 (Continued)

€ million	Fiscal Year Ended 31 December 2017
Other comprehensive income, net of tax	**2,305**
Total comprehensive income	**13,943**

Source: Volkswagen 2017 annual report.

Exhibit 6 shows total comprehensive income for 2017 was €13.9 billion, which is the sum of earnings after tax of €11.6 billion, reported on the income statement, and other comprehensive income of €2.3 billion. The items in OCI reflect changes in the company's equity that are not considered to be profit or loss, some of which may be reclassified as such in the future such as unrealized currency translation gains and losses. The statement of comprehensive income will be discussed in greater detail in a later reading.

5

STATEMENT OF CHANGES IN EQUITY AND CASH FLOW STATEMENT

b describe the roles of the statement of financial position, statement of comprehensive income, statement of changes in equity, and statement of cash flows in evaluating a company's performance and financial position;

The statement of changes in equity, sometimes called the "statement of changes in owners' equity" or "statement of changes in shareholders' equity," primarily serves to report changes in the owners' investment in the business over time. The basic components of owners' equity are paid-in capital and retained earnings. Retained earnings include the cumulative amount of the company's profits that have been retained in the company. In addition, non-controlling or minority interests and reserves that represent accumulated OCI items are included in equity. The latter items may be shown separately or included in retained earnings. Volkswagen includes reserves as components of retained earnings.

The statement of changes in equity is organized to present, for each component of equity, the beginning balance, any increases during the period, any decreases during the period, and the ending balance. For paid-in capital, an example of an increase is a new issuance of equity and an example of a decrease is a repurchase of previously issued stock. For retained earnings, income (both net income as reported on the income statement and OCI) is the most common increase and a dividend payment is the most common decrease.

Volkswagen's balance sheet in Exhibit 3 shows that equity at the end of 2017 totaled €109 billion, compared with €93 billion at the end of 2016. The company's statement of changes in equity presents additional detail on the change in each line item. Exhibit 7 presents an excerpt of the statement of changes in equity of the Volkswagen Group from its 2017 annual report. For purposes of brevity, several interim columns were excluded from the presentation.

Exhibit 7 Statement of Changes in Equity of the Volkswagen Group (Excerpt)

€ million	Subscribed capital	Capital reserves	Retained earnings	Total equity
Balance at 1 January 2017	1,283	14,551	70,446	92,910
Earnings after tax	—	—	11,354	11,638
Other comprehensive income, net of tax	—	—	586	2,305
Total comprehensive income	—	—	11,940	13,943
Capital increases	—	—	—	3,481
Dividends payment	—	—	−1,015	−1,332
Capital transactions involving a change in ownership interest	—	—	—	—
Other changes	—	—	−4	75
Balance at 31 December 2017	1,283	14,551	81,367	109,077

Note: Numbers are as shown in the annual report and may not add and cross-add because of the exclusion of columns and due to rounding.

In Exhibit 7, as shown in the far right column, total equity is increased during the year by total comprehensive income of €13.9 billion and by capital infusions of €3.5 billion; it is decreased by dividends of €1.3 billion; and finally, slightly increased by €75 million of other changes. Explanatory notes on equity are included in the notes to the consolidated financial statements.

5.1 Cash Flow Statement

Although the income statement and balance sheet provide measures of a company's success in terms of performance and financial position, cash flow is also vital to a company's long-term success. Disclosing the sources and uses of cash helps creditors, investors, and other statement users evaluate the company's liquidity, solvency, and financial flexibility. **Financial flexibility** is the ability of the company to react and adapt to financial adversity and opportunities.

The cash flow statement classifies all cash flows of the company into three categories: operating, investing, and financing. Cash flows from **operating activities** generally involve the cash effects of transactions involved in the determination of net income and, hence, comprise the day-to-day operations of the company. Cash flows from **investing activities** are associated with the acquisition and disposal of long-term assets, such as property and equipment. Cash flows from **financing activities** relate to obtaining or repaying capital to be used in the business. IFRS permit more flexibility than US GAAP in classifying dividend and interest receipts and payments within these categories.

Exhibit 8 presents Volkswagen's statement of cash flows for the fiscal years ended 31 December 2017 and 2016.

Exhibit 8 Cash Flow Statement of the Volkswagen Group: 1 January to 31 December

€ million	2017	2016
Cash and cash equivalents at beginning of period	**18,833**	**20,462**
Earnings before tax	13,913	7,292
Income taxes paid	(3,664)	(3,315)

(continued)

Exhibit 8 (Continued)

€ million	2017	2016
Depreciation and amortization of, and impairment losses on, intangible assets, property, plant and equipment, and investment property	10,562	10,100
Amortization of and impairment losses on capitalized development costs	3,734	3,586
Impairment losses on equity investments	136	130
Depreciation of and impairment losses on lease assets	7,734	7,107
Gain/loss on disposal of noncurrent assets and equity investments	(25)	(222)
Share of the result of equity-accounted investments	274	377
Other noncash expense/income	(480)	716
Change in inventories	(4,198)	(3,637)
Change in receivables (excluding financial services)	(1,660)	(2,155)
Change in liabilities (excluding financial liabilities)	5,302	5,048
Change in provisions	(9,443)	5,966
Change in lease assets	(11,478)	(12,074)
Change in financial services receivables	(11,891)	(9,490)
Cash flows from operating activities	**(1,185)**	**9,430**
Investments in intangible assets (excluding development costs), property, plant and equipment, and investment property	(13,052)	(13,152)
Additions to capitalized development costs	(5,260)	(5,750)
Acquisition of subsidiaries	(277)	(119)
Acquisition of other equity investments	(561)	(309)
Disposal of subsidiaries	496	(7)
Disposal of other equity investments	24	2,190
Proceeds from disposal of intangible assets, property, plant and equipment, and investment property	411	351
Change in investments in securities	1,376	(1,245)
Change in loans and time deposits	335	(2,638)
Cash flows from investing activities	**(16,508)**	**(20,679)**
Capital contributions	3,473	—
Dividends paid	(1,332)	(364)
Capital transactions with noncontrolling interest shareholders	—	(3)
Proceeds from issuance of bonds	30,279	14,262
Repayments of bonds	(17,877)	(23,601)
Changes in other financial liabilities	3,109	19,455
Lease payments	(28)	(36)
Cash flows from financing activities	**17,625**	**9,712**
Effect of exchange rate changes on cash and cash equivalents	(727)	(91)
Net change in cash and cash equivalents	**(796)**	**(1,628)**
Cash and cash equivalents at end of period	**18,038**	**18,833**

After showing the beginning cash balance, the operating activities section starts with profit before tax,[4] €13.9 billion for 2017, subtracts actual income tax payments, and then adjusts for the effects of non-cash transactions, accruals and deferrals, and transactions of an investing and financing nature to arrive at the amount of cash generated from operating activities: a negative €1.2 billion (a net use of cash). This approach to reporting cash flow from operating activities is termed the indirect method. The direct method of reporting cash flows from operating activities discloses major classes of gross cash receipts and gross cash payments. Examples are cash received from customers and cash paid to suppliers and employees.

The indirect method emphasizes the different perspectives of the income statement and cash flow statement. On the income statement, income is reported when earned, not necessarily when cash is received, and expenses are reported when incurred, not necessarily when paid. The cash flow statement presents another aspect of performance: the ability of a company to generate cash flow from running its business. Ideally, for an established company, the analyst would like to see that the primary source of cash flow is from operating activities as opposed to investing or financing activities.

The sum of the net cash flows from operating, investing, and financing activities and the effect of exchange rates on cash equals the net change in cash during the fiscal year. For Volkswagen, the sum of these four items was a negative €796 million in 2017, decreasing the company's cash and cash equivalents from €18.8 billion at the beginning of the period to €18.0 billion at the end of the period.

FINANCIAL NOTES, SUPPLEMENTARY SCHEDULES, AND MANAGEMENT COMMENTARY

6

c describe the importance of financial statement notes and supplementary information—including disclosures of accounting policies, methods, and estimates—and management's commentary;

The notes (also sometimes referred to as footnotes) that accompany the four financial statements are required and are an integral part of the complete set of financial statements. The notes provide information that is essential to understanding the information provided in the primary statements. Volkswagen's 2017 financial statements, for example, include more than 100 pages of notes.

The notes disclose the basis of preparation for the financial statements. For example, Volkswagen discloses that its fiscal year corresponds to the calendar year; its financial statements are prepared in accordance with IFRS as adopted by the European Union; the statements are prepared in compliance with German law; the statements are denominated in millions of euros unless otherwise specified; and the figures have been rounded, which might give rise to minor discrepancies when they are added. Volkswagen also states that its financial statements are on a consolidated basis—that is, including Volkswagen AG and all of the subsidiary companies it controls.

The notes also disclose information about the accounting policies, methods, and estimates used to prepare the financial statements. As will be discussed in later readings, both IFRS and US GAAP allow some flexibility in choosing among alternative policies and methods when accounting for certain items. This flexibility aims to meet the divergent needs of many businesses for reporting a variety of economic transactions. In addition to differences in accounting policies and methods, differences arise as a result of estimates needed to record and measure transactions, events, and financial statement line items.

---

[4] Other companies may choose to begin with net income.

Overall, flexibility in accounting choices is necessary because, ideally, a company will select those policies, methods, and estimates that are allowable and most relevant and that fairly reflect the unique economic environment of the company's business and industry. Flexibility can, however, create challenges for the analyst because the use of different policies, methods, and estimates reduces comparability across different companies' financial statements. Comparability occurs when different companies' information is measured and reported in a similar manner over time. Comparability helps the analyst identify and analyze the real economic differences across companies, rather than differences that arise solely from different accounting choices. Because comparability of financial statements is a critical requirement for objective financial analysis, an analyst should be aware of the potential for differences in accounting choices even when comparing two companies that use the same set of accounting standards.

For example, if a company acquires a piece of equipment to use in its operations, accounting standards require that the cost of the equipment be reported as an expense by allocating its cost, less any residual value, in a systematic manner over the equipment's useful life. This allocation of the cost is known as **depreciation**. Accounting standards permit flexibility, however, in determining the manner in which each year's expense is determined. Two companies may acquire similar equipment but use different methods and assumptions to record the expense over time. An analyst's ability to compare the companies' performance is hindered by the difference. Analysts must understand reporting choices in order to make appropriate adjustments when comparing companies' financial positions and performance.

A company's significant accounting choices (policies, methods, and estimates) must be discussed in the notes to the financial statements. For example, a note containing a summary of significant accounting policies includes how the company recognizes its revenues and depreciates its non-current tangible assets. Analysts must understand the accounting choices a company makes and determine whether they are similar to those of other companies identified and used as benchmarks or comparables. If the policies of the companies being compared are different, the analyst who understands accounting and financial reporting can often make necessary adjustments so that the financial statement data used are more comparable.

For many companies, the financial notes and supplemental schedules provide explanatory information about every line item (or almost every line item) on the balance sheet and income statement, as illustrated by the note references in Volkswagen's balance sheet and income statement in Exhibits 3 and 5. In addition, note disclosures include information about the following (this is not an exhaustive list):

- financial instruments and risks arising from financial instruments,
- commitments and contingencies,
- legal proceedings,
- related-party transactions,
- subsequent events (i.e., events that occur after the balance sheet date),
- business acquisitions and disposals, and
- operating segments' performance.

Exhibit 9 Notes to the Consolidated Financial Statements of the Volkswagen Group: Selected Data on Operating Segments (Excerpt)

€ million	Passenger Cars	Commercial Vehicles	Power Engineering	Financial Services	Total segments	Reconciliation	Volkswagen Group
Sales revenue from external customers	169,513	27,632	3,280	30,191	230,618	64	230,682
Intersegment sales revenue	18,892	7,568	3	3,541	30,004	−30,004	—
Total sales revenue	188,405	35,200	3,283	33,733	260,621	−29,939	230,682
Depreciation and amortization	11,363	2,557	371	6,797	21,089	−147	20,941
Impairment losses	704	2	0	574	1,280	0	1,280
Reversal of impairment losses	14	1	—	41	56	—	56
Segment result (operating result)	12,644	1,892	−55	2,673	17,153	−3,335	13,818
Share of the result of equity-accounted investments	3,390	83	1	9	3,482	—	3,482
Net interest result and other financial result	−1,920	−220	−2	−180	−2,321	−1,067	−3,388
Equity-accounted investments	6,724	753	18	710	8,205	—	8,205
Investments in intangible assets, property, plant and equipment, and investment property	15,713	1,915	159	421	18,208	104	18,313

An analyst uses a significant amount of judgment in deciding how to incorporate information from note disclosures into the analysis. For example, such information as financial instrument risk, contingencies, and legal proceedings can alert an analyst to risks that can affect a company's financial position and performance in the future and that require monitoring over time. As another example, information about a company's operating segments can be useful as a means of quickly understanding what a company does and how and where it earns money. The operating segment data shown in Exhibit 9 appear in the notes to the financial statements for Volkswagen. (The totals of the segment data do not equal the amounts reported in the company's financial statements because the financial statement data are adjusted for intersegment activities and unallocated items. The notes provide a complete reconciliation of the segment data to the reported data.) From the data in Exhibit 9, an analyst can quickly see that most of the company's revenues and operating profits come from the sale of passenger cars. Over 80 percent of the company's revenues was generated by this segment This segment accounted for over 90 percent of the company's total segment operating profits.

Experience using the disclosures made by a company and its competitors typically enhances an analyst's judgment about the relative importance of different disclosures and the ways in which they can be helpful.

6.1 Management Commentary or Management's Discussion and Analysis

Publicly held companies typically include a section in their annual reports where management discusses a variety of issues, including the nature of the business, past results, and future outlook. This section is referred to by a variety of names, including management report(ing), management commentary, operating and financial review, and management's discussion and analysis. Inclusion of a management report is recommended by the International Organization of Securities Commissions and frequently required by regulatory authorities, such as the US Securities and Exchange Commission (SEC) or the UK Financial Reporting Council (FRC). In Germany, management reporting has been required since 1931 and is audited.

The discussion by management is arguably one of the most useful parts of a company's annual report besides the financial statements themselves; however, other than excerpts from the financial statements, information included in the management commentary is typically unaudited. When using information from the management report, an analyst should be aware of whether the information is audited or unaudited.

To help improve the quality of the discussion by management, the International Accounting Standards Board (IASB) issued an IFRS Practice Statement "Management Commentary" includes a framework for the preparation and presentation of management commentary. The framework provides guidance rather than sets forth requirements in a standard. The framework identifies five content elements of a "decision-useful management commentary:" 1) the nature of the business; 2) management's objectives and strategies; 3) the company's significant resources, risks, and relationships; 4) results of operations; and 5) critical performance measures.

In the United States, the SEC requires listed companies to provide an MD&A and specifies the content.[5] Management must highlight any favorable or unfavorable trends and identify significant events and uncertainties that affect the company's liquidity, capital resources, and results of operations. The MD&A must also provide information about the effects of inflation, changing prices, or other material events and uncertainties that may cause the future operating results and financial condition to materially depart from the current reported financial information. In addition, the MD&A must provide information about off-balance-sheet obligations and about contractual commitments such as purchase obligations. Management should also discuss the critical accounting policies that require them to make subjective judgments and that have a significant impact on reported financial results.

The management commentary, or MD&A, is a good starting place for understanding information in the financial statements. In particular, the forward-looking disclosures, such as those about planned capital expenditures, new store openings, or divestitures, can be useful in projecting a company's future performance. However, the commentary is only one input for the analyst seeking an objective and independent perspective on a company's performance and prospects.

The management report in Volkswagen's 2017 annual report includes much information of potential interest to an analyst. The lengthy report contains sections such as Goals and Strategies, Internal Management and Key Performance Indicators, Structure and Business Activities, Corporate Governance, Remuneration, Executive Bodies, Disclosures Required under Takeover Law, Diesel Issue, Business Development, Shares and Bonds, Results of Operations, and Report on Risks and Opportunities.

---

[5] Relevant sections of SEC requirements are included for reference in the FASB ASC. The FASB ASC does not include sections of SEC requirements that deal with matters outside the basic financial statements, such as the MD&A.

AUDITOR'S REPORTS

7

d describe the objective of audits of financial statements, the types of audit reports, and the importance of effective internal controls;

Financial statements presented in companies' annual reports are generally required to be audited (examined) by an independent accounting firm in accordance with specified auditing standards. The independent auditor then provides a written opinion on the financial statements. This opinion is referred to as the audit report. Audit reports may vary in different jurisdictions, but the minimum components, including a specific statement of the auditor's opinion, are similar. Audits of financial statements may be required by contractual arrangement, law, or regulation.

International standards on auditing (ISAs) have been developed by the International Auditing and Assurance Standards Board (IAASB). This body has emerged from the International Federation of Accountants. ISAs have been adopted by many countries and are referenced in audit reports issued in those countries. Other countries, such as the United States, specify their own auditing standards. With the enactment of the Sarbanes–Oxley Act of 2002 in the United States, auditing standards for public companies are promulgated by the Public Company Accounting Oversight Board.

Under international standards for auditing (ISAs), the overall objectives of an auditor in conducting an audit of financial statements are

A To obtain reasonable assurance about whether the financial statements as a whole are free from material misstatement, whether due to fraud or error, thereby enabling the auditor to express an opinion on whether the financial statements are prepared, in all material respects, in accordance with an applicable financial reporting framework; and

B To report on the financial statements, and communicate as required by the ISAs, in accordance with the auditor's findings.[6]

Publicly traded companies may also have requirements set by regulators or stock exchanges, such as appointing an independent audit committee within its board of directors to oversee the audit process. The audit process provides a basis for the independent auditor to express an opinion on whether the information in the audited financial statements presents fairly the financial position, performance, and cash flows of the company in accordance with a specified set of accounting standards.

Audits are designed and conducted using sampling techniques, and financial statement line items may be based on estimates and assumptions. This means that the auditors cannot express an opinion that provides absolute assurance about the accuracy or precision of the financial statements. Instead, the independent audit report provides *reasonable assurance* that the financial statements are *fairly presented*, meaning that there is a high probability that the audited financial statements are free from *material* error, fraud, or illegal acts that have a direct effect on the financial statements.

The independent audit report expresses the auditor's opinion on the fairness of the audited financial statements, and specifies which financial statements were audited, the reporting entity, and the date. An *unqualified* audit opinion states that the financial statements give a "true and fair view" (international) or are "fairly presented" (international and US) in accordance with applicable accounting standards. This is also referred to as an "unmodified" or a "clean" opinion and is the one that analysts would like to see in a financial report. There are several other types of modified opinions. A *qualified* audit opinion is one in which there is some scope limitation or exception to accounting standards. Exceptions are described in the audit report with additional explanatory paragraphs so that the analyst can determine the importance of the exception. An *adverse* audit opinion is issued when an auditor determines that the financial statements materially depart from accounting standards and are not fairly presented. Finally, a *disclaimer of opinion* occurs when, for some reason, such as a scope limitation, the auditors are unable to issue an opinion.

---

[6] See the International Auditing and Assurance Standards Board (IAASB) *Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements.*

The audit report also describes the basis for the auditor's opinion and, for listed companies, includes a discussion of Key Audit Matters (international) and Critical Audit Matters (US).[7] Key Audit Matters are defined as issues that the auditor considers to be most important, such as those that have a higher risk of misstatement, involve significant management judgment, or report the effects of significant transactions during the period. Critical Audit Matters are defined as issues that involve "especially challenging, subjective, or complex auditor judgment" and similarly include areas with higher risk of misstatement or involving significant management judgement and estimates.

Exhibit 10 presents the independent auditor's report for Volkswagen. Note that Volkswagen received an unqualified audit opinion (i.e., clean or unmodified opinion) from PricewaterhouseCoopers for the company's fiscal year ended 31 December 2017.

Exhibit 10 Volkswagen's Independent Audit Report

On completion of our audit, we issued the following unqualified auditor's report dated February 23, 2018. This report was originally prepared in German. In case of ambiguities the German version takes precedence:

To VOLKSWAGEN AKTIENGESELLSCHAFT, Wolfsburg

REPORT ON THE AUDIT OF THE CONSOLIDATED FINANCIAL STATEMENTS AND OF THE GROUP MANAGEMENT REPORT

AUDIT OPINIONS

We have audited the consolidated financial statements of VOLKSWAGEN AKTIENGESELLSCHAFT, Wolfsburg, and its subsidiaries (the Group), which comprise the income statement and the statement of comprehensive income, the balance sheet, the statement of changes in equity and the cash flow statement for the financial year from January 1 to December 31, 2017, and notes to the consolidated financial statements, including a summary of significant accounting policies. In addition, we have audited the group management report of VOLKSWAGEN AKTIENGESELLSCHAFT, which is combined with the Company's management report, for the financial year from January 1 to December 31, 2017. We have not audited the content of those parts of the group management report listed in the "Other Information" section of our auditor's report in accordance with the German legal requirements.

In our opinion, on the basis of the knowledge obtained in the audit,

- the accompanying consolidated financial statements comply, in all material respects, with the IFRSs as adopted by the EU, and the additional requirements of German commercial law pursuant to § [Article] 315e Abs. [paragraph] 1 HGB [Handelsgesetzbuch: German Commercial Code] and, in compliance with these requirements, give a true and fair

---

[7] Discussion of Key Audit Matters in the auditor's report is required by the International Standard on Auditing (ISA) ISA 701, effective in 2017, issued by the International Audit and Assurance Standards Board. Discussion of Critical Audit Matters in the auditor's report is required by the Auditor Reporting Standard AS 3101, effective for large filers' fiscal years ending on or after June 30,2019, issued by the Public Company Accounting Oversight Board.

Exhibit 10 (Continued)

view of the assets, liabilities, and financial position of the Group as at December 31, 2017, and of its financial performance for the financial year from January 1 to December 31, 2017, and

- the accompanying group management report as a whole provides an appropriate view of the Group's position. In all material respects, this group management report is consistent with the consolidated financial statements, complies with German legal requirements and appropriately presents the opportunities and risks of future development. Our audit opinion on the group management report does not cover the content of those parts of the group management report listed in the "Other Information" section of our auditor's report.

Pursuant to § 322 Abs. 3 Satz [sentence] 1 HGB, we declare that our audit has not led to any reservations relating to the legal compliance of the consolidated financial statements and of the group management report.

BASIS FOR THE AUDIT OPINIONS

We conducted our audit of the consolidated financial statements and of the group management report in accordance with § 317 HGB and the EU Audit Regulation (No. 537/2014, referred to subsequently as "EU Audit Regulation") and in compliance with German Generally Accepted Standards for Financial Statement Audits promulgated by the Institut der Wirtschaftsprüfer [Institute of Public Auditors in Germany] (IDW). Our responsibilities under those requirements and principles are further described in the "Auditor's Responsibilities for the Audit of the Consolidated Financial Statements and of the Group Management Report" section of our auditor's report. We are independent of the group entities in accordance with the requirements of European law and German commercial and professional law, and we have fulfilled our other German professional responsibilities in accordance with these requirements. In addition, in accordance with Article 10 (2) point (f) of the EU Audit Regulation, we declare that we have not provided non-audit services prohibited under Article 5 (1) of the EU Audit Regulation. We believe that the audit evidence we have obtained is sufficient and appropriate to provide a basis for our audit opinions on the consolidated financial statements and on the group management report.

EMPHASIS OF MATTER – DIESEL ISSUE

We draw attention to the information provided and statements made in section "Key Events" of the notes to the consolidated financial statements and in section "Diesel Issue" of the group management report with regard to the diesel issue including information about the underlying causes, the noninvolvement of members of the board of management as well as the impact on these financial statements.

Based on the results of the various measures taken to investigate the issue presented so far, which underlie the consolidated financial statements and the group management report, there is still no evidence that members of the Company's board of management were aware of the deliberate manipulation of engine management software before summer 2015. Nevertheless, should as a result of the ongoing investigation new solid knowledge be obtained showing that members of the board of management were informed earlier about the diesel issue, this could eventually have an impact on the consolidated financial statements and on the group management report for financial year 2017 and prior years.

(continued)

Exhibit 10 (Continued)

The provisions for warranties and legal risks recorded so far are based on the presented state of knowledge. Due to the inevitable uncertainties associated with the current and expected litigation it cannot be excluded that a future assessment of the risks may be different.

Our opinions on the consolidated financial statements and on the group management report are not modified in respect of this matter.

Source: Volkswagen 2017 annual report

In the United States, under the Sarbanes–Oxley Act, the auditors must also express an opinion on the company's internal control systems. This information may be provided in a separate opinion or incorporated as a paragraph in the opinion related to the financial statements. The internal control system is the company's internal system that is designed, among other things, to ensure that the company's process for generating financial reports is sound. Although management has always been responsible for maintaining effective internal control, the Sarbanes–Oxley Act greatly increases management's responsibility for demonstrating that the company's internal controls are effective. Management of publicly traded companies in the United States are now required by securities regulators to explicitly accept responsibility for the effectiveness of internal control, evaluate the effectiveness of internal control using suitable control criteria, support the evaluation with sufficient competent evidence, and provide a report on internal control.

Although these reports and attestations provide some assurances to analysts, they are not infallible. The analyst must always use a degree of healthy skepticism when analyzing financial statements.

8 OTHER SOURCES OF INFORMATION

e identify and describe information sources that analysts use in financial statement analysis besides annual financial statements and supplementary information;

The information described in Sections 3–7 is generally provided to shareholders at least annually. In addition, companies also provide information on management and director compensation, company stock performance, and any potential conflicts of interest that may exist between management, the board, and shareholders. This information may appear in the company's annual report or other publicly available documents. Public companies often provide this information in proxy statements, which are distributed to shareholders about matters that are to be put to a vote at the annual (or special) meeting of shareholders.

Interim reports are also provided by the company either semiannually or quarterly, depending on the applicable regulatory requirements. Interim reports generally present the four primary financial statements and condensed notes but are not audited. These interim reports provide updated information on a company's performance and financial position since the last annual period.

Companies also provide relevant current information on their websites, in press releases, and in conference calls with analysts and investors. One type of press release, which analysts often consider to be particularly important, is the periodic earnings announcement. The earnings announcement often happens well before the company files its formal financial statements. Such earnings announcements are often followed by a conference call in which the company's senior executives describe the company's performance and answer questions posed by conference call participants. Following the earnings conference call, the investor relations portion of the company's website may post a recording of the call accompanied by slides and supplemental information that was discussed.

When performing financial statement analysis, analysts should review all these company sources of information as well as information from external sources regarding the economy, the industry, the company, and peer (comparable) companies. Information on the economy, industry, and peer companies is useful in putting the company's financial performance and position in perspective and in assessing the company's future. In most cases, information from sources apart from the company is crucial to an analyst's effectiveness. For example, an analyst studying a consumer-oriented company will typically seek direct experience with the products (taste the food or drink, use the shampoo or soap, visit the stores or hotels). An analyst following a highly regulated industry will study the existing and expected relevant regulations. An analyst following a highly technical industry will gain relevant expertise personally or seek input from a technical specialist. In sum, thorough research goes beyond financial reports.

The next section presents a framework for using all this information in financial statement analysis.

FINANCIAL STATEMENT ANALYSIS FRAMEWORK

9

f describe the steps in the financial statement analysis framework.

Analysts work in a variety of positions within the investment management industry. Some are equity analysts whose main objective is to evaluate potential investments in a company's equity securities (i.e., the shares or stock it issues) as a basis for deciding whether a prospective investment is attractive and what an appropriate purchase price might be. Others are credit analysts who evaluate the credit-worthiness of a company to decide whether (and on what terms) a loan should be made or what credit rating should be assigned. Analysts may also be involved in a variety of other tasks, such as evaluating the performance of a subsidiary company, evaluating a private equity investment, or finding stocks that are overvalued for purposes of taking a short position.

This section presents a generic framework for financial statement analysis that can be used in these various tasks. The framework is summarized in Exhibit 11.[8]

---

[8] Components of this framework have been adapted from van Greuning and Bratanovic (2003, p. 300) and from Benninga and Sarig (1997, pp. 134–156).

Exhibit 11 Financial Statement Analysis Framework

Phase	Sources of Information	Output
1 Articulate the purpose and context of the analysis.	■ The nature of the analyst's function, such as evaluating an equity or debt investment or issuing a credit rating. ■ Communication with client or supervisor on needs and concerns. ■ Institutional guidelines related to developing specific work product.	■ Statement of the purpose or objective of analysis. ■ A list (written or unwritten) of specific questions to be answered by the analysis. ■ Nature and content of report to be provided. ■ Timetable and budgeted resources for completion.
2 Collect input data.	■ Financial statements, other financial data, questionnaires, and industry/economic data. ■ Discussions with management, suppliers, customers, and competitors. ■ Company site visits (e.g., to production facilities or retail stores).	■ Organized financial statements. ■ Financial data tables. ■ Completed questionnaires, if applicable.
3 Process data.	■ Data from the previous phase.	■ Adjusted financial statements. ■ Common-size statements. ■ Ratios and graphs. ■ Forecasts.
4 Analyze/interpret the processed data.	■ Input data as well as processed data.	■ Analytical results.
5 Develop and communicate conclusions and recommendations (e.g., with an analysis report).	■ Analytical results and previous reports. ■ Institutional guidelines for published reports.	■ Analytical report answering questions posed in Phase 1. ■ Recommendation regarding the purpose of the analysis, such as whether to make an investment or grant credit.
6 Follow-up.	■ Information gathered by periodically repeating above steps as necessary to determine whether changes to holdings or recommendations are necessary.	■ Updated reports and recommendations.

The following sections discuss the individual phases of financial statement analysis.

9.1 Articulate the Purpose and Context of Analysis

Prior to undertaking any analysis, it is essential to understand the purpose of the analysis. An understanding of the purpose is particularly important in financial statement analysis because of the numerous available techniques and the substantial amount of data.

Some analytical tasks are well defined, in which case articulating the purpose of the analysis requires little decision making by the analyst. For example, a periodic credit review of an investment-grade debt portfolio or an equity analyst's report on a particular company may be guided by institutional norms such that the purpose of the analysis is given. Furthermore, the format, procedures, and/or sources of information may also be given.

For other analytical tasks, articulating the purpose of the analysis requires the analyst to make decisions. The purpose of an analysis guides further decisions about the approach, the tools, the data sources, the format in which to report the results of the analysis, and the relative importance of different aspects of the analysis.

When facing a substantial amount of data, a less experienced analyst may be tempted to just start making calculations and generating financial ratios without considering what is relevant for the decision at hand. It is generally advisable to resist this temptation and thus avoid unnecessary or pointless efforts. Consider the questions: If you could have all the calculations and ratios completed instantly, what conclusion would you be able to draw? What question would you be able to answer? What decision would your answer support?

The analyst should also define the context at this stage. Who is the intended audience? What is the end product—for example, a final report explaining conclusions and recommendations? What is the time frame (i.e., when is the report due)? What resources and resource constraints are relevant to completion of the analysis? Again, the context may be predefined (i.e., standard and guided by institutional norms).

Having clarified the purpose and context of the financial statement analysis, the analyst should next compile the specific questions to be answered by the analysis. For example, if the purpose of the financial statement analysis (or, more likely, the particular stage of a larger analysis) is to compare the historical performance of three companies operating in a particular industry, specific questions would include the following: What has been the relative growth rate of the companies, and what has been their relative profitability?

9.2 Collect Data

Next, the analyst obtains the data required to answer the specific questions. A key part of this step is obtaining an understanding of the company's business, financial performance, and financial position (including trends over time and in comparison with peer companies). For historical analyses, financial statement data alone are adequate in some cases. For example, to screen a large number of companies to find those with a minimum level of profitability, financial statement data alone would be adequate. But to address more in-depth questions, such as why and how one company performed better or worse than its competitors, additional information would be required. As another example, to compare the historical performance of two companies in a particular industry, the historical financial statements would be sufficient to determine which had faster-growing sales or earnings and which was more profitable. However, a broader comparison with overall industry growth and profitability would obviously require industry data.

Furthermore, information on the economy and industry is necessary to understand the environment in which the company operates. Analysts often take a top-down approach whereby they 1) gain an understanding of the macroeconomic environment, such as prospects for growth in the economy and inflation, 2) analyze the prospects of the industry in which the company operates, based on the expected macroeconomic environment, and 3) determine the prospects for the company given the expected industry and macroeconomic environments. For example, an analyst may need to forecast future growth in earnings for a company. Past company data provide the platform for statistical forecasting; however, an understanding of economic and industry conditions can improve the analyst's ability to forecast earnings.

9.3 Process Data

After obtaining the requisite financial statement and other information, the analyst processes these data using appropriate analytical tools. For example, processing the data may involve computing ratios or growth rates; preparing common-size financial statements; creating charts; performing statistical analyses, such as regressions or Monte Carlo simulations; performing equity valuation; performing sensitivity analyses; or using any other analytical tools or combination of tools that are available and appropriate for the task. A comprehensive financial analysis at this stage would include the following:

- Reading and evaluating financial statements for each company being analyzed. This includes reading the notes and understanding what accounting standards have been used (for example, IFRS or US GAAP), what accounting choices have been made (for example, when to report revenue on the income statement), and what operating decisions have been made that affect reported financial statements (for example, leasing versus purchasing equipment).
- Making any needed adjustments to the financial statements to facilitate comparison when the unadjusted statements of the subject companies reflect differences in accounting standards, accounting choices, or operating decisions. Note that commonly used databases do not always make such analyst adjustments.
- Preparing or collecting common-size financial statement data (which scale data to directly reflect percentages [e.g., of sales] or changes [e.g., from the prior year]) and financial ratios (which are measures of various aspects of corporate performance based on financial statement elements). On the basis of common-size financial statements and financial ratios, analysts can evaluate a company's relative profitability, liquidity, leverage, efficiency, and valuation in relation to past results and/or peers' results.

9.4 Analyze/Interpret the Processed Data

Once the data have been processed, the next step—critical to any analysis—is to interpret the output. The answer to a specific question is seldom the numerical answer alone. Rather, the answer relies on the analyst's interpretation of the output, and the use of this interpreted output to support a conclusion or recommendation. The answers to the specific analytical questions may themselves achieve the underlying purpose of the analysis, but usually, a conclusion or recommendation is required. For example, an equity analysis may require a buy, hold, or sell recommendation or a conclusion about the value of a share of stock. In support of the decision, the analysis would cite such information as target value, relative performance, expected future performance given a company's strategic position, quality of management, and whatever other information influenced the decision.

9.5 Develop and Communicate Conclusions/Recommendations

Communicating the conclusion or recommendation in an appropriate format is the next step. The appropriate format will vary by analytical task, by institution, and/or by audience. For example, an equity analyst's report would typically include the following components:

- summary and investment conclusion;
- earnings projections;

- valuation;
- business summary;
- risk, industry, and competitive analysis;
- historical performance; and
- forecasts.

The contents of reports may also be specified by regulatory agencies or professional standards. For example, the CFA Institute *Standards of Practice Handbook* (*Handbook*) dictates standards that must be followed in communicating recommendations. According to the *Handbook*:

> Standard V(B) states that members and candidates should communicate in a recommendation the factors that were instrumental in making the investment recommendation. A critical part of this requirement is to distinguish clearly between opinions and facts. In preparing a research report, the member or candidate must present the basic characteristics of the security(ies) being analyzed, which will allow the reader to evaluate the report and incorporate information the reader deems relevant to his or her investment decision making process.[9]

The *Handbook* requires that limitations to the analysis and any risks inherent to the investment be disclosed. Furthermore, it requires that any report include elements important to the analysis and conclusions so that readers can evaluate the conclusions themselves.

9.6 Follow-Up

The process does not end with the report. If an equity investment is made or a credit rating is assigned, periodic review is required to determine if the original conclusions and recommendations are still valid. In the case of a rejected investment, follow-up may not be necessary but may be useful in determining whether the analytical process is adequate or should be refined (for example, if a rejected investment turns out to be successful in the market, perhaps the rejection was due to inadequate analysis). Follow-up may involve repeating all the previous steps in the process on a periodic basis.

SUMMARY

The information presented in financial and other reports, including the financial statements, notes, and management's commentary, help the financial analyst to assess a company's performance and financial position. An analyst may be called on to perform a financial analysis for a variety of reasons, including the valuation of equity securities, the assessment of credit risk, the performance of due diligence on an acquisition, and the evaluation of a subsidiary's performance relative to other business units. Major considerations in both equity analysis and credit analysis are evaluating a company's financial position, its ability to generate profits and cash flow, and its potential to generate future growth in profits and cash flow.

---

[9] *Standards of Practice Handbook* (2014, p. 169).

This reading has presented an overview of financial statement analysis. Among the major points covered are the following:

- The primary purpose of financial reports is to provide information and data about a company's financial position and performance, including profitability and cash flows. The information presented in the reports —including the financial statements and notes and management's commentary or management's discussion and analysis—allows the financial analyst to assess a company's financial position and performance and trends in that performance.
- The primary financial statements are the statement of financial position (i.e., the balance sheet), the statement of comprehensive income (or two statements consisting of an income statement and a statement of comprehensive income), the statement of changes in equity, and the statement of cash flows.
- The balance sheet discloses what resources a company controls (assets) and what it owes (liabilities) at a specific point in time. Owners' equity represents the net assets of the company; it is the owners' residual interest in, or residual claim on, the company's assets after deducting its liabilities. The relationship among the three parts of the balance sheet (assets, liabilities, and owners' equity) may be shown in equation form as follows: Assets = Liabilities + Owners' equity.
- The income statement presents information on the financial results of a company's business activities over a period of time. The income statement communicates how much revenue and other income the company generated during a period and what expenses, including losses, it incurred in connection with generating that revenue and other income. The basic equation underlying the income statement is Revenue + Other income – Expenses = Net income.
- The statement of comprehensive income includes all items that change owners' equity except transactions with owners. Some of these items are included as part of net income, and some are reported as other comprehensive income (OCI).
- The statement of changes in equity provides information about increases or decreases in the various components of owners' equity.
- Although the income statement and balance sheet provide measures of a company's success, cash and cash flow are also vital to a company's long-term success. Disclosing the sources and uses of cash helps creditors, investors, and other statement users evaluate the company's liquidity, solvency, and financial flexibility.
- The notes (also referred to as footnotes) that accompany the financial statements are an integral part of those statements and provide information that is essential to understanding the statements. Analysts should evaluate note disclosures regarding the use of alternative accounting methods, estimates, and assumptions.
- In addition to the financial statements, a company provides other sources of information that are useful to the financial analyst. As part of his or her analysis, the financial analyst should read and assess this additional information, particularly that presented in the management commentary (also called management report[ing], operating and financial review, and management's discussion and analysis [MD&A]).
- A publicly traded company must have an independent audit performed on its annual financial statements. The auditor's report expresses an opinion on the financial statements and provides some assurance about whether the financial

statements fairly present a company's financial position, performance, and cash flows. In addition, for US publicly traded companies, auditors must also express an opinion on the company's internal control systems.

- Information on the economy, industry, and peer companies is useful in putting the company's financial performance and position in perspective and in assessing the company's future. In most cases, information from sources apart from the company are crucial to an analyst's effectiveness.
- The financial statement analysis framework provides steps that can be followed in any financial statement analysis project. These steps are:
  - articulate the purpose and context of the analysis;
  - collect input data;
  - process data;
  - analyze/interpret the processed data;
  - develop and communicate conclusions and recommendations; and
  - follow up.

REFERENCES

Benninga, Simon Z., and Oded H. Sarig. 1997. *Corporate Finance: A Valuation Approach*. New York: McGraw-Hill Publishing.

International Auditing and Assurance Standards Board (IAASB). 2018. *Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements*, Standard 200, available at www.ifac.org/IAASB.

van Greuning, Hennie, and Sonja Brajovic Bratanovic. 2003. *Analyzing and Managing Banking Risk: A Framework for Assessing Corporate Governance and Financial Risk*. Washington, DC: World Bank.

PRACTICE PROBLEMS

1 Providing information about the performance and financial position of companies so that users can make economic decisions *best* describes the role of:

A auditing.

B financial reporting.

C financial statement analysis.

2 Which of the following *best* describes the role of financial statement analysis?

A To provide information about a company's performance

B To provide information about a company's changes in financial position

C To form expectations about a company's future performance and financial position

3 The role of financial statement analysis is *best* described as:

A providing information useful for making investment decisions.

B evaluating a company for the purpose of making economic decisions.

C using financial reports prepared by analysts to make economic decisions.

4 A company's financial position would *best* be evaluated using the:

A balance sheet.

B income statement.

C statement of cash flows.

5 A company's profitability for a period would *best* be evaluated using the:

A balance sheet.

B income statement.

C statement of cash flows.

6 The financial statement that presents a shareholder's residual claim on assets is the:

A balance sheet.

B income statement.

C cash flow statement.

7 A company's profitability over a period of time is *best* evaluated using the:

A balance sheet.

B income statement.

C cash flow statement.

8 The income statement is *best* used to evaluate a company's:

A financial position.

B sources of cash flow.

C financial results from business activities.

9 Accounting policies, methods, and estimates used in preparing financial statements are *most likely* to be found in the:

A auditor's report.

B management commentary.

C notes to the financial statements.

© 2017 CFA Institute. All rights reserved.

10 Information about management and director compensation are *least likely* to be found in the:

A auditor's report.

B proxy statement.

C notes to the financial statements.

11 Information about a company's objectives, strategies, and significant risks are *most likely* to be found in the:

A auditor's report.

B management commentary.

C notes to the financial statements.

12 Which of the following *best* describes why the notes that accompany the financial statements are required? The notes:

A permit flexibility in statement preparation.

B standardize financial reporting across companies.

C provide information necessary to understand the financial statements.

13 What type of audit opinion is preferred when analyzing financial statements?

A Qualified.

B Adverse.

C Unqualified.

14 An auditor determines that a company's financial statements are prepared in accordance with applicable accounting standards except with respect to inventory reporting. This exception is *most likely* to result in an audit opinion that is:

A adverse.

B qualified.

C unqualified.

15 An independent audit report is *most likely* to provide:

A absolute assurance about the accuracy of the financial statements.

B reasonable assurance that the financial statements are fairly presented.

C a qualified opinion with respect to the transparency of the financial statements.

16 Interim financial reports released by a company are *most likely* to be:

A monthly.

B unaudited.

C unqualified.

17 Which of the following sources of information used by analysts is found outside a company's annual report?

A Auditor's report

B Peer company analysis

C Management's discussion and analysis

18 Ratios are an input into which step in the financial statement analysis framework?

A Process data.

B Collect input data.

C Analyze/interpret the processed data.

19 Which phase in the financial statement analysis framework is *most likely* to involve producing updated reports and recommendations?

A Follow-up

B Analyze/interpret the processed data

C Develop and communicate conclusions and recommendations

SOLUTIONS

1 B is correct. This is the role of financial reporting. The role of financial statement analysis is to evaluate the financial reports.

2 C is correct. In general, analysts seek to examine the past and current performance and financial position of a company in order to form expectations about its future performance and financial position.

3 B is correct. The primary role of financial statement analysis is to use financial reports prepared by companies to evaluate their past, current, and potential performance and financial position for the purpose of making investment, credit, and other economic decisions.

4 A is correct. The balance sheet portrays the company's financial position on a specified date. The income statement and statement of cash flows present different aspects of performance during the period.

5 B is correct. Profitability is the performance aspect measured by the income statement. The balance sheet portrays the financial position. The statement of cash flows presents a different aspect of performance.

6 A is correct. Owners' equity is the owners' residual interest in (i.e., residual claim on) the company's assets after deducting its liabilities, which is information presented on the balance sheet.

7 B is correct. A company's profitability is best evaluated using the income statement. The income statement presents information on the financial results of a company's business activities over a period of time by communicating how much revenue was generated and the expenses incurred to generate that revenue.

8 C is correct. A company's revenues and expenses are presented on the income statement, which is used to evaluate a company's financial results (or profitability) from business activities over a period of time. A company's financial position is best evaluated by using the balance sheet. A company's sources of cash flow are best evaluated using the cash flow statement.

9 C is correct. The notes disclose choices in accounting policies, methods, and estimates.

10 A is correct. Information about management and director compensation is not found in the auditor's report. Disclosure of management compensation is required in the proxy statement, and some aspects of management compensation are disclosed in the notes to the financial statements.

11 B is correct. These are components of management commentary.

12 C is correct. The notes provide information that is essential to understanding the information provided in the primary statements.

13 C is correct. An unqualified opinion is a "clean" opinion and indicates that the financial statements present the company's performance and financial position fairly in accordance with a specified set of accounting standards.

14 B is correct. A qualified audit opinion is one in which there is some scope limitation or exception to accounting standards. Exceptions are described in the audit report with additional explanatory paragraphs so that the analyst can determine the importance of the exception.

15 B is correct. The independent audit report provides reasonable assurance that the financial statements are fairly presented, meaning that there is a high probability that the audited financial statements are free from material error, fraud, or illegal acts that have a direct effect on the financial statements.

16 B is correct. Interim reports are typically provided semiannually or quarterly and present the four basic financial statements and condensed notes. They are not audited. Unqualified is a type of audit opinion

17 B is correct. When performing financial statement analysis, analysts should review all company sources of information as well as information from external sources regarding the economy, the industry, the company, and peer (comparable) companies.

18 C is correct. Ratios are an output of the process data step but are an input into the analyze/interpret data step.

19 A is correct. The follow-up phase involves gathering information and repeating the analysis to determine whether it is necessary to update reports and recommendations.

READING

16

Financial Reporting Standards

by Elaine Henry, PhD, CFA, J. Hennie van Greuning, DCom, CFA, and Thomas R. Robinson, PhD, CFA, CAIA

Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). J. Hennie van Greuning, DCom, CFA, is at BIBD (Brunei). Thomas R. Robinson, PhD, CFA, CAIA, Robinson Global Investment Management LLC, (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation;
☐	**b.** describe the roles of financial reporting standard-setting bodies and regulatory authorities in establishing and enforcing reporting standards;
☐	**c.** describe the International Accounting Standards Board's conceptual framework, including qualitative characteristics of financial reports, constraints on financial reports, and required reporting elements;
☐	**d.** describe general requirements for financial statements under International Financial Reporting Standards (IFRS);
☐	**e.** describe implications for financial analysis of alternative financial reporting systems and the importance of monitoring developments in financial reporting standards.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION

Financial reporting standards provide principles for preparing financial reports and determine the types and amounts of information that must be provided to users of financial statements, including investors and creditors, so that they may make informed decisions. This reading focuses on the context within which these standards are created. An understanding of the underlying framework of financial reporting standards, which is broader than knowledge of specific accounting rules, will allow an analyst to assess the valuation implications of financial statement elements and transactions—including transactions, such as those that represent new developments, which are not specifically addressed by the standards.

Section 2 of this reading discusses the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation. Sections 3 and 4 describe the roles of financial reporting standard-setting bodies and regulatory authorities and several of the financial reporting standard-setting bodies and regulatory authorities. Sections 5–7 describe the International Financial Reporting Standards (IFRS) framework[1] and general requirements for financial statements. Section 8 compares IFRS and alternative reporting systems, and Section 9 discusses the importance of monitoring developments in financial reporting standards. A summary of the key points concludes the reading.

2 THE OBJECTIVE OF FINANCIAL REPORTING

a describe the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation;

The financial reports of a company include financial statements and other supplemental disclosures necessary to assess a company's financial position and periodic financial performance. Financial reporting is based on a simple premise. The International Accounting Standards Board (IASB), which sets financial reporting standards that have been adopted in many countries, said in its *Conceptual Framework for Financial Reporting* (*Conceptual Framework*) that the objective of financial reporting is to provide financial information that is useful to users in making decisions about providing resources to the reporting entity, where those decisions relate to equity and debt instruments, or loans or other forms of credit, and in influencing management's actions that affect the use of the entity's economic resources.[2]

A fully articulated framework is an essential first step to guide the development of a set of standards. Previously, financial reporting standards were primarily developed independently by each country's standard-setting body. This independent standard setting process created a wide range of standards, some of which were quite comprehensive and complex (rules-based standards), and others that were more general (principles-based standards). The globalization of capital flows and various accounting scandals increased awareness of the need for more uniform, high quality global financial reporting standards and provided the impetus for stronger coordination among the major standard-setting bodies. Such coordination is also a natural outgrowth of the increased globalization of capital markets.

---

[1] The body of standards issued by the International Accounting Standards Board (IASB) is referred to as International Financial Reporting Standards.

[2] In March 2018, the IASB updated the *Conceptual Framework for Financial Reporting* originally adopted in 2010.

Developing financial reporting standards is complicated because the underlying economic reality is complicated. The financial transactions and financial position that companies aim to present in their financial reports are also complex. Furthermore, uncertainty about various aspects of transactions often results in the need for accruals and estimates, both of which necessitate judgment. Judgment varies from one preparer to the next. Accordingly, standards are needed to achieve some amount of consistency in these judgments. Even with such standards, there usually will be no single correct answer to the question of how to reflect economic reality in financial reports. Nevertheless, financial reporting standards try to limit the range of acceptable answers to increase consistency in financial reports.

The IASB and the US-based Financial Accounting Standards Board (FASB) have developed similar financial reporting frameworks which specify the overall objective and qualities of information to be provided. Financial reports are intended to provide information to many users, including investors, creditors, employees, customers, and others. As a result, financial reports are *not* designed solely with asset valuation in mind. However, financial reports provide important inputs into the process of valuing a company or the securities a company issues. Understanding the financial reporting framework—including how and when judgments and estimates can affect the numbers reported—enables an analyst to evaluate the information reported and to use the information appropriately when assessing a company's financial performance. Clearly, such an understanding is also important in assessing the financial impact of business decisions by, and in making comparisons across, entities.

ACCOUNTING STANDARDS BOARDS

3

a describe the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation;

b describe the roles of financial reporting standard-setting bodies and regulatory authorities in establishing and enforcing reporting standards;

A distinction must be made between standard-setting bodies and regulatory authorities. Standard-setting bodies, such as the IASB and FASB, are typically private sector, self-regulated organizations with board members who are experienced accountants, auditors, users of financial statements, and academics. The requirement to prepare financial reports in accordance with specified accounting standards is the responsibility of regulatory authorities. Regulatory authorities, such as the Accounting and Corporate Regulatory Authority in Singapore, the Securities and Exchange Commission (SEC) in the United States, and the Securities and Exchange Commission of Brazil have the legal authority to enforce financial reporting requirements and exert other controls over entities that participate in the capital markets within their jurisdiction.

In other words, *generally*, standard-setting bodies set the standards and regulatory authorities recognise and enforce the standards. Without the recognition of the standards by the regulatory authorities, the private sector standard-setting bodies would have no authority. Note, however, that regulators often retain the legal authority to establish financial reporting standards in their jurisdiction and can overrule the private sector standard-setting bodies.

This section provides a brief overview of the IASB and the FASB. The overview is followed by descriptions of the International Organization of Securities Commissions, the US Securities and Exchange Commission, and capital markets regulation in the European Union. The topics covered in these overviews were chosen to serve as examples of standard-setting boards, securities commissions, and capital market regulation. After reading these descriptions, the reader should be able to describe the functioning and roles of standard-setting bodies and regulatory authorities in more detail than is given in the introduction to this section.

3.1 Accounting Standards Boards

Accounting standards boards exist in virtually every national market. These boards are typically independent, private, not-for-profit organizations. There are certain attributes that are typically common to these standard setters—the IASB and the FASB are discussed in this section as primary examples.

3.1.1 *International Accounting Standards Board*

The IASB is the independent standard-setting body of the IFRS Foundation, an independent, not-for-profit private sector organization. The Trustees of the IFRS Foundation reflect a diversity of geographical and professional backgrounds. The Trustees appoint the members of the IASB and are accountable to a monitoring board composed of public authorities that include representatives from the European Commission, IOSCO, the Japan Financial Services Agency, and the US SEC, with the chairman of the Basel Committee on Banking Supervision as an observer.

The Trustees of the IFRS Foundation make a commitment to act in the public interest. The principle objectives of the IFRS Foundation are to develop and promote the use and adoption of a single set of high quality financial standards; to ensure the standards result in transparent, comparable, and decision-useful information while taking into account the needs of a range of sizes and types of entities in diverse economic settings; and to promote the convergence of national accounting standards and IFRS. The Trustees are responsible for ensuring that the IASB is and is perceived as independent.

The members of the IASB are appointed by the Trustees on the basis of professional competence and practical experience and reflect a diversity of geographical and professional backgrounds. The members deliberate, develop, and issue international financial reporting standards,[3] assisted by advice on the standards, and their application, from advisory bodies whose members represent a wide range of organizations and individuals that are affected by and interested in international financial reporting.

The IASB has a basic process that it goes through when deliberating, developing, and issuing international financial reporting standards. A simplified version of the typical process is as follows. An issue is identified as a priority for consideration and placed on the IASB's agenda. After considering an issue, which may include soliciting advice from others including national standard-setters, the IASB may publish an exposure draft for public comment. After reviewing the input of others, the IASB may issue a new or revised financial reporting standard. These standards are authoritative to the extent that they are recognised and adopted by regulatory authorities.

---

[3] Although the name of the IASB incorporates "Accounting Standards" and early standards were titled International Accounting Standards (IAS), the term "International Financial Reporting Standards" (IFRS) is being used for new standards. The use of the words "financial reporting" recognizes the importance of disclosures outside of the core financial statements, such as management discussion of the business, risks, and future plans.

3.1.2 *Financial Accounting Standards Board*

The FASB and its predecessor organizations have been issuing financial reporting standards in the United States since the 1930s. The FASB operates within a structure similar to that of the IASB. The Financial Accounting Foundation oversees, administers, and finances the organization. The Foundation ensures the independence of the standard-setting process and appoints members to the FASB and related advisory entities.

The FASB issues new and revised standards to improve standards of financial reporting so that decision-useful information is provided to users of financial reports. This is done through a thorough and independent process that seeks input from stakeholders and is overseen by the Financial Accounting Foundation. The steps in the process are similar to those described for the IASB. The outputs of the standard-setting process are contained in the FASB Accounting Standards Codification™ (Codification).[4] The Codification, organized by topic, is the source of authoritative US generally accepted accounting principles to be applied to non-governmental entities.

US GAAP, as established by the FASB, is officially recognized as authoritative by the SEC, however the SEC retains the authority to establish standards. Although it has rarely overruled the FASB, the SEC does issue authoritative financial reporting guidance including Staff Accounting Bulletins. These bulletins reflect the SEC's views regarding accounting-related disclosure practices and can be found on the SEC website. Certain portions—but not all portions—of the SEC regulations, releases, interpretations, and guidance are included for reference in the FASB Codification.

REGULATORY AUTHORITIES

4

a describe the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation;

b describe the roles of financial reporting standard-setting bodies and regulatory authorities in establishing and enforcing reporting standards;

The requirement to prepare financial reports in accordance with specified accounting standards is the responsibility of regulatory authorities. Regulatory authorities are governmental entities that have the legal authority to enforce financial reporting requirements and exert other controls over entities that participate in the capital markets within their jurisdiction. Regulatory authorities may require that financial reports be prepared in accordance with one specific set of accounting standards or may specify acceptable accounting standards. For example in Switzerland, Swiss-based companies listed on the main board of the Swiss Exchange have to prepare their financial statements in accordance with either IFRS or US GAAP if they are multinational.[5] Other registrants in Switzerland could use IFRS, US GAAP, or Swiss GAAP.

---

[4] The Codification combines literature issued by various standard setters, including the FASB, the Emerging Issues Task Force (EITF), the Derivative Implementation Group (DIG), and the American Institute of Certified Public Accountants (AICPA).

[5] https://www.iasplus.com/en/jurisdictions/europe/switzerland.

4.1 International Organization of Securities Commissions

While technically not a regulatory authority, the International Organization of Securities Commissions (IOSCO) regulate a significant portion of the world's financial capital markets. This organization has established objectives and principles to guide securities and capital market regulation.

IOSCO was formed in 1983 and consists of ordinary members, associate members, and affiliate members. Ordinary members are the securities commission or similar governmental regulatory authority with primary responsibility for securities regulation in the member country.[6] The members regulate more than 95 percent of the world's financial capital markets in more than 115 jurisdictions, and securities regulators in emerging markets account for 75% of its ordinary membership.

IOSCO's comprehensive set of *Objectives and Principles of Securities Regulation* is updated as required and is recognized as an international benchmark for all markets. The principles of securities regulation are based upon three core objectives:[7]

- protecting investors;
- ensuring that markets are fair, efficient, and transparent; and
- reducing systemic risk.

IOSCO's principles are grouped into ten categories, including principles for regulators, for enforcement, for auditing, and for issuers, among others. Within the category "Principles for Issuers," two principles relate directly to financial reporting:

- There should be full, accurate, and timely disclosure of financial results, risk, and other information which is material to investors' decisions.
- Accounting standards used by issuers to prepare financial statements should be of a high and internationally acceptable quality.

Historically, regulation and related financial reporting standards were developed within individual countries and were often based on the cultural, economic, and political norms of each country. As financial markets have become more global, it has become desirable to establish comparable financial reporting standards internationally. Ultimately, laws and regulations are established by individual jurisdictions, so this also requires cooperation among regulators. Another IOSCO principle deals with the use of self-regulatory organizations (SROs), which exercise some direct oversight for their areas of competence and should be subject to the oversight of the relevant regulator and observe fairness and confidentiality[8].

To ensure consistent application of international financial standards (such as the Basel Committee on Banking Supervision's standards and IFRS), it is important to have uniform regulation and enforcement across national boundaries. IOSCO assists in attaining this goal of uniform regulation as well as cross-border co-operation in combating violations of securities and derivatives laws.

4.2 The Securities and Exchange Commission (US)

The US SEC has primary responsibility for securities and capital markets regulation in the United States and is an ordinary member of IOSCO. Any company issuing securities within the United States, or otherwise involved in US capital markets, is subject to the rules and regulations of the SEC. The SEC, one of the oldest and most developed regulatory authorities, originated as a result of reform efforts made after the stock market crash of 1929, sometimes referred to as simply the "Great Crash."

---

[6] The names of the primary securities regulator vary from country to country. For example: China Securities Regulatory Commission, Egyptian Financial Supervisory Authority, Securities and Exchange Board of India, Kingdom of Saudi Arabia Capital Market Authority, and Banco Central del Uruguay.

[7] *Objectives and Principles of Securities Regulation*, IOSCO, May 2017.

[8] *Objectives and Principles of Securities Regulation*, IOSCO, May 2017.

A number of laws affect reporting companies, broker/dealers, and other market participants. From a financial reporting and analysis perspective, the most significant pieces of legislation are the Securities Acts of 1933 and 1934 and the Sarbanes–Oxley Act of 2002.

- **Securities Act of 1933** (The 1933 Act): This act specifies the financial and other significant information that investors must receive when securities are sold, prohibits misrepresentations, and requires initial registration of all public issuances of securities.
- **Securities Exchange Act of 1934** (The 1934 Act): This act created the SEC, gave the SEC authority over all aspects of the securities industry, and empowered the SEC to require periodic reporting by companies with publicly traded securities.
- **Sarbanes–Oxley Act of 2002**: This act created the Public Company Accounting Oversight Board (PCAOB) to oversee auditors. The SEC is responsible for carrying out the requirements of the act and overseeing the PCAOB. The act addresses auditor independence (it prohibits auditors from providing certain non-audit services to the companies they audit); strengthens corporate responsibility for financial reports (it requires executive management to certify that the company's financial reports fairly present the company's condition); and requires management to report on the effectiveness of the company's internal control over financial reporting (including obtaining external auditor confirmation of the effectiveness of internal control).

Companies comply with these acts principally through the completion and submission (i.e., filing) of standardized forms issued by the SEC. There are more than 50 different types of SEC forms that are used to satisfy reporting requirements; the discussion herein will be limited to those forms most relevant for financial analysts.

Most of the SEC filings are required to be made electronically, so filings that an analyst would be interested in can be retrieved online from one of many websites, including the SEC's own website. Some filings are required on the initial offering of securities, whereas others are required on a periodic basis thereafter. The following are some of the more common information sources used by analysts.

- **Securities Offerings Registration Statement**: The 1933 Act requires companies offering securities to file a registration statement. New issuers as well as previously registered companies that are issuing new securities are required to file these statements. Required information and the precise form vary depending upon the size and nature of the offering. Typically, required information includes: 1) disclosures about the securities being offered for sale, 2) the relationship of these new securities to the issuer's other capital securities, 3) the information typically provided in the annual filings, 4) recent audited financial statements, and 5) risk factors involved in the business.
- **Forms 10-K, 20-F, and 40-F**: These are forms that companies are required to file *annually*. Form 10-K is for US registrants, Form 40-F is for certain Canadian registrants, and Form 20-F is for all other non-US registrants. These forms require a comprehensive overview, including information concerning a company's business, financial disclosures, legal proceedings, and information related to management. The financial disclosures include a historical summary of financial data (usually 10 years), management's discussion and analysis (MD&A) of the company's financial condition and results of operations, and audited financial statements.

- **Annual Report**: In addition to the SEC's annual filings (e.g., Form 10-K), most companies prepare an annual report to shareholders. This is not a requirement of the SEC. The annual report is usually viewed as one of the most significant opportunities for a company to present itself to shareholders and other external parties; accordingly, it is often a highly polished marketing document with photographs, an opening letter from the chief executive officer, financial data, market segment information, research and development activities, and future corporate goals. In contrast, the Form 10-K is a more legal type of document with minimal marketing emphasis. Although the perspectives vary, there is considerable overlap between a company's annual report and its Form 10-K. Some companies elect to prepare just the Form 10-K or a document that integrates both the 10-K and annual report.
- **Proxy Statement/Form DEF-14A**: The SEC requires that shareholders of a company receive a proxy statement prior to a shareholder meeting. A proxy is an authorization from the shareholder giving another party the right to cast its vote. Shareholder meetings are held at least once a year, but any special meetings also require a proxy statement. Proxies, especially annual meeting proxies, contain information that is often useful to financial analysts. Such information typically includes proposals that require a shareholder vote, details of security ownership by management and principal owners, biographical information on directors, and disclosure of executive compensation. Proxy statement information is filed with the SEC as Form DEF-14A.
- **Forms 10-Q and 6-K**: These are forms that companies are required to submit for interim periods (quarterly for US companies on Form 10-Q, semiannually for many non-US companies on Form 6-K). The filing requires certain financial information, including unaudited financial statements and a MD&A for the interim period covered by the report. Additionally, if certain types of non-recurring events—such as the adoption of a significant accounting policy, commencement of significant litigation, or a material limitation on the rights of any holders of any class of registered securities—take place during the period covered by the report, these events must be included in the Form 10-Q report. Companies may provide the 10-Q report to shareholders or may prepare a separate, abbreviated, quarterly report to shareholders.

EXAMPLE 1

Initial Registration Statement

In 2004, Google filed a registration statement with the US SEC to register its initial public offering of securities (Class A Common Stock). In addition to a large amount of financial and business information, the registration statement provided a 20-page discussion Google's business and industry.

Which of the following is *most likely* to have been included in Google's registration statement?

A Interim unaudited financial statements.

B Assessment of risk factors involved in the business.

C Projected cash flows and earnings for the business.

Solution:

B is correct. Information provided by companies in registration statements typically includes disclosures about the securities being offered for sale; the relationship of these new securities to the issuer's other capital securities; the information typically provided in the annual filings; recent audited financial statements; and risk factors involved in the business.. Companies provide information useful in developing projected cash flows and earnings but do not typically include these in the registration statement, nor do they provide unaudited interim statements in the initial registration statement.

A company or its officers make other SEC filings—either periodically, or, if significant events or transactions have occurred, in between the periodic reports noted above. By their nature, these forms sometimes contain the most interesting and timely information and may have significant valuation implications.

- **Form 8-K:** In addition to filing annual and interim reports, SEC registrants must report material corporate events on a more current basis. Form 8-K (6-K for non-US registrants) is the "current report" companies must file with the SEC to announce such major events as acquisitions or disposals of corporate assets, changes in securities and trading markets, matters related to accountants and financial statements, corporate governance and management changes, and Regulation FD disclosures.[9]
- **Forms 3, 4, 5 and 144**: Forms 3, 4 and 5 are required to report beneficial ownership of securities. These filings are required for any director or officer of a registered company as well as beneficial owners of greater than 10 percent of a class of registered equity securities. Form 3 is the initial statement, Form 4 reports changes, and Form 5 is the annual report. Form 144 is notice of the proposed sale of restricted securities or securities held by an affiliate of the issuer. These forms can be used to examine purchases and sales of securities by officers, directors, and other affiliates of the company, who collectively are regarded as corporate insiders.
- **Form 11-K**: This is the annual report of employee stock purchase, savings, and similar plans. It might be of interest to analysts for companies with significant employee benefit plans because it contains more information than that disclosed in the company's financial statements.

In jurisdictions other than the United States, similar legislation exists for the purpose of regulating securities and capital markets. Regulatory authorities are responsible for enforcing regulation, and securities regulation is intended to be consistent with the IOSCO objectives described in the previous section. Within each jurisdiction, regulators will either establish or, more typically, recognize and adopt a specified set or sets of accounting standards. The regulators will also establish reporting and filing requirements. IOSCO members have agreed to cooperate in the development, implementation, and enforcement of internationally recognised and consistent standards of regulation.

4.3 Capital Markets Regulation in Europe

Each individual member state of the European Union (EU) regulates capital markets in its jurisdiction. There are, however, certain regulations that have been adopted at the EU level. Importantly the EU agreed that from 2005 consolidated accounts of EU listed companies would use International Financial Reporting Standards. The endorsement process by which newly issued IFRS are adopted by the EU reflects the balance between the individual member state's autonomy and the need for cooperation and convergence. When the IASB issues a new standard, the European Financial Reporting Advisory Group advises the European Commission on the standard, and the Standards Advice Review Group provides the Commission an opinion about that advice. Based on the input from these two entities, the Commission prepares a draft endorsement regulation. The Accounting Regulatory Committee votes on the proposal; and if the vote is favorable, the proposal proceeds to the European Parliament and the Council of the European Union for approval.[10]

---

[9] Regulation Fair Disclosure (FD) provides that when an issuer discloses material non-public information to certain individuals or entities—generally, securities market professionals such as stock analysts or holders of the issuer's securities who may trade on the basis of the information—the issuer must make public disclosure of that information. In this way, the rule aims to promote full and fair disclosure.

Two bodies related to securities regulation established by the European Commission are the European Securities Committee (ESC) and the European Securities and Market Authority (ESMA). The ESC consists of high-level representatives of member states and advises the European Commission on securities policy issues. ESMA is an EU cross-border supervisor established to co-ordinate supervision of the EU market. As noted earlier, regulation still rests with the individual member states and, therefore, requirements for registering shares and filing periodic financial reports vary from country to country. ESMA is one of three European supervisory authorities; the two others supervise the banking and insurance industries.

5

THE INTERNATIONAL FINANCIAL REPORTING STANDARDS FRAMEWORK

a describe the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation;

c describe the International Accounting Standards Board's conceptual framework, including qualitative characteristics of financial reports, constraints on financial reports, and required reporting elements;

As previously discussed, the IASB *Conceptual Framework for Financial Reporting* sets forth the concepts that underlie the preparation and presentation of financial statements for external users. The framework is designed to: assist standard setters in developing and reviewing standards; assist preparers of financial statements in applying standards and in dealing with issues not specifically covered by standard; assist auditors in forming an opinion on financial statements; and assist users in interpreting financial statement information. The objective of financial reporting is the provision of financial information that is useful to current and potential providers of resources in making decisions and all other aspects of the framework flow from that central objective.

The providers of resources are considered to be the primary users of financial reports and include investors, lenders, and other creditors. The purpose of providing the financial information is to be useful in making decisions about providing resources. Other users may find the financial information useful for making economic decisions. The types of economic decisions differ by users, so the specific information needed differs as well. However, although these users may have unique information needs, some information needs are common across all users. Information is needed about the company's financial position: its resources and its financial obligations. Information is needed about a company's financial performance; this information explains how and why the company's financial position changed in the past and can be useful in evaluating potential changes in the future. The third common information need is the need for information about a company's cash. How did the company obtain cash (by selling its products and services, borrowing, other)? How did the company use cash (by paying expenses, investing in new equipment, paying dividends, other)?

---

[10] Source: European Commission. http://ec.europa.eu/internal_market/accounting/legal_framework/ias_regulation_en.htm.

Information that is helpful to users in assessing future net cash inflows to the entity includes information about the economic resources of (assets) and claims against (liabilities and equity) the entity, and about how well the management and governing board have utilized the resources of the entity. Users need to consider information from other sources as well in making their decisions. Although financial reports do not show the value of an entity they are useful in estimating the value of an entity.

5.1 Qualitative Characteristics of Financial Reports

Flowing from the central objective of providing information that is *useful* to providers of resources, the *Conceptual Framework* identifies two fundamental qualitative characteristics that make financial information useful: relevance and faithful representation.[11] The concept of materiality is discussed within the context of relevance.

1. *Relevance*: Information is relevant if it would potentially affect or make a difference in users' decisions. The information can have predictive value (useful in making forecasts), confirmatory value (useful to evaluate past decisions or forecasts), or both. In other words, relevant information helps users of financial information to evaluate past, present, and future events, or to confirm or correct their past evaluations in a decision-making context. *Materiality:* Information is considered to be material if omission or misstatement of the information could influence users' decisions. Materiality is a function of the nature and/or magnitude of the information.
2. *Faithful representation*: Information that faithfully represents an economic phenomenon that it purports to represent is ideally complete, neutral, and free from error. Complete means that all information necessary to understand the phenomenon is depicted. Neutral means that information is selected and presented without bias. In other words, the information is not presented in such a manner as to bias the users' decisions. Free from error means that there are no errors of commission or omission in the description of the economic phenomenon, and that an appropriate process to arrive at the reported information was selected and was adhered to without error. Faithful representation maximizes the qualities of complete, neutral, and free from error to the extent possible.

Relevance and faithful representation are the fundamental, most critical characteristics of useful financial information. In addition the *Conceptual Framework* identifies four enhancing qualitative characteristics: comparability, verifiability, timeliness, and understandability.

1. *Comparability*: Comparability allows users "to identify and understand similarities and differences of items." Information presented in a consistent manner over time and across entities enables users to make comparisons more easily than information with variations in how similar economic phenomena are represented.
2. *Verifiability*: Verifiability means that different knowledgeable and independent observers would agree that the information presented faithfully represents the economic phenomena it purports to represent.

---

[11] *Conceptual Framework for Financial Reporting.*

3 *Timeliness*: Timely information is available to decision makers prior to their making a decision.

4 *Understandability*: Clear and concise presentation of information enhances understandability. Information should be prepared for and be understandable by users who have a reasonable knowledge of business and economic activities, and who are willing to study the information with diligence. Information that is useful should not be excluded simply because it is difficult to understand and it may be necessary for users to seek assistance to understand information about complex economic phenomena.

Financial information exhibiting these qualitative characteristics—fundamental and enhancing—should be useful for making economic decisions.

5.2 Constraints on Financial Reports

Although it would be ideal for financial statements to exhibit all of these qualitative characteristics and thus achieve maximum usefulness, it may be necessary to make tradeoffs across the enhancing characteristics. The application of the enhancing characteristics follows no set order of priority and each enhancing characteristic may take priority over the others. The aim is an appropriate balance among the enhancing characteristics.

A pervasive constraint on useful financial reporting is the cost of providing and using this information. Optimally, benefits derived from information should exceed the costs of providing and using it. Again, the aim is a balance between costs and benefits.

A limitation of financial reporting involves information that is not included. Financial statements, by necessity, omit information that is non-quantifiable. For example, the creativity, innovation, and competence of a company's work force are not directly captured in the financial statements. Similarly, customer loyalty, a positive corporate culture, environmental responsibility, and many other aspects about a company may not be directly reflected in the financial statements. Of course, to the extent that these items result in superior financial performance, a company's financial reports will reflect the results.

EXAMPLE 2

Balancing Qualitative Characteristics of Useful Information

A tradeoff between enhancing qualitative characteristics often occurs. For example, when a company records sales revenue, it is required to simultaneously estimate and record an expense for potential bad debts (uncollectible accounts). Including this estimated expense is considered to represent the economic event faithfully and to provide relevant information about the net profits for the accounting period. The information is timely and understandable; but because bad debts may not be known with certainty until a later period, inclusion of this estimated expense involves a sacrifice of verifiability. The bad debt expense is simply an estimate. It is apparent that it is not always possible to simultaneously fulfill all qualitative characteristics.

Companies are *most likely* to make tradeoffs between which of the following when preparing financial reports?

A Relevance and materiality.

B Timeliness and verifiability.

C Relevance and faithful representation.

Solution:

B is correct. Providing timely information implies a shorter time frame between the economic event and the information preparation; however, fully verifying information may require a longer time frame. Relevance and faithful representation are fundamental qualitative characteristics that make financial information useful. Both characteristics are required; there is no tradeoff between these. Materiality is an aspect of relevance.

THE ELEMENTS OF FINANCIAL STATEMENTS

6

c describe the International Accounting Standards Board's conceptual framework, including qualitative characteristics of financial reports, constraints on financial reports, and required reporting elements;

d describe general requirements for financial statements under International Financial Reporting Standards (IFRS);

Financial statements portray the financial effects of transactions and other events by grouping them into broad classes (elements) according to their economic characteristics. Three elements of financial statements are directly related to the measurement of financial position: assets, liabilities, and equity.

- **Assets**: A present economic resource controlled by the entity as a result of past events. An economic resource is a right that has the potential to produce economic benefits. Assets are what a company owns (e.g., inventory and equipment).
- **Liabilities**:. A present obligation of the entity to transfer an economic resource as a result of past events. An obligation is a duty or responsibility that the entity has no practical ability to avoid. Liabilities are what a company owes (e.g., bank borrowings).
- **Equity** (for public companies, also known as "shareholders' equity" or "stockholders' equity"): Assets less liabilities. Equity is the residual interest in the assets after subtracting the liabilities.

The elements of financial statements directly related to the measurement of performance (profit and related measures) are income and expenses.

- **Income**: Increases in assets, or decreases in liabilities, that result in increases in equity, other than those relating to contributions from holders of equity claims. Income includes both revenues and gains. Revenues represent income from the ordinary activities of the enterprise (e.g., the sale of products or provision of services). Gains may result from ordinary activities or other activities (the sale of surplus equipment).
- **Expenses**: Decreases in assets, or increases in liabilities, that result in decreases in equity, other than those relating to distributions to holders of equity claims. Expenses include those related to operating activities such as cost of good sold and operating expenses including wages, rents and other items. Losses are also considered expenses and can result from the sale of assets at less than their carrying values, impairments of asset values and a variety of other items.

6.1 Underlying Assumptions in Financial Statements

Two important assumptions underlie financial statements: accrual accounting and going concern. These assumptions determine how financial statement elements are recognized and measured.

The use of "accrual accounting" assumes that financial statements should reflect transactions in the period when they actually occur, not necessarily when cash movements occur. For example, a company reports revenues *when they are earned (when the performance obligations have been satisfied)*, regardless of whether the company received cash before or after delivering the product, at the time of delivery.

"Going concern" refers to the assumption that the company will continue in business for the foreseeable future. To illustrate, consider the value of a company's inventory if it is assumed that the inventory can be sold over a normal period of time versus the value of that same inventory if it is assumed that the inventory must all be sold in a day (or a week). Companies with the intent to liquidate or materially curtail operations would require different information for a fair presentation.

6.2 Recognition of Financial Statement Elements

Recognition means that an item is included in the balance sheet or income statement. Recognition occurs if the item meets the definition of an element and satisfies the criteria for recognition. Recognition is appropriate if it results in both relevant information about assets, liabilities, equity, income and expenses and a faithful representation of those items, because the aim is to provide information that is useful to investors, lenders and other creditors.

6.3 Measurement of Financial Statement Elements

Measurement is the process of determining the monetary amounts at which the elements of the financial statements are to be recognized and carried in the balance sheet and income statement. The following alternative bases of measurement are used to different degrees and in varying combinations to measure assets and liabilities:

- **Historical cost**: Historical cost is simply the amount of cash or cash equivalents paid to purchase an asset, including any costs of acquisition and/or preparation. If the asset was not bought for cash, historical cost is the fair value of whatever was given in order to buy the asset. When referring to liabilities, the historical cost basis of measurement means the amount of proceeds received in exchange for the obligation.
- **Amortised cost**: Historical cost adjusted for amortisation, depreciation, or depletion and/or impairment.
- **Current cost**: In reference to assets, current cost is the amount of cash or cash equivalents that would have to be paid to buy the same or an equivalent asset today. In reference to liabilities, the current cost basis of measurement means the undiscounted amount of cash or cash equivalents that would be required to settle the obligation today.
- **Realizable (settlement) value**: In reference to assets, realizable value is the amount of cash or cash equivalents that could currently be obtained by selling the asset in an orderly disposal. For liabilities, the equivalent to realizable value is called "settlement value"—that is, settlement value is the undiscounted amount of cash or cash equivalents expected to be paid to satisfy the liabilities in the normal course of business.

- **Present value (PV)**: For assets, present value is the present discounted value of the future net cash inflows that the asset is expected to generate in the normal course of business. For liabilities, present value is the present discounted value of the future net cash outflows that are expected to be required to settle the liabilities in the normal course of business.
- **Fair value**: is defined as an exit price, the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants at the measurement date. This may involve either market measures or present value measures depending on the availability of information.

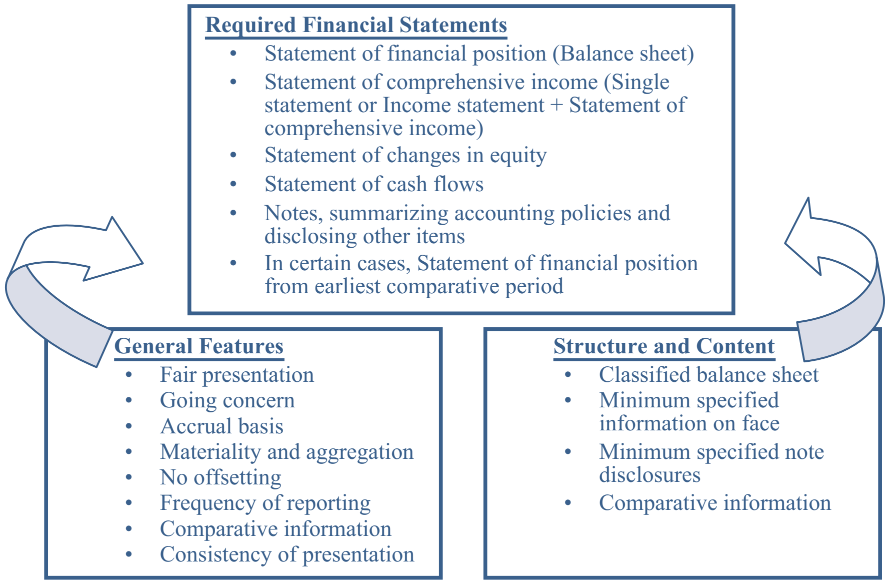
GENERAL REQUIREMENTS FOR FINANCIAL STATEMENTS

7

d describe general requirements for financial statements under International Financial Reporting Standards (IFRS);

The *Conceptual Framework* provides a basis for establishing standards and the elements of financial statements. However, the framework does not address the general contents of the financial statements, which are addressed in International Accounting Standard (IAS) No. 1, *Presentation of Financial Statements.* IAS No. 1 specifies the required financial statements, general features of financial statements, and structure and content of financial statements.[12] These general requirements are illustrated in Exhibit 1 and described in the subsections below.

Exhibit 1 IASB General Requirements for Financial Statements



Required Financial Statements

- Statement of financial position (Balance sheet)
- Statement of comprehensive income (Single statement or Income statement + Statement of comprehensive income)
- Statement of changes in equity
- Statement of cash flows
- Notes, summarizing accounting policies and disclosing other items
- In certain cases, Statement of financial position from earliest comparative period

General Features

- Fair presentation
- Going concern
- Accrual basis
- Materiality and aggregation
- No offsetting
- Frequency of reporting
- Comparative information
- Consistency of presentation

Structure and Content

- Classified balance sheet
- Minimum specified information on face
- Minimum specified note disclosures
- Comparative information

---

[12] For US GAAP, financial statement presentation is covered in Sections 205 through 280 of the Accounting Standards Codification.

In the following sections, we discuss the required financial statements, the general features underlying the preparation of financial statements, and the specified structure and content in greater detail.

7.1 Required Financial Statements

Under IAS No. 1, a complete set of financial statements includes:[13]

- a statement of financial position (balance sheet);
- a statement of comprehensive income (a single statement of comprehensive income or two statements, an income statement and a statement of comprehensive income that begins with profit or loss from the income statement);
- a statement of changes in equity, separately showing changes in equity resulting from profit or loss, each item of other comprehensive income, and transactions with owners in their capacity as owners;[14]
- a statement of cash flows; and
- notes comprising a summary of significant accounting policies and other explanatory notes that disclose information required by IFRS and not presented elsewhere and that provide information relevant to an understanding of the financial statements.

Entities are encouraged to furnish other related financial and non-financial information in addition to that required. Financial statements need to present fairly the financial position, financial performance, and cash flows of an entity.

7.2 General Features of Financial Statements

A company that applies IFRS is required to state explicitly in the notes to its financial statements that it is in compliance with the standards. Such a statement is only made when a company is in compliance with *all* requirements of IFRS. In extremely rare circumstances, a company may deviate from a requirement of IFRS if management concludes that complying with IFRS would result in misleading financial statements. In this case, management must disclose details of the departure from IFRS.

IAS No. 1 specifies a number of general features underlying the preparation of financial statements. These features clearly reflect the *Conceptual Framework*.

- *Fair Presentation*: The application of IFRS is presumed to result in financial statements that achieve a fair presentation. The IAS describes fair presentation as follows:

  > Fair presentation requires the faithful representation of the effects of transactions, other events and conditions in accordance with the definitions and recognition criteria for assets, liabilities, income and expenses set out in the *Framework*.[15]

- *Going Concern*: Financial statements are prepared on a going concern basis unless management either intends to liquidate the entity or to cease trading, or has no realistic alternative but to do so. If not presented on a going concern basis, the fact and rationale should be disclosed.

---

[13] IAS No. 1, *Presentation of Financial Statements*, paragraph 10.

[14] Examples of transactions with owners acting in their capacity as owners include sale of equity securities to investors, distributions of earnings to investors, and repurchases of equity securities from investors.

[15] IAS No. 1, *Presentation of Financial Statements*, paragraph 15.

- *Accrual Basis*: Financial statements (except for cash flow information) are to be prepared using the accrual basis of accounting.
- *Materiality and Aggregation*: Omissions or misstatements of items are material if they could, individually or collectively, influence the economic decisions that users make on the basis of the financial statements. Each material class of similar items is presented separately. Dissimilar items are presented separately unless they are immaterial.
- *No Offsetting*: Assets and liabilities, and income and expenses, are not offset unless required or permitted by an IFRS.
- *Frequency of Reporting*: Financial statements must be prepared at least annually.
- *Comparative Information*: Financial statements must include comparative information from the previous period. The comparative information of prior periods is disclosed for all amounts reported in the financial statements, unless an IFRS requires or permits otherwise.
- *Consistency*: The presentation and classification of items in the financial statements are usually retained from one period to the next.

7.3 Structure and Content Requirements

IAS No. 1 also specifies structure and content of financial statements. These requirements include the following:

- *Classified Statement of Financial Position (Balance Sheet)*: requires the balance sheet to distinguish between current and non-current assets, and between current and non-current liabilities unless a presentation based on liquidity provides more relevant and reliable information (e.g., in the case of a bank or similar financial institution).
- *Minimum Information on the Face of the Financial Statements*: specifies the minimum line item disclosures on the face of, or in the notes to, the financial statements. For example, companies are specifically required to disclose the amount of their plant, property, and equipment as a line item on the face of the balance sheet. The major line items included in financial statements are covered in other readings.
- *Minimum Information in the Notes* (or on the face of financial statements): specifies disclosures about information to be presented in the financial statements. This information must be provided in a systematic manner and cross-referenced from the face of the financial statements to the notes. The required information is summarized in Exhibit 2.
- *Comparative Information*: For all amounts reported in a financial statement, comparative information should be provided for the previous period unless another standard requires or permits otherwise. Such comparative information allows users to better understand reported amounts.

Exhibit 2 Summary of IFRS Required Disclosures in the Notes to the Financial Statements

Disclosure of Accounting Policies	■ Measurement bases used in preparing financial statements ■ Significant accounting policies used ■ Judgments made in applying accounting policies that have the most significant effect on the amounts recognized in the financial statements
Sources of Estimation Uncertainty	■ Key assumptions about the future and other key sources of estimation uncertainty that have a significant risk of causing material adjustment to the carrying amount of assets and liabilities within the next year
Other Disclosures	■ Information about capital and about certain financial instruments classified as equity ■ Dividends not recognized as a distribution during the period, including dividends declared before the financial statements were issued and any cumulative preference dividends ■ Description of the entity, including its domicile, legal form, country of incorporation, and registered office or business address ■ Nature of operations and principal activities ■ Name of parent and ultimate parent

8

COMPARISON OF IFRS WITH ALTERNATIVE REPORTING SYSTEMS

e describe implications for financial analysis of alternative financial reporting systems and the importance of monitoring developments in financial reporting standards.

The adoption of IFRS as the required financial reporting standard by the EU and other countries has advanced the goal of global convergence. Nevertheless, there are still significant differences in financial reporting in the global capital markets. Arguably, the most critical are the differences that exist between IFRS and US GAAP. Following the EU adoption of IFRS (in 2005), a significant number of the world's listed companies use one of these two reporting standards.

In general, the IASB and FASB work together to coordinate changes to accounting standards and reduce differences between the standards. A joint IASB/FASB project initiated in 2004 aimed to develop an improved, common conceptual framework. In late 2010, convergence of the conceptual frameworks was put on hold. In December 2012, the IASB reactivated the conceptual framework as an IASB project. As of year-end 2018, there are no new projects on the convergence agenda.

As more countries adopt IFRS, the need for analysts to consider other financial reporting systems will be reduced. Analysts are likely to encounter financial statements that are prepared on a basis other than IFRS. Differences between IFRS and US GAAP remain and affect the framework as well as numerous financial reporting standards. Curriculum readings on individual financial statements and specific topics provide a more detailed review of related differences in IFRS and US GAAP.

Exhibit 3 Differences between IFRS and US GAAP

Basis for Comparison	US GAAP	IFRS
Developed by	Financial Accounting Standard Board (FASB).	International Accounting Standard Board (IASB).
Based on	Rules	Principles
Inventory valuation	FIFO, LIFO and Weighted Average Method.	FIFO and Weighted Average Method.
Extraordinary items	Shown below (at the bottom of the income statement).	Not segregated in the income statement.
Development cost	Treated as an expense	Capitalized, only if certain conditions are satisfied.
Reversal of Inventory	Prohibited	Permissible, if specified conditions are met.

Source: https://keydifferences.com/difference-between-gaap-and-ifrs.html

For analyzing financial statements created under different standards, reconciliation schedules and disclosures regarding the significant differences between the reporting bases—historically available in some jurisdictions—were particularly helpful. For example, the SEC historically required reconciliation for foreign private issuers that did not prepare financial statements in accordance with US GAAP. In 2007, however, the SEC eliminated the reconciliation requirement for companies that prepared their financial statements according to IFRS. Although the disclosures related to any such differences were sometimes dauntingly long, the numerical reconciliations of net income and shareholders' equity appeared in charts that were relatively easy to use. Because reconciliation disclosures are no longer generally available, an analyst comparing two companies that use different reporting standards must be aware of areas where accounting standards have not converged. In many cases, a user of financial statements prepared under different accounting standards does not have enough information to make specific adjustments required to achieve comparability. Instead, an analyst must maintain caution in interpreting comparative financial measures produced under different accounting standards and monitor significant developments in financial reporting standards, as this can have important implications for comparing the performance of companies and security valuation.

MONITORING DEVELOPMENTS IN FINANCIAL REPORTING STANDARDS

9

a describe the objective of financial reporting and the importance of financial reporting standards in security analysis and valuation;

e describe implications for financial analysis of alternative financial reporting systems and the importance of monitoring developments in financial reporting standards.

In studying financial reporting and financial statement analysis in general, the analyst must be aware that reporting standards are evolving. Analysts need to monitor ongoing developments in financial reporting and assess their implications for security analysis and valuation. The need to monitor developments in financial reporting standards does not mean that analysts should be accountants. An accountant monitors these developments from a preparer's perspective; an analyst needs to monitor from a user's perspective. More specifically, analysts need to know how these developments will affect financial reports.

Analysts can remain aware of developments in financial reporting standards by monitoring new products or transactions, actions of standard setters and other groups representing users of financial statements (such as CFA Institute), and company disclosures regarding critical accounting policies and estimates.

9.1 New Products or Types of Transactions

New products and new types of transactions can have unusual or unique elements to them such that no explicit guidance in the financial reporting standards exists. New products or transactions typically arise from economic events, such as new businesses (e.g., fintech), or from a newly developed financial instrument or financial structure. Financial instruments, whether exchange traded or not, are typically designed to enhance a company's business or to mitigate inherent risks. However, at times, financial instruments or structured transactions have been developed primarily for purposes of financial report "window dressing."

Although companies might discuss new products and transactions in their financial reports, the analyst can also monitor business journals and the capital markets to identify such items. Additionally, when one company in an industry develops a new product or transaction, other companies in the industry often do the same. Once new products, financial instruments, or structured transactions are identified, it is helpful to gain an understanding of the business purpose. If necessary, an analyst can obtain further information from a company's management, which should be able to describe the economic purpose, the financial statement reporting, significant estimates, judgments applied in determining the reporting, and future cash flow implications for these items.

9.2 Evolving Standards and the Role of CFA Institute

The actions of standard setters and regulators are unlikely to be helpful in identifying new products and transactions, given the lag between new product development and regulatory action. Monitoring the actions of these authorities is nonetheless important for another reason: Changes in regulations can affect companies' financial reports and, thus, valuations. This is particularly true if the financial reporting standards change to require more explicit identification of matters affecting asset/liability valuation or financial performance. For example, one regulatory change required companies to include a provision for expenses associated with the grant and vesting of employee stock option grants as an expense in the income statement. Prior to the required expensing, an analyst could only assess the dilutive effect to shareholders associated with stock option grants by reviewing information disclosed in the notes to the financial statements.

To the extent that some market participants do not examine financial statement details and thus ignore some items when valuing a company's securities, more explicit identification could affect the value of the company's securities. Additionally, it is plausible to believe that management is more attentive to and rigorous in any calculations/estimates of items that appear in the financial statements, compared to items that are only disclosed in the notes.

The IASB (www.iasb.org) and FASB (www.fasb.org) provide a great deal of information on their websites regarding new standards and proposals for future changes in standards. In addition, the IASB and FASB seek input from the financial analyst community—those who regularly use financial statements in making investment and credit decisions. When a new standard is proposed, an exposure draft is made available and users of financial statements can draft comment letters and position papers for submission to the IASB and FASB in order to evaluate the proposal.

CFA Institute is active in supporting improvements to financial reporting. Volunteer members of CFA Institute serve on several liaison committees that meet regularly to make recommendations to the IASB and FASB on proposed standards and to draft comment letters and position papers. The comment letters and position papers of these groups on financial reporting issues are available at www.cfainstitute.org/advocacy.

In 2007, CFA Institute issued a position paper titled *A Comprehensive Business Reporting Model: Financial Reporting for Investors*, which provides a suggested model for significantly improving financial reporting. The position paper remains relevant in stating:

> Corporate financial statements and their related disclosures are fundamental to sound investment decision making. The well-being of the world's financial markets, and of the millions of investors who entrust their financial present and future to those markets, depends directly on the information financial statements and disclosures provide. Consequently, the quality of the information drives global financial markets. The quality, in turn, depends directly on the principles and standards managers apply when recognizing and measuring the economic activities and events affecting their companies' operations....
>
> Investors require timeliness, transparency, comparability, and consistency in financial reporting. Investors have a preference for decision relevance over reliability..."analysts need to know economic reality—what is really going on—to the greatest extent it can be depicted by accounting numbers." Corporate financial statements that fail to reflect this economic reality undermine the investment decision-making process.[16]

Among other principles, the proposed model stresses the importance of information regarding the current fair value of assets and liabilities, of neutrality in financial reporting, and of providing detailed information on cash flows to investors through the choice of the so-called direct format for the cash flow statement.

In summary, analysts can improve their investment decision making by keeping current on financial reporting standards. In addition, analysts can contribute to improving financial reporting by sharing their perspective as users with standard-setting bodies, which typically invite comments concerning proposed changes.

SUMMARY

An awareness of financial reporting and underlying financial reporting standards can assist in security valuation and other financial analysis. This reading describes the conceptual objectives of financial reporting standards, the parties involved in standard-setting processes, and the implication for analysts in monitoring developments in reporting standards.

---

[16] *A Comprehensive Business Reporting Model: Financial Reporting for Investors*, CFA Institute Centre for Financial Market Integrity, July 2007, p. 1, 2.

Some key points of the reading are summarized below:

- The objective of financial reporting is to provide financial information about the reporting entity that is useful to existing and potential investors, lenders, and other creditors in making decisions about providing resources to the entity.
- Financial reporting requires policy choices and estimates. These choices and estimates require judgment, which can vary from one preparer to the next. Accordingly, standards are needed to ensure increased consistency in these judgments.
- Private sector standard setting bodies and regulatory authorities play significant but different roles in the standard setting process. In general, standard setting bodies make the rules, and regulatory authorities enforce the rules. However, regulators typically retain legal authority to establish financial reporting standards in their jurisdiction.
- The IFRS framework sets forth the concepts that underlie the preparation and presentation of financial statements for external users.
- The objective of fair presentation of useful information is the center of the IASB's *Conceptual Framework*. The qualitative characteristics of useful information include fundamental and enhancing characteristics. Information must exhibit the fundamental characteristics of relevance and faithful representation to be useful. The enhancing characteristics identified are comparability, verifiability, timeliness, and understandability.
- *IFRS Financial Statements*: IAS No. 1 prescribes that a complete set of financial statements includes a statement of financial position (balance sheet), a statement of comprehensive income (either two statements—one for net income and one for comprehensive income—or a single statement combining both net income and comprehensive income), a statement of changes in equity, a cash flow statement, and notes. The notes include a summary of significant accounting policies and other explanatory information.
- Financial statements need to reflect certain basic features: fair presentation, going concern, accrual basis, materiality and aggregation, and no offsetting.
- Financial statements must be prepared at least annually, must include comparative information from the previous period, and must be consistent.
- Financial statements must follow certain presentation requirements including a classified statement of financial position (balance sheet) and minimum information on both the face of the financial statements and in the notes.
- A significant number of the world's listed companies report under either IFRS or US GAAP.
- In many cases, a user of financial statements will lack the information necessary to make specific adjustments required to achieve comparability between companies that use IFRS and companies that use US GAAP. Instead, an analyst must maintain general caution in interpreting comparative financial measures produced under different accounting standards and monitor significant developments in financial reporting standards.
- Analysts can remain aware of ongoing developments in financial reporting by monitoring new products or types of transactions; actions of standard setters, regulators, and other groups; and company disclosures regarding critical accounting policies and estimates.

PRACTICE PROBLEMS

1 Which of the following is *most likely* not an objective of financial statements?
   - **A** To provide information about the performance of an entity.
   - **B** To provide information about the financial position of an entity.
   - **C** To provide information about the users of an entity's financial statements.

2 International financial reporting standards are currently developed by which entity?
   - **A** The IFRS Foundation.
   - **B** The International Accounting Standards Board.
   - **C** The International Organization of Securities Commissions.

3 US generally accepted accounting principles are currently developed by which entity?
   - **A** The Securities and Exchange Commission.
   - **B** The Financial Accounting Standards Board.
   - **C** The Public Company Accounting Oversight Board.

4 A core objective of the International Organization of Securities Commissions is to:
   - **A** eliminate systemic risk.
   - **B** protect users of financial statements.
   - **C** ensure that markets are fair, efficient, and transparent.

5 According to the *Conceptual Framework for Financial Reporting*, which of the following is *not* an enhancing qualitative characteristic of information in financial statements?
   - **A** Accuracy.
   - **B** Timeliness.
   - **C** Comparability.

6 Which of the following is *not* a constraint on the financial statements according to the *Conceptual Framework*?
   - **A** Understandability.
   - **B** Benefit versus cost.
   - **C** Balancing of qualitative characteristics.

7 The assumption that an entity will continue to operate for the foreseeable future is called:
   - **A** accrual basis.
   - **B** comparability.
   - **C** going concern.

8 The assumption that the effects of transactions and other events are recognized when they occur, not when the cash flows occur, is called:
   - **A** relevance.
   - **B** accrual basis.
   - **C** going concern.

© 2019 CFA Institute. All rights reserved.

9 Neutrality of information in the financial statements most closely contributes to which qualitative characteristic?

A Relevance.

B Understandability.

C Faithful representation.

10 Valuing assets at the amount of cash or equivalents paid or the fair value of the consideration given to acquire them at the time of acquisition most closely describes which measurement of financial statement elements?

A Current cost.

B Historical cost.

C Realizable value.

11 The valuation technique under which assets are recorded at the amount that would be received in an orderly disposal is:

A current cost.

B present value.

C realizable value.

12 Which of the following is *not* a required financial statement according to IAS No. 1?

A Statement of financial position.

B Statement of changes in income.

C Statement of comprehensive income.

13 Which of the following elements of financial statements is *most* closely related to measurement of performance?

A Assets.

B Expenses.

C Liabilities.

14 Which of the following elements of financial statements is *most* closely related to measurement of financial position?

A Equity.

B Income.

C Expenses.

15 Which of the following disclosures regarding new accounting standards provides the *most* meaningful information to an analyst?

A The impact of adoption is discussed.

B The standard will have no material impact.

C Management is still evaluating the impact.

SOLUTIONS

1 C is correct. Financial statements provide information, including information about the entity's financial position, performance, and changes in financial position, to users. They do not typically provide information about users.

2 B is correct. The IASB is currently charged with developing International Financial Reporting Standards.

3 B is correct. The FASB is responsible for the Accounting Standards Codification™, the single source of nongovernmental authoritative US generally accepted accounting principles.

4 C is correct. A core objective of IOSCO is to ensure that markets are fair, efficient, and transparent. The other core objectives are to reduce, not eliminate, systemic risk and to protect investors, not all users of financial statements.

5 A is correct. Accuracy is not an enhancing qualitative characteristic. Faithful representation, not accuracy, is a fundamental qualitative characteristic.

6 A is correct. Understandability is an enhancing qualitative characteristic of financial information—not a constraint.

7 C is correct. The *Conceptual Framework* identifies two important underlying assumptions of financial statements: accrual basis and going concern. Going concern is the assumption that the entity will continue to operate for the foreseeable future. Enterprises with the intent to liquidate or materially curtail operations would require different information for a fair presentation.

8 B is correct. Accrual basis reflects the effects of transactions and other events being recognized when they occur, not when the cash flows. These effects are recorded and reported in the financial statements of the periods to which they relate.

9 C is correct. The fundamental qualitative characteristic of faithful representation is contributed to by completeness, neutrality, and freedom from error.

10 B is correct. Historical cost is the consideration paid to acquire an asset.

11 C is correct. The amount that would be received in an orderly disposal is realizable value.

12 B is correct. There is no statement of changes in income. Under IAS No. 1, a complete set of financial statements includes a statement of financial position, a statement of comprehensive income, a statement of changes in equity, a statement of cash flows, and notes comprising a summary of significant accounting policies and other explanatory information.

13 B is correct. The elements of financial statements related to the measure of performance are income and expenses.

14 A is correct. The elements of financial statements related to the measurement of financial position are assets, liabilities, and equity.

15 A is correct. A discussion of the impact would be the most meaningful, although B would also be useful.

Glossary

A priori probability A probability based on logical analysis rather than on observation or personal judgment.

Abandonment option The ability to terminate an investment at some future time if the financial results are disappointing.

Abnormal return The amount by which a security's actual return differs from its expected return, given the security's risk and the market's return.

Absolute advantage A country's ability to produce a good or service at a lower absolute cost than its trading partner.

Absolute dispersion The amount of variability present without comparison to any reference point or benchmark.

Absolute frequency The actual number of observations counted for each unique value of the variable (also called raw frequency).

Accelerated book build An offering of securities by an investment bank acting as principal that is accomplished in only one or two days.

Accelerated methods Depreciation methods that allocate a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Accounting costs Monetary value of economic resources used in performing an activity. These can be explicit, out-of-pocket, current payments, or an allocation of historical payments (depreciation) for resources. They do not include implicit opportunity costs.

Accounting profit Income as reported on the income statement, in accordance with prevailing accounting standards, before the provisions for income tax expense. Also called *income before taxes* or *pretax income*.

Accounts payable Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Accrued expenses Liabilities related to expenses that have been incurred but not yet paid as of the end of an accounting period—an example of an accrued expense is rent that has been incurred but not yet paid, resulting in a liability "rent payable." Also called *accrued liabilities*.

Accrued interest Interest earned but not yet paid.

Acquisition method A method of accounting for a business combination where the acquirer is required to measure each identifiable asset and liability at fair value. This method was the result of a joint project of the IASB and FASB aiming at convergence in standards for the accounting of business combinations.

Action lag Delay from policy decisions to implementation.

Active investment An approach to investing in which the investor seeks to outperform a given benchmark.

Active return The return on a portfolio minus the return on the portfolio's benchmark.

Activity ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory. Also called *asset utilization ratios* or *operating efficiency ratios*.

Add-on rates Bank certificates of deposit, repos, and indexes such as Libor and Euribor are quoted on an add-on rate basis (bond equivalent yield basis).

Addition rule for probabilities A principle stating that the probability that A or B occurs (both occur) equals the probability that A occurs, plus the probability that B occurs, minus the probability that both A and B occur.

Agency bond See *quasi-government bond*.

Agency costs Costs associated with the conflict of interest present between principals and agents when a company is managed by non-owners. Agency costs result from the inherent conflicts of interest between managers, bondholders, and equity owners.

Agency costs of debt Costs arising from conflicts of interest between managers and debtholders.

Agency costs of equity The smaller the stake managers have in the company, the less their share in bearing the cost of excessive perquisite consumption—consequently, the less their desire to give their best efforts in running the company.

Agency RMBS In the United States, securities backed by residential mortgage loans and guaranteed by a federal agency or guaranteed by either of the two GSEs (Fannie Mae and Freddie Mac).

Aggregate demand The quantity of goods and services that households, businesses, government, and non-domestic customers want to buy at any given level of prices.

Aggregate demand curve Inverse relationship between the price level and real output.

Aggregate income The value of all the payments earned by the suppliers of factors used in the production of goods and services.

Aggregate output The value of all the goods and services produced during a specified period.

Aggregate supply The quantity of goods and services producers are willing to supply at any given level of price.

Aggregate supply curve The level of domestic output that companies will produce at each price level.

All-or-nothing (AON) orders An order that includes the instruction to trade only if the trade fills the entire quantity (size) specified.

Allocationally efficient A characteristic of a market, a financial system, or an economy that promotes the allocation of resources to their highest value uses.

Alternative data Non-traditional data types generated by the use of electronic devices, social media, satellite and sensor networks, and company exhaust.

Alternative hypothesis The hypothesis that is accepted if the null hypothesis is rejected.

Alternative investment markets Market for investments other than traditional securities investments (i.e., traditional common and preferred shares and traditional fixed income instruments). The term usually encompasses direct and indirect investment in real estate (including timberland and farmland) and commodities (including precious metals); hedge funds, private equity, and other investments requiring specialized due diligence.

Alternative trading systems Trading venues that function like exchanges but that do not exercise regulatory authority over their subscribers except with respect to the conduct

of the subscribers' trading in their trading systems. Also called *electronic communications networks* or *multilateral trading facilities*.

American depository receipt A US dollar-denominated security that trades like a common share on US exchanges.

American depository share The underlying shares on which American depository receipts are based. They trade in the issuing company's domestic market.

American-style Type of option contract that can be exercised at any time up to the option's expiration date.

Amortisation The process of allocating the cost of intangible long-term assets having a finite useful life to accounting periods; the allocation of the amount of a bond premium or discount to the periods remaining until bond maturity.

Amortised cost The historical cost (initially recognised cost) of an asset, adjusted for amortisation and impairment.

Amortizing bond Bond with a payment schedule that calls for periodic payments of interest and repayments of principal.

Amortizing loans Loans with a payment schedule that calls for periodic payments of interest and repayments of principal.

Analysis of variance (ANOVA) The analysis that breaks the total variability of a dataset (such as observations on the dependent variable in a regression) into components representing different sources of variation. With reference to regression, ANOVA provides the inputs for an *F*-test of the significance of the regression as a whole, as well as the inputs for the coefficient of determination and the standard error of the estimate.

Analytical duration The use of mathematical formulas to estimate the impact of benchmark yield-to-maturity changes on bond prices.

Anchoring and adjustment bias An information-processing bias in which the use of a psychological heuristic influences the way people estimate probabilities.

Annual percentage rate The cost of borrowing expressed as a yearly rate.

Annuity A finite set of level sequential cash flows.

Annuity due An annuity having a first cash flow that is paid immediately.

Anomalies Apparent deviations from market efficiency.

Antidilutive With reference to a transaction or a security, one that would increase earnings per share (EPS) or result in EPS higher than the company's basic EPS—antidilutive securities are not included in the calculation of diluted EPS.

Arbitrage 1) The simultaneous purchase of an undervalued asset or portfolio and sale of an overvalued but equivalent asset or portfolio, in order to obtain a riskless profit on the price differential. Taking advantage of a market inefficiency in a risk-free manner. 2) The condition in a financial market in which equivalent assets or combinations of assets sell for two different prices, creating an opportunity to profit at no risk with no commitment of money. In a well-functioning financial market, few arbitrage opportunities are possible. 3) A risk-free operation that earns an expected positive net profit but requires no net investment of money.

Arbitrage-free pricing The overall process of pricing derivatives by arbitrage and risk neutrality. Also called the *principle of no arbitrage*.

Arbitrageurs Traders who engage in arbitrage. See *arbitrage*.

Arithmetic mean The sum of the observations divided by the number of observations.

Artificial intelligence Computer systems that exhibit cognitive and decision-making ability comparable (or superior) to that of humans.

Ask The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size). See *offer*.

Ask size The maximum quantity of an asset that pertains to a specific ask price from a trader. For example, if the ask for a share issue is $30 for a size of 1,000 shares, the trader is offering to sell at $30 up to 1,000 shares.

Asset allocation The process of determining how investment funds should be distributed among asset classes.

Asset-backed securities A type of bond issued by a legal entity called a *special purpose entity* (SPE) on a collection of assets that the SPE owns. Also, securities backed by receivables and loans other than mortgages.

Asset-based valuation models Valuation based on estimates of the market value of a company's assets.

Asset beta The unlevered beta; which reflects the business risk of the assets; the assets' systematic risk.

Asset class A group of assets that have similar characteristics, attributes, and risk–return relationships.

Asset swap Converts the periodic fixed coupon of a specific bond to an MRR plus or minus a spread.

Asset utilization ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Assets Resources controlled by an enterprise as a result of past events and from which future economic benefits to the enterprise are expected to flow.

Assignment of accounts receivable The use of accounts receivable as collateral for a loan.

Asymmetric information The differential of information between corporate insiders and outsiders regarding the company's performance and prospects. Managers typically have more information about the company's performance and prospects than owners and creditors.

At the money An option in which the underlying's price equals the exercise price.

Auction A type of bond issuing mechanism often used for sovereign bonds that involves bidding.

Autarkic price The price of a good or service in an autarkic economy.

Autarky A state in which a country does not trade with other countries.

Automatic stabilizer A countercyclical factor that automatically comes into play as an economy slows and unemployment rises.

Availability bias An information-processing bias in which people take a heuristic approach to estimating the probability of an outcome based on how easily the outcome comes to mind.

Available-for-sale Under US GAAP, debt securities not classified as either held-to-maturity or held-for-trading securities. The investor is willing to sell but not actively planning to sell. In general, available-for-sale debt securities are reported at fair value on the balance sheet, with unrealized gains included as a component of other comprehensive income.

Average fixed cost Total fixed cost divided by quantity produced.

Average life See *weighted average life*.

Average product Measures the productivity of inputs on average and is calculated by dividing total product by the total number of units for a given input that is used to generate that output.

Average revenue Total revenue divided by quantity sold.

Average total cost Total cost divided by quantity produced.

Average variable cost Total variable cost divided by quantity produced.

Back-testing With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backfill bias A bias introduced if a fund backfills its past performance data to the database when it first starts reporting actual performance.

Backtesting With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backup line of credit A type of credit enhancement provided by a bank to an issuer of commercial paper to ensure that the issuer will have access to sufficient liquidity to repay maturing commercial paper if issuing new paper is not a viable option.

Backwardation A condition in the futures markets in which the spot price exceeds the futures price, the forward curve is downward sloping, and the convenience yield is high.

Balance of payments A double-entry bookkeeping system that summarizes a country's economic transactions with the rest of the world for a particular period of time, typically a calendar quarter or year.

Balance of trade deficit When the domestic economy is spending more on non-domestic goods and services than non-domestic economies are spending on domestic goods and services.

Balance sheet The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet). Also called *statement of financial position* or *statement of financial condition.*

Balance sheet ratios Financial ratios involving balance sheet items only.

Balanced With respect to a government budget, one in which spending and revenues (taxes) are equal.

Balloon payment Large payment required at maturity to retire a bond's outstanding principal amount.

Bar chart A chart for plotting the frequency distribution of categorical data, where each bar represents a distinct category and each bar's height is proportional to the frequency of the corresponding category. In technical analysis, a bar chart that plots four bits of data for each time interval—the high, low, opening, and closing prices. A vertical line connects the high and low prices. A cross-hatch left indicates the opening price and a cross-hatch right indicates the closing price.

Barter economy An economy where economic agents as house-holds, corporations, and governments "pay" for goods and services with another good or service.

Base-rate neglect A type of representativeness bias in which the base rate or probability of the categorization is not adequately considered.

Base rates The reference rate on which a bank bases lending rates to all other customers.

Basic EPS Net earnings available to common shareholders (i.e., net income minus preferred dividends) divided by the weighted average number of common shares outstanding.

Basis point Used in stating yield spreads, one basis point equals one-hundredth of a percentage point, or 0.01%.

Basket of listed depository receipts An exchange-traded fund (ETF) that represents a portfolio of depository receipts.

Bayes' formula The rule for updating the probability of an event of interest—given a set of prior probabilities for the event, information, and information given the event—if you receive new information.

Bearer bonds Bonds for which ownership is not recorded; only the clearing system knows who the bond owner is.

Bearish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from above; this movement is considered bearish. A **death cross** is a bearish crossover based on 50-day and 200-day moving averages.

Behavioral finance A field of finance that examines the psychological variables that affect and often distort the investment decision making of investors, analysts, and portfolio managers.

Behind the market Said of prices specified in orders that are worse than the best current price; e.g., for a limit buy order, a limit price below the best bid.

Benchmark A comparison portfolio; a point of reference or comparison.

Benchmark issue The latest sovereign bond issue for a given maturity. It serves as a benchmark against which to compare bonds that have the same features but that are issued by another type of issuer.

Benchmark rate Typically the yield-to-maturity on a government bond having the same or close to the same time-to-maturity.

Benchmark spread The yield spread over a specific benchmark, usually measured in basis points.

Bernoulli random variable A random variable having the outcomes 0 and 1.

Bernoulli trial An experiment that can produce one of two outcomes.

Best bid The highest bid in the market.

Best effort offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-efforts offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-in-class An ESG implementation approach that seeks to identify the most favorable companies in an industry based on ESG considerations.

Best offer The lowest offer (ask price) in the market.

Beta A measure of the sensitivity of a given investment or portfolio to movements in the overall market.

Bid The price at which a dealer or trader is willing to buy an asset, typically qualified by a maximum quantity.

Bid–ask spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid–offer spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid size The maximum quantity of an asset that pertains to a specific bid price from a trader.

Big Data The vast amount of data being generated by industry, governments, individuals, and electronic devices that arises from both traditional and non-traditional data sources.

Bilateral loan A loan from a single lender to a single borrower.

Bimodal A distribution that has two most frequently occurring values.

Binomial model A model for pricing options in which the underlying price can move to only one of two possible new prices.

Binomial random variable The number of successes in n Bernoulli trials for which the probability of success is constant for all trials and the trials are independent.

Binomial tree The graphical representation of a model of asset price dynamics in which, at each period, the asset moves up with probability p or down with probability $(1 - p)$.

Bitcoin A cryptocurrency using blockchain technology that was created in 2009.

Bivariate correlation See *Pearson correlation*.

Block brokers A broker (agent) that provides brokerage services for large-size trades.

Blockchain A type of digital ledger in which information is recorded sequentially and then linked together and secured using cryptographic methods.

Blue chip Widely held large market capitalization companies that are considered financially sound and are leaders in their respective industry or local stock market.

Bollinger Bands A price-based technical analysis indicator consisting of a line representing the moving average, a higher line representing the moving average plus a set number of standard deviations from the average (for the same number of periods as was used to calculate the moving average), and a lower line representing the moving average minus the same number of standard deviations.

Bond Contractual agreement between the issuer and the bondholders.

Bond equivalent yield A calculation of yield that is annualized using the ratio of 365 to the number of days to maturity. Bond equivalent yield allows for the restatement and comparison of securities with different compounding periods.

Bond indenture The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *trust deed*.

Bond market vigilantes Bond market participants who might reduce their demand for long-term bonds, thus pushing up their yields.

Bond yield plus risk premium approach An estimate of the cost of common equity that is produced by summing the before-tax cost of debt and a risk premium that captures the additional yield on a company's stock relative to its bonds. The additional yield is often estimated using historical spreads between bond yields and stock yields.

Bonus issue of shares A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Book building Investment bankers' process of compiling a "book" or list of indications of interest to buy part of an offering.

Book value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. Also called *carrying value*.

Boom An expansionary phase characterized by economic growth "testing the limits" of the economy.

Bootstrap A resampling method that repeatedly draws samples with replacement of the selected elements from the original observed sample. Bootstrap is usually conducted by using computer simulation and is often used to find standard error or construct confidence intervals of population parameters.

Bottom-up analysis An investment selection approach that focuses on company-specific circumstances rather than emphasizing economic cycles or industry analysis.

Box and whisker plot A graphic for visualizing the dispersion of data across quartiles. It consists of a "box" with "whiskers" connected to the box.

Breakdown A breakdown occurs when the price of an asset moves below a support level.

Breakeven point The number of units produced and sold at which the company's net income is zero (Revenues = Total cost); in the case of perfect competition, the quantity at which price, average revenue, and marginal revenue equal average total cost.

Breakout A breakout occurs when the price of an asset moves above a resistance level.

Bridge financing Interim financing that provides funds until permanent financing can be arranged.

Broad money Encompasses narrow money plus the entire range of liquid assets that can be used to make purchases.

Broker 1) An agent who executes orders to buy or sell securities on behalf of a client in exchange for a commission. 2) See *futures commission merchants*.

Broker–dealer A financial intermediary (often a company) that may function as a principal (dealer) or as an agent (broker) depending on the type of trade.

Brokered market A market in which brokers arrange trades among their clients.

Brownfield investment Investing in *existing* infrastructure assets.

Bubble line chart A line chart that uses varying-sized bubbles to represent a third dimension of the data. The bubbles are sometimes color-coded to present additional information.

Budget surplus/deficit The difference between government revenue and expenditure for a stated fixed period of time.

Bullet bond Bond in which the principal repayment is made entirely at maturity.

Bullish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from below; this movement is considered bullish. A **golden cross** is a bullish crossover based on 50-day and 200-day moving averages.

Business cycles Are recurrent expansions and contractions in economic activity affecting broad segments of the economy.

Business risk The risk associated with operating earnings. Operating earnings are uncertain because total revenues and many of the expenditures contributed to produce those revenues are uncertain.

Buy-side firm An investment management company or other investor that uses the services of brokers or dealers (i.e., the client of the sell side firms).

Buyback A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Call An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call market A market in which trades occur only at a particular time and place (i.e., when the market is called).

Call money rate The interest rate that buyers pay for their margin loan.

Call option An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call protection The time during which the issuer of the bond is not allowed to exercise the call option.

Callable bond A bond containing an embedded call option that gives the issuer the right to buy the bond back from the investor at specified prices on pre-determined dates.

Calmar ratio A comparison of the average annual compounded return to the maximum drawdown risk.

Candlestick chart A price chart with four bits of data for each time interval. A candle indicates the opening and closing price for the interval. The body of the candle is shaded if the opening price was higher than the closing price, and the body is white (or clear) if the opening price was lower than the closing price. Vertical lines known as wicks or shadows extend from the top and bottom of the candle to indicate, respectively, the high and low prices for the interval.

Cannibalization Cannibalization occurs when an investment takes customers and sales away from another part of the company.

Cap rate A metric by which real estate managers are often judged; the annual rent actually earned (net of any vacancies) divided by the price originally paid for the property.

Capacity The ability of the borrower to make its debt payments on time.

Capital account A component of the balance of payments account that measures transfers of capital.

Capital allocation line (CAL) A graph line that describes the combinations of expected return and standard deviation of return available to an investor from combining the optimal portfolio of risky assets with the risk-free asset.

Capital allocation process The process that companies use for decision making on capital investments—those projects with a life of one year or longer.

Capital asset pricing model (CAPM) An equation describing the expected return on any asset (or portfolio) as a linear function of its beta relative to the market portfolio.

Capital consumption allowance A measure of the wear and tear (depreciation) of the capital stock that occurs in the production of goods and services.

Capital deepening investment Increases the stock of capital relative to labor.

Capital expenditure Expenditure on physical capital (fixed assets).

Capital light Capital light businesses require little incremental investment in fixed assets or working capital to enable revenue growth.

Capital loss ratio The percentage of capital in deals that have been realized below cost, net of any recovered proceeds, divided by total invested capital.

Capital market expectations An investor's expectations concerning the risk and return prospects of asset classes.

Capital market line (CML) The line with an intercept point equal to the risk-free rate that is tangent to the efficient frontier of risky assets; represents the efficient frontier when a risk-free asset is available for investment.

Capital market securities Securities with maturities at issuance longer than one year.

Capital markets Financial markets that trade securities of longer duration, such as bonds and equities.

Capital rationing A capital rationing environment assumes that the company has a fixed amount of funds to invest.

Capital restrictions Controls placed on foreigners' ability to own domestic assets and/or domestic residents' ability to own foreign assets.

Capital stock The accumulated amount of buildings, machinery, and equipment used to produce goods and services.

Capital structure The mix of debt and equity that a company uses to finance its business; a company's specific mixture of long-term financing.

Carry The net of the costs and benefits of holding, storing, or "carrying" an asset.

Carrying amount The amount at which an asset or liability is valued according to accounting principles.

Carrying value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. For a bond, the purchase price plus (or minus) the amortized amount of the discount (or premium).

Cartel Participants in collusive agreements that are made openly and formally.

Cash collateral account Form of external credit enhancement whereby the issuer immediately borrows the credit-enhancement amount and then invests that amount, usually in highly rated short-term commercial paper.

Cash conversion cycle A financial metric that measures the length of time required for a company to convert cash invested in its operations to cash received as a result of its operations; equal to days of inventory on hand + days of sales outstanding – number of days of payables. Also called *net operating cycle*.

Cash flow additivity principle The principle that dollar amounts indexed at the same point in time are additive.

Cash flow from operating activities The net amount of cash provided from operating activities.

Cash flow from operations The net amount of cash provided from operating activities.

Cash flow yield The internal rate of return on a series of cash flows.

Cash market securities Money market securities settled on a "same day" or "cash settlement" basis.

Cash markets See *spot markets*.

Cash prices See *spot prices*.

Cash-settled forwards See *non-deliverable forwards*.

Catch-up clause A clause in an agreement that favors the GP. For a GP who earns a 20% performance fee, a catch-up clause allows the GP to receive 100% of the distributions above the hurdle rate *until* she receives 20% of the profits generated, and then every excess dollar is split 80/20 between the LPs and GP.

Categorical data Values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize (also called **qualitative data**).

CBOE Volatility Index (VIX) A measure of near-term market volatility as conveyed by S&P 500 stock index option prices.

Central bank funds market The market in which deposit-taking banks that have an excess reserve with their national central bank can lend money to banks that need funds for maturities ranging from overnight to one year. Called the Federal or Fed funds market in the United States.

Central bank funds rate Interest rate at which central bank funds are bought (borrowed) and sold (lent) for maturities ranging from overnight to one year. Called Federal or Fed funds rate in the United States.

Central banks The dominant bank in a country, usually with official or semi-official governmental status.

Central limit theorem The theorem that states the sum (and the mean) of a set of independent, identically distributed random variables with finite variances is normally distributed, whatever distribution the random variables follow.

Certificate of deposit An instrument that represents a specified amount of funds on deposit with a bank for a specified maturity and interest rate. CDs are issued in various denominations and can be negotiable or non-negotiable.

Change in polarity principle A tenet of technical analysis that states that once a support level is breached, it becomes a resistance level. The same holds true for resistance levels: Once breached, they become support levels.

Change of control put A covenant giving bondholders the right to require the issuer to buy back their debt, often at par or at some small premium to par value, in the event that the borrower is acquired.

Character The quality of a debt issuer's management.

Chartist An individual who uses charts or graphs of a security's historical prices or levels to forecast its future trends.

Chi-square test of independence A statistical test for detecting a potential association between categorical variables.

Classical cycle Refers to fluctuations in the level of economic activity when measured by GDP in volume terms.

Classified balance sheet A balance sheet organized so as to group together the various assets and liabilities into subcategories (e.g., current and noncurrent).

Clawback A requirement that the general partner return any funds distributed as incentive fees until the limited partners have received back their initial investment and a percentage of the total profit.

Clearing The process by which the exchange verifies the execution of a transaction and records the participants' identities.

Clearing instructions Instructions that indicate how to arrange the final settlement ("clearing") of a trade.

Clearinghouse An entity associated with a futures market that acts as middleman between the contracting parties and guarantees to each party the performance of the other.

Closed economy An economy that does not trade with other countries; an *autarkic economy*.

Closed-end fund A mutual fund in which no new investment money is accepted. New investors invest by buying existing shares, and investors in the fund liquidate by selling their shares to other investors.

Cluster sampling A procedure that divides a population into subpopulation groups (clusters) representative of the population and then randomly draws certain clusters to form a sample.

Clustered bar chart See *grouped bar chart*.

Co-investing In co-investing, the investor invests in assets *indirectly* through the fund but also possesses rights (known as co-investment rights) to invest *directly* in the same assets. Through co-investing, an investor is able to make an investment *alongside* a fund when the fund identifies deals.

Code of ethics An established guide that communicates an organization's values and overall expectations regarding member behavior. A code of ethics serves as a general guide for how community members should act.

Coefficient of determination The percentage of the variation of the dependent variable that is explained by the independent variable. Also referred to as the "R-squared" or "R^2."

Coefficient of variation The ratio of a set of observations' standard deviation to the observations' mean value.

Cognitive cost The effort involved in processing new information and updating beliefs.

Cognitive dissonance The mental discomfort that occurs when new information conflicts with previously held beliefs or cognitions.

Cognitive errors Behavioral biases resulting from faulty reasoning; cognitive errors stem from basic statistical, information-processing, or memory errors.

Coincident economic indicators Turning points that are usually close to those of the overall economy; they are believed to have value for identifying the economy's present state.

Collateral Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Collateral manager Buys and sells debt obligations for and from the CDO's portfolio of assets (i.e., the collateral) to generate sufficient cash flows to meet the obligations to the CDO bondholders.

Collateral trust bonds Bonds secured by securities, such as common shares, other bonds, or other financial assets.

Collateralized bond obligations A structured asset-backed security that is collateralized by a pool of bonds.

Collateralized debt obligation Generic term used to describe a security backed by a diversified pool of one or more debt obligations.

Collateralized loan obligations A structured asset-backed security that is collateralized by a pool of loans.

Collateralized mortgage obligation A security created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collateralized mortgage obligations Securities created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collaterals Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Combination A listing in which the order of the listed items does not matter.

Combination formula (binomial formula) The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is ${}_nC_r = \binom{n}{r} = \dfrac{n!}{(n-r)!r!}$.

Commercial paper A short-term, negotiable, unsecured promissory note that represents a debt obligation of the issuer.

Committed (Regular) lines of credit A bank commitment to extend credit up to a pre-specified amount; the commitment is considered a short-term liability and is usually in effect for 364 days (one day short of a full year).

Committed capital The amount that the limited partners have agreed to provide to the private equity fund.

Commodity swap A swap in which the underlying is a commodity such as oil, gold, or an agricultural product.

Common market Level of economic integration that incorporates all aspects of the customs union and extends it by allowing free movement of factors of production among members.

Common shares A type of security that represent an ownership interest in a company.

Common-size analysis The restatement of financial statement items using a common denominator or reference item that allows one to identify trends and major differences; an example is an income statement in which all items are expressed as a percent of revenue.

Common stock See *common shares.*

Company analysis Analysis of an individual company.

Comparable company A company that has similar business risk, usually in the same industry and preferably with a single line of business.

Comparative advantage A country's ability to produce a good or service at a lower relative cost, or opportunity cost, than its trading partner.

Competitive strategy A company's plans for responding to the threats and opportunities presented by the external environment.

Complement The event not-S, written S^C, given the event S. Note that $P(S) + P(S^C) = 1$.

Complements Goods that tend to be used together; technically, two goods whose cross-price elasticity of demand is negative.

Complete markets Informally, markets in which the variety of distinct securities traded is so broad that any desired payoff in a future state-of-the-world is achievable.

Component cost of capital The rate of return required by suppliers of capital for an individual source of a company's funding, such as debt or equity.

Compounding The process of accumulating interest on interest.

Comprehensive income The change in equity of a business enterprise during a period from nonowner sources; includes all changes in equity during a period except those resulting from investments by owners and distributions to owners; comprehensive income equals net income plus other comprehensive income.

Conditional expected value The expected value of a stated event given that another event has occurred.

Conditional pass-through covered bonds Covered bonds that convert to pass-through securities after the original maturity date if all bond payments have not yet been made and the sponsor is in default.

Conditional probability The probability of an event given (conditioned on) another event.

Conditional variances The variance of one variable, given the outcome of another.

Confidence level The complement of the level of significance.

Confirmation bias A belief perseverance bias in which people tend to look for and notice what confirms their beliefs, to ignore or undervalue what contradicts their beliefs, and to misinterpret information as support for their beliefs.

Confusion matrix A type of contingency table used for evaluating the performance of a classification model.

Conservatism bias A belief perseverance bias in which people maintain their prior views or forecasts by inadequately incorporating new information.

Consolidation The movement of a stock's price within a well-defined range of trading levels for a period of time. The price consolidates between a support level and a resistance level.

Constant-yield price trajectory A graph that illustrates the change in the price of a fixed-income bond over time assuming no change in yield-to-maturity. The trajectory shows the "pull to par" effect on the price of a bond trading at a premium or a discount to par value.

Constituent securities With respect to an index, the individual securities within an index.

Consumer surplus The difference between the value that a consumer places on units purchased and the amount of money that was required to pay for them.

Contango A condition in the futures markets in which the spot price is lower than the futures price, the forward curve is upward sloping, and there is little or no convenience yield.

Contingency provision Clause in a legal document that allows for some action if a specific event or circumstance occurs.

Contingency table A tabular format that displays the frequency distributions of two or more categorical variables simultaneously and is used for finding patterns between the variables. A contingency table for two categorical variables is also known as a two-way table.

Contingent claim A security whose value is based on some other underlying security.

Contingent claims Derivatives in which the payoffs occur if a specific event occurs; generally referred to as options.

Contingent convertible bonds Bonds that automatically convert into equity if a specific event or circumstance occurs, such as the issuer's equity capital falling below the minimum requirement set by the regulators. Also called *CoCos.*

Continuation pattern A type of pattern used in technical analysis to predict the resumption of a market trend that was in place prior to the formation of a pattern.

Continuous data Data that can be measured and can take on any numerical value in a specified range of values.

Continuous random variable A random variable for which the range of possible outcomes is the real line (all real numbers between $-\infty$ and $+\infty$) or some subset of the real line.

Continuous trading market A market in which trades can be arranged and executed any time the market is open.

Continuously compounded return The natural logarithm of 1 plus the holding period return, or equivalently, the natural logarithm of the ending price over the beginning price.

Contra account An account that offsets another account.

Contract rate See *mortgage rate.*

Contraction The period of a business cycle after the peak and before the trough; often called a *recession* or, if exceptionally severe, called a *depression*.

Contraction risk The risk that when interest rates decline, the security will have a shorter maturity than was anticipated at the time of purchase because borrowers refinance at the new, lower interest rates.

Contractionary Tending to cause the real economy to contract.

Contractionary fiscal policy A fiscal policy that has the objective to make the real economy contract.

Contracts for differences See *non-deliverable forwards.*

Contribution margin The amount available for fixed costs and profit after paying variable costs; revenue minus variable costs.

Controlling shareholders A particular shareholder or block of shareholders holding a percentage of shares that gives them significant voting power.

Convenience sampling A procedure of selecting an element from a population on the basis of whether or not it is accessible to a researcher or how easy it is for a researcher to access the element.

Convenience yield A non-monetary advantage of holding an asset.

Conventional bond See *plain vanilla bond*.

Conventional cash flows A conventional cash flow pattern is one with an initial outflow followed by a series of inflows.

Convergence The tendency for differences in output per capita across countries to diminish over time. In technical analysis, the term describes the case when an indicator moves in the same manner as the security being analyzed.

Conversion premium The difference between the convertible bond's price and its conversion value.

Conversion price For a convertible bond, the price per share at which the bond can be converted into shares.

Conversion ratio For a convertible bond, the number of common shares that each bond can be converted into.

Conversion value For a convertible bond, the current share price multiplied by the conversion ratio.

Convertible bond Bond that gives the bondholder the right to exchange the bond for a specified number of common shares in the issuing company.

Convertible preference shares A type of equity security that entitles shareholders to convert their shares into a specified number of common shares.

Convexity adjustment For a bond, one half of the annual or approximate convexity statistic multiplied by the change in the yield-to-maturity squared.

Core inflation Refers to the inflation rate calculated based on a price index of goods and services except food and energy.

Corporate governance The system of internal controls and procedures by which individual companies are managed.

Correlation A measure of the linear relationship between two random variables.

Correlation coefficient A number between −1 and +1 that measures the consistency or tendency for two investments to act in a similar way. It is used to determine the effect on portfolio risk when two assets are combined.

Cost averaging The periodic investment of a fixed amount of money.

Cost of capital The cost of financing to a company; the rate of return that suppliers of capital require as compensation for their contribution of capital.

Cost of carry See *carry*.

Cost of debt The cost of debt financing to a company, such as when it issues a bond or takes out a bank loan; the rate of return required by a company's bondholders.

Cost of equity The rate of return required by a company's common stockholders.

Cost of preferred stock The cost to a company of issuing preferred stock; the dividend yield that a company must commit to pay preferred stockholders.

Cost structure The mix of a company's variable costs and fixed costs.

Counterparty risk The risk that the other party to a contract will fail to honor the terms of the contract.

Coupon rate The interest rate promised in a contract; this is the rate used to calculate the periodic interest payments.

Cournot assumption Assumption in which each firm determines its profit-maximizing production level assuming that the other firms' output will not change.

Covariance A measure of the co-movement (linear association) between two random variables.

Covariance matrix A matrix or square array whose entries are covariances; also known as a variance–covariance matrix.

Covenants The terms and conditions of lending agreements that the issuer must comply with; they specify the actions that an issuer is obligated to perform (affirmative covenant) or prohibited from performing (negative covenant).

Covered bond Debt obligation secured by a segregated pool of assets called the cover pool. The issuer must maintain the value of the cover pool. In the event of default, bondholders have recourse against both the issuer and the cover pool.

Covered bonds Senior debt obligations issued by a financial institution and backed by a segregated pool of assets. Covered bond investors have recourse to both the underlying assets and the financial institution.

Credit analysis The evaluation of credit risk; the evaluation of the creditworthiness of a borrower or counterparty.

Credit default swap (CDS) A type of credit derivative in which one party, the credit protection buyer who is seeking credit protection against a third party, makes a series of regularly scheduled payments to the other party, the credit protection seller. The seller makes no payments until a credit event occurs.

Credit derivatives A contract in which one party has the right to claim a payment from another party in the event that a specific credit event occurs over the life of the contract.

Credit enhancements Provisions that may be used to reduce the credit risk of a bond issue.

Credit-linked coupon bond Bond for which the coupon changes when the bond's credit rating changes.

Credit-linked note (CLN) Fixed-income security in which the holder of the security has the right to withhold payment of the full amount due at maturity if a credit event occurs.

Credit migration risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *downgrade risk*.

Credit risk The risk of loss caused by a counterparty's or debtor's failure to make a promised payment. Also called *default risk*.

Credit spread option An option on the yield spread on a bond.

Credit tranching A structure used to redistribute the credit risk associated with the collateral; a set of bond classes created to allow investors a choice in the amount of credit risk that they prefer to bear.

Creditworthiness The perceived ability of the borrower to pay its debt obligations in a timely manner; it represents the ability of a company to withstand adverse impacts on its cash flows.

Critical values Values of the test statistic at which the decision changes from fail to reject the null hypothesis to reject the null hypothesis.

Cross-default Covenant or contract clause that specifies a borrower is considered in default if they default on another debt obligation.

Cross-default provisions Provisions whereby events of default, such as non-payment of interest on one bond, trigger default on all outstanding debt; implies the same default probability for all issues.

Cross-price elasticity of demand The percentage change in quantity demanded for a given percentage change in the price of another good; the responsiveness of the demand for Product A that is associated with the change in price of Product B.

Cross-sectional analysis Analysis that involves comparisons across individuals in a group over a given time period or at a given point in time.

Cross-sectional data A list of the observations of a specific variable from multiple observational units at a given point in time. The observational units can be individuals, groups, companies, trading markets, regions, etc.

Crossing networks Trading systems that match buyers and sellers who are willing to trade at prices obtained from other markets.

Crowding out The thesis that government borrowing may divert private sector investment from taking place.

Cryptocurrency An electronic medium of exchange that lacks physical form.

Cryptography An algorithmic process to encrypt data, making the data unusable if received by unauthorized parties.

Cumulative absolute frequency Cumulates (i.e., adds up) in a frequency distribution the absolute frequencies as one moves from the first bin to the last bin.

Cumulative distribution function A function giving the probability that a random variable is less than or equal to a specified value.

Cumulative frequency distribution chart A chart that plots either the cumulative absolute frequency or the cumulative relative frequency on the y-axis against the upper limit of the interval and allows one to see the number or the percentage of the observations that lie below a certain value.

Cumulative preference shares Preference shares for which any dividends that are not paid accrue and must be paid in full before dividends on common shares can be paid.

Cumulative relative frequency A sequence of partial sums of the relative frequencies in a frequency distribution.

Cumulative voting A voting process whereby shareholders can accumulate and vote all their shares for a single candidate in an election, as opposed to having to allocate their voting rights evenly among all candidates.

Currencies Monies issued by national monetary authorities.

Currency option bonds Bonds that give bondholders the right to choose the currency in which they want to receive interest payments and principal repayments.

Currency swap A swap in which each party makes interest payments to the other in different currencies.

Current account A component of the balance of payments account that measures the flow of goods and services.

Current assets Assets that are expected to be consumed or converted into cash in the near future, typically one year or less. *Also called liquid assets.*

Current cost With reference to assets, the amount of cash or cash equivalents that would have to be paid to buy the same or an equivalent asset today; with reference to liabilities, the undiscounted amount of cash or cash equivalents that would be required to settle the obligation today.

Current government spending With respect to government expenditures, spending on goods and services that are provided on a regular, recurring basis including health, education, and defense.

Current liabilities Short-term obligations, such as accounts payable, wages payable, or accrued liabilities, that are expected to be settled in the near future, typically one year or less.

Current yield The sum of the coupon payments received over the year divided by the flat price; also called the *income* or *interest yield* or *running yield*.

Curve duration The sensitivity of the bond price (or the market value of a financial asset or liability) with respect to a benchmark yield curve.

Customs union Extends the free trade area (FTA) by not only allowing free movement of goods and services among members, but also creating a common trade policy against nonmembers.

CVaR Conditional VaR, a tail loss measure. The weighted average of all loss outcomes in the statistical distribution that exceed the VaR loss.

Cyclical A cyclical company is one whose profits are strongly correlated with the strength of the overall economy.

Cyclical companies Companies with sales and profits that regularly expand and contract with the business cycle or state of economy.

Daily settlement See *mark to market* and *marking to market.*

Dark pools Alternative trading systems that do not display the orders that their clients send to them.

Data A collection of numbers, characters, words, and text—as well as images, audio, and video—in a raw or organized format to represent facts or information.

Data mining The practice of determining a model by extensive searching through a dataset for statistically significant patterns. Also called *data snooping.*

Data science An interdisciplinary field that brings computer science, statistics, and other disciplines together to analyze and produce insights from Big Data.

Data snooping The practice of determining a model by extensive searching through a dataset for statistically significant patterns.

Data table see **two-dimensional rectangular array**.

Day order An order that is good for the day on which it is submitted. If it has not been filled by the close of business, the order expires unfilled.

Days of inventory on hand An activity ratio equal to the number of days in the period divided by inventory turnover over the period.

Dealers A financial intermediary that acts as a principal in trades.

Dealing securities Securities held by banks or other financial intermediaries for trading purposes.

Death cross See **Bearish crossover**.

Debentures Type of bond that can be secured or unsecured.

Debt rating An independent measure of the quality and safety of a company's debt based upon an analysis of the company's ability to pay the promised cash flows in full and on time. It includes an analysis of any indentures.

Debt-rating approach A method for estimating a company's before-tax cost of debt based on the yield on comparably rated bonds for maturities that closely match that of the company's existing debt.

Debt tax shield The tax benefit from interest paid on debt being tax deductible from income, equal to the marginal tax rate multiplied by the value of the debt.

Debt-to-assets ratio A solvency ratio calculated as total debt divided by total assets.

Debt-to-capital ratio A solvency ratio calculated as total debt divided by total debt plus total shareholders' equity.

Debt-to-equity ratio A solvency ratio calculated as total debt divided by total shareholders' equity.

Deciles Quantiles that divide a distribution into 10 equal parts.

Declaration date The day that the corporation issues a statement declaring a specific dividend.

Decreasing returns to scale When a production process leads to increases in output that are proportionately smaller than the increase in inputs.

Deductible temporary differences Temporary differences that result in a reduction of or deduction from taxable income in a future period when the balance sheet item is recovered or settled.

Deep learning Machine learning using neural networks with many hidden layers.

Deep learning nets Machine learning using neural networks with many hidden layers.

Default probability The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default risk*.

Default risk The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default probability*.

Default risk premium An extra return that compensates investors for the possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount.

Defensive companies Companies with sales and profits that have little sensitivity to the business cycle or state of the economy.

Defensive interval ratio A liquidity ratio that estimates the number of days that an entity could meet cash needs from liquid assets; calculated as (cash + short-term marketable investments + receivables) divided by daily cash expenditures.

Deferred coupon bond Bond that pays no coupons for its first few years but then pays a higher coupon than it otherwise normally would for the remainder of its life. Also called *split coupon bond*.

Deferred income A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred tax assets A balance sheet asset that arises when an excess amount is paid for income taxes relative to accounting profit. The taxable income is higher than accounting profit and income tax payable exceeds tax expense. The company expects to recover the difference during the course of future operations when tax expense exceeds income tax payable.

Deferred tax liabilities A balance sheet liability that arises when a deficit amount is paid for income taxes relative to accounting profit. The taxable income is less than the accounting profit and income tax payable is less than tax expense. The company expects to eliminate the liability over the course of future operations when income tax payable exceeds tax expense.

Defined benefit pension plans Plans in which the company promises to pay a certain annual amount (defined benefit) to the employee after retirement. The company bears the investment risk of the plan assets.

Defined contribution pension plans Individual accounts to which an employee and typically the employer makes contributions during their working years and expect to draw on the accumulated funds at retirement. The employee bears the investment and inflation risk of the plan assets.

Deflation Negative inflation.

Degree of confidence The probability that a confidence interval includes the unknown population parameter.

Degree of financial leverage (DFL) The ratio of the percentage change in net income to the percentage change in operating income; the sensitivity of the cash flows available to owners when operating income changes.

Degree of operating leverage (DOL) The ratio of the percentage change in operating income to the percentage change in units sold; the sensitivity of operating income to changes in units sold.

Degree of total leverage The ratio of the percentage change in net income to the percentage change in units sold; the sensitivity of the cash flows to owners to changes in the number of units produced and sold.

Degrees of freedom The number of independent variables used in defining sample statistics, such as variance, and the probability distributions they measure.

Delta The sensitivity of the derivative price to a small change in the value of the underlying asset.

Demand curve Graph of the inverse demand function. A graph showing the demand relation, either the highest quantity willingly purchased at each price or the highest price willingly paid for each quantity.

Demand function A relationship that expresses the quantity demanded of a good or service as a function of own-price and possibly other variables.

Demand shock A typically unexpected disturbance to demand, such as an unexpected interruption in trade or transportation.

Dependent With reference to events, the property that the probability of one event occurring depends on (is related to) the occurrence of another event.

Dependent variable The variable whose variation about its mean is to be explained by the regression; the left-side variable in a regression equation. Also referred to as the *explained variable*.

Depository bank A bank that raises funds from depositors and other investors and lends it to borrowers.

Depository institutions Commercial banks, savings and loan banks, credit unions, and similar institutions that raise funds from depositors and other investors and lend it to borrowers.

Depository receipt A security that trades like an ordinary share on a local exchange and represents an economic interest in a foreign company.

Depreciation The process of systematically allocating the cost of long-lived (tangible) assets to the periods during which the assets are expected to provide economic benefits.

Derivative pricing rule A pricing rule used by crossing networks in which a price is taken (derived) from the price that is current in the asset's primary market.

Derivatives A financial instrument whose value depends on the value of some underlying asset or factor (e.g., a stock price, an interest rate, or exchange rate).

Descriptive statistics Measures that summarize central tendency and spread variation in the data's distribution.

Diffuse prior The assumption of equal prior probabilities.

Diffusion index Reflects the proportion of the index's components that are moving in a pattern consistent with the overall index.

Diluted EPS The EPS that would result if all dilutive securities were converted into common shares.

Diluted shares The number of shares that would be outstanding if all potentially dilutive claims on common shares (e.g., convertible debt, convertible preferred stock, and employee stock options) were exercised.

Diminishing balance method An accelerated depreciation method, i.e., one that allocates a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Diminishing marginal productivity Describes a state in which each additional unit of input produces less output than previously.

Direct format With reference to the cash flow statement, a format for the presentation of the statement in which cash flow from operating activities is shown as operating cash receipts less operating cash disbursements. Also called *direct method*.

Direct investing Occurs when an investor makes a direct investment in an asset without the use of an intermediary.

Direct method See *direct format.*

Direct taxes Taxes levied directly on income, wealth, and corporate profits.

Direct write-off method An approach to recognizing credit losses on customer receivables in which the company waits until such time as a customer has defaulted and only then recognizes the loss.

Discount To reduce the value of a future payment in allowance for how far away it is in time; to calculate the present value of some future amount. Also, the amount by which an instrument is priced below its face (par) value.

Discount margin See *required margin.*

Discount rates In general, the interest rates used to calculate present values. In the money market, however, a discount rate is a specific type of quoted rate.

Discounted cash flow models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security.

Discouraged worker A person who has stopped looking for a job or has given up seeking employment.

Discrete data Numerical values that result from a counting process; therefore, practically speaking, the data are limited to a finite number of values.

Discrete random variable A random variable that can take on at most a countable number of possible values.

Discriminatory pricing rule A pricing rule used in continuous markets in which the limit price of the order or quote that first arrived determines the trade price.

Diseconomies of scale Increase in cost per unit resulting from increased production.

Dispersion The variability of a population or sample of observations around the central tendency.

Display size The size of an order displayed to public view.

Disposition effect As a result of loss aversion, an emotional bias whereby investors are reluctant to dispose of losers. This reluctance results in an inefficient and gradual adjustment to deterioration in fundamental value.

Distributed ledger A type of database that may be shared among entities in a network.

Distributed ledger technology Technology based on a distributed ledger.

Divergence In technical analysis, a term that describes the case when an indicator moves differently from the security being analyzed.

Diversification ratio The ratio of the standard deviation of an equally weighted portfolio to the standard deviation of a randomly selected security.

Dividend A distribution paid to shareholders based on the number of shares owned.

Dividend discount model (DDM) A present value model that estimates the intrinsic value of an equity share based on the present value of its expected future dividends.

Dividend payout ratio The ratio of cash dividends paid to earnings for a period.

Divisor A number (denominator) used to determine the value of a price return index. It is initially chosen at the inception of an index and subsequently adjusted by the index provider, as necessary, to avoid changes in the index value that are unrelated to changes in the prices of its constituent securities.

Doji In the Japanese terminology used in candlestick charting, the doji signifies that after a full day of trading, the positive price influence of buyers and the negative price influence of sellers exactly counteracted each other—with opening and closing prices that are virtually equal—which suggests that the market under analysis is in balance.

Domestic content provisions Stipulate that some percentage of the value added or components used in production should be of domestic origin.

Double bottom In technical analysis, a reversal pattern that is formed when the price reaches a low, rebounds, and then declines back to the first low level. A double bottom is used to predict a change from a downtrend to an uptrend.

Double coincidence of wants A prerequisite to barter trades, in particular that both economic agents in the transaction want what the other is selling.

Double declining balance depreciation An accelerated depreciation method that involves depreciating the asset at double the straight-line rate. This rate is multiplied by the book value of the asset at the beginning of the period (a declining balance) to calculate depreciation expense.

Double top In technical analysis, a reversal pattern that is formed when an uptrend reverses twice at roughly the same high price level. A double top is used to predict a change from an uptrend to a downtrend.

Down transition probability The probability that an asset's value moves down in a model of asset price dynamics.

Downgrade risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *credit migration risk.*

Downside risk Risk of incurring returns below a specified value.

Downtrend A pattern that occurs when the price of an asset moves lower over a period of time.

Drag on liquidity When receipts lag, creating pressure from the decreased available funds.

Drawdown A percentage peak-to-trough reduction in net asset value.

Dual-currency bonds Bonds that make coupon payments in one currency and pay the par value at maturity in another currency.

DuPont analysis An approach to decomposing return on investment, e.g., return on equity, as the product of other financial ratios.

Duration A measure of the approximate sensitivity of a security to a change in interest rates (i.e., a measure of interest rate risk).

Duration gap A bond's Macaulay duration minus the investment horizon.

Dutch Book Theorem A result in probability theory stating that inconsistent probabilities create profit opportunities.

Early repayment option See *prepayment option*.

Earnings per share The amount of income earned during a period per share of common stock.

Earnings surprise The portion of a company's earnings that is unanticipated by investors and, according to the efficient market hypothesis, merits a price adjustment.

Economic costs All the remuneration needed to keep a productive resource in its current employment or to acquire the resource for productive use; the sum of total accounting costs and implicit opportunity costs.

Economic indicator A variable that provides information on the state of the overall economy.

Economic loss The amount by which accounting profit is less than normal profit.

Economic profit Equal to accounting profit less the implicit opportunity costs not included in total accounting costs; the difference between total revenue (TR) and total cost (TC). Also called *abnormal profit* or *supernormal profit*.

Economic stabilization Reduction of the magnitude of economic fluctuations.

Economic union Incorporates all aspects of a common market and in addition requires common economic institutions and coordination of economic policies among members.

Economies of scale Reduction in cost per unit resulting from increased production.

Effective annual rate The amount by which a unit of currency will grow in a year with interest on interest included.

Effective convexity A *curve convexity* statistic that measures the secondary effect of a change in a benchmark yield curve on a bond's price.

Effective duration The sensitivity of a bond's price to a change in a benchmark yield curve.

Effective interest rate The borrowing rate or market rate that a company incurs at the time of issuance of a bond.

Efficient market A market in which asset prices reflect new information quickly and rationally.

Elastic Said of a good or service when the magnitude of elasticity is greater than one.

Elasticity The percentage change in one variable for a percentage change in another variable; a general measure of how sensitive one variable is to a change in the value of another variable.

Elasticity of demand A measure of the sensitivity of quantity demanded to a change in a product's own price: $\%\Delta Q^D/\%\Delta P$.

Elasticity of supply A measure of the sensitivity of quantity supplied to a change in price: $\%\Delta Q^S/\%\Delta P$.

Electronic communications networks See *alternative trading systems*.

Embedded option Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Embedded options Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Emotional biases Behavioral biases resulting from reasoning influenced by feelings; emotional biases stem from impulse or intuition.

Empirical duration The use of statistical methods and historical bond prices to estimate the price–yield relationship for a specific bond or portfolio of bonds.

Empirical probability The probability of an event estimated as a relative frequency of occurrence.

Employed The number of people with a job.

Endowment bias An emotional bias in which people value an asset more when they hold rights to it than when they do not.

Engagement/active ownership An ESG investment approach that uses shareholder power to influence corporate behavior through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is directed by ESG guidelines.

Enterprise risk management An overall assessment of a company's risk position. A centralized approach to risk management sometimes called firmwide risk management.

Enterprise value A measure of a company's total market value from which the value of cash and short-term investments have been subtracted.

Equal weighting An index weighting method in which an equal weight is assigned to each constituent security at inception.

Equipment trust certificates Bonds secured by specific types of equipment or physical assets.

Equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Equity risk premium The expected return on equities minus the risk-free rate; the premium that investors demand for investing in equities.

Equity swap A swap transaction in which at least one cash flow is tied to the return to an equity portfolio position, often an equity index.

Error term The difference between an observation and its expected value, where the expected value is based on the true underlying population relation between the dependent and independent variables. Also known simply as the *error*.

ESG An acronym that encompasses environmental, social, and governance.

ESG integration An ESG investment approach that focuses on systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions for the purpose of achieving the product's stated investment objectives. Used interchangeably with **ESG investing**.

ESG investing The consideration of environmental, social, and governance factors in the investment process. Used interchangeably with **ESG integration**.

Estimate The particular value calculated from sample observations using an estimator.

Estimated parameters With reference to a regression analysis, the estimated values of the population intercept and population slope coefficients in a regression.

Estimator An estimation formula; the formula used to compute the sample mean and other sample statistics are examples of estimators.

Ethical principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Ethics The study of moral principles or of making good choices. Ethics encompasses a set of moral principles and rules of conduct that provide guidance for our behavior.

Eurobonds Type of bond issued internationally, outside the jurisdiction of the country in whose currency the bond is denominated.

European-style Said of an option contract that can only be exercised on the option's expiration date.

Event Any outcome or specified set of outcomes of a random variable.

Ex-dividend date The first date that a share trades without (i.e., "ex") the dividend.

Excess kurtosis Degree of kurtosis (fatness of tails) relative to the kurtosis of the normal distribution.

Exchanges Places where traders can meet to arrange their trades.

Execution instructions Instructions that indicate how to fill an order.

Exercise The process of using an option to buy or sell the underlying.

Exercise price The fixed price at which an option holder can buy or sell the underlying. Also called *strike price*, *striking price*, or *strike*.

Exercise value The value obtained if an option is exercised based on current conditions. Also known as *intrinsic value*.

Exhaustive Covering or containing all possible outcomes.

Expansion The period of a business cycle after its lowest point and before its highest point.

Expansionary Tending to cause the real economy to grow.

Expansionary fiscal policy Fiscal policy aimed at achieving real economic growth.

Expected inflation The level of inflation that economic agents expect in the future.

Expected loss Default probability times loss severity given default.

Expected return on the portfolio ($E(R_p)$) The weighted average of the expected returns (R_1 to R_n) on the component securities using their respective weights (w_1 to w_n).

Expected value The probability-weighted average of the possible outcomes of a random variable.

Expenses Outflows of economic resources or increases in liabilities that result in decreases in equity (other than decreases because of distributions to owners); reductions in net assets associated with the creation of revenues.

Export subsidy Paid by the government to the firm when it exports a unit of a good that is being subsidized.

Exports Goods and services that an economy sells to other countries.

Extension risk The risk that when interest rates rise, fewer prepayments will occur because homeowners are reluctant to give up the benefits of a contractual interest rate that now looks low. As a result, the security becomes longer in maturity than anticipated at the time of purchase.

Externalities Spillover effects of production and consumption activities onto others who are not directly involved in a particular transaction, activity, or decision.

Externality An effect of a market transaction that is borne by parties other than those who transacted.

Extra dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Extreme value theory A branch of statistics that focuses primarily on extreme outcomes.

Face value The amount of cash payable by a company to the bondholders when the bonds mature; the promised payment at maturity separate from any coupon payment.

Factoring arrangement When a company sells its accounts receivable to a lender (known as a factor) who assumes responsibility for the credit granting and collection process.

Fair value The amount at which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's-length transaction; the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants.

False discovery approach An adjustment in the *p*-values for tests performed multiple times.

False discovery rate The rate of Type I errors in testing a null hypothesis multiple times for a given level of significance.

Fat-Tailed Describes a distribution that has fatter tails than a normal distribution (also called leptokurtic).

Fed funds rate The US interbank lending rate on overnight borrowings of reserves.

Federal funds rate The US interbank lending rate on overnight borrowings of reserves.

Fiat money Money that is not convertible into any other commodity.

Fiduciary An entity designated to represent the rights and responsibilities of a beneficiary whose assets they are managing, such as a bond trustee acting on behalf of fixed-income investors.

Fiduciary call A combination of a European call and a risk-free bond that matures on the option expiration day and has a face value equal to the exercise price of the call.

FIFO method The first in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the order in which they were placed in inventory.

Fill or kill See *immediate or cancel order*.

Finance lease From the lessee perspective, under US GAAP, a type of lease which is more akin to the purchase of an asset by the lessee. From the lessor perspective, under IFRS, a lease which "transfers substantially all the risks and rewards incidental to ownership of an underlying asset."

Financial account A component of the balance of payments account that records investment flows.

Financial distress Heightened uncertainty regarding a company's ability to meet its various obligations because of diminished earnings power or actual current losses.

Financial flexibility The ability to react and adapt to financial adversity and opportunities.

Financial leverage The extent to which a company can effect, through the use of debt, a proportional change in the return on common equity that is greater than a given proportional change in operating income; also, short for the financial leverage ratio.

Financial leverage ratio A measure of financial leverage calculated as average total assets divided by average total equity.

Financial risk The risk that environmental, social, or governance risk factors will result in significant costs or other losses to a company and its shareholders; the risk arising from a company's obligation to meet required payments under its financing agreements.

Financing activities Activities related to obtaining or repaying capital to be used in the business (e.g., equity and long-term debt).

Fintech Technological innovation in the design and delivery of financial services and products in the financial industry.

Firm commitment offering See *underwritten offering*.

First-degree price discrimination Where a monopolist is able to charge each customer the highest price the customer is willing to pay.

First lien debt Debt secured by a pledge of certain assets that could include buildings, but it may also include property and equipment, licenses, patents, brands, etc.

First mortgage debt Debt secured by a pledge of a specific property.

Fiscal multiplier The ratio of a change in national income to a change in government spending.

Fiscal policy The use of taxes and government spending to affect the level of aggregate expenditures.

Fisher effect The thesis that the real rate of interest in an economy is stable over time so that changes in nominal interest rates are the result of changes in expected inflation.

Fixed charge coverage A solvency ratio measuring the number of times interest and lease payments are covered by operating income, calculated as (EBIT + lease payments) divided by (interest payments + lease payments).

Fixed costs Costs that remain at the same level regardless of a company's level of production and sales.

Fixed-for-floating interest rate swap An interest rate swap in which one party pays a fixed rate and the other pays a floating rate, with both sets of payments in the same currency. Also called *plain vanilla swap* or *vanilla swap*.

Fixed-rate perpetual preferred stock Nonconvertible, noncallable preferred stock that has a fixed dividend rate and no maturity date.

Flag A technical analysis continuation pattern formed by parallel trendlines, typically over a short period.

Flat price The full price of a bond minus the accrued interest; also called the *quoted* or *clean* price.

Float-adjusted market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by adjusting its market capitalization for its market float.

Floaters See *floating-rate notes*.

Floating-rate notes Notes on which interest payments are not fixed but instead vary from period to period depending on the current level of a reference interest rate.

Flotation cost Fees charged to companies by investment bankers and other costs associated with raising new capital.

Foreclosure Allows the lender to take possession of a mortgaged property if the borrower defaults and then sell it to recover funds.

Foreign currency reserves Holding by the central bank of non-domestic currency deposits and non-domestic bonds.

Foreign direct investment Direct investment by a firm in one country (the source country) in productive assets in a foreign country (the host country).

Foreign exchange gains (or losses) Gains (or losses) that occur when the exchange rate changes between the investor's currency and the currency that foreign securities are denominated in.

Foreign portfolio investment Shorter-term investment by individuals, firms, and institutional investors (e.g., pension funds) in foreign financial instruments such as foreign stocks and foreign government bonds.

Forward commitments Class of derivatives that provides the ability to lock in a price to transact in the future at a previously agreed-upon price.

Forward contract An agreement between two parties in which one party, the buyer, agrees to buy from the other party, the seller, an underlying asset at a later date for a price established at the start of the contract.

Forward curve A series of forward rates, each having the same time frame.

Forward market For future delivery, beyond the usual settlement time period in the cash market.

Forward price The fixed price or rate at which the transaction scheduled to occur at the expiration of a forward contract will take place. This price is agreed on at the initiation date of the contract.

Forward rate The interest rate on a bond or money market instrument traded in a forward market. A forward rate can be interpreted as an incremental, or marginal, return for extending the time-to-maturity for an additional time period.

Forward rate agreements A forward contract calling for one party to make a fixed interest payment and the other to make an interest payment at a rate to be determined at the contract expiration.

Fractile A value at or below which a stated fraction of the data lies. Also called quantile.

Fractional reserve banking Banking in which reserves constitute a fraction of deposits.

Framing bias An information-processing bias in which a person answers a question differently based on the way in which it is asked (framed).

Free cash flow The actual cash that would be available to the company's investors after making all investments necessary to maintain the company as an ongoing enterprise (also referred to as free cash flow to the firm); the internally generated funds that can be distributed to the company's investors (e.g., shareholders and bondholders) without impairing the value of the company.

Free cash flow hypothesis The hypothesis that higher debt levels discipline managers by forcing them to make fixed debt service payments and by reducing the company's free cash flow.

Free-cash-flow-to-equity models Valuation models based on discounting expected future free cash flow to equity.

Free float The number of shares that are readily and freely tradable in the secondary market.

Free trade When there are no government restrictions on a country's ability to trade.

Free trade areas One of the most prevalent forms of regional integration, in which all barriers to the flow of goods and services among members have been eliminated.

Frequency distribution A tabular display of data constructed either by counting the observations of a variable by distinct values or groups or by tallying the values of a numerical variable into a set of numerically ordered bins (also called a one-way table).

Frequency polygon A graph of a frequency distribution obtained by drawing straight lines joining successive points representing the class frequencies.

Frequency table A representation of the frequency of occurrence of two discrete variables.

Full price The price of a security with accrued interest; also called the *invoice* or *dirty* price.

Fund investing In fund investing, the investor invests in assets indirectly by contributing capital to a fund as part of a group of investors. Fund investing is available for all major alternative investment types.

Fundamental analysis The examination of publicly available information and the formulation of forecasts to estimate the intrinsic value of assets.

Fundamental value The underlying or true value of an asset based on an analysis of its qualitative and quantitative characteristics. Also called *intrinsic value*.

Fundamental weighting An index weighting method in which the weight assigned to each constituent security is based on its underlying company's size. It attempts to address the disadvantages of market-capitalization weighting by using measures that are independent of the constituent security's price.

Funds of hedge funds Funds that hold a portfolio of hedge funds, more commonly shortened to *funds of funds*.

Fungible Freely exchangeable, interchangeable, or substitutable with other things of the same type. Money and commodities are the most common examples.

Future value (FV) The amount to which a payment or series of payments will grow by a stated future date.

Futures contract A variation of a forward contract that has essentially the same basic definition but with some additional features, such as a clearinghouse guarantee against credit losses, a daily settlement of gains and losses, and an organized electronic or floor trading facility.

Futures price The agreed-upon price of a futures contract.

FX swap The combination of a spot and a forward FX transaction.

G-spread The yield spread in basis points over an actual or interpolated government bond.

Gains Asset inflows not directly related to the ordinary activities of the business.

Game theory The set of tools decision makers use to incorporate responses by rival decision makers into their strategies.

Gamma A numerical measure of how sensitive an option's delta (the sensitivity of the derivative's price) is to a change in the value of the underlying.

Gap opening A gap is an area of a chart where a security's price either rises or falls from the previous day's close with no trading occurring in between. A gap opening is the start of a new trading session with a gap.

GDP deflator A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

General partner (GP) The partner that runs the business and theoretically bears unlimited liability for the business's debts and obligations.

Geometric mean A measure of central tendency computed by taking the nth root of the product of n non-negative values.

Giffen goods Goods that are consumed more as the price of the good rises because it is a very inferior good whose income effect overwhelms its substitution effect when price changes.

Gilts Bonds issued by the UK government.

Global depository receipt A depository receipt that is issued outside of the company's home country and outside of the United States.

Global minimum-variance portfolio The portfolio on the minimum-variance frontier with the smallest variance of return.

Global registered share A common share that is traded on different stock exchanges around the world in different currencies.

Gold standard With respect to a currency, if a currency is on the gold standard a given amount can be converted into a prespecified amount of gold.

Golden Cross See **Bullish crossover**.

Good-on-close An execution instruction specifying that an order can only be filled at the close of trading. Also called *market on close*.

Good-on-open An execution instruction specifying that an order can only be filled at the opening of trading.

Good-till-cancelled order An order specifying that it is valid until the entity placing the order has cancelled it (or, commonly, until some specified amount of time such as 60 days has elapsed, whichever comes sooner).

Goodwill An intangible asset that represents the excess of the purchase price of an acquired company over the value of the net assets acquired.

Government equivalent yield A yield that restates a yield-to-maturity based on a 30/360 day count to one based on actual/actual.

Green bonds Bonds used in green finance whereby the proceeds are earmarked toward environmental-related products.

Green finance A type of finance that addresses environmental concerns while achieving economic growth.

Green loans Any loan instruments made available exclusively to finance or re-finance, in whole or in part, new and/or existing eligible green projects. Green loans are commonly aligned in the market with the Green Loan Principles.

Greenfield investment Investing in infrastructure assets *that are to be constructed*.

Grey market The forward market for bonds about to be issued. Also called "when issued" market.

Gross domestic product The market value of all final goods and services produced within the economy during a given period (output definition) or, equivalently, the aggregate income earned by all households, all companies, and the government within the economy during a given period (income definition).

Gross margin Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit margin The ratio of gross profit to revenues.

Grouped bar chart A bar chart for showing joint frequencies for two categorical variables (also known as a **clustered bar chart**).

Grouping by function With reference to the presentation of expenses in an income statement, the grouping together of expenses serving the same function, e.g. all items that are costs of goods sold.

Grouping by nature With reference to the presentation of expenses in an income statement, the grouping together of expenses by similar nature, e.g., all depreciation expenses.

Growth cycle Refers to fluctuations in economic activity around the long-term potential trend growth level, focusing on how much actual economic activity is below or above trend growth in economic activity.

Growth cyclical A term sometimes used to describe companies that are growing rapidly on a long-term basis but that still experience above-average fluctuation in their revenues and profits over the course of a business cycle.

Growth investors With reference to equity investors, investors who seek to invest in high-earnings-growth companies.

Growth option The ability to make additional investments in a project at some future time if the financial results are strong. Also called *expansion option*.

Growth rate cycle Refers to fluctuations in the growth rate of economic activity.

Guarantee certificate A type of structured financial instrument that provides investors with capital protection. It combines a zero-coupon bond and a call option on some underlying asset.

Haircut See *repo margin*.

Halo effect An emotional bias that extends a favorable evaluation of some characteristics to other characteristics.

Hard-bullet covered bonds Covered bonds for which a bond default is triggered and bond payments are accelerated in the event of sponsor default if payments do not occur according to the original maturity schedule.

Harmonic mean A type of weighted mean computed as the reciprocal of the arithmetic average of the reciprocals.

Head and shoulders pattern In technical analysis, a reversal pattern that is formed in three parts: a left shoulder, a head, and a right shoulder. A head and shoulders pattern is used to predict a change from an uptrend to a downtrend.

Headline inflation Refers to the inflation rate calculated based on the price index that includes all goods and services in an economy.

Heat map A type of graphic that organizes and summarizes data in a tabular format and represents it using a color spectrum.

Hedge funds Private investment vehicles that typically use leverage, derivatives, and long and short investment strategies.

Hedge portfolio A hypothetical combination of the derivative and its underlying that eliminates risk.

Held-to-maturity Debt (fixed-income) securities that a company intends to hold to maturity; these are presented at their original cost, updated for any amortisation of discounts or premiums.

Herding Clustered trading that may or may not be based on information.

Heteroskedasticity A violation of the simple linear regression assumption of homoskedasticity; with heteroskedasticity, the variability of the residuals is not equal across all observations.

Hidden order An order that is exposed not to the public but only to the brokers or exchanges that receive it.

High-frequency trading A form of algorithmic trading that makes use of vast quantities of data to execute trades on ultra-high-speed networks in fractions of a second.

High-water mark The highest value, net of fees, which a fund has reached in history. It reflects the highest cumulative return used to calculate an incentive fee.

Hindsight bias A bias with selective perception and retention aspects in which people may see past events as having been predictable and reasonable to expect.

Histogram A chart that presents the distribution of numerical data by using the height of a bar or column to represent the absolute frequency of each bin or interval in the distribution.

Historical cost In reference to assets, the amount paid to purchase an asset, including any costs of acquisition and/or preparation; with reference to liabilities, the amount of proceeds received in exchange in issuing the liability.

Historical equity risk premium approach An estimate of a country's equity risk premium that is based on the historical averages of the risk-free rate and the rate of return on the market portfolio.

Holder-of-record date The date that a shareholder listed on the corporation's books will be deemed to have ownership of the shares for purposes of receiving an upcoming dividend.

Holding period return The return that an investor earns during a specified holding period; a synonym for total return.

Home bias A preference for securities listed on the exchanges of one's home country.

Homogeneity of expectations The assumption that all investors have the same economic expectations and thus have the same expectations of prices, cash flows, and other investment characteristics.

Homoskedasticity An assumption of simple linear regression such that the variability of the residuals is equal across all observations.

Horizon yield The internal rate of return between the total return (the sum of reinvested coupon payments and the sale price or redemption amount) and the purchase price of the bond.

Horizontal analysis Common-size analysis that involves comparing a specific financial statement with that statement in prior or future time periods; also, cross-sectional analysis of one company with another.

Horizontal demand schedule Implies that at a given price, the response in the quantity demanded is infinite.

Hostile takeover An attempt by one entity to acquire a company without the consent of the company's management.

Household A person or a group of people living in the same residence, taken as a basic unit in economic analysis.

Human capital The accumulated knowledge and skill that workers acquire from education, training, or life experience and the corresponding present value of future earnings to be generated by said skilled individual.

Hurdle rate The minimum rate of return on investment that a fund must reach before a GP receives carried interest.

Hypothesis A proposed explanation or theory that can be tested.

Hypothesis testing The process of testing of hypotheses about one or more populations using statistical inference.

I-spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Iceberg order An order in which the display size is less than the order's full size.

If-converted method A method for accounting for the effect of convertible securities on earnings per share (EPS) that specifies what EPS would have been if the convertible securities had been converted at the beginning of the period, taking account of the effects of conversion on net income and the weighted average number of shares outstanding.

Illusion of control bias A bias in which people tend to believe that they can control or influence outcomes when, in fact, they cannot.

Immediate or cancel order An order that is valid only upon receipt by the broker or exchange. If such an order cannot be filled in part or in whole upon receipt, it cancels immediately. Also called *fill or kill*.

Impact investing Investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.

Impact lag The lag associated with the result of actions affecting the economy with delay.

Implicit price deflator for GDP A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

Implicit selection bias One type of selection bias introduced through the presence of a threshold that filters out some unqualified members.

Implied forward rates Calculated from spot rates, an implied forward rate is a breakeven reinvestment rate that links the return on an investment in a shorter-term zero-coupon bond to the return on an investment in a longer-term zero-coupon bond.

Implied volatility The volatility that option traders use to price an option, implied by the price of the option and a particular option-pricing model.

Import license Specifies the quantity of a good that can be imported into a country.

Imports Goods and services that a domestic economy (i.e., house-holds, firms, and government) purchases from other countries.

In the money Options that, if exercised, would result in the value received being worth more than the payment required to exercise.

Income Increases in economic benefits in the form of inflows or enhancements of assets, or decreases of liabilities that result in an increase in equity (other than increases resulting from contributions by owners).

Income elasticity of demand A measure of the responsiveness of demand to changes in income, defined as the percentage change in quantity demanded divided by the percentage change in income.

Income tax paid The actual amount paid for income taxes in the period; not a provision, but the actual cash outflow.

Income tax payable The income tax owed by the company on the basis of taxable income.

Increasing marginal returns When the marginal product of a resource increases as additional units of that input are employed.

Increasing returns to scale When a production process leads to increases in output that are proportionately larger than the increase in inputs.

Incremental cash flow The cash flow that is realized because of a decision; the changes or increments to cash flows resulting from a decision or action.

Indenture Legal contract that describes the form of a bond, the obligations of the issuer, and the rights of the bondholders. Also called the *trust deed*.

Independent With reference to events, the property that the occurrence of one event does not affect the probability of another event occurring. With reference to two random variables X and Y, they are independent if and only if $P(X,Y) = P(X)P(Y)$.

Independent projects Independent projects are capital investments whose cash flows are independent of each other.

Independent variable A variable used to explain the dependent variable in a regression; a right-side variable in a regression equation. Also referred to as the *explanatory variable*.

Independently and identically distributed With respect to random variables, the property of random variables that are independent of each other but follow the identical probability distribution.

Index-linked bond Bond for which coupon payments and/or principal repayment are linked to a specified index.

Indexing An investment strategy in which an investor constructs a portfolio to mirror the performance of a specified index.

Indicator variable A variable that takes on only one of two values, 0 or 1, based on a condition. In simple linear regression, the slope is the difference in the dependent variable for the two conditions. Also referred to as a *dummy variable*.

Indifference curve A curve representing all the combinations of two goods or attributes such that the consumer is entirely indifferent among them.

Indirect format With reference to cash flow statements, a format for the presentation of the statement which, in the operating cash flow section, begins with net income then shows additions and subtractions to arrive at operating cash flow. Also called *indirect method*.

Indirect method See *indirect format*.

Indirect taxes Taxes such as taxes on spending, as opposed to direct taxes.

Industry A group of companies offering similar products and/or services.

Industry analysis The analysis of a specific branch of manufacturing, service, or trade.

Inelastic Said of a good or service when the magnitude of elasticity is less than one. Insensitive to price changes.

Inferior goods A good whose consumption decreases as income increases.

Inflation The percentage increase in the general price level from one period to the next; a sustained rise in the overall level of prices in an economy.

Inflation-linked bond Type of index-linked bond that offers investors protection against inflation by linking the bond's coupon payments and/or the principal repayment to an index of consumer prices. Also called *linkers*.

Inflation premium An extra return that compensates investors for expected inflation.

Inflation reports A type of economic publication put out by many central banks.

Inflation uncertainty The degree to which economic agents view future rates of inflation as difficult to forecast.

Information cascade The transmission of information from those participants who act first and whose decisions influence the decisions of others.

Information-motivated traders Traders that trade to profit from information that they believe allows them to predict future prices.

Informationally efficient market A market in which asset prices reflect new information quickly and rationally.

Initial coin offering An unregulated process whereby companies raise capital by selling crypto tokens to investors in exchange for fiat money or another agreed-upon cryptocurrency.

Initial margin The amount that must be deposited in a clearinghouse account when entering into a futures contract.

Initial margin requirement The margin requirement on the first day of a transaction as well as on any day in which additional margin funds must be deposited.

Initial public offering (IPO) The first issuance of common shares to the public by a formerly private corporation.

Input productivity The amount of output produced by workers in a given period of time—for example, output per hour worked; measures the efficiency of labor.

Intangible assets Assets lacking physical substance, such as patents and trademarks.

Interbank market The market of loans and deposits between banks for maturities ranging from overnight to one year.

Intercept The expected value of the dependent variable when the independent variable in a simple linear regression is equal to zero.

Interest Payment for lending funds.

Interest coverage A solvency ratio calculated as EBIT divided by interest payments.

Interest-only mortgage A loan in which no scheduled principal repayment is specified for a certain number of years.

Interest rate A rate of return that reflects the relationship between differently dated cash flows; a discount rate.

Interest rate effect The effect through which price level changes, through demand for money, impact interest rate, which in turn impacts investment and consumption.

Interest rate swap A swap in which the underlying is an interest rate. Can be viewed as a currency swap in which both currencies are the same and can be created as a combination of currency swaps.

Intermarket analysis A field within technical analysis that combines analysis of the major categories of securities—namely, equities, bonds, currencies, and commodities—to identify market trends and possible inflections in trends.

Internal rate of return The discount rate that makes net present value equal 0; the discount rate that makes the present value of an investment's costs (outflows) equal to the present value of the investment's benefits (inflows).

Internet of Things A network arrangement of structures and devices whereby the objects on the network are able to interact and share information.

Interpolated spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Interquartile range The difference between the third and first quartiles of a dataset.

Interval With reference to grouped data, a set of values within which an observation falls.

Intrinsic value See *exercise value*.

Inventory investment Net change in business inventory.

Inventory turnover An activity ratio calculated as cost of goods sold divided by average inventory.

Inverse demand function A restatement of the demand function in which price is stated as a function of quantity.

Inverse floater A type of leveraged structured financial instrument. The cash flows are adjusted periodically and move in the opposite direction of changes in the reference rate.

Inverse transformation method A method using randomly generated numbers from the continuous uniform distribution to generate random observations from any distribution.

Investing activities Activities associated with the acquisition and disposal of property, plant, and equipment; intangible assets; other long-term assets; and both long-term and short-term investments in the equity and debt (bonds and loans) issued by other companies.

Investment banks Financial intermediaries that provide advice to their mostly corporate clients and help them arrange transactions such as initial and seasoned securities offerings.

Investment policy statement (IPS) A written planning document that describes a client's investment objectives and risk tolerance over a relevant time horizon, along with constraints that apply to the client's portfolio.

Investment property Property used to earn rental income or capital appreciation (or both).

Jackknife A resampling method that repeatedly draws samples by taking the original observed data sample and leaving out one observation at a time (without replacement) from the set.

January effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January. Also called *turn-of-the-year effect.*

Joint frequencies The entry in the cells of the contingency table that represent the joining of one variable from a row and the other variable from a column to count observations.

Joint probability The probability of the joint occurrence of stated events.

Joint probability function A function giving the probability of joint occurrences of values of stated random variables.

Judgmental Sampling A procedure of selectively handpicking elements from the population based on a researcher's knowledge and professional judgment.

Key rate duration A method of measuring the interest rate sensitivities of a fixed-income instrument or portfolio to shifts in key points along the yield curve.

Keynesians Economists who believe that fiscal policy can have powerful effects on aggregate demand, output, and employment when there is substantial spare capacity in an economy.

Kurtosis The statistical measure that indicates the combined weight of the tails of a distribution relative to the rest of the distribution.

Labor force The portion of the working age population (over the age of 16) that is employed or is available for work but not working (unemployed).

Labor productivity The quantity of goods and services (real GDP) that a worker can produce in one hour of work.

Lagging economic indicators Turning points that take place later than those of the overall economy; they are believed to have value in identifying the economy's past condition.

Law of demand The principle that as the price of a good rises, buyers will choose to buy less of it, and as its price falls, they will buy more.

Law of diminishing marginal returns The observation that a variable factor's marginal product must eventually fall as more of it is added to a fixed amount of the other factors.

Law of diminishing returns The smallest output that a firm can produce such that its long run average costs are minimized.

Law of one price The condition in a financial market in which two equivalent financial instruments or combinations of financial instruments can sell for only one price. Equivalent to the principle that no arbitrage opportunities are possible.

Lead underwriter The lead investment bank in a syndicate of investment banks and broker–dealers involved in a securities underwriting.

Leading economic indicators Turning points that usually precede those of the overall economy; they are believed to have value for predicting the economy's future state, usually near-term.

Lease A contract that conveys the right to use an asset for a period of time in exchange for consideration.

Legal tender Something that must be accepted when offered in exchange for goods and services.

Lender of last resort An entity willing to lend money when no other entity is ready to do so.

Leptokurtic Describes a distribution that has fatter tails than a normal distribution (also called fat-tailed).

Lessee The party obtaining the use of an asset through a lease.

Lessor The owner of an asset that grants the right to use the asset to another party.

Letter of credit Form of external credit enhancement whereby a financial institution provides the issuer with a credit line to reimburse any cash flow shortfalls from the assets backing the issue.

Level of significance The probability of a Type I error in testing a hypothesis.

Leverage In the context of corporate finance, leverage refers to the use of fixed costs within a company's cost structure. Fixed costs that are operating costs (such as depreciation or rent) create operating leverage. Fixed costs that are financial costs (such as interest expense) create financial leverage.

Leveraged buyout A transaction whereby the target company's management team converts the target to a privately held company by using heavy borrowing to finance the purchase of the target company's outstanding shares.

Liabilities Present obligations of an enterprise arising from past events, the settlement of which is expected to result in an outflow of resources embodying economic benefits; creditors' claims on the resources of a company.

LIFO layer liquidation With respect to the application of the LIFO inventory method, the liquidation of old, relatively low-priced inventory; happens when the volume of sales rises above the volume of recent purchases so that some sales are made from relatively old, low-priced inventory. Also called *LIFO liquidation*.

LIFO method The last in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the reverse order the items were placed in inventory (i.e., inventory produced or acquired last are assumed to be sold first).

LIFO reserve The difference between the reported LIFO inventory carrying amount and the inventory amount that would have been reported if the FIFO method had been used (in other words, the FIFO inventory value less the LIFO inventory value).

Likelihood The probability of an observation, given a particular set of conditions.

Limit down A limit move in the futures market in which the price at which a transaction would be made is at or below the lower limit.

Limit order Instructions to a broker or exchange to obtain the best price immediately available when filling an order, but in no event accept a price higher than a specified (limit) price when buying or accept a price lower than a specified (limit) price when selling.

Limit order book The book or list of limit orders to buy and sell that pertains to a security.

Limit up A limit move in the futures market in which the price at which a transaction would be made is at or above the upper limit.

Limitations on liens Meant to put limits on how much secured debt an issuer can have.

Limited partners Partners with limited liability. Limited partnerships in hedge and private equity funds are typically restricted to investors who are expected to understand and to be able to assume the risks associated with the investments.

Limited partnership agreement (LPA) A legal document that outlines the rules of the partnership and establishes the framework that ultimately guides the fund's operations throughout its life.

Lin-log model A regression model in which the independent variable is in logarithmic form.

Line chart A type of graph used to visualize ordered observations. In technical analysis, a plot of price data, typically closing prices, with a line connecting the points.

Linear interpolation The estimation of an unknown value on the basis of two known values that bracket it, using a straight line between the two known values.

Linear regression Regression that models the straight-line relationship between the dependent and independent variables. Also known as *least squares regression* and *ordinary least squares regression*.

Linear scale A scale in which equal distances correspond to equal absolute amounts. Also called an *arithmetic scale*.

Linker See *inflation-linked bond*.

Liquid market Said of a market in which traders can buy or sell with low total transaction costs when they want to trade.

Liquidation To sell the assets of a company, division, or subsidiary piecemeal, typically because of bankruptcy; the form of bankruptcy that allows for the orderly satisfaction of creditors' claims after which the company ceases to exist.

Liquidity The extent to which a company is able to meet its short-term obligations using cash flows and those assets that can be readily transformed into cash.

Liquidity premium An extra return that compensates investors for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

Liquidity ratios Financial ratios measuring the company's ability to meet its short-term obligations to creditors.

Liquidity risk The risk that a financial instrument cannot be purchased or sold without a significant concession in price due to the size of the market.

Liquidity trap A condition in which the demand for money becomes infinitely elastic (horizontal demand curve) so that injections of money into the economy will not lower interest rates or affect real activity.

Load fund A mutual fund in which, in addition to the annual fee, a percentage fee is charged to invest in the fund and/or for redemptions from the fund.

Loan-to-value ratio The amount of the property's mortgage divided by the purchase price

Locked limit A condition in the futures markets in which a transaction cannot take place because the price would be beyond the limits.

Lockup period The minimum holding period before investors are allowed to make withdrawals or redeem shares from a fund.

Log-lin model A regression model in which the dependent variable is in logarithmic form.

Log-log model A regression model in which both the dependent and independent variables are in logarithmic form. Also known as the *double-log model*.

Logarithmic scale A scale in which equal distances represent equal proportional changes in the underlying quantity.

London interbank offered rate (Libor) Collective name for multiple rates at which a select set of banks believes they could borrow unsecured funds from other banks in the London interbank market for different currencies and different borrowing periods ranging from overnight to one year.

Long The buyer of a derivative contract. Also refers to the position of owning a derivative.

Long-lived assets Assets that are expected to provide economic benefits over a future period of time, typically greater than one year. Also called *long-term assets*.

Long position A position in an asset or contract in which one owns the asset or has an exercisable right under the contract.

Long-run average total cost The curve describing average total cost when no costs are considered fixed.

Look-ahead bias A bias caused by using information that was unavailable on the test date.

Loss aversion The tendency of people to dislike losses more than they like comparable gains.

Loss-aversion bias A bias in which people tend to strongly prefer avoiding losses as opposed to achieving gains.

Loss severity Portion of a bond's value (including unpaid interest) an investor loses in the event of default.

Losses Asset outflows not directly related to the ordinary activities of the business.

Lower bound The lowest possible value of an option.

M^2 A measure of what a portfolio would have returned if it had taken on the same total risk as the market index.

M^2 alpha Difference between the risk-adjusted performance of the portfolio and the performance of the benchmark.

Macaulay duration The approximate amount of time a bond would have to be held for the market discount rate at purchase to be realized if there is a single change in interest rate. It indicates the point in time when the coupon reinvestment and price effects of a change in yield-to-maturity offset each other.

Machine learning Computer based techniques that seek to extract knowledge from large amounts of data by "learning" from known examples and then generating structure or predictions. ML algorithms aim to "find the pattern, apply the pattern."

Macroeconomics The branch of economics that deals with aggregate economic quantities, such as national output and national income.

Maintenance covenants Covenants in bank loan agreements that require the borrower to satisfy certain financial ratio tests while the loan is outstanding.

Maintenance margin The minimum amount that is required by a futures clearinghouse to maintain a margin account and to protect against default. Participants whose margin balances drop below the required maintenance margin must replenish their accounts.

Maintenance margin requirement The margin requirement on any day other than the first day of a transaction.

Management buy-in Leveraged buyout in which the current management team is being replaced and the acquiring team will be involved in managing the company.

Management buyout A leveraged buyout event in which a group of investors consisting primarily of the company's existing management purchases at least controlling interest in its outstanding shares. At the extreme, they may purchase all shares and take the company private.

Management fee A fee based on assets under management or committed capital, as applicable—also called a *base fee*.

MAR ratio A variation of the Calmar ratio that uses a full investment history and the average drawdown.

Margin The amount of money that a trader deposits in a margin account. The term is derived from the stock market practice in which an investor borrows a portion of the money required to purchase a certain amount of stock. In futures markets, there is no borrowing so the margin is more of a down payment or performance bond.

Margin bond A cash deposit required by the clearinghouse from the participants to a contract to provide a credit guarantee. Also called a *performance bond*.

Margin call A request for the short to deposit additional funds to bring their balance up to the initial margin.

Margin loan Money borrowed from a broker to purchase securities.

Marginal cost The cost of producing an additional unit of a good.

Marginal frequencies The sums determined by adding joint frequencies across rows or across columns in a contingency table.

Marginal product Measures the productivity of each unit of input and is calculated by taking the difference in total product from adding another unit of input (assuming other resource quantities are held constant).

Marginal propensity to consume The proportion of an additional unit of disposable income that is consumed or spent; the change in consumption for a small change in income.

Marginal propensity to save The proportion of an additional unit of disposable income that is saved (not spent).

Marginal revenue The change in total revenue divided by the change in quantity sold; simply, the additional revenue from selling one more unit.

Marginal value curve A curve describing the highest price consumers are willing to pay for each additional unit of a good.

Mark to market The revaluation of a financial asset or liability to its current market value or fair value.

Mark-to-market Refers to the current expected fair market value for which a given security would likely be available for purchase or sale if traded in current market conditions.

Market anomaly Change in the price or return of a security that cannot directly be linked to current relevant information known in the market or to the release of new information into the market.

Market bid–ask spread The difference between the best bid and the best offer.

Market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its market capitalization by the total market capitalization (sum of the market capitalization) of all securities in the index. Also called *value weighting*.

Market discount rate The rate of return required by investors given the risk of the investment in a bond; also called the *required yield* or the *required rate of return*.

Market float The number of shares that are available to the investing public.

Market liquidity risk The risk that the price at which investors can actually transact—buying or selling—may differ from the price indicated in the market.

Market model A regression equation that specifies a linear relationship between the return on a security (or portfolio) and the return on a broad market index.

Market multiple models Valuation models based on share price multiples or enterprise value multiples.

Market-on-close An execution instruction specifying that an order can only be filled at the close of trading.

Market order Instructions to a broker or exchange to obtain the best price immediately available when filling an order.

Market-oriented investors With reference to equity investors, investors whose investment disciplines cannot be clearly categorized as value or growth.

Market rate of interest The rate demanded by purchasers of bonds, given the risks associated with future cash payment obligations of the particular bond issue.

Market reference rate (MRR) Collective name for a set of rates covering different currencies for different maturities, ranging from overnight to one year.

Market risk The risk that arises from movements in interest rates, stock prices, exchange rates, and commodity prices.

Market value The price at which an asset or security can currently be bought or sold in an open market.

Marketable limit order A buy limit order in which the limit price is placed above the best offer, or a sell limit order in which the limit price is placed below the best bid. Such orders generally will partially or completely fill right away.

Markowitz efficient frontier The graph of the set of portfolios offering the maximum expected return for their level of risk (standard deviation of return).

Matching principle The accounting principle that expenses should be recognized in the same period in which the associated revenue is recognized.

Matrix pricing Process of estimating the market discount rate and price of a bond based on the quoted or flat prices of more frequently traded comparable bonds.

Maturity premium An extra return that compensates investors for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended.

Maturity structure A factor explaining the differences in yields on similar bonds; also called *term structure*.

Mean absolute deviation With reference to a sample, the mean of the absolute values of deviations from the sample mean.

Mean square error (MSE) The sum of squares error divided by the degrees of freedom, $n - k - 1$; in a simple linear regression, $n - k - 1 = n - 2$.

Mean square regression (MSR) The sum of squares regression divided by the number of independent variables k; in a simple linear regression, $k = 1$.

Mean–variance analysis An approach to portfolio analysis using expected means, variances, and covariances of asset returns.

Measure of central tendency A quantitative measure that specifies where data are centered.

Measure of value A standard for measuring value; a function of money.

Measures of location Quantitative measures that describe the location or distribution of data. They include not only measures of central tendency but also other measures, such as percentiles.

Median The value of the middle item of a set of items that has been sorted into ascending or descending order (i.e., the 50th percentile).

Medium of exchange Any asset that can be used to purchase goods and services or to repay debts; a function of money.

Medium-term note A corporate bond offered continuously to investors by an agent of the issuer, designed to fill the funding gap between commercial paper and long-term bonds.

Mental accounting bias An information-processing bias in which people treat one sum of money differently from another equal-sized sum based on which mental account the money is assigned to.

Menu costs A cost of inflation in which businesses constantly have to incur the costs of changing the advertised prices of their goods and services.

Mesokurtic Describes a distribution with kurtosis equal to that of the normal distribution, namely, kurtosis equal to three.

Microeconomics The branch of economics that deals with markets and decision making of individual economic units, including consumers and businesses.

Minimum efficient scale The smallest output that a firm can produce such that its long-run average total cost is minimized.

Minimum-variance portfolio The portfolio with the minimum variance for each given level of expected return.

Minority shareholders Particular shareholders or a block of shareholders holding a small proportion of a company's outstanding shares, resulting in a limited ability to exercise control in voting activities.

Minsky moment Named for Hyman Minksy. A point in a business cycle when, after individuals become overextended in borrowing to finance speculative investments, people start realizing that something is likely to go wrong and a panic ensues, leading to asset sell-offs.

Modal interval With reference to grouped data, the interval containing the greatest number of observations (i.e., highest frequency).

Mode The most frequently occurring value in a distribution.

Modern portfolio theory (MPT) The analysis of rational portfolio choices based on the efficient use of risk.

Modified duration A measure of the percentage price change of a bond given a change in its yield-to-maturity.

Momentum oscillator A graphical representation of market sentiment that is constructed from price data and calculated so that it oscillates either between a low and a high or around some number.

Monetarists Economists who believe that the rate of growth of the money supply is the primary determinant of the rate of inflation.

Monetary policy Actions taken by a nation's central bank to affect aggregate output and prices through changes in bank reserves, reserve requirements, or its target interest rate.

Monetary transmission mechanism The process whereby a central bank's interest rate gets transmitted through the economy and ultimately affects the rate of increase of prices.

Monetary union An economic union in which the members adopt a common currency.

Money A generally accepted medium of exchange and unit of account.

Money convexity For a bond, the annual or approximate convexity multiplied by the full price.

Money creation The process by which changes in bank reserves translate into changes in the money supply.

Money duration A measure of the price change in units of the currency in which the bond is denominated given a change in its yield-to-maturity.

Money market The market for short-term debt instruments (one-year maturity or less).

Money market securities Fixed-income securities with maturities at issuance of one year or less.

Money multiplier Describes how a change in reserves is expected to affect the money supply; in its simplest form, 1 divided by the reserve requirement.

Money neutrality The thesis that an increase in the money supply leads in the long-run to an increase in the price level, while leaving real variables like output and employment unaffected.

Money-weighted return The internal rate of return on a portfolio, taking account of all cash flows.

Moneyness The relationship between the price of the underlying and an option's exercise price.

Monopolistic competition Highly competitive form of imperfect competition; the competitive characteristic is a notably large number of firms, while the monopoly aspect is the result of product differentiation.

Monopoly In pure monopoly markets, there are no substitutes for the given product or service. There is a single seller, which exercises considerable power over pricing and output decisions.

Monte Carlo simulation A technique involving the use of computer software to represent the operation of a complex financial system. A characteristic feature is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Moral principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Mortgage-backed securities Debt obligations that represent claims to the cash flows from pools of mortgage loans, most commonly on residential property.

Mortgage loan A loan secured by the collateral of some specified real estate property that obliges the borrower to make a predetermined series of payments to the lender.

Mortgage pass-through security A security created when one or more holders of mortgages form a pool of mortgages and sell shares or participation certificates in the pool.

Mortgage rate The interest rate on a mortgage loan; also called *contract rate* or *note rate*.

Moving average The average of the closing price of a security over a specified number of periods. With each new period, the average is recalculated.

Moving-average convergence/divergence oscillator A momentum oscillator that is based on the difference between short-term and long-term moving averages of a security's price.

Multi-factor model A model that explains a variable in terms of the values of a set of factors.

Multi-market indexes Comprised of indexes from different countries, designed to represent multiple security markets.

Multi-step format With respect to the format of the income statement, a format that presents a subtotal for gross profit (revenue minus cost of goods sold).

Multilateral trading facilities See *alternative trading systems*.

Multinational corporation A company operating in more than one country or having subsidiary firms in more than one country.

Multinomial formula (general formula for labeling problems) The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by

$$\frac{n!}{n_1!n_2!\ldots n_k!}.$$

Multiple of invested capital (MOIC) A simplified calculation that measures the total value of all distributions and residual asset values relative to an initial total investment; also known as a *money multiple*.

Multiple testing problem The risk of getting statistically significant test results when performing a test multiple times.

Multiplication rule for counting If one task can be done in n_1 ways, and a second task, given the first, can be done in n_2 ways, and a third task, given the first two tasks, can be done in n_3 ways, and so on for k tasks, then the number of ways the k tasks can be done is $(n_1)(n_2)(n_3) \ldots (n_k)$.

Multiplication rule for expected value of the product of uncorrelated random variables This rule is $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated.

Multiplication rule for independent events The rule that when two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B.

Multiplication rule for probability The rule that the joint probability of events A and B equals the probability of A given B times the probability of B.

Multiplier models Valuation models based on share price multiples or enterprise value multiples.

Multivariate distribution A probability distribution that specifies the probabilities for a group of related random variables.

Multivariate normal distribution A probability distribution for a group of random variables that is completely defined by the means and variances of the variables plus all the correlations between pairs of the variables.

Muni See *municipal bond*.

Municipal bond A type of non-sovereign bond issued by a state or local government in the United States. It very often (but not always) offers income tax exemptions.

Mutual fund A comingled investment pool in which investors in the fund each have a pro-rata claim on the income and value of the fund.

Mutually exclusive Indicates that only one event can occur at a time.

Mutually exclusive projects Mutually exclusive projects compete directly with each other. For example, if Projects A and B are mutually exclusive, you can choose A or B but you cannot choose both.

***n* Factorial** For a positive integer n, the product of the first n positive integers; 0 factorial equals 1 by definition. n factorial is written as $n!$.

Narrow money The notes and coins in circulation in an economy, plus other very highly liquid deposits.

Nash equilibrium When two or more participants in a non-coop-erative game have no incentive to deviate from their respective equilibrium strategies given their opponent's strategies.

National income The income received by all factors of production used in the generation of final output. National income equals gross domestic product (or, in some countries, gross national product) minus the capital consumption allowance and a statistical discrepancy.

Natural language processing Computer programs developed to analyze and interpret human language.

Natural rate of unemployment Effective unemployment rate, below which pressure emerges in labor markets.

Negative screening An ESG investment style that focuses on the exclusion of certain sectors, companies, or practices in a fund or portfolio on the basis of specific ESG criteria.

Net book value The remaining (undepreciated) balance of an asset's purchase cost. For liabilities, the face value of a bond minus any unamortized discount, or plus any unamortized premium.

Net exports The difference between the value of a country's exports and the value of its imports (i.e., value of exports minus imports).

Net income The difference between revenue and expenses; what remains after subtracting all expenses (including depreciation, interest, and taxes) from revenue.

Net present value The present value of an investment's cash inflows (benefits) minus the present value of its cash outflows (costs).

Net profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also called *profit margin* or *return on sales*.

Net realisable value Estimated selling price in the ordinary course of business less the estimated costs necessary to make the sale.

Net revenue Revenue after adjustments (e.g., for estimated returns or for amounts unlikely to be collected).

Net tax rate The tax rate net of transfer payments.

Neural networks Computer programs based on how our own brains learn and process information.

Neutral rate of interest The rate of interest that neither spurs on nor slows down the underlying economy.

No-load fund A mutual fund in which there is no fee for investing in the fund or for redeeming fund shares, although there is an annual fee based on a percentage of the fund's net asset value.

Node Each value on a binomial tree from which successive moves or outcomes branch.

Nominal data Categorical values that are not amenable to being organized in a logical order. An example of nominal data is the classification of publicly listed stocks into sectors.

Nominal GDP The value of goods and services measured at current prices.

Nominal risk-free interest rate The sum of the real risk-free interest rate and the inflation premium.

Non-accelerating inflation rate of unemployment Effective unemployment rate below which pressure emerges in labor markets.

Non-agency RMBS In the United States, securities issued by private entities that are not guaranteed by a federal agency or a GSE.

Non-Bank lender Unlike typical banks, which make loans and take deposits, these lenders only make loans. Although they lend money to companies, they usually specialize in issuing mortgages to first-time homebuyers and parties looking to refinance an existing mortgage.

Non-cumulative preference shares Preference shares for which dividends that are not paid in the current or subsequent periods are forfeited permanently (instead of being accrued and paid at a later date).

Non-current assets Assets that are expected to benefit the company over an extended period of time (usually more than one year).

Non-current liabilities Obligations that broadly represent a probable sacrifice of economic benefits in periods generally greater than one year in the future.

Non-cyclical A company whose performance is largely independent of the business cycle. Also known as defensive.

Non-deliverable forwards Cash-settled forward contracts, used predominately with respect to foreign exchange forwards. Also called *contracts for differences*.

Non-financial risks Risks that arise from sources other than changes in the external financial markets, such as changes in accounting rules, legal environment, or tax rates.

Non-participating preference shares Preference shares that do not entitle shareholders to share in the profits of the company. Instead, shareholders are only entitled to receive a fixed dividend payment and the par value of the shares in the event of liquidation.

Non-probability sampling A sampling plan dependent on factors other than probability considerations, such as a sampler's judgment or the convenience to access data.

Non-recourse loan A loan in which the lender does not have a shortfall claim against the borrower, so the lender can look only to the property to recover the outstanding mortgage balance.

Non-renewable resources Finite resources that are depleted once they are consumed, such as oil and coal.

Non-sovereign bond A bond issued by a government below the national level, such as a province, region, state, or city.

Nonconventional cash flows In a nonconventional cash flow pattern, the initial outflow is not followed by inflows only, but the cash flows can flip from positive (inflows) to negative (outflows) again or even change signs several times.

Nonparametric test A test that is not concerned with a parameter or that makes minimal assumptions about the population from which a sample comes.

Nonsystematic risk Unique risk that is local or limited to a particular asset or industry that need not affect assets outside of that asset class.

Normal distribution A continuous, symmetric probability distribution that is completely described by its mean and its variance.

Normal goods Goods that are consumed in greater quantities as income increases.

Normal profit The level of accounting profit needed to just cover the implicit opportunity costs ignored in accounting costs.

Notching Ratings adjustment methodology where specific issues from the same borrower may be assigned different credit ratings.

Note rate See *mortgage rate*.

Notice period The length of time (typically 30 to 90 days) in advance that investors may be required to notify a fund of their intent to redeem some or all of their investment.

Notional principal An imputed principal amount.

Null hypothesis The hypothesis that is tested.

Numerical data Values that represent measured or counted quantities as a number. Also called **quantitative data**.

Objective probabilities Probabilities that generally do not vary from person to person; includes a priori and empirical probabilities.

Observation The value of a specific variable collected at a point in time or over a specified period of time.

Odds against *E* The reciprocal of odds for *E*.

Odds for *E* The probability of *E* divided by 1 minus the probability of *E*.

Off-the-run Seasoned government bonds are off-the-run securities; they are not the most recently issued or the most actively traded.

Offer The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size).

Official interest rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks. Also called *official policy rate* or *policy rate*.

Official policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Oligopoly Market structure with a relatively small number of firms supplying the market.

On-the-run The most recently issued and most actively traded sovereign securities.

One-dimensional array The simplest format for representing a collection of data of the same data type.

One-sided hypothesis test A test in which the null hypothesis is rejected only if the evidence indicates that the population parameter is greater than or less than the hypothesized parameter; occurs when the alternative hypothesis is stated either as greater than or less than the hypothesized population parameter.

Open economy An economy that trades with other countries.

Open-end fund A mutual fund that accepts new investment money and issues additional shares at a value equal to the net asset value of the fund at the time of investment.

Open interest The number of outstanding contracts in a clearinghouse at any given time. The open interest figure changes daily as some parties open up new positions, while other parties offset their old positions.

Open market operations The purchase or sale of bonds by the national central bank to implement monetary policy. The bonds traded are usually sovereign bonds issued by the national government.

Operating activities Activities that are part of the day-to-day business functioning of an entity, such as selling inventory and providing services.

Operating breakeven The number of units produced and sold at which the company's operating profit is zero (revenues = operating costs).

Operating cash flow The net amount of cash provided from operating activities.

Operating efficiency ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Operating lease A type of lease which is more akin to the rental of the underlying asset.

Operating leverage The use of fixed costs in operations.

Operating profit A company's profits on its usual business activities before deducting taxes. Also called *operating income*.

Operating profit margin A profitability ratio calculated as operating income (i.e., income before interest and taxes) divided by revenue. Also called *operating margin*.

Operating risk The risk attributed to the operating cost structure, in particular the use of fixed costs in operations; the risk arising from the mix of fixed and variable costs; the risk that a company's operations may be severely affected by environmental, social, and governance risk factors.

Operational independence A bank's ability to execute monetary policy and set interest rates in the way it thought would best meet the inflation target.

Operational risk The risk that arises from inadequate or failed people, systems, and internal policies, procedures, and processes, as well as from external events that are beyond the control of the organization but that affect its operations.

Operationally efficient Said of a market, a financial system, or an economy that has relatively low transaction costs.

Opportunity cost The value that investors forgo by choosing a particular course of action; the value of something in its best alternative use.

Optimal capital structure The capital structure at which the value of the company is maximized.

Option A financial instrument that gives one party the right, but not the obligation, to buy or sell an underlying asset from or to another party at a fixed price over a specific period of time. Also referred to as *contingent claim* or *option contract.*

Option-adjusted price The value of the embedded option plus the flat price of the bond.

Option-adjusted spread OAS = Z-spread – Option value (in basis points per year).

Option-adjusted yield The required market discount rate whereby the price is adjusted for the value of the embedded option.

Option contract See *option*.

Option premium The amount of money a buyer pays and seller receives to engage in an option transaction.

Order A specification of what instrument to trade, how much to trade, and whether to buy or sell.

Order-driven markets A market (generally an auction market) that uses rules to arrange trades based on the orders that traders submit; in their pure form, such markets do not make use of dealers.

Order precedence hierarchy With respect to the execution of orders to trade, a set of rules that determines which orders execute before other orders.

Ordinal data Categorical values that can be logically ordered or ranked.

Ordinary annuity An annuity with a first cash flow that is paid one period from the present.

Ordinary shares Equity shares that are subordinate to all other types of equity (e.g., preferred equity). Also called *common stock* or *common shares.*

Organized exchange A securities marketplace where buyers and seller can meet to arrange their trades.

Other comprehensive income Items of comprehensive income that are not reported on the income statement; comprehensive income minus net income.

Out-of-sample test A test of a strategy or model using a sample outside the period on which the strategy or model was developed.

Out of the money Options that, if exercised, would require the payment of more money than the value received and therefore would not be currently exercised.

Outcome A possible value of a random variable.

Over-the-counter (OTC) market A decentralized market where buy and sell orders initiated from various locations are matched through a communications network.

Overbought A market condition in which market sentiment is thought to be unsustainably bullish.

Overcollateralization Form of internal credit enhancement that refers to the process of posting more collateral than needed to obtain or secure financing.

Overconfidence bias A bias in which people demonstrate unwarranted faith in their own intuitive reasoning, judgments, and/or cognitive abilities.

Overfitting An undesirable result from fitting a model so closely to a dataset that it does not perform well on new data.

Oversold A market condition in which market sentiment is thought to be unsustainably bearish.

Own price The price of a good or service itself (as opposed to the price of something else).

Own-price elasticity of demand The percentage change in quantity demanded for a percentage change in good's own price, holding all other things constant.

Owners' equity The excess of assets over liabilities; the residual interest of shareholders in the assets of an entity after deducting the entity's liabilities. Also called *shareholders' equity* or *shareholders' funds*.

***p*-Value** The smallest level of significance at which the null is rejected.

Paired comparisons test See *test of the mean of the differences.*

Panel data A mix of time-series and cross-sectional data that contains observations through time on characteristics of across multiple observational units.

Par curve A sequence of yields-to-maturity such that each bond is priced at par value. The bonds are assumed to have the same currency, credit risk, liquidity, tax status, and annual yields stated for the same periodicity.

Par value The amount of principal on a bond.

Parallel shift A parallel yield curve shift implies that all rates change by the same amount in the same direction.

Parameter A descriptive measure computed from or used to describe a population of data, conventionally represented by Greek letters.

Parametric test Any test (or procedure) concerned with parameters or whose validity depends on assumptions concerning the population generating the sample.

Pari passu Covenant or contract clause that ensures a debt obligation is treated the same as the borrower's other senior debt instruments and is not subordinated to similar obligations.

Partial duration See *key rate duration.*

Participating preference shares Preference shares that entitle shareholders to receive the standard preferred dividend plus the opportunity to receive an additional dividend if the company's profits exceed a pre-specified level.

Pass-through rate The coupon rate of a mortgage pass-through security.

Passive investment A buy and hold approach in which an investor does not make portfolio changes based on short-term expectations of changing market or security performance.

Payable date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payment date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payments system The system for the transfer of money.

Pearson correlation A parametric measure of the relationship between two variables.

Pecking order theory The theory that managers consider how their actions might be interpreted by outsiders and thus order their preferences for various forms of corporate financing. Forms of financing that are least visible to outsiders (e.g., internally generated funds) are most preferable to managers, and those that are most visible (e.g., equity) are least preferable.

Peer company See *comparable company.*

Peer group A group of companies engaged in similar business activities whose economics and valuation are influenced by closely related factors.

Pennant A technical analysis continuation pattern formed by trendlines that converge to form a triangle, typically over a short period.

Per capita real GDP Real GDP divided by the size of the population, often used as a measure of a country's average standard of living.

Per unit contribution margin The amount that each unit sold contributes to covering fixed costs—that is, the difference between the price per unit and the variable cost per unit.

Percentiles Quantiles that divide a distribution into 100 equal parts that sum to 100.

Perfect capital markets Markets in which, by assumption, there are no taxes, transaction costs, or bankruptcy costs and in which all investors have equal ("symmetric") information.

Perfect competition A market structure in which the individual firm has virtually no impact on market price, because it is assumed to be a very small seller among a very large number of firms selling essentially identical products.

Perfectly elastic When the quantity demanded or supplied of a given good is infinitely sensitive to a change in the value of a specified variable (e.g., price).

Perfectly inelastic When the quantity demanded or supplied of a given good is completely insensitive to a change in the value of a specified variable (e.g., price).

Performance bond See *margin bond.*

Performance evaluation The measurement and assessment of the outcomes of investment management decisions.

Performance fee Fees paid to the general partner from the limited partner(s) based on realized net profits.

Period costs Costs (e.g., executives' salaries) that cannot be directly matched with the timing of revenues and which are thus expensed immediately.

Periodicity The assumed number of periods in the year; typically matches the frequency of coupon payments.

Permanent differences Differences between tax and financial reporting of revenue (expenses) that will not be reversed at some future date. These result in a difference between the company's effective tax rate and statutory tax rate and do not result in a deferred tax item.

Permissioned networks Networks that are fully open only to select participants on a DLT network.

Permissionless networks Networks that are fully open to any user on a DLT network.

Permutation An ordered listing.

Permutation formula The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is ${}_nP_r = \dfrac{n!}{(n-r)!}$.

Perpetual bonds Bonds with no stated maturity date.

Perpetuity A perpetual annuity, or a set of never-ending level sequential cash flows, with the first cash flow occurring one period from now. A bond that does not mature.

Personal income A broad measure of household income that includes all income received by households, whether earned or unearned; measures the ability of consumers to make purchases.

Pet projects Investments in which influential managers want the corporation to invest. Often, unfortunately, pet projects are selected without undergoing normal capital allocation analysis.

Plain vanilla bond Bond that makes periodic, fixed coupon payments during the bond's life and a lump-sum payment of principal at maturity. Also called *conventional bond*.

Platykurtic Describes a distribution that has relatively less weight in the tails than the normal distribution (also called thin-tailed).

Point estimate A single numerical estimate of an unknown quantity, such as a population parameter.

Policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Population All members of a specified group.

Portfolio company In private equity, the company in which the private equity fund is investing.

Portfolio demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments.

Portfolio planning The process of creating a plan for building a portfolio that is expected to satisfy a client's investment objectives.

Position The quantity of an asset that an entity owns or owes.

Positive screening The process of including sectors or companies based on specific ESG criteria, typically ESG performance relative to industry peers

Posterior probability An updated probability that reflects or comes after new information.

Potential GDP The level of real GDP that can be produced at full employment; measures the productive capacity of the economy.

Power of a test The probability of correctly rejecting the null—that is, rejecting the null hypothesis when it is false.

Precautionary money balances Money held to provide a buffer against unforeseen events that might require money.

Preference shares A type of equity interest which ranks above common shares with respect to the payment of dividends and the distribution of the company's net assets upon liquidation. They have characteristics of both debt and equity securities. Also called *preferred stock*.

Preferred stock See *preference shares*.

Premium In the case of bonds, premium refers to the amount by which a bond is priced above its face (par) value. In the case of an option, the amount paid for the option contract.

Prepaid expense A normal operating expense that has been paid in advance of when it is due.

Prepayment option Contractual provision that entitles the borrower to prepay all or part of the outstanding mortgage principal prior to the scheduled due date when the principal must be repaid. Also called *early repayment option*.

Prepayment penalty mortgages Mortgages that stipulate a monetary penalty if a borrower prepays within a certain time period after the mortgage is originated.

Prepayment risk The uncertainty that the timing of the actual cash flows will be different from the scheduled cash flows as set forth in the loan agreement due to the borrowers' ability to alter payments, usually to take advantage of interest rate movements.

Present value (PV) The present discounted value of future cash flows: For assets, the present discounted value of the future net cash inflows that the asset is expected to generate; for liabilities, the present discounted value of the future net cash outflows that are expected to be required to settle the liabilities.

Present value models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security. Also called *discounted cash flow models*.

Pretax margin A profitability ratio calculated as earnings before taxes divided by revenue.

Price elasticity of demand Measures the percentage change in the quantity demanded, given a percentage change in the price of a given product.

Price index Represents the average prices of a basket of goods and services.

Price limits Limits imposed by a futures exchange on the price change that can occur from one day to the next.

Price multiple A ratio that compares the share price with some sort of monetary flow or value to allow evaluation of the relative worth of a company's stock.

Price priority The principle that the highest priced buy orders and the lowest priced sell orders execute first.

Price relative A ratio of an ending price over a beginning price; it is equal to 1 plus the holding period return on the asset.

Price return Measures *only* the price appreciation or percentage change in price of the securities in an index or portfolio.

Price return index An index that reflects *only* the price appreciation or percentage change in price of the constituent securities. Also called *price index*.

Price-setting option The option to adjust prices when demand varies from what is forecast.

Price stability In economics, refers to an inflation rate that is low on average and not subject to wide fluctuation.

Price takers Producers that must accept whatever price the market dictates.

Price to book value A valuation ratio calculated as price per share divided by book value per share.

Price to cash flow A valuation ratio calculated as price per share divided by cash flow per share.

Price to earnings ratio (P/E ratio or P/E) The ratio of share price to earnings per share.

Price to sales A valuation ratio calculated as price per share divided by sales per share.

Price value of a basis point A version of money duration, it is an estimate of the change in the full price of a bond given a 1 basis point change in the yield-to-maturity.

Price weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its price by the sum of all the prices of the constituent securities.

Priced risk Risk for which investors demand compensation for bearing (e.g., equity risk, company-specific factors, macroeconomic factors).

Primary bond market A market in which issuers first sell bonds to investors to raise capital.

Primary capital markets (primary markets) The market where securities are first sold and the issuers receive the proceeds.

Primary dealer Financial institution that is authorized to deal in new issues of sovereign bonds and that serves primarily as a trading counterparty of the office responsible for issuing sovereign bonds.

Primary market The market where securities are first sold and the issuers receive the proceeds.

Prime brokers Brokers that provide services that commonly include custody, administration, lending, short borrowing, and trading.

Principal The amount of funds originally invested in a project or instrument; the face value to be paid at maturity.

Principal–agent relationship A relationship in which a principal hires an agent to perform a particular task or service; also known as an *agency relationship*.

Principal amount Amount that an issuer agrees to repay the debtholders on the maturity date.

Principal business activity The business activity from which a company derives a majority of its revenues and/or earnings.

Principal value Amount that an issuer agrees to repay the debtholders on the maturity date.

Principle of no arbitrage See *arbitrage-free pricing*.

Prior probabilities Probabilities reflecting beliefs prior to the arrival of new information.

Priority of claims Priority of payment, with the most senior or highest ranking debt having the first claim on the cash flows and assets of the issuer.

Private equity fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *venture capital fund*.

Private equity funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *venture capital funds*.

Private equity securities Securities that are not listed on public exchanges and have no active secondary market. They are issued primarily to institutional investors via non-public offerings, such as private placements.

Private investment in public equity (PIPE) An investment in the equity of a publicly traded firm that is made at a discount to the market value of the firm's shares.

Private placement Typically, a non-underwritten, unregistered offering of securities that are sold only to an investor or a small group of investors. It can be accomplished directly between the issuer and the investor(s) or through an investment bank.

Probability A number between 0 and 1 describing the chance that a stated event will occur.

Probability density function A function with non-negative values such that probability can be described by areas under the curve graphing the function.

Probability distribution A distribution that specifies the probabilities of a random variable's possible outcomes.

Probability function A function that specifies the probability that the random variable takes on a specific value.

Probability sampling A sampling plan that allows every member of the population to have an equal chance of being selected.

Probability tree diagram A diagram with branches emanating from nodes representing either mutually exclusive chance events or mutually exclusive decisions.

Production-flexibility options The options to alter production when demand varies from what is forecast.

Production function Provides the quantitative link between the levels of output that the economy can produce and the inputs used in the production process.

Productivity The amount of output produced by workers during a given period—for example, output per hour worked measures the efficiency of labor.

Profession An occupational group that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition.

Profit The return that owners of a company receive for the use of their capital and the assumption of financial risk when making their investments.

Profit and loss (P&L) statement A financial statement that provides information about a company's profitability over a stated period of time. Also called the *income statement*.

Profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses.

Profitability ratios Ratios that measure a company's ability to generate profitable sales from its resources (assets).

Project sequencing To defer the decision to invest in a future project until the outcome of some or all of a current investment is known. Investments are sequenced over time, so that making an investment creates the option to invest in future projects.

Promissory note A written promise to pay a certain amount of money on demand.

Property, plant, and equipment Tangible assets that are expected to be used for more than one period in either the production or supply of goods or services, or for administrative purposes.

Prospectus The document that describes the terms of a new bond issue and helps investors perform their analysis on the issue.

Protective put An option strategy in which a long position in an asset is combined with a long position in a put.

Proxy contest Corporate takeover mechanism in which shareholders are persuaded to vote for a group seeking a controlling position on a company's board of directors.

Proxy voting A process that enables shareholders who are unable to attend a meeting to authorize another individual to vote on their behalf.

Public offer See *public offering*.

Public offering An offering of securities in which any member of the public may buy the securities. Also called *public offer*.

Public–private partnership (PPP) An agreement between the public sector and the private sector to finance, build, and operate public infrastructure, such as hospitals and toll roads.

Pull on liquidity When disbursements are paid too quickly or trade credit availability is limited, requiring companies to expend funds before they receive funds from sales that could cover the liability.

Pure discount bonds See *zero-coupon bond*.

Put An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Put–call–forward parity The relationship among puts, calls, and forward contracts.

Put–call parity An equation expressing the equivalence (parity) of a portfolio of a call and a bond with a portfolio of a put and the underlying, which leads to the relationship between put and call prices.

Put/call ratio A technical analysis indicator that evaluates market sentiment based on the volume of put options traded divided by the volume of call options traded for a particular financial instrument.

Put option An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Putable bonds Bonds that give the bondholder the right to sell the bond back to the issuer at a predetermined price on specified dates.

Qualitative data see **categorical data**.

Quantile A value at or below which a stated fraction of the data lies. Also referred to as a fractile.

Quantitative data see **numerical data**.

Quantitative easing An expansionary monetary policy based on aggressive open market purchase operations.

Quantity equation of exchange An expression that over a given period, the amount of money used to purchase all goods and services in an economy, $M \times V$, is equal to monetary value of this output, $P \times Y$.

Quantity theory of money Asserts that total spending (in money terms) is proportional to the quantity of money.

Quartiles Quantiles that divide a distribution into four equal parts.

Quasi-fixed cost A cost that stays the same over a range of production but can change to another constant level when production moves outside of that range.

Quasi-government bond A bond issued by an entity that is either owned or sponsored by a national government. Also called *agency bond*.

Quintiles Quantiles that divide a distribution into five equal parts.

Quota rents Profits that foreign producers can earn by raising the price of their goods higher than they would without a quota.

Quotas Government policies that restrict the quantity of a good that can be imported into a country, generally for a specified period of time.

Quote-driven market A market in which dealers acting as principals facilitate trading.

Quoted interest rate A quoted interest rate that does not account for compounding within the year. Also called *stated annual interest rate*.

Quoted margin The specified yield spread over the reference rate, used to compensate an investor for the difference in the credit risk of the issuer and that implied by the reference rate.

Random number An observation drawn from a uniform distribution.

Random number generator An algorithm that produces uniformly distributed random numbers between 0 and 1.

Random variable A quantity whose future outcomes are uncertain.

Range The difference between the maximum and minimum values in a dataset.

Raw data Data available in their original form as collected.

Real exchange rate effect The effect through which changing price level impacts real exchange rate which in turn impacts net exports and aggregate demand.

Real GDP The value of goods and services produced, measured at base year prices.

Real income Income adjusted for the effect of inflation on the purchasing power of money. Also known as the *purchasing power of income*. If income remains constant and a good's price falls, real income is said to rise, even though the number of monetary units (e.g., dollars) remains unchanged.

Real interest rate Nominal interest rate minus the expected rate of inflation.

Real risk-free interest rate The single-period interest rate for a completely risk-free security if no inflation were expected.

Realizable (settlement) value With reference to assets, the amount of cash or cash equivalents that could currently be obtained by selling the asset in an orderly disposal; with reference to liabilities, the undiscounted amount of cash or cash equivalents expected to be paid to satisfy the liabilities in the normal course of business.

Rebalancing Adjusting the weights of the constituent securities in an index.

Rebalancing policy The set of rules that guide the process of restoring a portfolio's asset class weights to those specified in the strategic asset allocation.

Recession A period during which real GDP decreases (i.e., negative growth) for at least two successive quarters, or a period of significant decline in total output, income, employment, and sales usually lasting from six months to a year.

Recognition lag The lag in government response to an economic problem resulting from the delay in confirming a change in the state of the economy.

Recourse loan A loan in which the lender has a claim against the borrower for any shortfall between the outstanding mortgage balance and the proceeds received from the sale of the property.

Redemption yield See *yield-to-maturity*.

Redemptions Withdrawals of funds by investors, as allowed by the notice period and other terms in the partnership agreement.

Refinancing rate A type of central bank policy rate.

Registered bonds Bonds for which ownership is recorded by either name or serial number.

Regression analysis A tool for examining whether a variable is useful for explaining another variable.

Regression coefficients The intercept and slope coefficient of a regression.

Regret The feeling that an opportunity has been missed; typically, an expression of *hindsight bias*.

Regret-aversion bias An emotional bias in which people tend to avoid making decisions that will result in action out of fear that the decision will turn out poorly.

Relative dispersion The amount of dispersion relative to a reference value or benchmark.

Relative frequency The absolute frequency of each unique value of the variable divided by the total number of observations of the variable.

Relative price The price of a specific good or service in comparison with those of other goods and services.

Relative strength analysis A comparison of the performance of one asset with the performance of another asset or a benchmark, based on changes in the ratio of the two assets' prices over time.

Relative strength index (RSI) A technical analysis momentum oscillator that compares a security's gains with its losses over a set period.

Renewable resources Resources that can be replenished, such as a forest.

Rent Payment for the use of property.

Reorganization Agreements made by a company in bankruptcy under which a company's capital structure is altered and/or alternative arrangements are made for debt repayment; US Chapter 11 bankruptcy. The company emerges from bankruptcy as a going concern.

Replication The creation of an asset or portfolio from another asset, portfolio, and/or derivative.

Repo A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy back the same security from the purchaser at an agreed-on price and future date. The party who sells the

security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repo margin The difference between the market value of the security used as collateral and the value of the loan. Also called *haircut*.

Repo rate The interest rate on a repurchase agreement.

Representativeness bias A belief perseverance bias in which people tend to classify new information based on past experiences and classifications.

Repurchase agreement A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy the same security back from the purchaser at an agreed-on price and future date. The party who sells the security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repurchase date The date when the party who sold the security at the inception of a repurchase agreement buys back the security from the cash lending counterparty.

Repurchase price The price at which the party who sold the security at the inception of the repurchase agreement buys back the security from the cash lending counterparty.

Required margin The yield spread over or under the reference rate such that an FRN is priced at par value on a rate reset date.

Required rate of return See *market discount rate*.

Required yield See *market discount rate*.

Required yield spread The difference between the yield-to-maturity on a new bond and the benchmark rate; additional compensation required by investors for the difference in risk and tax status of a bond relative to a government bond. Sometimes called the *spread over the benchmark*.

Resampling A statistical method that repeatedly draws samples from the original observed data sample for the statistical inference of population parameters.

Reserve accounts Form of internal credit enhancement that relies on creating accounts and depositing in these accounts cash that can be used to absorb losses. Also called *reserve funds*.

Reserve funds See *reserve accounts*.

Reserve requirement The requirement for banks to hold reserves in proportion to the size of deposits.

Residual The difference between an observation and its predicted value, where the predicted value is based on the estimated linear relation between the dependent and independent variables using sample data.

Resistance In technical analysis, a price range in which selling activity is sufficient to stop the rise in the price of a security.

Responsible investing An overall (umbrella) term for several investment strategies—ESG investing, SRI investing, thematic investing, and impact investing—that incorporate ESG analysis in their investment processes.

Restricted payments A bond covenant meant to protect creditors by limiting how much cash can be paid out to shareholders over time.

Retracement In technical analysis, a reversal in the movement of a security's price such that it is counter to the prevailing longer-term price trend.

Return-generating model A model that can provide an estimate of the expected return of a security given certain parameters and estimates of the values of the independent variables in the model.

Return on assets (ROA) A profitability ratio calculated as net income divided by average total assets; indicates a company's net profit generated per dollar invested in total assets.

Return on equity (ROE) A profitability ratio calculated as net income divided by average shareholders' equity.

Return on invested capital A measure of the profitability of a company relative to the amount of capital invested by the equity- and debtholders.

Return on sales An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also referred to as *net profit margin*.

Return on total capital A profitability ratio calculated as EBIT divided by the sum of short- and long-term debt and equity.

Revaluation model Under IFRS, the process of valuing long-lived assets at fair value, rather than at cost less accumulated depreciation. Any resulting profit or loss is either reported on the income statement and/or through equity under revaluation surplus.

Revenue The amount charged for the delivery of goods or services in the ordinary activities of a business over a stated period; the inflows of economic resources to a company over a stated period.

Reversal pattern A type of pattern used in technical analysis to predict the end of a trend and a change in the direction of a security's price.

Reverse repo A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse repurchase agreement A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse stock split A reduction in the number of shares outstanding with a corresponding increase in share price, but no change to the company's underlying fundamentals.

Revolving credit agreements (also known as revolvers) The most reliable form of short-term bank borrowing facilities; they are in effect for multiple years (e.g., three to five years) and can have optional medium-term loan features.

Rho The sensitivity of the option price to the risk-free rate.

Ricardian equivalence An economic theory that implies that it makes no difference whether a government finances a deficit by increasing taxes or issuing debt.

Risk Exposure to uncertainty. The chance of a loss or adverse outcome as a result of an action, inaction, or external event.

Risk averse The assumption that an investor will choose the least risky alternative.

Risk aversion The degree of an investor's inability and unwillingness to take risk.

Risk budgeting The establishment of objectives for individuals, groups, or divisions of an organization that takes into account the allocation of an acceptable level of risk.

Risk exposure The state of being exposed or vulnerable to a risk. The extent to which an organization is sensitive to underlying risks.

Risk governance The top-down process and guidance that directs risk management activities to align with and support the overall enterprise.

Risk management The process of identifying the level of risk an organization wants, measuring the level of risk the organization currently has, taking actions that bring the actual

level of risk to the desired level of risk, and monitoring the new actual level of risk so that it continues to be aligned with the desired level of risk.

Risk management framework The infrastructure, process, and analytics needed to support effective risk management in an organization.

Risk-neutral pricing Sometimes said of derivatives pricing, uses the fact that arbitrage opportunities guarantee that a risk-free portfolio consisting of the underlying and the derivative must earn the risk-free rate.

Risk-neutral probabilities Weights that are used to compute a binomial option price. They are the probabilities that would apply if a risk-neutral investor valued an option.

Risk premium An extra return expected by investors for bearing some specified risk.

Risk shifting Actions to change the distribution of risk outcomes.

Risk tolerance The amount of risk an investor is willing and able to bear to achieve an investment goal.

Risk transfer Actions to pass on a risk to another party, often, but not always, in the form of an insurance policy.

Robo-adviser A machine-based analytical tool or service that provides technology-driven investment solutions through online platforms.

Rule of 72 The principle that the approximate number of years necessary for an investment to double is 72 divided by the stated interest rate.

Running yield See *current yield*.

Safety-first rules Rules for portfolio selection that focus on the risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Sales Generally, a synonym for revenue; "sales" is generally understood to refer to the sale of goods, whereas "revenue" is understood to include the sale of goods or services.

Sales risk Uncertainty with respect to the quantity of goods and services that a company is able to sell and the price it is able to achieve; the risk related to the uncertainty of revenues.

Sample A subset of a population.

Sample correlation coefficient A standardized measure of how two variables in a sample move together. It is the ratio of the sample covariance to the product of the two variables' standard deviations.

Sample covariance A measure of how two variables in a sample move together.

Sample excess kurtosis A sample measure of the degree of a distribution's kurtosis in excess of the normal distribution's kurtosis.

Sample mean The sum of the sample observations divided by the sample size.

Sample selection bias Bias introduced by systematically excluding some members of the population according to a particular attribute—for example, the bias introduced when data availability leads to certain observations being excluded from the analysis.

Sample-size neglect A type of representativeness bias in which financial market participants incorrectly assume that small sample sizes are representative of populations (or "real" data).

Sample skewness A sample measure of the degree of asymmetry of a distribution.

Sample standard deviation The positive square root of the sample variance.

Sample statistic A quantity computed from or used to describe a sample.

Sample variance The sum of squared deviations around the mean divided by the degrees of freedom.

Sampling The process of obtaining a sample.

Sampling distribution The distribution of all distinct possible values that a statistic can assume when computed from samples of the same size randomly drawn from the same population.

Sampling error The difference between the observed value of a statistic and the estimate resulting from using subsets of the population.

Sampling plan The set of rules used to select a sample.

Say on pay A process whereby shareholders may vote on executive remuneration (compensation) matters.

Scatter plot A chart in which two variables are plotted along the axis and points on the chart represent pairs of the two variables. In regression, the dependent variable is plotted on the vertical axis and the independent variable is plotted along the horizontal axis. Also known as a scattergram and a *scatter diagram*.

Scatter plot matrix A tool for organizing scatter plots between pairs of variables, making it easy to inspect all pairwise relationships in one combined visual.

Scenario analysis Analysis that shows the changes in key financial quantities that result from given (economic) events, such as the loss of customers, the loss of a supply source, or a catastrophic event; a risk management technique involving examination of the performance of a portfolio under specified situations. Closely related to stress testing.

Screening The application of a set of criteria to reduce a set of potential investments to a smaller set having certain desired characteristics.

Seasoned offering An offering in which an issuer sells additional units of a previously issued security.

Second-degree price discrimination When the monopolist charges different per-unit prices using the quantity purchased as an indicator of how highly the customer values the product.

Second lien A secured interest in the pledged assets that ranks below first lien debt in both collateral protection and priority of payment.

Secondary bond markets Markets in which existing bonds are traded among investors.

Secondary market The market where securities are traded among investors.

Secondary precedence rules Rules that determine how to rank orders placed at the same time.

Sector A group of related industries (GICS definition).

Sector indexes Indexes that represent and track different economic sectors—such as consumer goods, energy, finance, health care, and technology—on either a national, regional, or global basis.

Secured ("asset-based") loan A loan that is backed by specific, secured company assets.

Secured bonds Bonds secured by assets or financial guarantees pledged to ensure debt repayment in case of default.

Secured debt Debt in which the debtholder has a direct claim—a pledge from the issuer—on certain assets and their associated cash flows.

Securitization A process that involves moving assets into a special legal entity, which then uses the assets as guarantees to secure a bond issue.

Securitized assets Assets that are typically used to create asset-backed bonds; for example, when a bank securitizes a pool of loans, the loans are said to be securitized.

Security characteristic line A plot of the excess return of a security on the excess return of the market.

Security market index A portfolio of securities representing a given security market, market segment, or asset class.

Security market line (SML) The graph of the capital asset pricing model.

Security selection The process of selecting individual securities; typically, security selection has the objective of generating superior risk-adjusted returns relative to a portfolio's benchmark.

Self-attribution bias A bias in which people take too much credit for successes (*self-enhancing*) and assign responsibility to others for failures (*self-protecting*).

Self-control bias A bias in which people fail to act in pursuit of their long-term, overarching goals because of a lack of self-discipline.

Self-investment limits With respect to investment limitations applying to pension plans, restrictions on the percentage of assets that can be invested in securities issued by the pension plan sponsor.

Sell-side firm A broker/dealer that sells securities and provides independent investment research and recommendations to their clients (i.e., buy-side firms).

Semi-strong-form efficient market A market in which security prices reflect all publicly known and available information.

Semiannual bond basis yield An annual rate having a periodicity of two; also known as a *semiannual bond equivalent yield*.

Semiannual bond equivalent yield See *semiannual bond basis yield*.

Seniority ranking Priority of payment of various debt obligations.

Sensitivity analysis Analysis that shows the range of possible outcomes as specific assumptions are changed.

Separately managed account (SMA) An investment portfolio managed exclusively for the benefit of an individual or institution.

Serial maturity structure Structure for a bond issue in which the maturity dates are spread out during the bond's life; a stated number of bonds mature and are paid off each year before final maturity.

Settlement The process that occurs after a trade is completed, the securities are passed to the buyer, and payment is received by the seller.

Settlement date Date when the buyer makes cash payment and the seller delivers the security.

Settlement price The official price, designated by the clearinghouse, from which daily gains and losses will be determined and marked to market.

Share repurchase A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Shareholder activism Strategies used by shareholders to attempt to compel a company to act in a desired manner.

Shareholder engagement The process whereby companies engage with their shareholders.

Shareholders' equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Sharpe ratio The average return in excess of the risk-free rate divided by the standard deviation of return; a measure of the average excess return earned per unit of standard deviation of return.

Shelf registration A type of public offering that allows the issuer to file a single, all-encompassing offering circular that covers a series of bond issues.

Short The seller of an asset or derivative contract. Also refers to the position of being short an asset or derivative contract.

Short position A position in an asset or contract in which one has sold an asset one does not own, or in which a right under a contract can be exercised against oneself.

Short-run average total cost The curve describing average total cost when some costs are considered fixed.

Short selling A transaction in which borrowed securities are sold with the intention to repurchase them at a lower price at a later date and return them to the lender.

Shortfall risk The risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Shutdown point The point at which average revenue is equal to the firm's average variable cost.

Side letters Side agreements created between the GP and specific LPs. These agreements exist *outside* the LPA. These agreements provide additional terms and conditions related to the investment agreement.

Simple interest The interest earned each period on the original investment; interest calculated on the principal only.

Simple linear regression (SLR) A regression that summarizes the relation between the dependent variable and a single independent variable.

Simple random sample A subset of a larger population created in such a way that each element of the population has an equal probability of being selected to the subset.

Simple random sampling The procedure of drawing a sample to satisfy the definition of a simple random sample.

Simple yield The sum of the coupon payments plus the straight-line amortized share of the gain or loss, divided by the flat price.

Simulation Computer-generated sensitivity or scenario analysis that is based on probability models for the factors that drive outcomes.

Simulation trial A complete pass through the steps of a simulation.

Single-step format With respect to the format of the income statement, a format that does not subtotal for gross profit (revenue minus cost of goods sold).

Sinking fund arrangement Provision that reduces the credit risk of a bond issue by requiring the issuer to retire a portion of the bond's principal outstanding each year.

Situational influences External factors, such as environmental or cultural elements, that shape our behavior.

Skewed Not symmetrical.

Skewness A quantitative measure of skew (lack of symmetry); a synonym of skew. It is computed as the average cubed deviation from the mean standardized by dividing by the standard deviation cubed.

Slope coefficient The coefficient of an independent variable that represents the average change in the dependent variable for a one-unit change in the independent variable.

Small country A country that is a price taker in the world market for a product and cannot influence the world market price.

Smart beta Involves the use of simple, transparent, rules-based strategies as a basis for investment decisions.

Smart contract A computer program that is designed to self-execute on the basis of pre-specified terms and conditions agreed to by parties to a contract.

Socially responsible investing An investment approach that excludes investments in companies or industries that deviate from an organization's beliefs and sometimes includes investments with favorable environmental or social profiles.

Soft-bullet covered bonds Covered bonds for which bond default and payment acceleration of bond cash flows may be delayed upon sponsor default until a new final maturity date is reached.

Solvency With respect to financial statement analysis, the ability of a company to fulfill its long-term obligations.

Solvency ratios Ratios that measure a company's ability to meet its long-term obligations.

Solvency risk The risk that an organization does not survive or succeed because it runs out of cash, even though it might otherwise be solvent.

Sovereign bond A bond issued by a national government. Also called "Sovereign."

Spearman rank correlation coefficient A measure of correlation applied to ranked data.

Special dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Special purpose entity A non-operating entity created to carry out a specified purpose, such as leasing assets or securitizing receivables; can be a corporation, partnership, trust, or limited liability partnership formed to facilitate a specific type of business activity. Also called *special purpose vehicle*, *special purpose company*, or *variable interest entity*.

Special purpose vehicle See *special purpose entity*.

Specific identification method An inventory accounting method that identifies which specific inventory items were sold and which remained in inventory to be carried over to later periods.

Speculative demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments. Also called *portfolio demand for money*.

Speculative money balances Monies held in anticipation that other assets will decline in value.

Split coupon bond See *deferred coupon bond*.

Sponsored A type of depository receipt in which the foreign company whose shares are held by the depository has a direct involvement in the issuance of the receipts.

Spot curve A sequence of yields-to-maturity on zero-coupon bonds. Sometimes called *zero* or *strip curve* (because coupon payments are "stripped" off the bonds).

Spot markets Markets in which assets are traded for immediate delivery.

Spot prices The price of an asset for immediately delivery.

Spot rates A sequence of market discount rates that correspond to the cash flow dates; yields-to-maturity on zero-coupon bonds maturing at the date of each cash flow.

Spread In general, the difference in yield between different fixed-income securities. Often used to refer to the difference between the yield-to-maturity and the benchmark.

Spread over the benchmark See *required yield spread*.

Spread risk Bond price risk arising from changes in the yield spread on credit-risky bonds; reflects changes in the market's assessment and/or pricing of credit migration (or downgrade) risk and market liquidity risk.

Spurious correlation Refers to: 1) correlation between two variables that reflects chance relationships in a particular dataset; 2) correlation induced by a calculation that mixes each of two variables with a third variable; and 3) correlation between two variables arising not from a direct relation between them but from their relation to a third variable.

Stacked bar chart An alternative form for presenting the frequency distribution of two categorical variables, where bars representing the sub-groups are placed on top of each other to form a single bar. Each sub-section is shown in a different color to represent the contribution of each sub-group, and the overall height of the stacked bar represents the marginal frequency for the category.

Stackelberg model A prominent model of strategic decision making in which firms are assumed to make their decisions sequentially.

Stagflation The combination of a high inflation rate with a high level of unemployment and a slowdown of the economy.

Staggered boards Election process whereby directors are typically divided into multiple classes that are elected separately in consecutive years—that is, one class every year.

Stakeholder management The identification, prioritization, and understanding of the interests of stakeholder groups and managing the company's relationships with these groups.

Stakeholders Individuals or groups of individuals who may be affected either directly or indirectly by a decision and thus have an interest, or stake, in the decision.

Standard deviation The positive square root of the variance; a measure of dispersion in the same units as the original data.

Standard error of the estimate A measure of the fit of a regression line, calculated as the square root of the mean square error. Also known as the *standard error of the regression* and the *root mean square error*.

Standard error of the forecast A measure of the uncertainty associated with a forecasted value of the dependent variable that depends on the standard error of the estimate, the variability of the independent variable, the deviation of the forecasted independent variable from the mean in the regression, and the number of observations.

Standard error of the slope coefficient The standard error of the slope, which in a simple linear regression is the ratio of the model's standard error of the estimate (s_e) to the square root of the variation of the independent variable.

Standard normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Standardizing A transformation that involves subtracting the mean and dividing the result by the standard deviation.

Standards of conduct Behaviors required by a group; established benchmarks that clarify or enhance a group's code of ethics.

Standing limit orders A limit order at a price below market and which therefore is waiting to trade.

Stated annual interest rate A quoted interest rate that does not account for compounding within the year. Also called *quoted interest rate*.

Statement of changes in equity (statement of owners' equity) A financial statement that reconciles the beginning-of-period and end-of-period balance sheet values of

shareholders' equity; provides information about all factors affecting shareholders' equity. Also called *statement of owners' equity*.

Statement of financial condition The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of financial position The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of operations A financial statement that provides information about a company's profitability over a stated period of time.

Static trade-off theory of capital structure A theory pertaining to a company's optimal capital structure. The optimal level of debt is found at the point where additional debt would cause the costs of financial distress to increase by a greater amount than the benefit of the additional tax shield.

Statistic A summary measure of a sample of observations.

Statistically significant A result indicating that the null hypothesis can be rejected; with reference to an estimated regression coefficient, frequently understood to mean a result indicating that the corresponding population regression coefficient is different from zero.

Status quo bias An emotional bias in which people do nothing (i.e., maintain the status quo) instead of making a change.

Statutory voting A common method of voting where each share represents one vote.

Step-up coupon bond Bond for which the coupon, fixed or floating, increases by specified margins at specified dates.

Stochastic oscillator A momentum indicator that compares a particular closing price of a security to a range of the security's prices over a certain period of time.

Stock dividend A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Stock split An increase in the number of shares outstanding with a consequent decrease in share price, but no change to the company's underlying fundamentals.

Stop-loss order See *stop order*.

Stop order An order in which a trader has specified a stop price condition. Also called *stop-loss order*.

Store of value The quality of tending to preserve value.

Store of wealth Goods that depend on the fact that they do not perish physically over time, and on the belief that others would always value the good.

Straight-line method A depreciation method that allocates evenly the cost of a long-lived asset less its estimated residual value over the estimated useful life of the asset.

Straight voting A shareholder voting process in which shareholders receive one vote for each share owned.

Strategic analysis Analysis of the competitive environment with an emphasis on the implications of the environment for corporate strategy.

Strategic asset allocation The set of exposures to IPS-permissible asset classes that is expected to achieve the client's long-term objectives given the client's investment constraints.

Stratified random sampling A procedure that first divides a population into subpopulations (strata) based on classification criteria and then randomly draws samples from each stratum in sizes proportional to that of each stratum in the population.

Street convention A yield measure that neglects weekends and holidays; the internal rate of return on cash flows assuming payments are made on the scheduled dates, even when the scheduled date falls on a weekend or holiday.

Stress testing A specific type of scenario analysis that estimates losses in rare and extremely unfavorable combinations of events or scenarios.

Strong-form efficient market A market in which security prices reflect all public and private information.

Structural (or cyclically adjusted) budget deficit The deficit that would exist if the economy was at full employment (or full potential output).

Structural subordination Arises in a holding company structure when the debt of operating subsidiaries is serviced by the cash flow and assets of the subsidiaries before funds can be passed to the holding company to service debt at the parent level.

Structured data Data that are highly organized in a predefined manner, usually with repeating patterns.

Structured financial instrument A financial instrument that shares the common attribute of repackaging risks. Structured financial instruments include asset-backed securities, collateralized debt obligations, and other structured financial instruments such as capital protected, yield enhancement, participation, and leveraged instruments.

Subjective probability A probability drawing on personal or subjective judgment.

Subordinated debt A class of unsecured debt that ranks below a firm's senior unsecured obligations.

Subordination Form of internal credit enhancement that relies on creating more than one bond tranche and ordering the claim priorities for ownership or interest in an asset between the tranches. The ordering of the claim priorities is called a senior/subordinated structure, where the tranches of highest seniority are called senior followed by subordinated or junior tranches. Also called *credit tranching*.

Substitutes Said of two goods or services such that if the price of one increases the demand for the other tends to increase, holding all other things equal (e.g., butter and margarine).

Sum of squares error (SSE) The sum of the squared deviations of (1) the value of the dependent variable and (2) the value of the dependent variable based on the estimated regression line. Also referred to as the *residual sum of squares*.

Sum of squares regression (SSR) The sum of the squared deviations of (1) the value of the dependent variable based on the estimated regression line and (2) the mean of the dependent variable.

Sum of squares total (SST) The sum of the squared deviations of the dependent variable from its mean; the variation of the dependent variable. Also referred to as the *total sum of squares*.

Sunk cost A cost that has already been incurred.

Supervised learning A machine learning approach that makes use of labeled training data.

Supply shock A typically unexpected disturbance to supply.

Support In technical analysis, a price range in which buying activity is sufficient to stop the decline in the price of a security.

Support tranches Classes or tranches in CMOs that protect PAC tranches from prepayment risk.

Supranational bond A bond issued by a supranational agency such as the World Bank.

Surety bond Form of external credit enhancement whereby a rated and regulated insurance company guarantees to reimburse bondholders for any losses incurred up to a maximum amount if the issuer defaults.

Survey approach An estimate of the equity risk premium that is based on estimates provided by a panel of finance experts.

Survivorship bias Relates to the inclusion of only current investment funds in a database. As such, the returns of funds that are no longer available in the marketplace (have been liquidated) are excluded from the database. In addition, *backfill bias* is another problem, whereby certain surviving hedge funds may be added to databases and various hedge fund indexes only after they are initially successful and start to report their returns.

Sustainability linked loans These are any types of loan instruments and/or contingent facilities (such as bonding lines, guarantee lines, or letters of credit) that incentivize the borrower's achievement of ambitious, pre-determined sustainability performance objectives.

Sustainable growth rate The rate of dividend (and earnings) growth that can be sustained over time for a given level of return on equity, keeping the capital structure constant and without issuing additional common stock.

Sustainable investing A term used in a similar context to responsible investing, but its key focus is on factoring in sustainability issues while investing.

Sustainable rate of economic growth The rate of increase in the economy's productive capacity or potential GDP.

Swap contract An agreement between two parties to exchange a series of future cash flows.

Syndicated loan A loan from a group of lenders to a single borrower.

Syndicated offering A bond issue underwritten by a group of investment banks.

Systematic risk Risk that affects the entire market or economy; it cannot be avoided and is inherent in the overall market. Systematic risk is also known as non-diversifiable or market risk.

Systematic sampling A procedure of selecting every *k*th member until reaching a sample of the desired size. The sample that results from this procedure should be approximately random.

Tactical asset allocation The decision to deliberately deviate from the strategic asset allocation in an attempt to add value based on forecasts of the near-term relative performance of asset classes.

Tactical asset allocation (TAA) A portfolio strategy that shifts the percentages of assets held in various asset classes (or categories) to take advantage of market opportunities. Allocation shifts can occur within an asset class or across asset classes.

Tag cloud see **word cloud**.

Target capital structure A company's chosen proportions of debt and equity.

Target independent A bank's ability to determine the definition of inflation that they target, the rate of inflation that they target, and the horizon over which the target is to be achieved.

Target semideviation A measure of downside risk, calculated as the square root of the average of the squared deviations of observations below the target (also called target downside deviation).

Tariffs Taxes that a government levies on imported goods.

Tax base The amount at which an asset or liability is valued for tax purposes.

Tax expense An aggregate of an entity's income tax payable (or recoverable in the case of a tax benefit) and any changes in deferred tax assets and liabilities. It is essentially the income tax payable or recoverable if these had been determined based on accounting profit rather than taxable income.

Tax loss carry forward A taxable loss in the current period that may be used to reduce future taxable income.

Taxable income The portion of an entity's income that is subject to income taxes under the tax laws of its jurisdiction.

Taxable temporary differences Temporary differences that result in a taxable amount in a future period when determining the taxable profit as the balance sheet item is recovered or settled.

Technical analysis A form of security analysis that uses price and volume data, often displayed graphically, in decision making.

Technology The process a company uses to transform inputs into outputs.

Tender offer Corporate takeover mechanism that involves shareholders selling their interests directly to the group seeking to gain control.

Tenor The time-to-maturity for a bond or derivative contract. Also called *term to maturity*.

Term maturity structure Structure for a bond issue in which the bond's notional principal is paid off in a lump sum at maturity.

Term structure See *maturity structure*.

Term structure of credit spreads The relationship between the spreads over the "risk-free" (or benchmark) rates and times-to-maturity.

Term structure of yield volatility The relationship between the volatility of bond yields-to-maturity and times-to-maturity.

Terminal stock value The expected value of a share at the end of the investment horizon—in effect, the expected selling price. Also called *terminal value*.

Terminal value The expected value of a share at the end of the investment horizon—in effect, the expected selling price.

Terms of trade The ratio of the price of exports to the price of imports, representing those prices by export and import price indexes, respectively.

Test of the mean of the differences A statistical test for differences based on paired observations drawn from samples that are dependent on each other.

Text analytics The use of computer programs to analyze and derive meaning from typically large, unstructured text- or voice-based datasets.

Thematic investing An ESG investment approach that focuses on investing in themes or assets specifically relating to ESG factors, such as clean energy, green technology, or sustainable agriculture.

Thin-Tailed Describes a distribution that has relatively less weight in the tails than the normal distribution (also called platykurtic).

Third-degree price discrimination When the monopolist segregates customers into groups based on demographic or other characteristics and offers different pricing to each group.

Time-period bias The possibility that when we use a time-series sample, our statistical conclusion may be sensitive to the starting and ending dates of the sample.

Time-series data A sequence of observations for a single observational unit of a specific variable collected over time and at discrete and typically equally spaced intervals of time (such as daily, weekly, monthly, annually, or quarterly).

Time tranching The creation of classes or tranches in an ABS/MBS that possess different (expected) maturities.

Time value The difference between the market price of the option and its intrinsic value.

Time value decay Said of an option when, at expiration, no time value remains and the option is worth only its exercise value.

Time value of money The principles governing equivalence relationships between cash flows with different dates.

Time-weighted rate of return The compound rate of growth of one unit of currency invested in a portfolio during a stated measurement period; a measure of investment performance that is not sensitive to the timing and amount of withdrawals or additions to the portfolio.

Tokenization The process of representing ownership rights to physical assets on a blockchain or distributed ledger.

Top-down analysis An investment selection approach that begins with consideration of macroeconomic conditions and then evaluates markets and industries based upon such conditions.

Total comprehensive income The change in equity during a period resulting from transaction and other events, other than those changes resulting from transactions with owners in their capacity as owners.

Total cost The summation of all costs, for which costs are classified as fixed or variable.

Total factor productivity A scale factor that reflects the portion of growth unaccounted for by explicit factor inputs (e.g., capital and labor).

Total fixed cost The summation of all expenses that do not change as the level of production varies.

Total invested capital The sum of market value of common equity, book value of preferred equity, and face value of debt.

Total probability rule A rule explaining the unconditional probability of an event in terms of probabilities of the event conditional on mutually exclusive and exhaustive scenarios.

Total probability rule for expected value A rule explaining the expected value of a random variable in terms of expected values of the random variable conditional on mutually exclusive and exhaustive scenarios.

Total return Measures the price appreciation, or percentage change in price of the securities in an index or portfolio, plus any income received over the period.

Total return index An index that reflects the price appreciation or percentage change in price of the constituent securities plus any income received since inception.

Total return swap A swap in which one party agrees to pay the total return on a security. Often used as a credit derivative, in which the underlying is a bond.

Total variable cost The summation of all variable expenses.

Tracking error The standard deviation of the differences between a portfolio's returns and its benchmark's returns; a synonym of active risk.

Tracking risk The standard deviation of the differences between a portfolio's returns and its benchmarks returns. Also called *tracking error*.

Trade creation When regional integration results in the replacement of higher cost domestic production by lower cost imports from other members.

Trade credit A spontaneous form of credit in which a purchaser of the goods or service is financing its purchase by delaying the date on which payment is made.

Trade diversion When regional integration results in lower-cost imports from non-member countries being replaced with higher-cost imports from members.

Trade payables Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Trade protection Government policies that impose restrictions on trade, such as tariffs and quotas.

Trade surplus (deficit) When the value of exports is greater (less) than the value of imports.

Trading securities Under US GAAP, a category of debt securities held by a company with the intent to trade them. Also called *held-for-trading securities*.

Traditional investment markets Markets for traditional investments, which include all publicly traded debts and equities and shares in pooled investment vehicles that hold publicly traded debts and/or equities.

Transactions money balances Money balances that are held to finance transactions.

Transfer payments Welfare payments made through the social security system that exist to provide a basic minimum level of income for low-income households.

Transparency Said of something (e.g., a market) in which information is fully disclosed to the public and/or regulators.

Treasury stock method A method for accounting for the effect of options (and warrants) on earnings per share (EPS) that specifies what EPS would have been if the options and warrants had been exercised and the company had used the proceeds to repurchase common stock.

Tree-Map Another graphical tool for displaying categorical data. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group.

Trend A long-term pattern of movement in a particular direction.

Treynor ratio A measure of risk-adjusted performance that relates a portfolio's excess returns to the portfolio's beta.

Triangle pattern In technical analysis, a continuation chart pattern that forms as the range between high and low prices narrows, visually forming a triangle.

Trimmed mean A mean computed after excluding a stated small percentage of the lowest and highest observations.

Trimodal A distribution that has the three most frequently occurring values.

Triple bottom In technical analysis, a reversal pattern that results when the price forms three troughs at roughly the same price level. A triple bottom is used to predict a change from a downtrend to an uptrend.

Triple top In technical analysis, a reversal pattern that results when the price forms three peaks at roughly the same price level. A triple top is used to predict a change from an uptrend to a downtrend.

True yield The internal rate of return on cash flows using the actual calendar, including weekends and bank holidays.

Trust deed The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *bond indenture*.

Turn-of-the-year effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January.

Two-dimensional rectangular array A popular form for organizing data for processing by computers or for presenting data visually. It is comprised of columns and rows to hold multiple variables and multiple observations, respectively (also called a **data table**).

Two-fund separation theorem The theory that all investors regardless of taste, risk preferences, and initial wealth will hold a combination of two portfolios or funds: a risk-free asset and an optimal portfolio of risky assets.

Two-sided hypothesis test A test in which the null hypothesis is rejected in favor of the alternative hypothesis if the evidence indicates that the population parameter is either smaller or larger than a hypothesized value; occurs when the alternative hypothesis is stated as not equal to the hypothesized population parameters.

Two-way table See *contingency table*.

Two-week repo rate The interest rate on a two-week repurchase agreement; may be used as a policy rate by a central bank.

Type I error The error of rejecting a true null hypothesis; a false positive.

Type II error The error of not rejecting a false null hypothesis; false negative.

Unanticipated (unexpected) inflation The component of inflation that is a surprise.

Uncommitted lines of credit The least reliable form of bank borrowing in which a bank offers, without formal commitment, a line of credit for an extended period of time but reserves the right to refuse any request for its use.

Unconditional probability The probability of an event *not* conditioned on another event.

Underemployed A person who has a job but has the qualifications to work a significantly higher-paying job.

Underlying An asset that trades in a market in which buyers and sellers meet, decide on a price, and the seller then delivers the asset to the buyer and receives payment. The underlying is the asset or other derivative on which a particular derivative is based. The market for the underlying is also referred to as the *spot market*.

Underwriter A firm, usually an investment bank, that takes the risk of buying the newly issued securities from the issuer and then reselling them to investors or to dealers, thus guaranteeing the sale of the securities at the offering price negotiated with the issuer.

Underwritten offering A type of securities issue mechanism in which the investment bank guarantees the sale of the securities at an offering price that is negotiated with the issuer. Also known as *firm commitment offering*.

Unearned revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service. Also called *deferred revenue* or *deferred income*.

Unemployed People who are actively seeking employment but are currently without a job.

Unemployment rate The ratio of unemployed to the labor force.

Unexpected inflation The component of inflation that is a surprise.

Unimodal A distribution with a single value that is most frequently occurring.

Unit elastic An elasticity with a magnitude of negative one. Also called *unitary elastic*.

Unit labor cost The average labor cost to produce one unit of output.

Unit normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Unitranche debt Consists of a hybrid or blended loan structure that combines different tranches of secured and unsecured debt into a single loan with a single, blended interest rate.

Units-of-production method A depreciation method that allocates the cost of a long-lived asset based on actual usage during the period.

Univariate distribution A distribution that specifies the probabilities for a single random variable.

Universal owners Long-term investors, such as pension funds, that have significant assets invested in globally diversified portfolios.

Unlimited funds An unlimited funds environment assumes that the company can raise the funds it wants for all profitable investments simply by paying the required rate of return.

Unsecured debt Debt that gives the debtholder only a general claim on an issuer's assets and cash flow.

Unsponsored A type of depository receipt in which the foreign company whose shares are held by the depository has no involvement in the issuance of the receipts.

Unstructured data Data that do not follow any conventionally organized forms.

Unsupervised learning A machine learning approach that does not make use of labeled training data.

Up transition probability The probability that an asset's value moves up.

Uptrend A pattern that occurs when the price of an asset moves higher over a period of time.

Validity instructions Instructions which indicate when the order may be filled.

Valuation allowance A reserve created against deferred tax assets, based on the likelihood of realizing the deferred tax assets in future accounting periods.

Valuation ratios Ratios that measure the quantity of an asset or flow (e.g., earnings) in relation to the price associated with a specified claim (e.g., a share or ownership of the enterprise).

Value at risk A money measure of the minimum value of losses expected during a specified time period at a given level of probability.

Value investors With reference to equity investors, investors who are focused on paying a relatively low share price in relation to earnings or assets per share.

VaR See *value at risk*.

Variable A characteristic or quantity that can be measured, counted, or categorized and that is subject to change (also called a field, an attribute, or a feature).

Variable costs Costs that fluctuate with the level of production and sales.

Variance The expected value (the probability-weighted average) of squared deviations from a random variable's expected value.

Variation margin Additional margin that must be deposited in an amount sufficient to bring the balance up to the initial margin requirement.

Veblen goods Goods that increase in desirability with increasing price.

Vega A measure of the sensitivity of an option's price to changes in the underlying's volatility.

Venture capital Investments that provide "seed" or startup capital, early-stage financing, or later-stage financing (including mezzanine-stage financing) to companies that are in early development stages and require additional capital for expansion or preparation for an initial public offering.

Venture capital fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *private equity fund*.

Venture capital funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *private equity funds.*

Vertical analysis Common-size analysis using only one reporting period or one base financial statement; for example, an income statement in which all items are stated as percentages of sales.

Vertical demand schedule Implies that some fixed quantity is demanded, regardless of price.

Visual technique The most common and readily available method of initial data assessment. Experts in pattern recognition maintain that the visual (or "eyeball") technique is still the most effective way of searching for recognizable patterns.

Visualization The presentation of data in a pictorial or graphical format for the purpose of increasing understanding and for gaining insights into the data.

Volatility The standard deviation of the continuously compounded returns on the underlying asset.

Voluntarily unemployed A person voluntarily outside the labor force, such as a jobless worker refusing an available vacancy.

Voluntary export restraint A trade barrier under which the exporting country agrees to limit its exports of the good to its trading partners to a specific number of units.

Vote by proxy A mechanism that allows a designated party—such as another shareholder, a shareholder representative, or management—to vote on the shareholder's behalf.

Warrant Attached option that gives its holder the right to buy the underlying stock of the issuing company at a fixed exercise price until the expiration date.

Waterfall Represents the distribution method that defines the order in which allocations are made to LPs and GPs. There are two major types of waterfall: *deal by deal* (or *American*) and *whole of fund* (or *European*).

Weak-form efficient market hypothesis The belief that security prices fully reflect all past market data, which refers to all historical price and volume trading information.

Wealth effect An increase (decrease) in household wealth increases (decreases) consumer spending out of a given level of current income.

Web-based lender A lender that operates primarily on the internet, offering loans in relatively small amounts, typically to small businesses in need of cash.

Weighted average cost method An inventory accounting method that averages the total cost of available inventory items over the total units available for sale.

Weighted average cost of capital A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average cost of capital (WACC) A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average coupon rate Weighting the mortgage rate of each mortgage loan in the pool by the percentage of the mortgage outstanding relative to the outstanding amount of all the mortgages in the pool.

Weighted average life A measure that gives investors an indication of how long they can expect to hold the MBS before it is paid off; the convention-based average time to receipt of all principal repayments. Also called *average life*.

Weighted average maturity Weighting the remaining number of months to maturity for each mortgage loan in the pool by the amount of the outstanding mortgage balance.

Weighted mean An average in which each observation is weighted by an index of its relative importance.

Winsorized mean A mean computed after assigning a stated percentage of the lowest values equal to one specified low value and a stated percentage of the highest values equal to one specified high value.

Word cloud A visual device for representing textual data, which consists of words extracted from a source of textual data. The size of each distinct word is proportional to the frequency with which it appears in the given text (also known as **tag cloud**).

Working capital The difference between current assets and current liabilities.

Working capital management The management of a company's short-term assets (such as inventory) and short-term liabilities (such as money owed to suppliers).

World price The price prevailing in the world market.

Yield duration The sensitivity of the bond price with respect to the bond's own yield-to-maturity.

Yield to maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield to redemption* or *redemption yield*.

Yield-to-maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield-to-redemption* or *redemption yield*.

Yield-to-redemption See *yield-to-maturity*.

Yield-to-worst The lowest of the sequence of yields-to-call and the yield-to-maturity.

Zero-coupon bond A bond that does not pay interest during its life. It is issued at a discount to par value and redeemed at par. Also called *pure discount bond*.

Zero-volatility spread (Z-spread) Calculates a constant yield spread over a government (or interest rate swap) spot curve.

CFA Institute®
CFA Program

FINANCIAL STATEMENT ANALYSIS AND CORPORATE ISSUERS

CFA® Program Curriculum
2022 • LEVEL I • VOLUME 3

WILEY

CONTENTS

How to Use the CFA Program Curriculum xi
Background on the CBOK xi
Organization of the Curriculum xii
Features of the Curriculum xii
Designing Your Personal Study Program xiii
CFA Institute Learning Ecosystem (LES) xiv
Prep Providers xv
Feedback xvi

Financial Statement Analysis

Study Session 6 **Financial Statement Analysis (2)** 3

Reading 17 **Understanding Income Statements** 5
Introduction 6
Components and Format of the Income Statement 6
Revenue Recognition 12
  General Principles 13
  Accounting Standards for Revenue Recognition 14
Expense Recognition: General Principles 18
  General Principles 18
Issues in Expense Recognition: Doubtful Accounts, Warranties 22
  Doubtful Accounts 22
  Warranties 22
Issues in Expense Recognition: Depreciation and Amortization 23
Implications for Financial Analysts: Expense Recognition 27
Non-Recurring Items and Non-Operating Items: Discontinued Operations and Unusual or Infrequent items 28
  Discontinued Operations 28
  Unusual or Infrequent Items 29
Non-Recurring Items: Changes in Accounting Policy 30
Non-Operating Items 33
Earnings Per Share and Capital Structure and Basic EPS 34
  Simple versus Complex Capital Structure 35
  Basic EPS 36
Diluted EPS: the If-Converted Method 37
  Diluted EPS When a Company Has Convertible Preferred Stock Outstanding 38
  Diluted EPS When a Company Has Convertible Debt Outstanding 39
Diluted EPS: the Treasury Stock Method 40
Other Issues with Diluted EPS and Changes in EPS 43
  Changes in EPS 44
Common-Size Analysis of the Income Statement 44
  Common-Size Analysis of the Income Statement 45
Income Statement Ratios 47

O indicates an optional segment

Comprehensive Income 49
Summary 52
Practice Problems 55
Solutions 60

Reading 18 **Understanding Balance Sheets** 63
Introduction and Components of the Balance Sheet 64
  Components and Format of the Balance Sheet 64
  Balance Sheet Components 65
Current and Non-Current Classification 67
Liquidity-Based Presentation 68
Current Assets: Cash and Cash Equivalents, Marketable Securities and Trade Receivables 69
  Current Assets 69
Current Assets: Inventories and Other Current Assets 74
  Other Current Assets 74
Current liabilities 75
Non-Current Assets: Property, Plant and Equipment and Investment Property 79
  Property, Plant, and Equipment 80
  Investment Property 81
Non-Current Assets: Intangible Assets 81
  Identifiable Intangibles 82
Non-Current Assets: Goodwill 84
Non-Current Assets: Financial Assets 87
Non-Current Assets: Deferred Tax Assets 91
Non-Current Liabilities 91
  Long-term Financial Liabilities 93
  Deferred Tax Liabilities 93
Components of Equity 94
  Components of Equity 94
Statement of Changes in Equity 97
Common Size Analysis of Balance Sheet 98
  Common-Size Analysis of the Balance Sheet 98
Balance Sheet Ratios 106
Summary 108
Practice Problems 111
Solutions 116

Reading 19 **Understanding Cash Flow Statements** 119
Introduction 120
Classification Of Cash Flows and Non-Cash Activities 121
  Classification of Cash Flows and Non-Cash Activities 121

Cash Flow Statement: Differences Between IFRS and US GAAP 123
Cash Flow Statement: Direct and Indirect Methods for Reporting Cash Flow from Operating Activities 125
Cash Flow Statement: Indirect Method Under IFRS 126
Cash Flow Statement: Direct Method Under IFRS 129
Cash Flow Statement: Direct Method Under US GAAP 131
Cash Flow Statement: Indirect Method Under US GAAP 133
Linkages of Cash Flow Statement with the Income Statement and Balance Sheet 135
    Linkages of the Cash Flow Statement with the Income Statement and Balance Sheet 135
Preparing the Cash Flow Statement: The Direct Method for Operating Activities 137
    Operating Activities: Direct Method 138
Preparing the Cash Flow Statement: Investing Activities 142
Preparing the Cash Flow Statement: Financing Activities 144
    Long-Term Debt and Common Stock 144
    Dividends 145
Preparing the Cash Flow Statement: Overall Statement of Cash Flows Under the Direct Method 145
Preparing the Cash Flow Statement: Overall Statement of Cash Flows Under the Indirect Method 146
Conversion of Cash Flows from the Indirect to Direct Method 149
Cash Flow Statement Analysis: Evaluation of Sources and Uses of Cash 150
    Evaluation of the Sources and Uses of Cash 150
Cash Flow Statement Analysis: Common Size Analysis 155
Cash Flow Statement Analysis: Free Cash Flow to Firm and Free Cash Flow to Equity 160
Cash Flow Statement Analysis: Cash Flow Ratios 162
Summary 164
Practice Problems 165
Solutions 171

Reading 20 **Financial Analysis Techniques** 175
Introduction 176
    The Financial Analysis Process 177
Analytical Tools and Techniques 181
Financial Ratio Analysis 183
    The Universe of Ratios 184
    Value, Purposes, and Limitations of Ratio Analysis 186
    Sources of Ratios 187
Common Size Balance Sheets and Income Statements 188
    Common-Size Analysis of the Balance Sheet 189
    Common-Size Analysis of the Income Statement 189
Cross-Sectional, Trend Analysis & Relationships in Financial Statements 191
    Trend Analysis 192
    Relationships among Financial Statements 194
The Use of Graphs and Regression Analysis 195
    Regression Analysis 197

Common Ratio Categories & Interpretation and Context 197
    Interpretation and Context 198
Activity Ratios 199
    Calculation of Activity Ratios 199
    Interpretation of Activity Ratios 201
Liquidity Ratios 205
    Calculation of Liquidity Ratios 206
    Interpretation of Liquidity Ratios 207
Solvency Ratios 210
    Calculation of Solvency Ratios 211
    Interpretation of Solvency Ratios 212
Profitability Ratios 215
    Calculation of Profitability Ratios 215
    Interpretation of Profitability Ratios 216
Integrated Financial Ratio Analysis 218
    The Overall Ratio Picture: Examples 219
DuPont Analysis: The Decomposition of ROE 221
Equity Analysis and Valuation Ratios 226
    Valuation Ratios 226
Industry-Specific Financial Ratios 229
Research on Financial Ratios in Credit and Equity Analysis 231
Credit Analysis 232
    The Credit Rating Process 233
    Historical Research on Ratios in Credit Analysis 234
Business and Geographic Segments 234
    Segment Reporting Requirements 235
    Segment Ratios 236
Model Building and Forecasting 238
Summary 239
Practice Problems 241
Solutions 247

Study Session 7 **Financial Statement Analysis (3)** 251

Reading 21 **Inventories** 253
Introduction 254
Cost of inventories 255
Inventory valuation methods 256
    Specific Identification 257
    First-In, First-Out (FIFO) 257
    Weighted Average Cost 258
    Last-In, First-Out (LIFO) 258
Calculations of cost of sales, gross profit, and ending inventory 258
Periodic versus perpetual inventory systems 260
Comparison of inventory valuation methods 263
The LIFO method and LIFO reserve 265
    LIFO Reserve 266

LIFO liquidations 266
Inventory method changes 274
Inventory adjustments 275
Evaluation of inventory management: Disclosures & ratios 282
  Presentation and Disclosure 282
  Inventory Ratios 283
Illustrations of inventory analysis: Adjusting LIFO to FIFO 284
Illustrations of inventory analysis: Impacts of writedowns 288
Summary 294
Practice Problems 297
Solutions 313

Reading 22 **Long-Lived Assets** 321
Introduction & Acquisition of Property, Plant and Equipment 322
  Introduction 322
  Acquisition of Long-Lived Assets 323
  Property, Plant, and Equipment 323
Acquisition of Intangible Assets 326
  Intangible Assets Purchased in Situations Other Than Business Combinations 326
  Intangible Assets Developed Internally 327
  Intangible Assets Acquired in a Business Combination 328
Capitalization versus Expensing: Impact on Financial Statements and Ratios 330
Capitalisation of Interest Costs 335
Capitalisation of Interest and Internal Development Costs 338
Depreciation of Long-Lived Assets: Methods and Calculation 342
  Depreciation Methods and Calculation of Depreciation Expense 343
Amortisation of Long-Lived Assets: Methods and Calculation 351
The Revaluation Model 352
Impairment of Assets 356
  Impairment of Property, Plant, and Equipment 357
  Impairment of Intangible Assets with a Finite Life 359
  Impairment of Intangibles with Indefinite Lives 359
  Impairment of Long-Lived Assets Held for Sale 359
  Reversals of Impairments of Long-Lived Assets 360
Derecognition 360
  Sale of Long-Lived Assets 360
  Long-Lived Assets Disposed of Other Than by a Sale 361
Presentation and Disclosure Requirements 363
Using Disclosures in Analysis 370
Investment Property 374
Summary 377
Practice Problems 380
Solutions 392

Reading 23 **Income Taxes** **397**
- Introduction **398**
- Differences Between Accounting Profit and Taxable Income **398**
- Current and Deferred Tax Assets and Liabilities **399**
  - Deferred Tax Assets and Liabilities **400**
- Determining the Tax Base of Assets and Liabilities **403**
  - Determining the Tax Base of an Asset **404**
  - Determining the Tax Base of a Liability **405**
- Changes in Income Tax Rates **407**
- Temporary and Permanent Differences Between Taxable and Accounting Profit **408**
  - Taxable Temporary Differences **409**
  - Deductible Temporary Differences **410**
  - Examples of Taxable and Deductible Temporary Differences **410**
- Exceptions to the Usual Rules for Temporary Differences **412**
  - Business Combinations and Deferred Taxes **413**
  - Investments in Subsidiaries, Branches, Associates and Interests in Joint Ventures **413**
- Unused Tax Losses and Tax Credits **413**
- Recognition and Measurement of Current and Deferred Tax **414**
  - Recognition of a Valuation Allowance **415**
  - Recognition of Current and Deferred Tax Charged Directly to Equity **415**
- Presentation and Disclosure **418**
- Comparison of IFRS and US GAAP **424**
- *Summary* **426**
- *Practice Problems* **428**
- *Solutions* **433**

Reading 24 **Non-Current (Long-Term) Liabilities** **435**
- Introduction **436**
- Bonds Payable & Accounting for Bond Issuance **436**
  - Accounting for Bond Issuance **436**
- Accounting for Bond Amortisation, Interest Expense, and Interest Payments **440**
- Accounting for Bonds at Fair Value **445**
- Derecognition of Debt **448**
- Debt Covenants **451**
- Presentation and Disclosure of Long-Term Debt **453**
- Leases **456**
  - Examples of Leases **457**
  - Advantages of Leasing **457**
  - Lease Classification as Finance or Operating **457**
  - Financial Reporting of Leases **459**
  - Lessee Accounting—IFRS **459**
  - Lessee Accounting—US GAAP **461**
  - Lessor Accounting **463**
- Introduction to Pensions and Other Post-Employment Benefits **465**
- Evaluating Solvency: Leverage and Coverage Ratios **469**

Summary 472
Practice Problems 475
Solutions 480

Study Session 8 **Financial Statement Analysis (4)** 487

Reading 25 **Financial Reporting Quality** 489
Introduction & Conceptual Overview 490
  Conceptual Overview 491
GAAP, Decision Useful Financial Reporting 492
  GAAP, Decision-Useful, but Sustainable? 493
Biased Accounting Choices 494
  Within GAAP, but "Earnings Management" 503
Departures from GAAP 503
Differentiate between Conservative and Aggressive Accounting 505
  Conservatism in Accounting Standards 506
  Bias in the Application of Accounting Standards 507
Context for Assessing Financial Reporting Quality: Motivations and Conditions Conducive to Issuing Low Quality Financial Reports 508
  Motivations 509
  Conditions Conducive to Issuing Low-Quality Financial Reports 509
Mechanisms That Discipline Financial Reporting Quality 510
  Market Regulatory Authorities 510
  Auditors 512
  Private Contracting 516
Detection of Financial Reporting Quality Issues: Introduction & Presentation Choices 518
  Presentation Choices 519
Accounting Choices and Estimates and How Accounting Choices and Estimates Affect Earnings and Balance Sheets 525
  How Accounting Choices and Estimates Affect Earnings and Balance Sheets 526
How Choices that Affect the Cash Flow Statement 537
Choices that Affect Financial Reporting 540
Warning Signs 544
  1) Pay attention to revenue. 544
  2) Pay attention to signals from inventories. 545
  3) Pay attention to capitalization policies and deferred costs. 546
  4) Pay attention to the relationship of cash flow and net income. 546
  5) Other potential warnings signs. 546
Summary 549
Practice Problems 552
Solutions 556

Reading 26 **Applications of Financial Statement Analysis** 561
Introduction & Evaluating Past Financial Performance 562
  Application: Evaluating Past Financial Performance 563

Application: Projecting Future Financial Performance as an Input to Market Based Valuation 566
Projecting Performance: An Input to Market-Based Valuation 567
Projecting Multiple-Period Performance 572
Application: Assessing Credit Risk 576
Screening for Potential Equity Investments 578
Framework for Analyst Adjustments & Adjustments to Investments & Adjustments to Inventory 581
A Framework for Analyst Adjustments 582
Analyst Adjustments Related to Investments 582
Analyst Adjustments Related to Inventory 582
Adjustments Related to Property, Plant, and Equipment 586
Adjustments Related to Goodwill 588
Summary 590
Practice Problems 592
Solutions 594

Corporate Issuers

Study Session 9 **Corporate Issuers (1)** 597

Reading 27 **Introduction to Corporate Governance and Other ESG Considerations** 599
Introduction and Overview of Corporate Governance 600
Corporate Governance Overview 600
Stakeholder Groups 602
Stakeholder Groups 602
Principal–Agent and Other Relationships in Corporate Governance 605
Shareholder and Manager/Director Relationships 606
Controlling and Minority Shareholder Relationships 606
Manager and Board Relationships 607
Shareholder versus Creditor Interests 607
Other Stakeholder Conflicts 608
Overview and Mechanisms of Stakeholder Management 608
Overview of Stakeholder Management 609
Mechanisms of Stakeholder Management 609
Mechanisms to Mitigate Associated Stakeholder Risks 613
Employee Laws and Contracts 614
Contractual Agreements with Customers and Suppliers 615
Laws and Regulations 615
Company Boards and Committees 615
Composition of the Board of Directors 616
Functions and Responsibilities of the Board 617
Board of Directors Committees 617
Relevant Factors in Analyzing Corporate Governance and Stakeholder Management 620
Market Factors 620
Non-Market Factors 622
Risks and Benefits of Corporate Governance and Stakeholder Management 623
Risks of Poor Governance and Stakeholder Management 623

Benefits of Effective Governance and Stakeholder Management 625
Factors Relevant to Corporate Governance and Stakeholder Management Analysis 626
Economic Ownership and Voting Control 627
Board of Directors Representation 627
Remuneration and Company Performance 628
Investors in the Company 629
Strength of Shareholders' Rights 629
Managing Long-Term Risks 630
Summary of Analyst Considerations 630
ESG Considerations for Investors and Analysts 632
Introduction to Environmental, Social, and Governance issues 632
ESG Investment Strategies 633
ESG Investment Approaches 634
Catalysts for Growth in ESG Investing 636
ESG Market Overview 637
ESG Factors in Investment Analysis 638
Summary 640
Practice Problems 642
Solutions 644

Reading 28 **Uses of Capital** 645
Introduction 645
The Capital Allocation Process 646
Investment Decision Criteria 651
Net Present Value 651
Internal Rate of Return 652
Corporate Usage of Capital Allocation Methods 654
Real Options 656
Timing Options 657
Sizing Options 657
Flexibility Options 657
Fundamental Options 658
Common Capital Allocation Pitfalls 659
Summary 660
Practice Problems 662
Solutions 666

Reading 29 **Sources of Capital** 669
Introduction 669
Corporate Financing Options 670
Internal Financing 671
Financial Intermediaries 672
Capital Markets 673
Other Financing 676
Considerations Affecting Financing Choices 676
Managing and Measuring Liquidity 680
Defining Liquidity Management 681
Measuring Liquidity 684

Evaluating Short-Term Financing Choices **688**
Summary **691**
Practice Problems **692**
Solutions **695**

Glossary **G-1**

How to Use the CFA Program Curriculum

Congratulations on your decision to enter the Chartered Financial Analyst (CFA®) Program. This exciting and rewarding program of study reflects your desire to become a serious investment professional. You are embarking on a program noted for its high ethical standards and the breadth of knowledge, skills, and abilities (competencies) it develops. Your commitment should be educationally and professionally rewarding.

The credential you seek is respected around the world as a mark of accomplishment and dedication. Each level of the program represents a distinct achievement in professional development. Successful completion of the program is rewarded with membership in a prestigious global community of investment professionals. CFA charterholders are dedicated to life-long learning and maintaining currency with the ever-changing dynamics of a challenging profession. CFA Program enrollment represents the first step toward a career-long commitment to professional education.

The CFA exam measures your mastery of the core knowledge, skills, and abilities required to succeed as an investment professional. These core competencies are the basis for the Candidate Body of Knowledge (CBOK™). The CBOK consists of four components:

- A broad outline that lists the major CFA Program topic areas (www.cfainstitute.org/programs/cfa/curriculum/cbok);
- Topic area weights that indicate the relative exam weightings of the top-level topic areas (www.cfainstitute.org/programs/cfa/curriculum);
- Learning outcome statements (LOS) that advise candidates about the specific knowledge, skills, and abilities they should acquire from readings covering a topic area (LOS are provided in candidate study sessions and at the beginning of each reading); and
- CFA Program curriculum that candidates receive upon exam registration.

Therefore, the key to your success on the CFA exams is studying and understanding the CBOK. The following sections provide background on the CBOK, the organization of the curriculum, features of the curriculum, and tips for designing an effective personal study program.

BACKGROUND ON THE CBOK

CFA Program is grounded in the practice of the investment profession. CFA Institute performs a continuous practice analysis with investment professionals around the world to determine the competencies that are relevant to the profession, beginning with the Global Body of Investment Knowledge (GBIK®). Regional expert panels and targeted surveys are conducted annually to verify and reinforce the continuous feedback about the GBIK. The practice analysis process ultimately defines the CBOK. The CBOK reflects the competencies that are generally accepted and applied by investment professionals. These competencies are used in practice in a generalist context and are expected to be demonstrated by a recently qualified CFA charterholder.

© 2021 CFA Institute. All rights reserved.

The CFA Institute staff—in conjunction with the Education Advisory Committee and Curriculum Level Advisors, who consist of practicing CFA charterholders—designs the CFA Program curriculum in order to deliver the CBOK to candidates. The exams, also written by CFA charterholders, are designed to allow you to demonstrate your mastery of the CBOK as set forth in the CFA Program curriculum. As you structure your personal study program, you should emphasize mastery of the CBOK and the practical application of that knowledge. For more information on the practice analysis, CBOK, and development of the CFA Program curriculum, please visit www.cfainstitute.org.

ORGANIZATION OF THE CURRICULUM

The Level I CFA Program curriculum is organized into 10 topic areas. Each topic area begins with a brief statement of the material and the depth of knowledge expected. It is then divided into one or more study sessions. These study sessions should form the basic structure of your reading and preparation. Each study session includes a statement of its structure and objective and is further divided into assigned readings. An outline illustrating the organization of these study sessions can be found at the front of each volume of the curriculum.

The readings are commissioned by CFA Institute and written by content experts, including investment professionals and university professors. Each reading includes LOS and the core material to be studied, often a combination of text, exhibits, and in-text examples and questions. End of Reading Questions (EORQs) followed by solutions help you understand and master the material. The LOS indicate what you should be able to accomplish after studying the material. The LOS, the core material, and the EORQs are dependent on each other, with the core material and EORQs providing context for understanding the scope of the LOS and enabling you to apply a principle or concept in a variety of scenarios.

The entire readings, including the EORQs, are the basis for all exam questions and are selected or developed specifically to teach the knowledge, skills, and abilities reflected in the CBOK.

You should use the LOS to guide and focus your study because each exam question is based on one or more LOS and the core material and practice problems associated with the LOS. As a candidate, you are responsible for the entirety of the required material in a study session.

We encourage you to review the information about the LOS on our website (www.cfainstitute.org/programs/cfa/curriculum/study-sessions), including the descriptions of LOS “command words” on the candidate resources page at www.cfainstitute.org.

FEATURES OF THE CURRICULUM

End of Reading Questions/Solutions *All End of Reading Questions (EORQs) as well as their solutions are part of the curriculum and are required material for the exam.* In addition to the in-text examples and questions, these EORQs help demonstrate practical applications and reinforce your understanding of the concepts presented. Some of these EORQs are adapted from past CFA exams and/or may serve as a basis for exam questions.

Glossary For your convenience, each volume includes a comprehensive Glossary. Throughout the curriculum, a **bolded** word in a reading denotes a term defined in the Glossary.

Note that the digital curriculum that is included in your exam registration fee is searchable for key words, including Glossary terms.

LOS Self-Check We have inserted checkboxes next to each LOS that you can use to track your progress in mastering the concepts in each reading.

Source Material The CFA Institute curriculum cites textbooks, journal articles, and other publications that provide additional context or information about topics covered in the readings. As a candidate, you are not responsible for familiarity with the original source materials cited in the curriculum.

Note that some readings may contain a web address or URL. The referenced sites were live at the time the reading was written or updated but may have been deactivated since then.

Some readings in the curriculum cite articles published in the *Financial Analysts Journal*®, which is the flagship publication of CFA Institute. Since its launch in 1945, the *Financial Analysts Journal* has established itself as the leading practitioner-oriented journal in the investment management community. Over the years, it has advanced the knowledge and understanding of the practice of investment management through the publication of peer-reviewed practitioner-relevant research from leading academics and practitioners. It has also featured thought-provoking opinion pieces that advance the common level of discourse within the investment management profession. Some of the most influential research in the area of investment management has appeared in the pages of the *Financial Analysts Journal*, and several Nobel laureates have contributed articles.

Candidates are not responsible for familiarity with *Financial Analysts Journal* articles that are cited in the curriculum. But, as your time and studies allow, we strongly encourage you to begin supplementing your understanding of key investment management issues by reading this, and other, CFA Institute practice-oriented publications through the Research & Analysis webpage (www.cfainstitute.org/en/research).

Errata The curriculum development process is rigorous and includes multiple rounds of reviews by content experts. Despite our efforts to produce a curriculum that is free of errors, there are times when we must make corrections. Curriculum errata are periodically updated and posted by exam level and test date online (www.cfainstitute.org/en/programs/submit-errata). If you believe you have found an error in the curriculum, you can submit your concerns through our curriculum errata reporting process found at the bottom of the Curriculum Errata webpage.

DESIGNING YOUR PERSONAL STUDY PROGRAM

Create a Schedule An orderly, systematic approach to exam preparation is critical. You should dedicate a consistent block of time every week to reading and studying. Complete all assigned readings and the associated problems and solutions in each study session. Review the LOS both before and after you study each reading to ensure that

you have mastered the applicable content and can demonstrate the knowledge, skills, and abilities described by the LOS and the assigned reading. Use the LOS self-check to track your progress and highlight areas of weakness for later review.

Successful candidates report an average of more than 300 hours preparing for each exam. Your preparation time will vary based on your prior education and experience, and you will probably spend more time on some study sessions than on others.

You should allow ample time for both in-depth study of all topic areas and additional concentration on those topic areas for which you feel the least prepared.

CFA INSTITUTE LEARNING ECOSYSTEM (LES)

As you prepare for your exam, we will email you important exam updates, testing policies, and study tips. Be sure to read these carefully.

Your exam registration fee includes access to the CFA Program Learning Ecosystem (LES). This digital learning platform provides access, even offline, to all of the readings and End of Reading Questions found in the print curriculum organized as a series of shorter online lessons with associated EORQs. This tool is your one-stop location for all study materials, including practice questions and mock exams.

The LES provides the following supplemental study tools:

Structured and Adaptive Study Plans The LES offers two ways to plan your study through the curriculum. The first is a structured plan that allows you to move through the material in the way that you feel best suits your learning. The second is an adaptive study plan based on the results of an assessment test that uses actual practice questions.

Regardless of your chosen study path, the LES tracks your level of proficiency in each topic area and presents you with a dashboard of where you stand in terms of proficiency so that you can allocate your study time efficiently.

Flashcards and Game Center The LES offers all the Glossary terms as Flashcards and tracks correct and incorrect answers. Flashcards can be filtered both by curriculum topic area and by action taken—for example, answered correctly, unanswered, and so on. These Flashcards provide a flexible way to study Glossary item definitions.

The Game Center provides several engaging ways to interact with the Flashcards in a game context. Each game tests your knowledge of the Glossary terms a in different way. Your results are scored and presented, along with a summary of candidates with high scores on the game, on your Dashboard.

Discussion Board The Discussion Board within the LES provides a way for you to interact with other candidates as you pursue your study plan. Discussions can happen at the level of individual lessons to raise questions about material in those lessons that you or other candidates can clarify or comment on. Discussions can also be posted at the level of topics or in the initial Welcome section to connect with other candidates in your area.

Practice Question Bank The LES offers access to a question bank of hundreds of practice questions that are in addition to the End of Reading Questions. These practice questions, only available on the LES, are intended to help you assess your mastery of individual topic areas as you progress through your studies. After each practice question, you will receive immediate feedback noting the correct response and indicating the relevant assigned reading so you can identify areas of weakness for further study.

Mock Exams The LES also includes access to three-hour Mock Exams that simulate the morning and afternoon sessions of the actual CFA exam. These Mock Exams are intended to be taken after you complete your study of the full curriculum and take practice questions so you can test your understanding of the curriculum and your readiness for the exam. If you take these Mock Exams within the LES, you will receive feedback afterward that notes the correct responses and indicates the relevant assigned readings so you can assess areas of weakness for further study. We recommend that you take Mock Exams during the final stages of your preparation for the actual CFA exam. For more information on the Mock Exams, please visit www.cfainstitute.org.

PREP PROVIDERS

You may choose to seek study support outside CFA Institute in the form of exam prep providers. After your CFA Program enrollment, you may receive numerous solicitations for exam prep courses and review materials. When considering a prep course, make sure the provider is committed to following the CFA Institute guidelines and high standards in its offerings.

Remember, however, that there are no shortcuts to success on the CFA exams; reading and studying the CFA Program curriculum *is* the key to success on the exam. The CFA Program exams reference only the CFA Institute assigned curriculum; no prep course or review course materials are consulted or referenced.

SUMMARY

Every question on the CFA exam is based on the content contained in the required readings and on one or more LOS. Frequently, an exam question is based on a specific example highlighted within a reading or on a specific practice problem and its solution. To make effective use of the CFA Program curriculum, please remember these key points:

1. All pages of the curriculum are required reading for the exam.
2. All questions, problems, and their solutions are part of the curriculum and are required study material for the exam. These questions are found at the end of the readings in the print versions of the curriculum. In the LES, these questions appear directly after the lesson with which they are associated. The LES provides immediate feedback on your answers and tracks your performance on these questions throughout your study.
3. We strongly encourage you to use the CFA Program Learning Ecosystem. In addition to providing access to all the curriculum material, including EORQs, in the form of shorter, focused lessons, the LES offers structured and adaptive study planning, a Discussion Board to communicate with other candidates, Flashcards, a Game Center for study activities, a test bank of practice questions, and online Mock Exams. Other supplemental study tools, such as eBook and PDF versions of the print curriculum, and additional candidate resources are available at www.cfainstitute.org.
4. Using the study planner, create a schedule and commit sufficient study time to cover the study sessions. You should also plan to review the materials, answer practice questions, and take Mock Exams.
5. Some of the concepts in the study sessions may be superseded by updated rulings and/or pronouncements issued after a reading was published. Candidates are expected to be familiar with the overall analytical framework contained in the assigned readings. Candidates are not responsible for changes that occur after the material was written.

FEEDBACK

At CFA Institute, we are committed to delivering a comprehensive and rigorous curriculum for the development of competent, ethically grounded investment professionals. We rely on candidate and investment professional comments and feedback as we work to improve the curriculum, supplemental study tools, and candidate resources.

Please send any comments or feedback to info@cfainstitute.org. You can be assured that we will review your suggestions carefully. Ongoing improvements in the curriculum will help you prepare for success on the upcoming exams and for a lifetime of learning as a serious investment professional.

Financial Statement Analysis

STUDY SESSIONS

Study Session 5	Financial Statement Analysis (1)
Study Session 6	Financial Statement Analysis (2)
Study Session 7	Financial Statement Analysis (3)
Study Session 8	Financial Statement Analysis (4)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to demonstrate a thorough knowledge of financial reporting procedures and the standards that govern financial reporting disclosure. Emphasis is on basic financial statements and how alternative accounting methods affect those statements and the analysis of them.

Financial statement analysis is critical in assessing a company's overall financial position and associated risks over time. Security and business valuation, credit risk assessment, and acquisition due diligence all require an understanding of the major financial statements including general principles and reporting approaches. Because no set of accounting standards has universal acceptance, companies around the world may differ in reporting treatment based on their jurisdiction.

Financial statement analysis requires the ability to analyze a company's reported results with its economic reality, normalize differences in accounting treatment to make valid cross company comparisons, identify quality issues that may exist in reported financial statements, and discern evidence of financial statement manipulation by management.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2021 CFA Institute. All rights reserved.

Candidates should be familiar with the material covered in the following prerequisite reading available in Candidate Resources on the CFA Institute website:

- **Financial Reporting Mechanics**

FINANCIAL STATEMENT ANALYSIS
STUDY SESSION

6

Financial Statement Analysis (2)

This study session addresses the three major financial statements—the income statement, the balance sheet, and the cash flow statement—by examining each in turn. The purpose, elements of, construction, pertinent ratios, and common-size analysis are presented for each major financial statement. The session concludes with a discussion of financial analysis techniques including the use of ratios to evaluate corporate financial health.

READING ASSIGNMENTS

Reading 17	Understanding Income Statements by Elaine Henry, PhD, CFA, and Thomas R. Robinson, PhD, CFA
Reading 18	Understanding Balance Sheets by Elaine Henry, PhD, CFA, and Thomas R. Robinson, PhD, CFA
Reading 19	Understanding Cash Flow Statements by Elaine Henry, PhD, CFA, Thomas R. Robinson, PhD, CFA, J. Hennie van Greuning, DCom, CFA, and Michael A. Broihahn, CPA, CIA, CFA
Reading 20	Financial Analysis Techniques by Elaine Henry, PhD, CFA, Thomas R. Robinson, PhD, CFA, and J. Hennie van Greuning, DCom, CFA

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2021 CFA Institute. All rights reserved.

READING

17

Understanding Income Statements

by Elaine Henry, PhD, CFA, and Thomas R. Robinson, PhD, CFA

Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). Thomas R. Robinson, PhD, CFA, is at AACSB International (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the components of the income statement and alternative presentation formats of that statement;
☐	**b.** describe general principles of revenue recognition and accounting standards for revenue recognition;
☐	**c.** calculate revenue given information that might influence the choice of revenue recognition method;
☐	**d.** describe general principles of expense recognition, specific expense recognition applications, and implications of expense recognition choices for financial analysis;
☐	**e.** describe the financial reporting treatment and analysis of non-recurring items (including discontinued operations, unusual or infrequent items) and changes in accounting policies;
☐	**f.** contrast operating and non-operating components of the income statement;
☐	**g.** describe how earnings per share is calculated and calculate and interpret a company's earnings per share (both basic and diluted earnings per share) for both simple and complex capital structures;
☐	**h.** contrast dilutive and antidilutive securities and describe the implications of each for the earnings per share calculation;
☐	**i.** formulate income statements into common-size income statements;
☐	**j.** evaluate a company's financial performance using common-size income statements and financial ratios based on the income statement;
☐	**k.** describe, calculate, and interpret comprehensive income;
☐	**l.** describe other comprehensive income and identify major types of items included in it.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION

The income statement presents information on the financial results of a company's business activities over a period of time. The income statement communicates how much revenue the company generated during a period and what costs it incurred in connection with generating that revenue. The basic equation underlying the income statement, ignoring gains and losses, is Revenue minus Expenses equals Net income. The income statement is also sometimes referred to as the "statement of operations," "statement of earnings," or "profit and loss (P&L) statement." Under both International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP), the income statement may be presented as a separate statement followed by a statement of comprehensive income that begins with the profit or loss from the income statement or as a section of a single statement of comprehensive income.[1] This reading focuses on the income statement, and the term *income statement* will be used to describe either the separate statement that reports profit or loss used for earnings per share calculations or that section of a statement of comprehensive income that reports the same profit or loss. The reading also includes a discussion of comprehensive income (profit or loss from the income statement plus other comprehensive income).

Investment analysts intensely scrutinize companies' income statements. Equity analysts are interested in them because equity markets often reward relatively high- or low-earnings growth companies with above-average or below-average valuations, respectively, and because inputs into valuation models often include estimates of earnings. Fixed-income analysts examine the components of income statements, past and projected, for information on companies' abilities to make promised payments on their debt over the course of the business cycle. Corporate financial announcements frequently emphasize information reported in income statements, particularly earnings, more than information reported in the other financial statements.

This reading is organized as follows: Section 2 describes the components of the income statement and its format. Section 3 describes basic principles and selected applications related to the recognition of revenue, and Sections 4–7 describe basic principles and selected applications related to the recognition of expenses. Sections 8–10 cover non-recurring items and non-operating items. Sections 11–14 explain the calculation of earnings per share. Sections 15–16 introduce income statement analysis, and Section 17 explains comprehensive income and its reporting. A summary of the key points and practice problems in the CFA Institute multiple choice format complete the reading.

2 COMPONENTS AND FORMAT OF THE INCOME STATEMENT

a describe the components of the income statement and alternative presentation formats of that statement;

---

[1] International Accounting Standard (IAS) 1, *Presentation of Financial Statements*, establishes the presentation and minimum content requirements of financial statements and guidelines for the structure of financial statements under IFRS. Under US GAAP, the Financial Accounting Standards Board Accounting Standards Codification ASC Section 220-10-45 [Comprehensive Income–Overall–Other Presentation Matters] discusses acceptable formats in which to present income, other comprehensive income, and comprehensive income.

Exhibits 1, 2, and 3 show the income statements for Anheuser-Busch InBev SA/NV (AB InBev), a multinational beverage company based in Belgium, Molson Coors Brewing Company (Molson Coors), a US-based multinational brewing company, and Groupe Danone (Danone), a French food manufacturer.[2] AB InBev and Danone report under IFRS, and Molson Coors reports under US GAAP. Note that both AB InBev and Molson Coors show three years' income statements and list the years in chronological order with the most recent year listed in the left-most column. In contrast, Danone shows two years of income statements and lists the years in chronological order from left to right with the most recent year in the right-most column. Different orderings of chronological information are common.

On the top line of the income statement, companies typically report revenue. **Revenue** generally refers to the amount charged for the delivery of goods or services in the *ordinary activities* of a business. Revenue may also be called sales or turnover.[3] For the year ended 31 December 2017, AB InBev reports $56.44 billion of revenue, Molson Coors reports $13.47 billion of revenue (labeled "sales"), and Danone reports €24.68 billion of revenue (labeled "sales").

Revenue is reported after adjustments (e.g., for cash or volume discounts, or for other reductions), and the term **net revenue** is sometimes used to specifically indicate that the revenue has been adjusted (e.g., for estimated returns). For all three companies in Exhibits 1 through 3, footnotes to their financial statements (not shown here) state that revenues are stated net of such items as returns, customer rebates, trade discounts, or volume-based incentive programs for customers.

In a comparative analysis, an analyst may need to reference information disclosed elsewhere in companies' annual reports—typically the notes to the financial statements and the Management Discussion and Analysis (MD&A)—to identify the appropriately comparable revenue amounts. For example, excise taxes represent a significant expenditure for brewing companies. On its income statement, Molson Coors reports $13.47 billion of revenue (labeled "sales") and $11.00 billion of net revenue (labeled "net sales"), which equals sales minus $2.47 billion of excise taxes. Unlike Molson Coors, AB InBev does not show the amount of excise taxes on its income statement. However, in its disclosures, AB InBev notes that excise taxes (amounting to $15.4 billion in 2017) have been deducted from the revenue amount shown on its income statement. Thus, the amount on AB InBev's income statement labeled "revenue" is more comparable to the amount on Molson Coors' income statement labeled "net sales."

Exhibit 1 Anheuser-Busch InBev SA/NV Consolidated Income Statement (in Millions of US Dollars) [Excerpt]

	12 Months Ended December 31		
	2017	**2016**	**2015**
Revenue	$56,444	$45,517	$43,604
Cost of sales	(21,386)	(17,803)	(17,137)
Gross profit	35,058	27,715	26,467
Distribution expenses	(5,876)	(4,543)	(4,259)
Sales and marketing expenses	(8,382)	(7,745)	(6,913)

(continued)

---

[2] Following net income, the income statement also presents **earnings per share**, the amount of earnings per common share of the company. Earnings per share will be discussed in detail later in this reading, and the per-share display has been omitted from these exhibits to focus on the core income statement.

[3] **Sales** is sometimes understood to refer to the sale of goods, whereas *revenue* can include the sale of goods or services; however, the terms are often used interchangeably. In some countries, the term "turnover" may be used in place of revenue.

Exhibit 1 (Continued)

	12 Months Ended December 31		
	2017	2016	2015
Administrative expenses	(3,841)	(2,883)	(2,560)
Other operating income/(expenses)	854	732	1,032
Restructuring	(468)	(323)	(171)
Business and asset disposal	(39)	377	524
Acquisition costs business combinations	(155)	(448)	(55)
Impairment of assets	—	—	(82)
Judicial settlement	—	—	(80)
Profit from operations	17,152	12,882	13,904
Finance cost	(6,885)	(9,216)	(3,142)
Finance income	378	652	1,689
Net finance income/(cost)	(6,507)	(8,564)	(1,453)
Share of result of associates and joint ventures	430	16	10
Profit before tax	11,076	4,334	12,461
Income tax expense	(1,920)	(1,613)	(2,594)
Profit from continuing operations	9,155	2,721	9,867
Profit from discontinued operations	28	48	
Profit of the year	9,183	2,769	9,867
Profit from continuing operations attributable to:			
Equity holders of AB InBev	7,968	1,193	8,273
Non-controlling interest	1,187	1,528	1,594
Profit of the year attributable to:			
Equity holders of AB InBev	7,996	1,241	8,273
Non-controlling interest	$1,187	$1,528	$1,594

Note: reported total amounts may have slight discrepancies due to rounding

Exhibit 2 Molson Coors Brewing Company Consolidated Statement of Operations (in Millions of US Dollars) [Excerpt]

	12 Months Ended		
	Dec. 31, 2017	Dec. 31, 2016	Dec. 31, 2015
Sales	$13,471.5	$6,597.4	$5,127.4
Excise taxes	(2,468.7)	(1,712.4)	(1,559.9)
Net sales	11,002.8	4,885.0	3,567.5
Cost of goods sold	(6,217.2)	(2,987.5)	(2,131.6)
Gross profit	4,785.6	1,897.5	1,435.9
Marketing, general and administrative expenses	(3,032.4)	(1,589.8)	(1,038.3)
Special items, net	(28.1)	2,522.4	(346.7)
Equity Income in MillerCoors	0	500.9	516.3
Operating income (loss)	1,725.1	3,331.0	567.2

Exhibit 2 (Continued)

	12 Months Ended		
	Dec. 31, 2017	Dec. 31, 2016	Dec. 31, 2015
Other income (expense), net			
Interest expense	(349.3)	(271.6)	(120.3)
Interest income	6.0	27.2	8.3
Other income (expense), net	(0.1)	(29.7)	0.9
Total other income (expense), net	(343.4)	(274.1)	(111.1)
Income (loss) from continuing operations before income taxes	1,381.7	3,056.9	456.1
Income tax benefit (expense)	53.2	(1,055.2)	(61.5)
Net income (loss) from continuing operations	1,434.9	2,001.7	394.6
Income (loss) from discontinued operations, net of tax	1.5	(2.8)	3.9
Net income (loss) including noncontrolling interests	1,436.4	1,998.9	398.5
Net (income) loss attributable to noncontrolling interests	(22.2)	(5.9)	(3.3)
Net income (loss) attributable to Molson Coors Brewing Company	$1,414.2	$1,993.0	$395.2

Exhibit 3 Groupe Danone Consolidated Income Statement (in Millions of Euros) [Excerpt]

	Year Ended 31 December	
	2016	2017
Sales	21,944	24,677
Cost of goods sold	(10,744)	(12,459)
Selling expense	(5,562)	(5,890)
General and administrative expense	(2,004)	(2,225)
Research and development expense	(333)	(342)
Other income (expense)	(278)	(219)
Recurring operating income	**3,022**	**3,543**
Other operating income (expense)	(99)	192
Operating income	**2,923**	**3,734**
Interest income on cash equivalents and short-term investments	130	151
Interest expense	(276)	(414)
Cost of net debt	(146)	(263)
Other financial income	67	137
Other financial expense	(214)	(312)
Income before tax	**2,630**	**3,296**
Income tax expense	(804)	(842)
Net income from fully consolidated companies	**1,826**	**2,454**
Share of profit of associates	1	109
Net income	1,827	2,563

(continued)

Exhibit 3 (Continued)

	Year Ended 31 December	
	2016	**2017**
Net income – Group share	**1,720**	**2,453**
Net income – Non-controlling interests	107	110

Differences in presentations of items, such as expenses, are also common. **Expenses** reflect outflows, depletions of assets, and incurrences of liabilities in the course of the activities of a business. Expenses may be grouped and reported in different formats, subject to some specific requirements.

At the bottom of the income statement, companies report net income (companies may use other terms such as "net earnings" or "profit or loss"). For 2017, AB InBev reports $9,183 million "Profit of the year", Molson Coors reports $1,436.4 million of net income including noncontrolling interests, and Danone reports €2,563 million of net income. Net income is often referred to as the "bottom line." The basis for this expression is that net income is the final—or bottom—line item in an income statement. Because net income is often viewed as the single most relevant number to describe a company's performance over a period of time, the term "bottom line" sometimes is used in business to refer to any final or most relevant result.

Despite this customary terminology, note that each company presents additional items below net income: information about how much of that net income is attributable to the company itself and how much of that income is attributable to noncontrolling interests, also known as minority interests. The companies consolidate subsidiaries over which they have control. Consolidation means that they include all of the revenues and expenses of the subsidiaries even if they own less than 100 percent. Noncontrolling interest represents the portion of income that "belongs" to the minority shareholders of the consolidated subsidiaries, as opposed to the parent company itself. For AB InBev, $7,996 million of the total profit is attributable to the shareholders of AB InBev, and $1,187 million is attributable to noncontrolling interests. For Molson Coors, $1,414.2 million is attributable to the shareholders of Molson Coors, and $22.2 million is attributable to noncontrolling interests. For Danone, €2,453 million of the net income amount is attributable to shareholders of Groupe Danone and €110 million is attributable to noncontrolling interests.

Net income also includes **gains** and **losses**, which are increases and decreases in economic benefits, respectively, which may or may not arise in the ordinary activities of the business. For example, when a manufacturing company sells its products, these transactions are reported as revenue, and the costs incurred to generate these revenues are expenses and are presented separately. However, if a manufacturing company sells surplus land that is not needed, the transaction is reported as a gain or a loss. The amount of the gain or loss is the difference between the carrying value of the land and the price at which the land is sold. For example, in Exhibit 1, AB InBev reports a loss (proceeds, net of carrying value) of $39 million on disposals of businesses and assets in fiscal 2017, and gains of $377 million and $524 million in 2016 and 2015, respectively. Details on these gains and losses can typically be found in the companies' disclosures. For example, AB InBev discloses that the $377 million gain in 2016 was mainly from selling one of its breweries in Mexico.

The definition of income encompasses both revenue and gains and the definition of expenses encompasses both expenses that arise in the ordinary activities of the business and losses.[4] Thus, **net income** (profit or loss) can be defined as: a) income minus expenses, or equivalently b) revenue plus other income plus gains minus expenses, or equivalently c) revenue plus other income plus gains minus expenses in the ordinary activities of the business minus other expenses, and minus losses. The last definition can be rearranged as follows: net income equals (i) revenue minus expenses in the ordinary activities of the business, plus (ii) other income minus other expenses, plus (iii) gains minus losses.

In addition to presenting the net income, income statements also present items, including subtotals, which are significant to users of financial statements. Some of the items are specified by IFRS but other items are not specified.[5] Certain items, such as revenue, finance costs, and tax expense, are required to be presented separately on the face of the income statement. IFRS additionally require that line items, headings, and subtotals relevant to understanding the entity's financial performance should be presented even if not specified. Expenses may be grouped together either by their nature or function. Grouping together expenses such as depreciation on manufacturing equipment and depreciation on administrative facilities into a single line item called "depreciation" is an example of a **grouping by nature** of the expense. An example of **grouping by function** would be grouping together expenses into a category such as cost of goods sold, which may include labour and material costs, depreciation, some salaries (e.g., salespeople's), and other direct sales related expenses.[6] All three companies in Exhibits 1 through 3 present their expenses by function, which is sometimes referred to "cost of sales" method.

One subtotal often shown in an income statement is **gross profit** or **gross margin** (that is revenue less cost of sales). When an income statement shows a gross profit subtotal, it is said to use a **multi-step format** rather than a **single-step format**. The AB InBev and Molson Coors income statements are examples of the multi-step format, whereas the Groupe Danone income statement is in a single-step format. For manufacturing and merchandising companies, gross profit is a relevant item and is calculated as revenue minus the cost of the goods that were sold. For service companies, gross profit is calculated as revenue minus the cost of services that were provided. In summary, gross profit is the amount of revenue available after subtracting the costs of delivering goods or services. Other expenses related to running the business are subtracted after gross profit.

Another important subtotal which may be shown on the income statement is **operating profit** (or, synonymously, operating income). Operating profit results from deducting operating expenses such as selling, general, administrative, and research and development expenses from gross profit. Operating profit reflects a company's profits on its business activities before deducting taxes, and for non-financial companies, before deducting interest expense. For financial companies, interest expense would be included in operating expenses and subtracted in arriving at operating profit because it relates to the operating activities for such companies. For some companies composed of a number of separate business segments, operating profit can be useful in evaluating the performance of the individual business segments, because interest and tax expenses may be more relevant at the level of the overall company rather than an individual segment level. The specific calculations of gross profit and operating profit may vary by company, and a reader of financial statements can consult the notes to the statements to identify significant variations across companies.

---

[4] IASB *Conceptual Framework for Financial Reporting (2010)*, paragraphs 4.29 to 4.32.

[5] Requirements are presented in IAS 1, *Presentation of Financial Statements.*

[6] Later readings will provide additional information about alternative methods to calculate cost of goods sold.

Operating profit is sometimes referred to as EBIT (earnings before interest and taxes). However, operating profit and EBIT are not necessarily the same. Note that in each of the Exhibits 1 through 3, interest and taxes do not represent the only differences between earnings (net income, net earnings) and operating income. For example, AB InBev separately reports its share of associates' and joint ventures' income and Molson Coors separately reports some income from discontinued operations.

Exhibit 4 shows an excerpt from the income statement of CRA International, a company providing management consulting services. Accordingly, CRA deducts cost of services (rather than cost of goods) from revenues to derive gross profit. CRA's fiscal year ends on the Saturday nearest December 31st. Because of this fiscal year timeframe, CRA's fiscal year occasionally comprises 53 weeks rather than 52 weeks. Although the extra week is likely immaterial in computing year-to-year growth rates, it may have a material impact on a quarter containing the extra week. In general, an analyst should be alert to the effect of an extra week when making historical comparisons and forecasting future performance.

Exhibit 4 CRA International Inc. Consolidated Statements of Operations (Excerpt) (in Thousands of Dollars)

	Fiscal Year Ended		
	Dec. 30, 2017	**Dec. 31, 2016**	**Jan. 02, 2016**
Revenues	$370,075	$324,779	$303,559
Costs of services (exclusive of depreciation and amortization)	258,829	227,380	207,650
Selling, general and administrative expenses	86,537	70,584	72,439
Depreciation and amortization	8,945	7,896	6,552
GNU goodwill impairment	—	—	4,524
Income from operations	15,764	18,919	12,394

Note: Remaining items omitted

Exhibits 1 through 4 illustrate basic points about the income statement, including variations across the statements—some of which depend on the industry and/or country, and some of which reflect differences in accounting policies and practices of a particular company. In addition, some differences within an industry are primarily differences in terminology, whereas others are more fundamental accounting differences. Notes to the financial statements are helpful in identifying such differences.

Having introduced the components and format of an income statement, the next objective is to understand the actual reported numbers in it. To accurately interpret reported numbers, the analyst needs to be familiar with the principles of revenue and expense recognition—that is, how revenue and expenses are measured and attributed to a given accounting reporting period.

3

REVENUE RECOGNITION

b Describe general principles of revenue recognition and accounting standards for revenue recognition;

c calculate revenue given information that might influence the choice of revenue recognition method;

Revenue is the top line in an income statement, so we begin the discussion of line items in the income statement with revenue recognition. Accounting standards for revenue recognition (which we discuss later in this section) became effective at the beginning of 2018 and are nearly identical under IFRS and US GAAP. The revenue recognition standards for IFRS and US GAAP (IFRS 15 and ASC Topic 606, respectively) were issued in 2014 and resulted from an effort to achieve convergence, consistency, and transparency in revenue recognition globally.

A first task is to explain some relevant accounting terminology. The terms revenue, sales, gains, losses, and net income (profit, net earnings) have been briefly defined. The IASB *Conceptual Framework for Financial Reporting* (2010),[7] referred to hereafter as the *Conceptual Framework*, further defines and discusses these income statement items. The *Conceptual Framework* explains that profit is a frequently used measure of performance and is composed of income and expenses.[8] It defines **income** as follows:

> Income is increases in economic benefits during the accounting period in the form of inflows or enhancements of assets or decreases of liabilities that result in increases in equity, other than those relating to contributions from equity participants.[9]

In IFRS, the term "income" includes revenue and gains. Gains are similar to revenue, but they typically arise from secondary or peripheral activities rather than from a company's primary business activities. For example, for a restaurant, the sale of surplus restaurant equipment for more than its carrying value is referred to as a gain rather than as revenue. Similarly, a loss typically arises from secondary activities. Gains and losses may be considered part of operating activities (e.g., a loss due to a decline in the value of inventory) or may be considered part of non-operating activities (e.g., the sale of non-trading investments).

In the following simple hypothetical scenario, revenue recognition is straightforward: a company sells goods to a buyer for cash and does not allow returns, so the company recognizes revenue when the exchange of goods for cash takes place and measures revenue at the amount of cash received. In practice, however, determining when revenue should be recognized and at what amount is considerably more complex for reasons discussed in the following sections.

3.1 General Principles

An important aspect concerning revenue recognition is that it can occur independently of cash movements. For example, assume a company sells goods to a buyer on credit, so does not actually receive cash until some later time. A fundamental principle of accrual accounting is that revenue is recognized (reported on the income statement) when it is earned, so the company's financial records reflect revenue from the sale when the risk and reward of ownership is transferred; this is often when the company delivers the goods or services. If the delivery was on credit, a related asset, such as trade or accounts receivable, is created. Later, when cash changes hands, the company's financial records simply reflect that cash has been received to settle an account receivable. Similarly, there are situations when a company receives cash in advance and actually delivers the product or service later, perhaps over a period of time. In this case, the company would record a liability for **unearned revenue** when the cash is initially received, and revenue would be recognized as being earned over time as products and services are delivered. An example would be a subscription payment received for a publication that is to be delivered periodically over time.

---

[7] The IASB is currently in the process of updating its *Conceptual Framework for Financial Reporting*.

[8] *Conceptual Framework*, paragraph 4.24. The text on the elements of financial statements and their recognition and measurement is the same in the IASB *Conceptual Framework for Financial Reporting* (2010) and the IASB *Framework for the Preparation and Presentation of Financial Statements* (1989).

[9] Ibid., paragraph 4.25(a).

3.2 Accounting Standards for Revenue Recognition

The converged accounting standards issued by the IASB and FASB in May 2014 introduced some changes to the basic principles of revenue recognition and should enhance comparability.[10] The content of the two standards is nearly identical, and this discussion pertains to both, unless specified otherwise. Issuance of this converged standard is significant because of the differences between IFRS and US GAAP on revenue recognition prior to the converged standard. The converged standard aims to provide a principles-based approach to revenue recognition that can be applied to many types of revenue-generating activities.

The core principle of the converged standard is that revenue should be recognized to "depict the transfer of promised goods or services to customers in an amount that reflects the consideration to which the entity expects to be entitled in an exchange for those goods or services." To achieve the core principle, the standard describes the application of five steps in recognizing revenue:

1. Identify the contract(s) with a customer
2. Identify the separate or distinct performance obligations in the contract
3. Determine the transaction price
4. Allocate the transaction price to the performance obligations in the contract
5. Recognize revenue when (or as) the entity satisfies a performance obligation

According to the standard, a contract is an agreement and commitment, with commercial substance, between the contacting parties. It establishes each party's *obligations* and *rights,* including payment terms. In addition, a contract exists only if collectability is probable. Each standard uses the same wording, but the threshold for probable collectability differs. Under IFRS, probable means more likely than not, and under US GAAP it means likely to occur. As a result, economically similar contracts may be treated differently under IFRS and US GAAP.

The performance obligations within a contract represent promises to transfer distinct good(s) or service(s). A good or service is distinct if the customer can benefit from it on its own or in combination with readily available resources and if the promise to transfer it can be separated from other promises in the contract. Each identified performance obligation is accounted for separately.

The transaction price is what the seller estimates will be received in exchange for transferring the good(s) or service(s) identified in the contract. The transaction price is then allocated to each identified performance obligation. Revenue is recognized when a performance obligation is fulfilled. Steps three and four address amount, and step five addresses timing of recognition. The amount recognized reflects expectations about collectability and (if applicable) an allocation to multiple obligations within the same contract. Revenue is recognized when the obligation-satisfying transfer is made.

Revenue should only be recognized when it is highly probable that it will not be subsequently reversed. This may result in the recording of a minimal amount of revenue upon sale when an estimate of total revenue is not reliable. The balance sheet will be required to reflect the entire refund obligation as a liability and will include an asset for the "right to returned goods" based on the carrying amount of inventory less costs of recovery.

---

[10] *IFRS 15 Revenue from Contracts with Customers* and FASB ASC Topic 606 (Revenue from Contracts with Customers).

When revenue is recognized, a contract asset is presented on the balance sheet. It is only at the point when all performance obligations have been met except for payment that a receivable appears on the seller's balance sheet. If consideration is received in advance of transferring good(s) or service(s), the seller presents a contract liability.

The entity will recognize revenue when it is able to satisfy the performance obligation by transferring control to the customer. Factors to consider when assessing whether the customer has obtained control of an asset at a point in time:

- Entity has a present right to payment,
- Customer has legal title,
- Customer has physical possession,
- Customer has the significant risks and rewards of ownership, and
- Customer has accepted the asset.

For a simple contract with only one deliverable at a single point in time, completing the five steps is straight-forward. For more complex contracts—such as when the performance obligations are satisfied over time, when the terms of the multi-period contracts change, when the performance obligation includes various components of goods and services, or when the compensation is "variable"—accounting choices can be less obvious. The steps in the standards are intended to provide guidance that can be generalized to most situations.

In addition, the standard provides many specific examples. These examples are intended to provide guidance as to how to approach more complex contracts. Some of these examples are summarized in Exhibit 5. Note that the end result for many examples may not differ substantially from that under revenue recognition standards that were in effect prior to the adoption of the converged standard; instead it is the conceptual approach and, in some cases, the terminology that will differ.

Exhibit 5 Applying the Converged Revenue Recognition Standard

The references in this exhibit are to Examples in IFRS 15 *Revenue from Contracts with Customers* (and ASU 2014-09 (FASB ASC Topic 606)), on which these summaries are based.

Part 1 (*ref. Example 10*)

Builder Co. enters into a contract with Customer Co. to construct a commercial building. Builder Co. identifies various goods and services to be provided, such as pre-construction engineering, construction of the building's individual components, plumbing, electrical wiring, and interior finishes. With respect to "Identifying the Performance Obligation," should Builder Co. treat each specific item as a separate performance obligation to which revenue should be allocated?

The standard provides two criteria, which must be met, to determine if a good or service is distinct for purposes of identifying performance obligations. First, the customer can benefit from the good or service either on its own or together with other readily available resources. Second, the seller's "promise to transfer the good or service to the customer is separately identifiable from other promises in the contract." In this example, the second criterion is not met because it is the building for which the customer has contracted, not the separate goods and services. The seller will integrate all the goods and services into a combined output and each specific item should not be treated as a distinct good or service but accounted for together as a single performance obligation.

(continued)

Exhibit 5 (Continued)

Part 2 (*ref. Example 8*)

Builder Co.'s contract with Customer Co. to construct the commercial building specifies consideration of $1 million. Builder Co.'s expected total costs are $700,000. The Builder incurs $420,000 in costs in the first year. Assuming that costs incurred provide an appropriate measure of progress toward completing the contract, how much revenue should Builder Co. recognize for the first year?

The standard states that for performance obligations satisfied over time (e.g., where there is a long-term contract), revenue is recognized over time by measuring progress toward satisfying the obligation. In this case, the Builder has incurred 60% of the total expected costs ($420,000/$700,000) and will thus recognize $600,000 (60% × $1 million) in revenue for the first year.

This is the same amount of revenue that would be recognized using the "percentage-of-completion" method under previous accounting standards, but that term is not used in the converged standard. Instead, the standard refers to performance obligations satisfied over time and requires that progress toward complete satisfaction of the performance obligation be measured based on input method such as the one illustrated here (recognizing revenue based on the proportion of total costs that have been incurred in the period) or an output method (recognizing revenue based on units produced or milestones achieved).

Part 3 (*ref. Example 8*)

Assume that Builder Co.'s contract with Customer Co. to construct the commercial building specifies consideration of $1 million *plus* a bonus of $200,000 if the building is completed within 2 years. Builder Co. has only limited experience with similar types of contracts and knows that many factors outside its control (e.g., weather, regulatory requirements) could cause delay. Builder Co.'s expected total costs are $700,000. The Builder incurs $420,000 in costs in the first year. Assuming that costs incurred provide an appropriate measure of progress toward completing the contract, how much revenue should Builder Co. recognize for the first year?

The standard addresses so-called "variable consideration" as part of determining the transaction price. A company is only allowed to recognize variable consideration if it can conclude that it will not have to reverse the cumulative revenue in the future. In this case, Builder Co. does not recognize any of the bonus in year one because it cannot reach the non-reversible conclusion given its limited experience with similar contracts and potential delays from factors outside its control.

Part 4 (*ref. Example 8*)

Assume all facts from Part 3. In the beginning of year two, Builder Co. and Customer Co. agree to change the building floor plan and modify the contract. As a result the consideration will increase by $150,000, and the allowable time for achieving the bonus is extended by 6 months. Builder expects its costs will increase by $120,000. Also, given the additional 6 months to earn the completion bonus, Builder concludes that it now meets the criteria for including the $200,000 bonus in revenue. How should Builder account for this change in the contract?

Note that previous standards did not provide a general framework for contract modifications. The converged standard provides guidance on whether a change in a contract is a new contract or a modification of an existing contract. To be considered a new contract, the change would need to involve goods and services that are distinct from the goods and services already transferred.

Exhibit 5 (Continued)

In this case, the change does not meet the criteria of a new contract and is therefore considered a modification of the existing contract, which requires the company to reflect the impact on a cumulative catch-up basis. Therefore, the company must update its transaction price and measure of progress. Builder's total revenue on the transaction (transaction price) is now $1.35 million ($1 million original plus the $150,000 new consideration plus $200,000 for the completion bonus). Builder Co.'s progress toward completion is now 51.2% ($420,000 costs incurred divided by total expected costs of $820,000). Based on the changes in the contract, the amount of additional revenue to be recognized is $91,200, calculated as (51.2% × $1.35 million) minus the $600,000 already recognized. The additional $91,200 of revenue would be recognized as a "cumulative catch-up adjustment" on the date of the contract modification.

Part 5 (*ref. Example 15*)

Assume a Company operates a website that enables customers to purchase goods from various suppliers. The customers pay the Company in advance, and orders are nonrefundable. The *suppliers* deliver the goods directly to the customer, and the Company receives a 10% commission. Should the Company report Total Revenues equal to 100% of the sales amount (gross) or Total Revenues equal to 10% of the sales amount (net)? Revenues are reported gross if the Company is acting as a Principal and net if the Company is acting as an Agent.

In this example, the Company is an Agent because it isn't primarily responsible for fulfilling the contract, doesn't take any inventory risk or credit risk, doesn't have discretion in setting the price, and receives compensation in the form of a commission. Because the Company is acting as an Agent, it should report only the amount of commission as its revenue.

Some related costs require specific accounting treatment under the new standards. In particular, incremental costs of obtaining a contract and certain costs incurred to fulfill a contract must be capitalized under the new standards (i.e., reported as an asset on the balance sheet rather than as an expense on the income statement). If a company had previously expensed these incremental costs in the years prior to adopting the converged standard, all else equal, its profitability will initially appear higher under the converged standards.

The disclosure requirements are quite extensive. Companies are required at year end[11] to disclose information about contracts with customers disaggregated into different categories of contracts. The categories might be based on the type of product, the geographic region, the type of customer or sales channel, the type of contract pricing terms, the contract duration, or the timing of transfers. Companies are also required to disclose balances of any contract-related assets and liabilities and significant changes in those balances, remaining performance obligations and transaction price allocated to those obligations, and any significant judgments and changes in judgments related to revenue recognition. Significant judgments are those used in determining timing and amounts of revenue to be recognized.

The converged standard is expected to affect some industries more than others. For example, industries where bundled sales are common, such as the telecommunications and software industries, are expected to be significantly affected by the converged standard.

---

[11] Interim period disclosures are required under IFRS and US GAAP but differ between them.

4 EXPENSE RECOGNITION: GENERAL PRINCIPLES

d describe general principles of expense recognition, specific expense recognition applications, and implications of expense recognition choices for financial analysis;

Expenses are deducted against revenue to arrive at a company's net profit or loss. Under the IASB *Conceptual Framework*, **expenses** are "decreases in economic benefits during the accounting period in the form of outflows or depletions of assets or incurrences of liabilities that result in decreases in equity, other than those relating to distributions to equity participants."[12]

The IASB *Conceptual Framework* also states:

> The definition of expenses encompasses losses as well as those expenses that arise in the course of the ordinary activities of the enterprise. Expenses that arise in the course of the ordinary activities of the enterprise include, for example, cost of sales, wages and depreciation. They usually take the form of an outflow or depletion of assets such as cash and cash equivalents, inventory, property, plant and equipment.
>
> Losses represent other items that meet the definition of expenses and may, or may not, arise in the course of the ordinary activities of the enterprise. Losses represent decreases in economic benefits and as such they are no different in nature from other expenses. Hence, they are not regarded as a separate element in this *Conceptual Framework*.
>
> Losses include, for example, those resulting from disasters such as fire and flood, as well as those arising on the disposal of non-current assets.[13]

Similar to the issues with revenue recognition, in a simple hypothetical scenario, expense recognition would not be an issue. For instance, assume a company purchased inventory for cash and sold the entire inventory in the same period. When the company paid for the inventory, absent indications to the contrary, it is clear that the inventory cost has been incurred and when that inventory is sold, it should be recognized as an expense (cost of goods sold) in the financial records. Assume also that the company paid all operating and administrative expenses in cash within each accounting period. In such a simple hypothetical scenario, no issues of expense recognition would arise. In practice, however, as with revenue recognition, determining when expenses should be recognized can be somewhat more complex.

4.1 General Principles

In general, a company recognizes expenses in the period that it consumes (i.e., uses up) the economic benefits associated with the expenditure, or loses some previously recognized economic benefit.[14]

A general principle of expense recognition is the **matching principle**. Strictly speaking, IFRS do not refer to a "matching principle" but rather to a "matching concept" or to a process resulting in "matching of costs with revenues."[15] The distinction is relevant in certain standard setting deliberations. Under matching, a company recognizes some expenses (e.g., cost of goods sold) when associated revenues are recognized and thus, expenses and revenues are matched. Associated revenues and expenses are those that result directly and jointly from the same transactions or events. Unlike the simple scenario in which a company purchases inventory and sells all of the inventory within the same accounting period, in practice, it is more likely that some of the current period's sales are made from inventory purchased in a previous period or previous periods. It is also likely that some of the inventory purchased in the current period will remain unsold at the end of the current period and so will be sold in a following period. Matching requires that a company recognizes cost of goods sold in the same period as revenues from the sale of the goods.

---

[12] IASB *Conceptual Framework*, paragraph 4.25(b).
[13] Ibid., paragraphs 4.33–4.35.
[14] Ibid., paragraph 4.49.
[15] Ibid., paragraph 4.50.

Period costs, expenditures that less directly match revenues, are reflected in the period when a company makes the expenditure or incurs the liability to pay. Administrative expenses are an example of period costs. Other expenditures that also less directly match revenues relate more directly to future expected benefits; in this case, the expenditures are allocated systematically with the passage of time. An example is depreciation expense.

Examples 1 and 2 demonstrate matching applied to inventory and cost of goods sold.

EXAMPLE 1

The Matching of Inventory Costs with Revenues

Kahn Distribution Limited (KDL), a hypothetical company, purchases inventory items for resale. At the beginning of 2018, Kahn had no inventory on hand. During 2018, KDL had the following transactions:

Inventory Purchases		
First quarter	2,000	units at $40 per unit
Second quarter	1,500	units at $41 per unit
Third quarter	2,200	units at $43 per unit
Fourth quarter	1,900	units at $45 per unit
Total	7,600	units at a total cost of $321,600

KDL sold 5,600 units of inventory during the year at $50 per unit, and received cash. KDL determines that there were 2,000 remaining units of inventory and specifically identifies that 1,900 were those purchased in the fourth quarter and 100 were purchased in the third quarter. What are the revenue and expense associated with these transactions during 2018 based on specific identification of inventory items as sold or remaining in inventory? (Assume that the company does not expect any products to be returned.)

Solution:

The revenue for 2018 would be $280,000 (5,600 units × $50 per unit). Initially, the total cost of the goods purchased would be recorded as inventory (an asset) in the amount of $321,600. During 2018, the cost of the 5,600 units sold would be expensed (matched against the revenue) while the cost of the 2,000 remaining unsold units would remain in inventory as follows:

Cost of Goods Sold		
From the first quarter	2,000 units at $40 per unit =	$80,000
From the second quarter	1,500 units at $41 per unit =	$61,500
From the third quarter	2,100 units at $43 per unit =	$90,300
Total cost of goods sold		$231,800
Cost of Goods Remaining in Inventory		
From the third quarter	100 units at $43 per unit =	$4,300

(continued)

From the fourth quarter	1,900 units at $45 per unit =	$85,500
Total remaining (or ending) inventory cost		$89,800

To confirm that total costs are accounted for: $231,800 + $89,800 = $321,600. The cost of the goods sold would be expensed against the revenue of $280,000 as follows:

Revenue	$280,000
Cost of goods sold	231,800
Gross profit	$48,200

An alternative way to think about this is that the company created an asset (inventory) of $321,600 as it made its purchases. At the end of the period, the value of the company's inventory on hand is $89,800. Therefore, the amount of the Cost of goods sold expense recognized for the period should be the difference: $231,800.

The remaining inventory amount of $89,800 will be matched against revenue in a future year when the inventory items are sold.

EXAMPLE 2

Alternative Inventory Costing Methods

In Example 1, KDL was able to specifically identify which inventory items were sold and which remained in inventory to be carried over to later periods. This is called the **specific identification method** and inventory and cost of goods sold are based on their physical flow. It is generally not feasible to specifically identify which items were sold and which remain on hand, so accounting standards permit the assignment of inventory costs to costs of goods sold and to ending inventory using cost formulas (IFRS terminology) or cost flow assumptions (US GAAP). The cost formula or cost flow assumption determines which goods are assumed to be sold and which goods are assumed to remain in inventory. Both IFRS and US GAAP permit the use of the first in, first out (FIFO) method, and the weighted average cost method to assign costs.

Under the **FIFO method**, the oldest goods purchased (or manufactured) are assumed to be sold first and the newest goods purchased (or manufactured) are assumed to remain in inventory. Cost of goods in beginning inventory and costs of the first items purchased (or manufactured) flow into cost of goods sold first, as if the earliest items purchased sold first. Ending inventory would, therefore, include the most recent purchases. It turns out that those items specifically identified as sold in Example 1 were also the first items purchased, so in this example, under FIFO, the cost of goods sold would also be $231,800, calculated as above.

The **weighted average cost method** assigns the average cost of goods available for sale to the units sold and remaining in inventory. The assignment is based on the average cost per unit (total cost of goods available for sale/total units available for sale) and the number of units sold and the number remaining in inventory.

For KDL, the weighted average cost per unit would be

$321,600/7,600 units = $42.3158 per unit

Cost of goods sold using the weighted average cost method would be

5,600 units at $42.3158 = $236,968

Ending inventory using the weighted average cost method would be

2,000 units at $42.3158 = $84,632

Another method is permitted under US GAAP but is not permitted under IFRS. This is the last in, first out (LIFO) method. Under the **LIFO method**, the newest goods purchased (or manufactured) are assumed to be sold first and the oldest goods purchased (or manufactured) are assumed to remain in inventory. Costs of the latest items purchased flow into cost of goods sold first, as if the most recent items purchased were sold first. Although this may seem contrary to common sense, it is logical in certain circumstances. For example, lumber in a lumberyard may be stacked up with the oldest lumber on the bottom. As lumber is sold, it is sold from the top of the stack, so the last lumber purchased and put in inventory is the first lumber out. Theoretically, a company should choose a method linked to the physical inventory flows.[16] Under the LIFO method, in the KDL example, it would be assumed that the 2,000 units remaining in ending inventory would have come from the first quarter's purchases:[17]

Ending inventory 2,000 units at $40 per unit = $80,000

The remaining costs would be allocated to cost of goods sold under LIFO:

Total costs of $321,600 less $80,000 remaining in ending inventory = $241,600

Alternatively, the cost of the last 5,600 units purchased is allocated to cost of goods sold under LIFO:

1,900 units at $45 per unit + 2,200 units at $43 per unit + 1,500 units at $41 per unit = $241,600

An alternative way to think about expense recognition is that the company created an asset (inventory) of $321,600 as it made its purchases. At the end of the period, the value of the company's inventory is $80,000. Therefore, the amount of the Cost of goods sold expense recognized for the period should be the difference: $241,600.

Exhibit 6 summarizes and compares inventory costing methods.

---

[16] Practically, the reason some companies choose to use LIFO in the United States is to reduce taxes. When prices and inventory quantities are rising, LIFO will normally result in higher cost of goods sold and lower income and hence lower taxes. US tax regulations require that if LIFO is used on a company's tax return, it must also be used on the company's GAAP financial statements.

[17] If data on the precise timing of quarterly sales were available, the answer would differ because the cost of goods sold would be determined during the quarter rather than at the end of the quarter.

Exhibit 6 Summary Table on Inventory Costing Methods

Method	Description	Cost of Goods Sold When Prices Are Rising, Relative to Other Two Methods	Ending Inventory When Prices Are Rising, Relative to Other Two Methods
FIFO (first in, first out)	Costs of the earliest items purchased flow to cost of goods sold first	Lowest	Highest
LIFO (last in, first out)	Costs of the most recent items purchased flow to cost of goods sold first	Highest*	Lowest*
Weighted average cost	Averages total costs over total units available	Middle	Middle

*Assumes no LIFO layer liquidation. **LIFO layer liquidation** occurs when the volume of sales exceeds the volume of purchases in the period so that some sales are assumed to be made from existing, relatively low-priced inventory rather than from more recent purchases.

5 ISSUES IN EXPENSE RECOGNITION: DOUBTFUL ACCOUNTS, WARRANTIES

d describe general principles of expense recognition, specific expense recognition applications, and implications of expense recognition choices for financial analysis;

The following sections cover applications of the principles of expense recognition to certain common situations.

5.1 Doubtful Accounts

When a company sells its products or services on credit, it is likely that some customers will ultimately default on their obligations (i.e., fail to pay). At the time of the sale, it is not known which customer will default. (If it were known that a particular customer would ultimately default, presumably a company would not sell on credit to that customer.) One possible approach to recognizing credit losses on customer receivables would be for the company to wait until such time as a customer defaulted and only then recognize the loss (**direct write-off method**). Such an approach would usually not be consistent with generally accepted accounting principles.

Under the matching principle, at the time revenue is recognized on a sale, a company is required to record an estimate of how much of the revenue will ultimately be uncollectible. Companies make such estimates based on previous experience with uncollectible accounts. Such estimates may be expressed as a proportion of the overall amount of sales, the overall amount of receivables, or the amount of receivables overdue by a specific amount of time. The company records its estimate of uncollectible amounts as an expense on the income statement, not as a direct reduction of revenues.

5.2 Warranties

At times, companies offer warranties on the products they sell. If the product proves deficient in some respect that is covered under the terms of the warranty, the company will incur an expense to repair or replace the product. At the time of sale, the company does not know the amount of future expenses it will incur in connection with its warranties. One possible approach would be for a company to wait until actual expenses are incurred under the warranty and to reflect the expense at that time. However, this would not result in a matching of the expense with the associated revenue.

Under the matching principle, a company is required to estimate the amount of future expenses resulting from its warranties, to recognize an estimated warranty expense in the period of the sale, and to update the expense as indicated by experience over the life of the warranty.

ISSUES IN EXPENSE RECOGNITION: DEPRECIATION AND AMORTIZATION

6

d describe general principles of expense recognition, specific expense recognition applications, and implications of expense recognition choices for financial analysis;

Companies commonly incur costs to obtain long-lived assets. **Long-lived assets** are assets expected to provide economic benefits over a future period of time greater than one year. Examples are land (property), plant, equipment, and **intangible assets** (assets lacking physical substance) such as trademarks. The costs of most long-lived assets are allocated over the period of time during which they provide economic benefits. The two main types of long-lived assets whose costs are *not* allocated over time are land and those intangible assets with indefinite useful lives.

Depreciation is the process of systematically allocating costs of long-lived assets over the period during which the assets are expected to provide economic benefits. “Depreciation” is the term commonly applied to this process for physical long-lived assets such as plant and equipment (land is not depreciated), and **amortisation** is the term commonly applied to this process for intangible long-lived assets with a finite useful life.[18] Examples of intangible long-lived assets with a finite useful life include an acquired mailing list, an acquired patent with a set expiration date, and an acquired copyright with a set legal life. The term “amortisation” is also commonly applied to the systematic allocation of a premium or discount relative to the face value of a fixed-income security over the life of the security.

IFRS allow two alternative models for valuing property, plant, and equipment: the cost model and the revaluation model.[19] Under the cost model, the depreciable amount of that asset (cost less residual value) is allocated on a systematic basis over the remaining useful life of the asset. Under the cost model, the asset is reported at its cost less any accumulated depreciation. Under the revaluation model, the asset is reported at its fair value. The revaluation model is not permitted under US GAAP. Although the revaluation model is permitted under IFRS, as noted earlier, it is not as widely used and thus we focus on the cost model here. There are two other differences between IFRS and US GAAP to note: IFRS require each component of an asset to be depreciated separately and US GAAP do not require component depreciation; and IFRS require an annual review of residual value and useful life, and US GAAP do not explicitly require such a review.

---

[18] Intangible assets with indefinite life are not amortised. Instead, they are reviewed each period as to the reasonableness of continuing to assume an indefinite useful life and are tested at least annually for impairment (i.e., if the recoverable or fair value of an intangible asset is materially lower than its value in the company’s books, the value of the asset is considered to be impaired and its value must be decreased). IAS 38, *Intangible Assets* and FASB ASC Topic 350 [Intangibles–Goodwill and Other].

[19] IAS No. 16, *Property, Plant, and Equipment.*

The method used to compute depreciation should reflect the pattern over which the economic benefits of the asset are expected to be consumed. IFRS do not prescribe a particular method for computing depreciation but note that several methods are commonly used, such as the straight-line method, diminishing balance method (accelerated depreciation), and the units of production method (depreciation varies depending upon production or usage).

The **straight-line method** allocates evenly the cost of long-lived assets less estimated residual value over the estimated useful life of an asset. (The term "straight line" derives from the fact that the annual depreciation expense, if represented as a line graph over time, would be a straight line. In addition, a plot of the cost of the asset minus the cumulative amount of annual depreciation expense, if represented as a line graph over time, would be a straight line with a negative downward slope.) Calculating depreciation and amortisation requires two significant estimates: the estimated useful life of an asset and the estimated residual value (also known as "salvage value") of an asset. Under IFRS, the residual value is the amount that the company expects to receive upon sale of the asset at the end of its useful life. Example 3 assumes that an item of equipment is depreciated using the straight-line method and illustrates how the annual depreciation expense varies under different estimates of the useful life and estimated residual value of an asset. As shown, annual depreciation expense is sensitive to both the estimated useful life and to the estimated residual value.

EXAMPLE 3

Sensitivity of Annual Depreciation Expense to Varying Estimates of Useful Life and Residual Value

Using the straight-line method of depreciation, annual depreciation expense is calculated as:

$$\frac{\text{Cost} - \text{Residual value}}{\text{Estimated useful life}}$$

Assume the cost of an asset is $10,000. If, for example, the residual value of the asset is estimated to be $0 and its useful life is estimated to be 5 years, the annual depreciation expense under the straight-line method would be ($10,000 – $0)/5 years = $2,000. In contrast, holding the estimated useful life of the asset constant at 5 years but increasing the estimated residual value of the asset to $4,000 would result in annual depreciation expense of only $1,200 [calculated as ($10,000 – $4,000)/5 years]. Alternatively, holding the estimated residual value at $0 but increasing the estimated useful life of the asset to 10 years would result in annual depreciation expense of only $1,000 [calculated as ($10,000 – $0)/10 years]. Exhibit 7 shows annual depreciation expense for various combinations of estimated useful life and residual value.

Exhibit 7 Annual Depreciation Expense (in Dollars)

Estimated Useful Life (Years)	Estimated Residual Value					
	0	1,000	2,000	3,000	4,000	5,000
2	5,000	4,500	4,000	3,500	3,000	2,500

Exhibit 7 (Continued)

Estimated Useful Life (Years)	Estimated Residual Value					
4	2,500	2,250	2,000	1,750	1,500	1,250
5	2,000	1,800	1,600	1,400	1,200	1,000
8	1,250	1,125	1,000	875	750	625
10	1,000	900	800	700	600	500

Generally, alternatives to the straight-line method of depreciation are called **accelerated methods** of depreciation because they accelerate (i.e., speed up) the timing of depreciation. Accelerated depreciation methods allocate a greater proportion of the cost to the early years of an asset's useful life. These methods are appropriate if the plant or equipment is expected to be used up faster in the early years (e.g., an automobile). A commonly used accelerated method is the **diminishing balance method**, (also known as the declining balance method). The diminishing balance method is demonstrated in Example 4.

EXAMPLE 4

An Illustration of Diminishing Balance Depreciation

Assume the cost of computer equipment was $11,000, the estimated residual value is $1,000, and the estimated useful life is five years. Under the diminishing or declining balance method, the first step is to determine the straight-line rate, the rate at which the asset would be depreciated under the straight-line method. This rate is measured as 100 percent divided by the useful life or 20 percent for a five-year useful life. Under the straight-line method, 1/5 or 20 percent of the depreciable cost of the asset (here, $11,000 – $1,000 = $10,000) would be expensed each year for five years: The depreciation expense would be $2,000 per year.

The next step is to determine an acceleration factor that approximates the pattern of the asset's wear. Common acceleration factors are 150 percent and 200 percent. The latter is known as **double declining balance depreciation** because it depreciates the asset at double the straight-line rate. Using the 200 percent acceleration factor, the diminishing balance rate would be 40 percent (20 percent × 2.0). This rate is then applied to the remaining undepreciated balance of the asset each period (known as the **net book value**).

At the beginning of the first year, the net book value is $11,000. Depreciation expense for the first full year of use of the asset would be 40 percent of $11,000, or $4,400. Under this method, the residual value, if any, is generally not used in the computation of the depreciation each period (the 40 percent is applied to $11,000 rather than to $11,000 minus residual value). However, the company will stop taking depreciation when the salvage value is reached.

At the beginning of Year 2, the net book value is measured as

Asset cost	$11,000
Less: Accumulated depreciation	(4,400)
Net book value	$6,600

For the second full year, depreciation expense would be $6,600 × 40 percent, or $2,640. At the end of the second year (i.e., beginning of the third year), a total of $7,040 ($4,400 + $2,640) of depreciation would have been recorded. So, the remaining net book value at the beginning of the third year would be

Asset cost	$11,000
Less: Accumulated depreciation	(7,040)
Net book value	$3,960

For the third full year, depreciation would be $3,960 × 40 percent, or $1,584. At the end of the third year, a total of $8,624 ($4,400 + $2,640 + $1,584) of depreciation would have been recorded. So, the remaining net book value at the beginning of the fourth year would be

Asset cost	$11,000
Less: Accumulated depreciation	(8,624)
Net book value	$2,376

For the fourth full year, depreciation would be $2,376 × 40 percent, or $950. At the end of the fourth year, a total of $9,574 ($4,400 + $2,640 + $1,584 + $950) of depreciation would have been recorded. So, the remaining net book value at the beginning of the fifth year would be

Asset cost	$11,000
Less: Accumulated depreciation	(9,574)
Net book value	$1,426

For the fifth year, if deprecation were determined as in previous years, it would amount to $570 ($1,426 × 40 percent). However, this would result in a remaining net book value of the asset below its estimated residual value of $1,000. So, instead, only $426 would be depreciated, leaving a $1,000 net book value at the end of the fifth year.

Asset cost	$11,000
Less: Accumulated depreciation	(10,000)
Net book value	$1,000

Companies often use a zero or small residual value, which creates problems for diminishing balance depreciation because the asset never fully depreciates. In order to fully depreciate the asset over the initially estimated useful life when a zero or small residual value is assumed, companies often adopt a depreciation policy that combines the diminishing balance and straight-line methods. An example would be a deprecation policy of using double-declining balance depreciation and switching to the straight-line method halfway through the useful life.

Under accelerated depreciation methods, there is a higher depreciation expense in early years relative to the straight-line method. This results in higher expenses and lower net income in the early depreciation years. In later years, there is a reversal with accelerated depreciation expense lower than straight-line depreciation. Accelerated depreciation is sometimes referred to as a conservative accounting choice because it results in lower net income in the early years of asset use.

For those intangible assets that must be amortised (those with an identifiable useful life), the process is the same as for depreciation; only the name of the expense is different. IFRS state that if a pattern cannot be determined over the useful life, then the straight-line method should be used.[20] In most cases under IFRS and US GAAP, amortisable intangible assets are amortised using the straight-line method with no residual value. **Goodwill**[21] and intangible assets with indefinite life are not amortised. Instead, they are tested at least annually for impairment (i.e., if the current value of an intangible asset or goodwill is materially lower than its value in the company's books, the value of the asset is considered to be impaired and its value in the company's books must be decreased).

In summary, to calculate depreciation and amortisation, a company must choose a method, estimate the asset's useful life, and estimate residual value. Clearly, different choices have a differing effect on depreciation or amortisation expense and, therefore, on reported net income.

IMPLICATIONS FOR FINANCIAL ANALYSTS: EXPENSE RECOGNITION

7

d describe general principles of expense recognition, specific expense recognition applications, and implications of expense recognition choices for financial analysis;

A company's estimates for doubtful accounts and/or for warranty expenses can affect its reported net income. Similarly, a company's choice of depreciation or amortisation method, estimates of assets' useful lives, and estimates of assets' residual values can affect reported net income. These are only a few of the choices and estimates that affect a company's reported net income.

As with revenue recognition policies, a company's choice of expense recognition can be characterized by its relative conservatism. A policy that results in recognition of expenses later rather than sooner is considered less conservative. In addition, many items of expense require the company to make estimates that can significantly affect net income. Analysis of a company's financial statements, and particularly comparison of one company's financial statements with those of another, requires an understanding of differences in these estimates and their potential impact.

If, for example, a company shows a significant year-to-year change in its estimates of uncollectible accounts as a percentage of sales, warranty expenses as a percentage of sales, or estimated useful lives of assets, the analyst should seek to understand the underlying reasons. Do the changes reflect a change in business operations (e.g., lower estimated warranty expenses reflecting recent experience of fewer warranty claims because of improved product quality)? Or are the changes seemingly unrelated to changes in business operations and thus possibly a signal that a company is manipulating estimates in order to achieve a particular effect on its reported net income?

As another example, if two companies in the same industry have dramatically different estimates for uncollectible accounts as a percentage of their sales, warranty expenses as a percentage of sales, or estimated useful lives as a percentage of assets, it is important to understand the underlying reasons. Are the differences consistent with differences in the two companies' business operations (e.g., lower uncollectible accounts for one company reflecting a different, more creditworthy customer base or possibly stricter credit policies)? Another difference consistent with differences in business operations would be a difference in estimated useful lives of assets if one of the companies employs newer equipment. Or, alternatively, are the differences seemingly inconsistent with differences in the two companies' business operations, possibly signaling that a company is manipulating estimates?

---

[20] IAS 38, *Intangible Assets.*

[21] Goodwill is recorded in acquisitions and is the amount by which the price to purchase an entity exceeds the amount of net identifiable assets acquired (the total amount of identifiable assets acquired less liabilities assumed).

Information about a company's accounting policies and significant estimates are described in the notes to the financial statements and in the management discussion and analysis section of a company's annual report.

When possible, the monetary effect of differences in expense recognition policies and estimates can facilitate more meaningful comparisons with a single company's historical performance or across a number of companies. An analyst can use the monetary effect to adjust the reported expenses so that they are on a comparable basis.

Even when the monetary effects of differences in policies and estimates cannot be calculated, it is generally possible to characterize the relative conservatism of the policies and estimates and, therefore, to qualitatively assess how such differences might affect reported expenses and thus financial ratios.

8

NON-RECURRING ITEMS AND NON-OPERATING ITEMS: DISCONTINUED OPERATIONS AND UNUSUAL OR INFREQUENT ITEMS

e describe the financial reporting treatment and analysis of non-recurring items (including discontinued operations, unusual or infrequent items) and changes in accounting policies;

From a company's income statements, we can see its earnings from last year and in the previous year. Looking forward, the question is: What will the company earn next year and in the years after?

To assess a company's future earnings, it is helpful to separate those prior years' items of income and expense that are likely to continue in the future from those items that are less likely to continue.[22] Some items from prior years are clearly not expected to continue in the future periods and are separately disclosed on a company's income statement. This is consistent with "An entity shall present additional line items, headings, and subtotals … when such presentation is relevant to an understanding of the entity's financial performance."[23] IFRS describe considerations that enter into the decision to present information other than that explicitly specified by a standard. Both IFRS and US GAAP specify that the results of discontinued operations should be reported separately from continuing operations. Other items that may be reported separately on a company's income statement, such as unusual items, items that occur infrequently, effects due to accounting changes, and non-operating income, require the analyst to make some judgments.

8.1 Discontinued Operations

When a company disposes of or establishes a plan to dispose of one of its component operations and will have no further involvement in the operation, the income statement reports separately the effect of this disposal as a "discontinued" operation under both IFRS and US GAAP. Financial standards provide various criteria for reporting the effect separately, which are generally that the discontinued component must be separable both physically and operationally.[24]

---

[22] In business writing, items expected to continue in the future are often described as "persistent" or "permanent," whereas those not expected to continue are described as "transitory."

[23] IAS No. 1, *Presentation of Financial Statements*, paragraph 85.

In Exhibit 1, AB InBev reported profit from discontinued operations of $28 million in 2017 and $48 million in 2016. In Exhibit 2, Molson Coors reported income from discontinued operations of $1.5 million and $3.9 million in 2017 and 2015, respectively, and a loss from discontinued operations of $2.8 million in 2016.

Because the discontinued operation will no longer provide earnings (or cash flow) to the company, an analyst may eliminate discontinued operations in formulating expectations about a company's future financial performance.

8.2 Unusual or Infrequent Items

IFRS require that items of income or expense that are material and/or relevant to the understanding of the entity's financial performance should be disclosed separately. Unusual or infrequent items are likely to meet these criteria. Under US GAAP, material items that are unusual or infrequent, and that are both as of reporting periods beginning after December 15, 2015, are shown as part of a company's continuing operations but are presented separately. For example, restructuring charges, such as costs to close plants and employee termination costs, are considered part of a company's ordinary activities. As another example, gains and losses arising when a company sells an asset or part of a business, for more or less than its carrying value, are also disclosed separately on the income statement. These sales are considered ordinary business activities.

Highlighting the unusual or infrequent nature of these items assists an analyst in judging the likelihood that such items will reoccur. This meets the IFRS criteria of disclosing items that are relevant to the understanding of an entity's financial performance. In Exhibit 2, Molson Coors' income statement showed a separate line item for "Special Items, net." The company's footnotes provide details on the amount and explain that this line includes revenues or expenses that either they "do not believe to be indicative of [their] core operations, or they believe are significant to [their] current operating results warranting separate classification". In Exhibit 3, the income statement of Danone shows an amount for "Recurring operating income" followed by a separate line item for "other operating income (expense)", which is not included as a component of recurring income. Exhibit 8 presents an excerpt from Danone's additional disclosure about this non-recurring amount.

Exhibit 8 Highlighting Infrequent Nature of Items—Excerpt from Groupe Danone footnotes to its 2017 financial statements

NOTE 6. Events and Transactions Outside the Group's Ordinary Activities [Excerpt]

> "Other operating income (expense) is defined under Recommendation 2013-03 of the French CNC relating to the format of consolidated financial statements prepared under international accounting standards, and comprises significant items that, because of their exceptional nature, cannot be viewed as inherent to Danone's current activities. These mainly include capital gains and losses on disposals of fully consolidated companies, impairment charges on goodwill, significant costs related to strategic restructuring and major external

(continued)

---

[24] IFRS No. 5, *Non-Current Assets Held for Sale and Discontinued Operations*, paragraphs 31–33.

Exhibit 8 (Continued)

growth transactions, and incurred or estimated costs related to major crises and major litigation. Furthermore, in connection with Revised IFRS 3 and Revised IAS 27, Danone also classifies in Other operating income (expense) (i) acquisition costs related to business combinations, (ii) revaluation profit or loss accounted for following a loss of control, and (iii) changes in earn-outs related to business combinations and subsequent to the acquisition date.

"In 2017, the net Other operating income of €192 million consisted mainly of the following items:

(in € millions)	Related income (expense)
Capital gain on disposal of Stonyfield	628
Compensation received following the decision of the Singapore arbitration court in the Fonterra case	105
Territorial risks, mainly in certain countries in the ALMA region	(148)
Costs associated with the integration of WhiteWave	(118)
Impairment of several intangible assets in Waters and Specialized Nutrition Reporting entities	(115)
Remainder of table omitted	

In Exhibit 8, Danone provides details on items considered to be "exceptional" items and not "inherent" to the company's current activities. The exceptional items include gains on asset disposals, receipts from a legal case, costs of integrating an acquisition, and impairment of intangible assets, among others. Generally, in forecasting future operations, an analyst would assess whether the items reported are likely to reoccur and also possible implications for future earnings. It is generally not advisable simply to ignore all unusual items.

9 NON-RECURRING ITEMS: CHANGES IN ACCOUNTING POLICY

e describe the financial reporting treatment and analysis of non-recurring items (including discontinued operations, unusual or infrequent items) and changes in accounting policies;

At times, standard setters issue new standards that require companies to change accounting policies. Depending on the standard, companies may be permitted to adopt the standards prospectively (in the future) or retrospectively (restate financial statements as though the standard existed in the past). In other cases, changes in accounting policies (e.g., from one acceptable inventory costing method to another) are made for other reasons, such as providing a better reflection of the company's performance. Changes in accounting policies are reported through retrospective application[25] unless it is impractical to do so.

Retrospective application means that the financial statements for all fiscal years shown in a company's financial report are presented as if the newly adopted accounting principle had been used throughout the entire period. Notes to the financial statements describe the change and explain the justification for the change. Because changes in accounting principles are retrospectively applied, the financial statements that appear within a financial report are comparable.

Example 5 presents an excerpt from Microsoft Corporation's Form 10-K for the fiscal year ended 30 June 2018 describing a change in accounting principle resulting from the new revenue recognition standard. Microsoft elected to adopt the new standard 1 July 2017, earlier than the required adoption date. Microsoft also elected to use the "full retrospective method," which requires companies to restate prior periods' results. On its income statement, both 2016 and 2017 are presented as if the new standard had been used throughout both years. In the footnotes to its financial statements, Microsoft discloses the impact of the new standard.

EXAMPLE 5

Microsoft Corporation Excerpt from Footnotes to the Financial Statements

> The most significant impact of the [new revenue recognition] standard relates to our accounting for software license revenue. Specifically, for Windows 10, we recognize revenue predominantly at the time of billing and delivery rather than ratably over the life of the related device. For certain multi-year commercial software subscriptions that include both distinct software licenses and SA, we recognize license revenue at the time of contract execution rather than over the subscription period. Due to the complexity of certain of our commercial license subscription contracts, the actual revenue recognition treatment required under the standard depends on contract-specific terms and in some instances may vary from recognition at the time of billing. Revenue recognition related to our hardware, cloud offerings (such as Office 365), LinkedIn, and professional services remains substantially unchanged. Refer to Impacts to Previously Reported Results below for the impact of adoption of the standard in our consolidated financial statements.

(In $ millions, except per share amounts)	As Previously Reported	New Revenue Standard Adjustment	As Restated
Income Statements			
Year Ended June 30, 2017			
Revenue	89,950	6,621	96,571
Provision for income taxes	1,945	2,467	4,412

(continued)

---

[25] IAS No. 8, *Accounting Policies, Changes in Accounting Estimates and Errors,* and FASB ASC Topic 250 [Accounting Changes and Error Corrections].

(In $ millions, except per share amounts)	As Previously Reported	New Revenue Standard Adjustment	As Restated
Net income	21,204	4,285	25,489
Diluted earnings per share	2.71	0.54	3.25
Year Ended June 30, 2016			
Revenue	85,320	5,834	91,154
Provision for income taxes	2,953	2,147	5,100
Net income	16,798	3,741	20,539
Diluted earnings per share	2.1	0.46	2.56

Question: Based on the above information, describe whether Microsoft's results appear better or worse under the new revenue recognition standard.

Solution:

Microsoft's results appear better under the new revenue recognition standard. Revenues and income are higher under the new standard. The net profit margin is higher under the new standard. For 2017, the net profit margin is 26.4% (= 25,489/96,571) under the new standard versus 23.6% (= 21,204/89,950) under the old standard. Reported revenue grew faster under the new standard. Revenue growth under the new standard was 5.9% [= (96,571/91,154) – 1] compared to 5.4% [= (89,950/85,320) – 1)] under the old standard.

Microsoft's presentation of the effects of the new revenue recognition enables an analyst to identify the impact of the change in accounting standards.

Note that the new revenue recognition standard also offered companies the option of using a "modified retrospective" method of adoption. Under the modified retrospective approach, companies were not required to revise previously reported financial statements. Instead, they adjusted opening balances of retained earnings (and other applicable accounts) for the cumulative impact of the new standard.

In contrast to changes in accounting policies (such as whether to expense the cost of employee stock options), companies sometimes make *changes in accounting estimates* (such as the useful life of a depreciable asset). Changes in accounting estimates are handled prospectively, with the change affecting the financial statements for the period of change and future periods. No adjustments are made to prior statements, and the adjustment is not shown on the face of the income statement. Significant changes should be disclosed in the notes. Exhibit 9 provides an excerpt from the annual Form 10-K of Catalent Inc., a US-based biotechnology company, that illustrates a change in accounting estimate.

Exhibit 9 Change in Accounting Estimate

Catalent Inc. discloses a change in the method it uses to calculate both annual expenses related to its defined benefit pension plans. Rather than use a single, weighted-average discount rate in its calculations, the company will use the spot rates applicable to each projected cash flow.

Post-Retirement and Pension Plans

Exhibit 9 (Continued)

> …The measurement of the related benefit obligations and the net periodic benefit costs recorded each year are based upon actuarial computations, which require management's judgment as to certain assumptions. These assumptions include the discount rates used in computing the present value of the benefit obligations and the net periodic benefit costs...
>
> Effective June 30, 2016, the approach used to estimate the service and interest components of net periodic benefit cost for benefit plans was changed to provide a more precise measurement of service and interest costs. Historically, the Company estimated these service and interest components utilizing a single weighted-average discount rate derived from the yield curve used to measure the benefit obligation at the beginning of the period. Going forward, the Company has elected to utilize an approach that discounts the individual expected cash flows using the applicable spot rates derived from the yield curve over the projected cash flow period. The Company has accounted for this change as a change in accounting estimate that is inseparable from a change in accounting principle and accordingly has accounted for it prospectively.

Another possible adjustment is a *correction of an error for a prior period* (e.g., in financial statements issued for an earlier year). This cannot be handled by simply adjusting the current period income statement. Correction of an error for a prior period is handled by restating the financial statements (including the balance sheet, statement of owners' equity, and cash flow statement) for the prior periods presented in the current financial statements.[26] Note disclosures are required regarding the error. These disclosures should be examined carefully because they may reveal weaknesses in the company's accounting systems and financial controls.

NON-OPERATING ITEMS

10

f contrast operating and non-operating components of the income statement;

Non-operating items are typically reported separately from operating income because they are material and/or relevant to the understanding of the entity's financial performance. Under IFRS, there is no definition of operating activities, and companies that choose to report operating income or the results of operating activities should ensure that these represent activities that are normally regarded as operating. Under US GAAP, operating activities generally involve producing and delivering goods and providing services and include all transactions and other events that are not defined as investing or financing activities.[27] For example, if a non-financial service company invests in equity or debt securities issued by another company, any interest, dividends, or profits from sales of these securities will be shown as non-operating income. In general, for non-financial services companies,[28] non-operating income that is disclosed separately on the income statement (or in the notes) includes amounts earned through investing activities.

---

[26] Ibid.

[27] FASB ASC *Master Glossary.*

Among non-operating items on the income statement (or accompanying notes), non-financial service companies also disclose the interest expense on their debt securities, including amortisation of any discount or premium. The amount of interest expense is related to the amount of a company's borrowings and is generally described in the notes to the financial statements. For financial service companies, interest income and expense are likely components of operating activities. (Note that the characterization of interest and dividends as non-operating items on the income statement is not necessarily consistent with the classification on the statement of cash flows. Specifically, under IFRS, interest and dividends received can be shown either as operating or as investing on the statement of cash flows, while under US GAAP interest and dividends received are shown as operating cash flows. Under IFRS, interest and dividends paid can be shown either as operating or as financing on the statement of cash flows, while under US GAAP, interest paid is shown as operating and dividends paid are shown as financing.)

In practice, companies often disclose the interest expense and income separately, along with a net amount. For example, in Exhibit 1, ABN InBev's 2017 income statement shows finance cost of $6,885 million, finance income of $378 million, and net finance cost of $6,507 million. Similarly, in Exhibit 3, Danone's 2017 income statement shows interest income of €130, interest expense of €276, and cost of net debt of €146.

For purposes of assessing a company's future performance, the amount of financing expense will depend on the company's financing policy (target capital structure) and borrowing costs. The amount of investing income will depend on the purpose and success of investing activities. For a non-financial company, a significant amount of financial income would typically warrant further exploration. What are the reasons underlying the company's investments in the securities of other companies? Is the company simply investing excess cash in short-term securities to generate income higher than cash deposits, or is the company purchasing securities issued by other companies for strategic reasons, such as access to raw material supply or research?

11

EARNINGS PER SHARE AND CAPITAL STRUCTURE AND BASIC EPS

g describe how earnings per share is calculated and calculate and interpret a company's earnings per share (both basic and diluted earnings per share) for both simple and complex capital structures;

One metric of particular importance to an equity investor is earnings per share (EPS). EPS is an input into ratios such as the price/earnings ratio. Additionally, each shareholder in a company owns a different number of shares. IFRS require the presentation of EPS on the face of the income statement for net profit or loss (net income) and profit or loss (income) from continuing operations.[29] Similar presentation is required under US GAAP.[30] This section outlines the calculations for EPS and explains how the calculation differs for a simple versus complex capital structure.

---

[28] Examples of financial services companies are insurance companies, banks, brokers, dealers, and investment companies.

[29] IAS No. 33, *Earnings Per Share.*

[30] FASB ASC Topic 260 [Earnings Per Share].

11.1 Simple versus Complex Capital Structure

A company's capital is composed of its equity and debt. Some types of equity have preference over others, and some debt (and other instruments) may be converted into equity. Under IFRS, the type of equity for which EPS is presented is referred to as ordinary. **Ordinary shares** are those equity shares that are subordinate to all other types of equity. The ordinary shareholders are basically the owners of the company—the equity holders who are paid last in a liquidation of the company and who benefit the most when the company does well. Under US GAAP, this ordinary equity is referred to as **common stock** or **common shares**, reflecting US language usage. The terms "ordinary shares," "common stock," and "common shares" are used interchangeably in the following discussion.

When a company has issued any financial instruments that are potentially convertible into common stock, it is said to have a complex capital structure. Examples of financial instruments that are potentially convertible into common stock include convertible bonds, convertible preferred stock, employee stock options, and warrants.[31] If a company's capital structure does not include such potentially convertible financial instruments, it is said to have a simple capital structure.

The distinction between simple versus complex capital structure is relevant to the calculation of EPS because financial instruments that are potentially convertible into common stock could, as a result of conversion or exercise, potentially dilute (i.e., decrease) EPS. Information about such a potential dilution is valuable to a company's current and potential shareholders; therefore, accounting standards require companies to disclose what their EPS would be if all dilutive financial instruments were converted into common stock. The EPS that would result if all dilutive financial instruments were converted is called **diluted EPS**. In contrast, **basic EPS** is calculated using the reported earnings available to common shareholders of the parent company and the weighted average number of shares outstanding.

Companies are required to report both basic and diluted EPS as well as amounts for continuing operations. Exhibit 10 shows the per share amounts reported by AB InBev at the bottom of its income statement that was presented in Exhibit 1. The company's basic EPS ("before dilution") was $4.06, and diluted EPS ("after dilution") was $3.98 for 2017. In addition, in the same way that AB InBev's income statement shows income from continuing operations separately from total income, EPS from continuing operations is also shown separately from total EPS. For 2017, the basic and diluted EPS from continuing operations were $4.04 and $3.96, respectively. Across all measures, AB InBev's EPS was much higher in 2017 than in 2016. An analyst would seek to understand the causes underlying the changes in EPS, a topic we will address following an explanation of the calculations of both basic and diluted EPS.

Exhibit 10 AB InBev's Earnings Per Share

	12 Months Ended December 31		
	2017	**2016**	**2015**
Basic earnings per share	$4.06	$0.72	$5.05
Diluted earnings per share	3.98	0.71	4.96

(continued)

---

[31] A warrant is a call option typically attached to securities issued by a company, such as bonds. A warrant gives the holder the right to acquire the company's stock from the company at a specified price within a specified time period. IFRS and US GAAP standards regarding earnings per share apply equally to call options, warrants, and equivalent instruments.

Exhibit 10 (Continued)

	12 Months Ended December 31		
	2017	**2016**	**2015**
Basic earnings per share from continuing operations	4.04	0.69	5.05
Diluted earnings per share from continuing operations	\$3.96	\$0.68	\$4.96

11.2 Basic EPS

Basic EPS is the amount of income available to common shareholders divided by the weighted average number of common shares outstanding over a period. The amount of income available to common shareholders is the amount of net income remaining after preferred dividends (if any) have been paid. Thus, the formula to calculate basic EPS is:

$$\text{Basic EPS} = \frac{\text{Net income} - \text{Preferred dividends}}{\text{Weighted average number of shares outstanding}} \tag{1}$$

The weighted average number of shares outstanding is a time weighting of common shares outstanding. For example, assume a company began the year with 2,000,000 common shares outstanding and repurchased 100,000 common shares on 1 July. The weighted average number of common shares outstanding would be the sum of 2,000,000 shares × 1/2 year + 1,900,000 shares × 1/2 year, or 1,950,000 shares. So the company would use 1,950,000 shares as the weighted average number of shares in calculating its basic EPS.

If the number of shares of common stock increases as a result of a stock dividend or a stock split, the EPS calculation reflects the change retroactively to the beginning of the period.

Examples 6, 7, and 8 illustrate the computation of basic EPS.

EXAMPLE 6

A Basic EPS Calculation (1)

For the year ended 31 December 2018, Shopalot Company had net income of \$1,950,000. The company had 1,500,000 shares of common stock outstanding, no preferred stock, and no convertible financial instruments. What is Shopalot's basic EPS?

Solution:

Shopalot's basic EPS is \$1.30 (\$1,950,000 divided by 1,500,000 shares).

EXAMPLE 7

A Basic EPS Calculation (2)

For the year ended 31 December 2018, Angler Products had net income of \$2,500,000. The company declared and paid \$200,000 of dividends on preferred stock. The company also had the following common stock share information:

Shares outstanding on 1 January 2018	1,000,000
Shares issued on 1 April 2018	200,000
Shares repurchased (treasury shares) on 1 October 2018	(100,000)
Shares outstanding on 31 December 2018	1,100,000

1 What is the company's weighted average number of shares outstanding?

2 What is the company's basic EPS?

Solution to 1:

The weighted average number of shares outstanding is determined by the length of time each quantity of shares was outstanding:

1,000,000 × (3 months/12 months) =	250,000
1,200,000 × (6 months/12 months) =	600,000
1,100,000 × (3 months/12 months) =	275,000
Weighted average number of shares outstanding	1,125,000

Solution to 2:

Basic EPS = (Net income – Preferred dividends)/Weighted average number of shares = ($2,500,000 – $200,000)/1,125,000 = $2.04

EXAMPLE 8

A Basic EPS Calculation (3)

Assume the same facts as Example 7 except that on 1 December 2018, a previously declared 2-for-1 stock split took effect. Each shareholder of record receives two shares in exchange for each current share that he or she owns. What is the company's basic EPS?

Solution:

For EPS calculation purposes, a stock split is treated as if it occurred at the beginning of the period. The weighted average number of shares would, therefore, be 2,250,000, and the basic EPS would be $1.02 [= ($2,500,000 – $200,000)/2,250,000].

DILUTED EPS: THE IF-CONVERTED METHOD

12

g describe how earnings per share is calculated and calculate and interpret a company's earnings per share (both basic and diluted earnings per share) for both simple and complex capital structures;

If a company has a simple capital structure (in other words, one that includes no potentially dilutive financial instruments), then its basic EPS is equal to its diluted EPS. However, if a company has potentially dilutive financial instruments, its diluted EPS may differ from its basic EPS. Diluted EPS, by definition, is always equal to or less than basic EPS. The sections below describe the effects of three types of potentially dilutive financial instruments on diluted EPS: convertible preferred, convertible debt, and employee stock options. The final section explains why not all potentially dilutive financial instruments actually result in a difference between basic and diluted EPS.

12.1 Diluted EPS When a Company Has Convertible Preferred Stock Outstanding

When a company has convertible preferred stock outstanding, diluted EPS is calculated using the **if-converted method**. The if-converted method is based on what EPS would have been if the convertible preferred securities had been converted at the beginning of the period. In other words, the method calculates what the effect would have been if the convertible preferred shares converted at the beginning of the period. If the convertible shares had been converted, there would be two effects. First, the convertible preferred securities would no longer be outstanding; instead, additional common stock would be outstanding. Thus, under the if-converted method, the weighted average number of shares outstanding would be higher than in the basic EPS calculation. Second, if such a conversion had taken place, the company would not have paid preferred dividends. Thus, under the if-converted method, the net income available to common shareholders would be higher than in the basic EPS calculation.

Diluted EPS using the if-converted method for convertible preferred stock is equal to net income divided by the weighted average number of shares outstanding from the basic EPS calculation plus the additional shares of common stock that would be issued upon conversion of the preferred. Thus, the formula to calculate diluted EPS using the if-converted method for preferred stock is:

$$\text{Diluted EPS} = \frac{(\text{Net income})}{\begin{array}{l}(\text{Weighted average number of shares} \\ \text{outstanding} + \text{New common shares that} \\ \text{would have been issued at conversion})\end{array}} \tag{2}$$

A diluted EPS calculation using the if-converted method for preferred stock is provided in Example 9.

EXAMPLE 9

A Diluted EPS Calculation Using the If-Converted Method for Preferred Stock

For the year ended 31 December 2018, Bright-Warm Utility Company (fictitious) had net income of $1,750,000. The company had an average of 500,000 shares of common stock outstanding, 20,000 shares of convertible preferred, and no other potentially dilutive securities. Each share of preferred pays a dividend of $10 per share, and each is convertible into five shares of the company's common stock. Calculate the company's basic and diluted EPS.

Solution:

If the 20,000 shares of convertible preferred had each converted into 5 shares of the company's common stock, the company would have had an additional 100,000 shares of common stock (5 shares of common for each of the 20,000 shares of preferred). If the conversion had taken place, the company would not have paid preferred dividends of $200,000 ($10 per share for each of the 20,000 shares of preferred). As shown in Exhibit 11, the company's basic EPS was $3.10 and its diluted EPS was $2.92.

Exhibit 11 Calculation of Diluted EPS for Bright-Warm Utility Company Using the If-Converted Method: Case of Preferred Stock

	Basic EPS	Diluted EPS Using If-Converted Method
Net income	$1,750,000	$1,750,000
Preferred dividend	–200,000	0
Numerator	$1,550,000	$1,750,000
Weighted average number of shares outstanding	500,000	500,000
Additional shares issued if preferred converted	0	100,000
Denominator	500,000	600,000
EPS	**$3.10**	**$2.92**

12.2 Diluted EPS When a Company Has Convertible Debt Outstanding

When a company has convertible debt outstanding, the diluted EPS calculation also uses the if-converted method. Diluted EPS is calculated as if the convertible debt had been converted at the beginning of the period. If the convertible debt had been converted, the debt securities would no longer be outstanding; instead, additional shares of common stock would be outstanding. Also, if such a conversion had taken place, the company would not have paid interest on the convertible debt, so the net income available to common shareholders would increase by the after-tax amount of interest expense on the debt converted.

Thus, the formula to calculate diluted EPS using the if-converted method for convertible debt is:

$$\text{Diluted EPS} = \frac{\left(\begin{array}{l}\text{Net income} + \text{After-tax interest on}\\ \text{convertible debt} - \text{Preferred dividends}\end{array}\right)}{\left(\begin{array}{l}\text{Weighted average number of shares}\\ \text{outstanding} + \text{Additional common}\\ \text{shares that would have been}\\ \text{issued at conversion}\end{array}\right)} \tag{3}$$

A diluted EPS calculation using the if-converted method for convertible debt is provided in Example 10.

EXAMPLE 10

A Diluted EPS Calculation Using the If-Converted Method for Convertible Debt

Oppnox Company (fictitious) reported net income of $750,000 for the year ended 31 December 2018. The company had a weighted average of 690,000 shares of common stock outstanding. In addition, the company has only one potentially dilutive security: $50,000 of 6 percent convertible bonds, convertible into a total of 10,000 shares. Assuming a tax rate of 30 percent, calculate Oppnox's basic and diluted EPS.

Solution:

If the debt securities had been converted, the debt securities would no longer be outstanding and instead, an additional 10,000 shares of common stock would be outstanding. Also, if the debt securities had been converted, the company would not have paid interest of $3,000 on the convertible debt, so net income available to common shareholders would have increased by $2,100 [= $3,000(1 – 0.30)] on an after-tax basis. Exhibit 12 illustrates the calculation of diluted EPS using the if-converted method for convertible debt.

Exhibit 12 Calculation of Diluted EPS for Oppnox Company Using the If-Converted Method: Case of a Convertible Bond

	Basic EPS	Diluted EPS Using If-Converted Method
Net income	$750,000	$750,000
After-tax cost of interest		2,100
Numerator	$750,000	$752,100
Weighted average number of shares outstanding	690,000	690,000
If converted	0	10,000
Denominator	690,000	700,000
EPS	**$1.09**	**$1.07**

13 DILUTED EPS: THE TREASURY STOCK METHOD

g describe how earnings per share is calculated and calculate and interpret a company's earnings per share (both basic and diluted earnings per share) for both simple and complex capital structures;

When a company has stock options, warrants, or their equivalents[32] outstanding, diluted EPS is calculated as if the financial instruments had been exercised and the company had used the proceeds from exercise to repurchase as many shares of common stock as possible at the average market price of common stock during the period. The weighted average number of shares outstanding for diluted EPS is thus increased by the number of shares that would be issued upon exercise minus the number of shares that would have been purchased with the proceeds. This method is called the **treasury stock method** under US GAAP because companies typically hold repurchased shares as treasury stock. The same method is used under IFRS but is not named.

---

[32] Hereafter, options, warrants, and their equivalents will be referred to simply as "options" because the accounting treatment for EPS calculations is interchangeable for these instruments under IFRS and US GAAP.

For the calculation of diluted EPS using this method, the assumed exercise of these financial instruments would have the following effects:

- The company is assumed to receive cash upon exercise and, in exchange, to issue shares.
- The company is assumed to use the cash proceeds to repurchase shares at the weighted average market price during the period.

As a result of these two effects, the number of shares outstanding would increase by the incremental number of shares issued (the difference between the number of shares issued to the holders and the number of shares assumed to be repurchased by the company). For calculating diluted EPS, the incremental number of shares is weighted based upon the length of time the financial instrument was outstanding in the year. If the financial instrument was issued prior to the beginning of the year, the weighted average number of shares outstanding increases by the incremental number of shares. If the financial instruments were issued during the year, then the incremental shares are weighted by the amount of time the financial instruments were outstanding during the year.

The assumed exercise of these financial instruments would not affect net income. For calculating EPS, therefore, no change is made to the numerator. The formula to calculate diluted EPS using the treasury stock method (same method as used under IFRS but not named) for options is:

$$\text{Diluted EPS} = \frac{(\text{Net income} - \text{Preferred dividends})}{\begin{aligned}&[\text{Weighted average number of shares}\\ &\text{outstanding} + (\text{New shares that would}\\ &\text{have been issued at option exercise} -\\ &\text{Shares that could have been purchased}\\ &\text{with cash received upon exercise}) \times\\ &(\text{Proportion of year during which the}\\ &\text{financial instruments were outstanding})]\end{aligned}} \tag{4}$$

A diluted EPS calculation using the treasury stock method for options is provided in Example 11.

EXAMPLE 11

A Diluted EPS Calculation Using the Treasury Stock Method for Options

Hihotech Company (fictitious) reported net income of $2.3 million for the year ended 30 June 2018 and had a weighted average of 800,000 common shares outstanding. At the beginning of the fiscal year, the company has outstanding 30,000 options with an exercise price of $35. No other potentially dilutive financial instruments are outstanding. Over the fiscal year, the company's market price has averaged $55 per share. Calculate the company's basic and diluted EPS.

Solution:

Using the treasury stock method, we first calculate that the company would have received $1,050,000 ($35 for each of the 30,000 options exercised) if all the options had been exercised. The options would no longer be outstanding; instead, 30,000 shares of common stock would be outstanding. Under the treasury stock method, we assume that shares would be repurchased with the cash received upon exercise of the options. At an average market price of $55 per share, the $1,050,000 proceeds from option exercise, the company could have repurchased 19,091 shares. Therefore, the incremental number of shares issued is 10,909 (calculated as 30,000 minus 19,091). For the diluted EPS calculation, no change is made to the numerator. As shown in Exhibit 13, the company's basic EPS was $2.88 and the diluted EPS was $2.84.

Exhibit 13 Calculation of Diluted EPS for Hihotech Company Using the Treasury Stock Method: Case of Stock Options

	Basic EPS	Diluted EPS Using Treasury Stock Method
Net income	$2,300,000	$2,300,000
Numerator	$2,300,000	$2,300,000
Weighted average number of shares outstanding	800,000	800,000
If converted	0	10,909
Denominator	800,000	810,909
EPS	**$2.88**	**$2.84**

As noted, IFRS require a similar computation but does not refer to it as the "treasury stock method." The company is required to consider that any assumed proceeds are received from the issuance of new shares at the average market price for the period. These new "inferred" shares would be disregarded in the computation of diluted EPS, but the excess of the new shares that would be issued under options contracts minus the new inferred shares would be added to the weighted average number of shares outstanding. The results are the same as the treasury stock method, as shown in Example 12.

EXAMPLE 12

Diluted EPS for Options under IFRS

Assuming the same facts as in Example 11, calculate the weighted average number of shares outstanding for diluted EPS under IFRS.

Solution:

If the options had been exercised, the company would have received $1,050,000. If this amount had been received from the issuance of new shares at the average market price of $55 per share, the company would have issued 19,091 shares. IFRS refer to the 19,091 shares the company would have issued at market prices as the inferred shares. The number of shares issued under options (30,000) minus the number of inferred shares (19,091) equals 10,909. This amount is added to the weighted average number of shares outstanding of 800,000 to get diluted shares of 810,909. Note that this is the same result as that obtained under US GAAP; it is just derived in a different manner.

OTHER ISSUES WITH DILUTED EPS AND CHANGES IN EPS

14

h contrast dilutive and antidilutive securities and describe the implications of each for the earnings per share calculation;

It is possible that some potentially convertible securities could be **antidilutive** (i.e., their inclusion in the computation would result in an EPS higher than the company's basic EPS). Under IFRS and US GAAP, antidilutive securities are not included in the calculation of diluted EPS. Diluted EPS should reflect the maximum potential dilution from conversion or exercise of potentially dilutive financial instruments. Diluted EPS will always be less than or equal to basic EPS. Example 13 provides an illustration of an antidilutive security.

EXAMPLE 13

An Antidilutive Security

For the year ended 31 December 2018, Dim-Cool Utility Company (fictitious) had net income of $1,750,000. The company had an average of 500,000 shares of common stock outstanding, 20,000 shares of convertible preferred, and no other potentially dilutive securities. Each share of preferred pays a dividend of $10 per share, and each is convertible into three shares of the company's common stock. What was the company's basic and diluted EPS?

Solution:

If the 20,000 shares of convertible preferred had each converted into 3 shares of the company's common stock, the company would have had an additional 60,000 shares of common stock (3 shares of common for each of the 20,000 shares of preferred). If the conversion had taken place, the company would not have paid preferred dividends of $200,000 ($10 per share for each of the 20,000 shares of preferred). The effect of using the if-converted method would be EPS of $3.13, as shown in Exhibit 14. Because this is greater than the company's basic EPS of $3.10, the securities are said to be antidilutive and the effect of their conversion would not be included in diluted EPS. Diluted EPS would be the same as basic EPS (i.e., $3.10).

Exhibit 14 Calculation for an Antidilutive Security

	Basic EPS	Diluted EPS Using If-Converted Method
Net income	$1,750,000	$1,750,000
Preferred dividend	–200,000	0
Numerator	$1,550,000	$1,750,000
Weighted average number of shares outstanding	500,000	500,000
If converted	0	60,000
Denominator	500,000	560,000

(continued)

Exhibit 14 (Continued)

	Basic EPS	Diluted EPS Using If-Converted Method	
EPS	**$3.10**	$3.13	←Exceeds basic EPS; security is antidilutive and, therefore, **not** included. **Reported diluted EPS = $3.10**.

14.1 Changes in EPS

Having explained the calculations of both basic and diluted EPS, we return to an examination of changes in EPS. As noted above, AB InBev's fully diluted EPS from continuing operations increased from $0.68 in 2016 to $3.96 in 2017. In general, an increase in EPS results from an increase in net income, a decrease in the number of shares outstanding, or a combination of both. In the notes to its financial statements (not shown), AB InBev discloses that the weighted average number of shares for both the basic and fully-diluted calculations was greater in 2017 than in 2016. Thus, for AB InBev, the improvement in EPS from 2016 to 2017 was driven by an increase in net income. Changes in the numerator and denominator explain the changes in EPS arithmetically. To understand the business drivers of those changes requires further research. The next section presents analytical tools that an analyst can use to highlight areas for further examination.

15

COMMON-SIZE ANALYSIS OF THE INCOME STATEMENT

i. formulate income statements into common-size income statements;

j. evaluate a company's financial performance using common-size income statements and financial ratios based on the income statement;

In this section, we apply two analytical tools to analyze the income statement: common-size analysis and income statement ratios. The objective of this analysis is to assess a company's performance over a period of time—compared with its own past performance or the performance of another company.

15.1 Common-Size Analysis of the Income Statement

Common-size analysis of the income statement can be performed by stating each line item on the income statement as a percentage of revenue.[33] Common-size statements facilitate comparison across time periods (time series analysis) and across companies (cross-sectional analysis) because the standardization of each line item removes the effect of size.

To illustrate, Panel A of Exhibit 15 presents an income statement for three hypothetical companies in the same industry. Company A and Company B, each with $10 million in sales, are larger (as measured by sales) than Company C, which has only $2 million in sales. In addition, Companies A and B both have higher operating profit: $2 million and $1.5 million, respectively, compared with Company C's operating profit of only $400,000.

How can an analyst meaningfully compare the performance of these companies? By preparing a common-size income statement, as illustrated in Panel B, an analyst can readily see that the percentages of Company C's expenses and profit relative to its sales are exactly the same as for Company A. Furthermore, although Company C's operating profit is lower than Company B's in absolute dollars, it is higher in percentage terms (20 percent for Company C compared with only 15 percent for Company B). For each $100 of sales, Company C generates $5 more operating profit than Company B. In other words, Company C is relatively more profitable than Company B based on this measure.

The common-size income statement also highlights differences in companies' strategies. Comparing the two larger companies, Company A reports significantly higher gross profit as a percentage of sales than does Company B (70 percent compared with 25 percent). Given that both companies operate in the same industry, why can Company A generate so much higher gross profit? One possible explanation is found by comparing the operating expenses of the two companies. Company A spends significantly more on research and development and on advertising than Company B. Expenditures on research and development likely result in products with superior technology. Expenditures on advertising likely result in greater brand awareness. So, based on these differences, it is likely that Company A is selling technologically superior products with a better brand image. Company B may be selling its products more cheaply (with a lower gross profit as a percentage of sales) but saving money by not investing in research and development or advertising. In practice, differences across companies are more subtle, but the concept is similar. An analyst, noting significant differences, would do more research and seek to understand the underlying reasons for the differences and their implications for the future performance of the companies.

Exhibit 15

Panel A: Income Statements for Companies A, B, and C ($)

	A	B	C
Sales	$10,000,000	$10,000,000	$2,000,000
Cost of sales	3,000,000	7,500,000	600,000
Gross profit	7,000,000	2,500,000	1,400,000
Selling, general, and administrative expenses	1,000,000	1,000,000	200,000

(continued)

---

[33] This format can be distinguished as "vertical common-size analysis." As the reading on financial statement analysis discusses, there is another type of common-size analysis, known as "horizontal common-size analysis," that states items in relation to a selected base year value. Unless otherwise indicated, text references to "common-size analysis" refer to vertical analysis.

Exhibit 15 (Continued)

Panel A: Income Statements for Companies A, B, and C ($)

	A	B	C
Research and development	2,000,000	—	400,000
Advertising	2,000,000	—	400,000
Operating profit	2,000,000	1,500,000	400,000

Panel B: Common-Size Income Statements for Companies A, B, and C (%)

	A	B	C
Sales	100%	100%	100%
Cost of sales	30	75	30
Gross profit	70	25	70
Selling, general, and administrative expenses	10	10	10
Research and development	20	0	20
Advertising	20	0	20
Operating profit	20	15	20

Note: Each line item is expressed as a percentage of the company's sales.

For most expenses, comparison to the amount of sales is appropriate. However, in the case of taxes, it is more meaningful to compare the amount of taxes with the amount of pretax income. Using note disclosure, an analyst can then examine the causes for differences in effective tax rates. To project the companies' future net income, an analyst would project the companies' pretax income and apply an estimated effective tax rate determined in part by the historical tax rates.

Vertical common-size analysis of the income statement is particularly useful in cross-sectional analysis—comparing companies with each other for a particular time period or comparing a company with industry or sector data. The analyst could select individual peer companies for comparison, use industry data from published sources, or compile data from databases based on a selection of peer companies or broader industry data. For example, Exhibit 16 presents median common-size income statement data compiled for the components of the S&P 500 classified into the 10 S&P/MSCI Global Industrial Classification System (GICS) sectors using 2017 data. Note that when compiling aggregate data such as this, some level of aggregation is necessary and less detail may be available than from peer company financial statements. The performance of an individual company can be compared with industry or peer company data to evaluate its relative performance.

Exhibit 16 Median Common-Size Income Statement Statistics for the S&P 500 Classified by S&P/MSCI GICS Sector Data for 2017

	Energy	Materials	Industrials	Consumer Discretionary	Consumer Staples	Health Care
Number of observations	34	27	69	81	34	59
Gross Margin	37.7%	33.0%	36.8%	37.6%	43.4%	59.0%
Operating Margin	6.4%	14.9%	13.5%	11.0%	17.2%	17.4%
Net Profit Margin	4.9%	9.9%	8.8%	6.0%	10.9%	7.2%

Exhibit 16 (Continued)

	Financials	Information Technology	Telecommunication Services	Utilities	Real Estate
Number of observations	63	64	4	29	29
Gross Margin	40.5%	62.4%	56.4%	34.3%	39.8%
Operating Margin	36.5%	21.1%	15.4%	21.7%	30.1%
Net Profit Margin	18.5%	11.3%	13.1%	10.1%	21.3%

Source: Based on data from Compustat. Operating margin based on EBIT (earnings before interest and taxes.)

INCOME STATEMENT RATIOS

16

j evaluate a company's financial performance using common-size income statements and financial ratios based on the income statement;

One aspect of financial performance is profitability. One indicator of profitability is **net profit margin**, also known as **profit margin** and **return on sales**, which is calculated as net income divided by revenue (or sales).[34]

$$\text{Net profit margin} = \frac{\text{Net income}}{\text{Revenue}}$$

Net profit margin measures the amount of income that a company was able to generate for each dollar of revenue. A higher level of net profit margin indicates higher profitability and is thus more desirable. Net profit margin can also be found directly on the common-size income statements.

For AB InBev, net profit margin based on continuing operations for 2017 was 16.2 percent (calculated as profit from continuing operations of $9,155 million, divided by revenue of $56,444 million). To judge this ratio, some comparison is needed. AB InBev's profitability can be compared with that of another company or with its own previous performance. Compared with previous years, AB InBev's profitability is higher than in 2016 but lower than 2015. In 2016, net profit margin based on continuing operations was 6.0 percent, and in 2015, it was 22.9 percent.

Another measure of profitability is the gross profit margin. Gross profit (gross margin) is calculated as revenue minus cost of goods sold, and the **gross profit margin** is calculated as the gross profit divided by revenue.

$$\text{Gross profit margin} = \frac{\text{Gross profit}}{\text{Revenue}}$$

The gross profit margin measures the amount of gross profit that a company generated for each dollar of revenue. A higher level of gross profit margin indicates higher profitability and thus is generally more desirable, although differences in gross profit margins across companies reflect differences in companies' strategies. For example, consider a company pursuing a strategy of selling a differentiated product (e.g., a product differentiated based on brand name, quality, superior technology, or patent protection). The company would likely be able to sell the differentiated product at a higher price than a similar, but undifferentiated, product and, therefore, would likely show a higher gross profit margin than a company selling an undifferentiated product. Although a company selling a differentiated product would likely show a higher gross profit margin, this may take time. In the initial stage of the strategy, the company would likely incur costs to create a differentiated product, such as advertising or research and development, which would not be reflected in the gross margin calculation.

---

[34] In the definition of margin ratios of this type, "sales" is often used interchangeably with "revenue." "Return on sales" has also been used to refer to a class of profitability ratios having revenue in the denominator.

AB InBev's gross profit (shown in Exhibit 1) was $35,058 million in 2017, $27,715 million in 2016, and $26,467 million in 2015. Expressing gross profit as a percentage of revenues, we see that the gross profit margin was 62.1 percent in 2017, 60.9 percent in 2016, and 60.7 percent in 2015. In absolute terms, AB InBev's gross profit was higher in 2016 than in 2015. However, AB InBev's gross profit *margin* was approximately constant between 2015 and 2016.

Exhibit 17 presents a common-size income statement for AB InBev, and highlights certain profitability ratios. The net profit margin and gross profit margin described above are just two of the many subtotals that can be generated from common-size income statements. Other "margins" used by analysts include the **operating profit margin** (profit from operations divided by revenue) and the **pretax margin** (profit before tax divided by revenue).

Exhibit 17 AB InBev's Margins: Abbreviated Common-Size Income Statement

	12 Months Ended December 31					
	2017		2016		2015	
	$	%	$	%	$	%
Revenue	56,444	100.0	45,517	100.0	43,604	100.0
Cost of sales	(21,386)	(37.9)	(17,803)	(39.1)	(17,137)	(39.3)
Gross profit	**35,058**	**62.1**	**27,715**	**60.9**	**26,467**	**60.7**
Distribution expenses	(5,876)	(10.4)	(4,543)	(10.0)	(4,259)	(9.8)
Sales and marketing expenses	(8,382)	(14.9)	(7,745)	(17.0)	(6,913)	(15.9)
Administrative expenses	(3,841)	(6.8)	(2,883)	(6.3)	(2,560)	(5.9)
	Portions omitted					
Profit from operations	**17,152**	**30.4**	**12,882**	**28.3**	**13,904**	**31.9**
Finance cost	(6,885)	(12.2)	(9,382)	(20.6)	(3,142)	(7.2)
Finance income	378	0.7	818	1.8	1,689	3.9
Net finance income/(cost)	(6,507)	(11.5)	(8,564)	(18.8)	(1,453)	(3.3)
Share of result of associates and joint ventures	430	0.8	16	0.0	10	0.0
Profit before tax	**11,076**	**19.6**	**4,334**	**9.5**	**12,461**	**28.6**
Income tax expense	(1,920)	(3.4)	(1,613)	(3.5)	(2,594)	(5.9)
Profit from continuing operations	**9,155**	**16.2**	**2,721**	**6.0**	**9,867**	**22.6**
Profit from discontinued operations	28	0.0	48	0.1	—	—
Profit of the year	**9,183**	**16.3**	**2,769**	**6.1**	**9,867**	**22.6**

The profitability ratios and the common-size income statement yield quick insights about changes in a company's performance. For example, AB InBev's decrease in profitability in 2016 was not driven by a decrease in gross profit margin. Gross profit margin in 2016 was actually slightly higher than in 2015. The company's decrease in profitability in 2016 was driven in part by higher operating expenses and, in particular, by a significant increase in finance costs. The increased finance costs resulted from the 2016 merger with SABMiller. Valued at more than $100 billion, the acquisition was one of the largest in history. The combination of AB InBev and SABMiller also explains the increase in revenue from around $45 billion to over $56 billion. The profitability ratios and the common-size income statement thus serve to highlight areas about which an analyst might wish to gain further understanding.

COMPREHENSIVE INCOME

17

k describe, calculate, and interpret comprehensive income;

l describe other comprehensive income and identify major types of items included in it.

The general expression for net income is revenue minus expenses. There are, however, certain items of revenue and expense that, by accounting convention, are excluded from the net income calculation. To understand how reported shareholders' equity of one period links with reported shareholders' equity of the next period, we must understand these excluded items, known as **other comprehensive income**. Under IFRS, other comprehensive income includes items of income and expense that are "not recognized in profit or loss as required or permitted by other IFRS." **Total comprehensive income** is "the change in equity during a period resulting from transaction and other events, other than those changes resulting from transactions with owners in their capacity as owners."[35]

Under US GAAP, **comprehensive income** is defined as "the change in equity [net assets] of a business enterprise during a period from transactions and other events and circumstances from non-owner sources. It includes all changes in equity during a period except those resulting from investments by owners and distributions to owners."[36] While the wording differs, comprehensive income is conceptually the same under IFRS and US GAAP.

Comprehensive income includes *both* net income and other revenue and expense items that are excluded from the net income calculation (collectively referred to as Other Comprehensive Income). Assume, for example, a company's beginning shareholders' equity is €110 million, its net income for the year is €10 million, its cash dividends for the year are €2 million, and there was no issuance or repurchase of common stock. If the company's actual ending shareholders' equity is €123 million, then €5 million [€123 – (€110 + €10 – €2)] has bypassed the net income calculation by being classified as other comprehensive income. If the company had no other comprehensive income, its ending shareholders' equity would have been €118 million [€110 + €10 – €2].

Four types of items are treated as other comprehensive income under both IFRS and US GAAP. (The specific treatment of some of these items differs between the two sets of standards, but these types of items are common to both.)

- Foreign currency translation adjustments. In consolidating the financial statements of foreign subsidiaries, the effects of translating the subsidiaries' balance sheet assets and liabilities at current exchange rates are included as other comprehensive income.

---

[35] IAS 1, *Presentation of Financial Statements.*

[36] FASB ASC Section 220-10-05 [Comprehensive Income–Overall–Overview and Background].

- Unrealized gains or losses on derivatives contracts accounted for as hedges. Changes in the fair value of derivatives are recorded each period, but certain changes in value are treated as other comprehensive income and thus bypass the income statement.
- Unrealized holding gains and losses on a certain category of investment securities, namely, available-for-sale debt securities under US GAAP and securities designated as "fair value through other comprehensive income" under IFRS. (Note: IFRS, but not US GAAP, also includes a category of equity investments designated at fair value through other comprehensive income.)
- Certain costs of a company's defined benefit post-retirement plans that are not recognized in the current period.

In addition, under IFRS, other comprehensive income includes certain changes in the value of long-lived assets that are measured using the revaluation model rather than the cost model. Also, under IFRS, companies are not permitted to reclassify certain items of other comprehensive income to profit or loss, and companies must present separately the items of other comprehensive income that will and will not be reclassified subsequently to profit or loss.

The third type of item listed above is perhaps the simplest to illustrate. Holding gains on securities arise when a company owns securities over an accounting period, during which time the securities' value increases. Similarly, holding losses on securities arise when a company owns securities over a period during which time the securities' value decreases. If the company has not sold the securities (i.e., has not realized the gain or loss), its holding gain or loss is said to be unrealized. The question is: Should the company exclude unrealized gains and losses from income; reflect these unrealized holding gains and losses in its income statement (i.e., statement of profit and loss); or reflect these unrealized holding gains as other comprehensive income?

According to accounting standards, the answer depends on how the company has categorized the securities. Categorization depends on what the company intends to do with the securities (i.e., the business model for managing the asset) and on the cash flows of the security. Unrealized gains and losses are excluded from income for debt securities that the company intends to hold to maturity. These held-to-maturity debt securities are reported at their amortized cost, so no unrealized gains or losses are reported. For other securities reported at fair value, the unrealized gains or losses are reflected either in the income statement or as other comprehensive income.

Under US GAAP, unrealized gains and losses are reflected in the income statement for: (a) debt securities designated as **trading securities**; and (b) all investments in equity securities (other than investments giving rise to ownership positions that confer significant influence over the investee). The trading securities category pertains to a debt security that is acquired with the intent of selling it rather than holding it to collect the interest and principal payments. Also, under US GAAP, unrealized gains and losses are reflected as other comprehensive income for debt securities designated as **available-for-sale** securities. Available-for-sale debt securities are those not designated as either held-to-maturity or trading.

Under IFRS, unrealized gains and losses are reflected in the income statement for: (a) investments in equity investments, unless the company makes an irrevocable election otherwise; and (b) debt securities, if the securities do not fall into the other measurement categories or if the company makes an irrevocable election to show gains and losses on the income statement. These debt and equity investments are referred to as being measured at *fair value through profit or loss.* Also under IFRS, unrealized gains and losses are reflected as other comprehensive income for: (a) "debt securities held within a business model whose objective is achieved both by collecting contractual cash flows and selling financial assets"; and (b) equity investments for which the company makes an irrevocable election at initial recognition to show gains and losses as part of other comprehensive income. These debt and equity investments are referred to as being measured at *fair value through other comprehensive income*. Accounting for these securities is similar to accounting for US GAAP's available-for-sale debt securities.

Even where unrealized holding gains and losses are excluded from a company's net income (profit and loss), they are *included* in other comprehensive income and thus form a part of a company's comprehensive income.

EXAMPLE 14

Other Comprehensive Income

Assume a company's beginning shareholders' equity is €200 million, its net income for the year is €20 million, its cash dividends for the year are €3 million, and there was no issuance or repurchase of common stock. The company's actual ending shareholders' equity is €227 million.

1 What amount has bypassed the net income calculation by being classified as other comprehensive income?

- **A** €0.
- **B** €7 million.
- **C** €10 million.

2 Which of the following statements *best* describes other comprehensive income?

- **A** Income earned from diverse geographic and segment activities.
- **B** Income that increases stockholders' equity but is not reflected as part of net income.
- **C** Income earned from activities that are not part of the company's ordinary business activities.

Solution to 1:

C is correct. If the company's actual ending shareholders' equity is €227 million, then €10 million [€227– (€200 + €20 – €3)] has bypassed the net income calculation by being classified as other comprehensive income.

Solution to 2:

B is correct. Answers A and C are not correct because they do not specify whether such income is reported as part of net income and shown in the income statement.

EXAMPLE 15

Other Comprehensive Income in Analysis

An analyst is looking at two comparable companies. Company A has a lower price/earnings (P/E) ratio than Company B, and the conclusion that has been suggested is that Company A is undervalued. As part of examining this conclusion, the analyst decides to explore the question: What would the company's P/E look like if total comprehensive income per share—rather than net income per share—were used as the relevant metric?

	Company A	Company B
Price	$35	$30
EPS	$1.60	$0.90
P/E ratio	21.9×	33.3×
Other comprehensive income (loss) $ million	($16.272)	$(1.757)
Shares (millions)	22.6	25.1

Solution:

As shown in the following table, part of the explanation for Company A's lower P/E ratio may be that its significant losses—accounted for as other comprehensive income (OCI)—are not included in the P/E ratio.

	Company A	Company B
Price	$35	$30
EPS	$1.60	$0.90
OCI (loss) $ million	($16.272)	$(1.757)
Shares (millions)	22.6	25.1
OCI (loss) per share	$(0.72)	$(0.07)
Comprehensive EPS = EPS + OCI per share	$ 0.88	$0.83
Price/Comprehensive EPS ratio	39.8×	36.1×

Both IFRS and US GAAP allow companies two alternative presentations. One alternative is to present two statements—a separate income statement and a second statement additionally including other comprehensive income. The other alternative is to present a single statement of other comprehensive income. Particularly in comparing financial statements of two companies, it is relevant to examine significant differences in comprehensive income.

SUMMARY

This reading has presented the elements of income statement analysis. The income statement presents information on the financial results of a company's business activities over a period of time; it communicates how much revenue the company generated during a period and what costs it incurred in connection with generating that revenue. A company's net income and its components (e.g., gross margin, operating earnings, and pretax earnings) are critical inputs into both the equity and credit analysis processes. Equity analysts are interested in earnings because equity markets often reward relatively high- or low-earnings growth companies with above-average or below-average valuations, respectively. Fixed-income analysts examine the components of income statements, past and projected, for information on companies' abilities to make promised payments on their debt over the course of the business cycle. Corporate financial announcements frequently emphasize income statements more than the other financial statements.

Key points to this reading include the following:

- The income statement presents revenue, expenses, and net income.

- The components of the income statement include: revenue; cost of sales; sales, general, and administrative expenses; other operating expenses; non-operating income and expenses; gains and losses; non-recurring items; net income; and EPS.
- An income statement that presents a subtotal for gross profit (revenue minus cost of goods sold) is said to be presented in a multi-step format. One that does not present this subtotal is said to be presented in a single-step format.
- Revenue is recognized in the period it is earned, which may or may not be in the same period as the related cash collection. Recognition of revenue when earned is a fundamental principal of accrual accounting.
- An analyst should identify differences in companies' revenue recognition methods and adjust reported revenue where possible to facilitate comparability. Where the available information does not permit adjustment, an analyst can characterize the revenue recognition as more or less conservative and thus qualitatively assess how differences in policies might affect financial ratios and judgments about profitability.
- As of the beginning of 2018, revenue recognition standards have converged. The core principle of the converged standards is that revenue should be recognized to "depict the transfer of promised goods or services to customers in an amount that reflects the consideration to which the entity expects to be entitled in an exchange for those goods or services."
- To achieve the core principle, the standard describes the application of five steps in recognizing revenue. The standard also specifies the treatment of some related contract costs and disclosure requirements.
- The general principles of expense recognition include a process to match expenses either to revenue (such as, cost of goods sold) or to the time period in which the expenditure occurs (period costs such as, administrative salaries) or to the time period of expected benefits of the expenditures (such as, depreciation).
- In expense recognition, choice of method (i.e., depreciation method and inventory cost method), as well as estimates (i.e., uncollectible accounts, warranty expenses, assets' useful life, and salvage value) affect a company's reported income. An analyst should identify differences in companies' expense recognition methods and adjust reported financial statements where possible to facilitate comparability. Where the available information does not permit adjustment, an analyst can characterize the policies and estimates as more or less conservative and thus qualitatively assess how differences in policies might affect financial ratios and judgments about companies' performance.
- To assess a company's future earnings, it is helpful to separate those prior years' items of income and expense that are likely to continue in the future from those items that are less likely to continue.
- Under IFRS, a company should present additional line items, headings, and subtotals beyond those specified when such presentation is relevant to an understanding of the entity's financial performance. Some items from prior years clearly are not expected to continue in future periods and are separately disclosed on a company's income statement. Under US GAAP, unusual and/or infrequently occurring items, which are material, are presented separately within income from continuing operations.
- Non-operating items are reported separately from operating items on the income statement. Under both IFRS and US GAAP, the income statement reports separately the effect of the disposal of a component operation as a "discontinued" operation.

- Basic EPS is the amount of income available to common shareholders divided by the weighted average number of common shares outstanding over a period. The amount of income available to common shareholders is the amount of net income remaining after preferred dividends (if any) have been paid.
- If a company has a simple capital structure (i.e., one with no potentially dilutive securities), then its basic EPS is equal to its diluted EPS. If, however, a company has dilutive securities, its diluted EPS is lower than its basic EPS.
- Diluted EPS is calculated using the if-converted method for convertible securities and the treasury stock method for options.
- Common-size analysis of the income statement involves stating each line item on the income statement as a percentage of sales. Common-size statements facilitate comparison across time periods and across companies of different sizes.
- Two income-statement-based indicators of profitability are net profit margin and gross profit margin.
- Comprehensive income includes *both* net income and other revenue and expense items that are excluded from the net income calculation.

PRACTICE PROBLEMS

1 Expenses on the income statement may be grouped by:

A nature, but not by function.

B function, but not by nature.

C either function or nature.

2 An example of an expense classification by function is:

A tax expense.

B interest expense.

C cost of goods sold.

3 Denali Limited, a manufacturing company, had the following income statement information:

Revenue	$4,000,000
Cost of goods sold	$3,000,000
Other operating expenses	$500,000
Interest expense	$100,000
Tax expense	$120,000

Denali's gross profit is equal to:

A $280,000.

B $500,000.

C $1,000,000.

4 Under IFRS, income includes increases in economic benefits from:

A increases in liabilities not related to owners' contributions.

B enhancements of assets not related to owners' contributions.

C increases in owners' equity related to owners' contributions.

5 Fairplay had the following information related to the sale of its products during 2009, which was its first year of business:

Revenue	$1,000,000
Returns of goods sold	$100,000
Cash collected	$800,000
Cost of goods sold	$700,000

Under the accrual basis of accounting, how much net revenue would be reported on Fairplay's 2009 income statement?

A $200,000.

B $900,000.

C $1,000,000.

6 Apex Consignment sells items over the internet for individuals on a consignment basis. Apex receives the items from the owner, lists them for sale on the internet, and receives a 25 percent commission for any items sold. Apex collects the full amount from the buyer and pays the net amount after commission to the owner. Unsold items are returned to the owner after 90 days. During 2009, Apex had the following information:

© 2019 CFA Institute. All rights reserved.

- Total sales price of items sold during 2009 on consignment was €2,000,000.
- Total commissions retained by Apex during 2009 for these items was €500,000.

How much revenue should Apex report on its 2009 income statement?

A €500,000.

B €2,000,000.

C €1,500,000.

7 A company previously expensed the incremental costs of obtaining a contract. All else being equal, adopting the May 2014 IASB and FASB converged accounting standards on revenue recognition makes the company's profitability initially appear:

A lower.

B unchanged.

C higher.

8 During 2009, Accent Toys Plc., which began business in October of that year, purchased 10,000 units of a toy at a cost of £10 per unit in October. The toy sold well in October. In anticipation of heavy December sales, Accent purchased 5,000 additional units in November at a cost of £11 per unit. During 2009, Accent sold 12,000 units at a price of £15 per unit. Under the first in, first out (FIFO) method, what is Accent's cost of goods sold for 2009?

A £120,000.

B £122,000.

C £124,000.

9 Using the same information as in Question 8, what would Accent's cost of goods sold be under the weighted average cost method?

A £120,000.

B £122,000.

C £124,000.

10 Which inventory method is least likely to be used under IFRS?

A First in, first out (FIFO).

B Last in, first out (LIFO).

C Weighted average.

11 At the beginning of 2009, Glass Manufacturing purchased a new machine for its assembly line at a cost of $600,000. The machine has an estimated useful life of 10 years and estimated residual value of $50,000. Under the straight-line method, how much depreciation would Glass take in 2010 for financial reporting purposes?

A $55,000.

B $60,000.

C $65,000.

12 Using the same information as in Question 11, how much depreciation would Glass take in 2009 for financial reporting purposes under the double-declining balance method?

A $60,000.

B $110,000.

C $120,000.

13 Which combination of depreciation methods and useful lives is most conservative in the year a depreciable asset is acquired?

A Straight-line depreciation with a short useful life.

B Declining balance depreciation with a long useful life.

C Declining balance depreciation with a short useful life.

14 Under IFRS, a loss from the destruction of property in a fire would most likely be classified as:

A continuing operations.

B discontinued operations.

C other comprehensive income.

15 A company chooses to change an accounting policy. This change requires that, if practical, the company restate its financial statements for:

A all prior periods.

B current and future periods.

C prior periods shown in a report.

16 For 2009, Flamingo Products had net income of $1,000,000. At 1 January 2009, there were 1,000,000 shares outstanding. On 1 July 2009, the company issued 100,000 new shares for $20 per share. The company paid $200,000 in dividends to common shareholders. What is Flamingo's basic earnings per share for 2009?

A $0.80.

B $0.91.

C $0.95.

17 For its fiscal year-end, Calvan Water Corporation (CWC) reported net income of $12 million and a weighted average of 2,000,000 common shares outstanding. The company paid $800,000 in preferred dividends and had 100,000 options outstanding with an average exercise price of $20. CWC's market price over the year averaged $25 per share. CWC's diluted EPS is *closest* to:

A $5.33.

B $5.54.

C $5.94.

18 A company with no debt or convertible securities issued publicly traded common stock three times during the current fiscal year. Under both IFRS and US GAAP, the company's:

A basic EPS equals its diluted EPS.

B capital structure is considered complex at year-end.

C basic EPS is calculated by using a simple average number of shares outstanding.

19 Laurelli Builders (LB) reported the following financial data for year-end 31 December:

Common shares outstanding, 1 January	2,020,000
Common shares issued as stock dividend, 1 June	380,000
Warrants outstanding, 1 January	500,000
Net income	$3,350,000
Preferred stock dividends paid	$430,000
Common stock dividends paid	$240,000

Which statement about the calculation of LB's EPS is *most* accurate?

A LB's basic EPS is \$1.12.

B LB's diluted EPS is equal to or less than its basic EPS.

C The weighted average number of shares outstanding is 2,210,000.

20 Cell Services Inc. (CSI) had 1,000,000 average shares outstanding during all of 2009. During 2009, CSI also had 10,000 options outstanding with exercise prices of \$10 each. The average stock price of CSI during 2009 was \$15. For purposes of computing diluted earnings per share, how many shares would be used in the denominator?

A 1,003,333.

B 1,006,667.

C 1,010,000.

21 For its fiscal year-end, Sublyme Corporation reported net income of \$200 million and a weighted average of 50,000,000 common shares outstanding. There are 2,000,000 convertible preferred shares outstanding that paid an annual dividend of \$5. Each preferred share is convertible into two shares of the common stock. The diluted EPS is *closest to*:

A \$3.52.

B \$3.65.

C \$3.70.

22 When calculating diluted EPS, which of the following securities in the capital structure increases the weighted average number of common shares outstanding without affecting net income available to common shareholders?

A Stock options

B Convertible debt that is dilutive

C Convertible preferred stock that is dilutive

23 Which statement is *most* accurate? A common size income statement:

A restates each line item of the income statement as a percentage of net income.

B allows an analyst to conduct cross-sectional analysis by removing the effect of company size.

C standardizes each line item of the income statement but fails to help an analyst identify differences in companies' strategies.

24 Selected year-end financial statement data for Workhard are shown below.

	\$ millions
Beginning shareholders' equity	475
Ending shareholders' equity	493
Unrealized gain on available-for-sale securities	5
Unrealized loss on derivatives accounted for as hedges	–3
Foreign currency translation gain on consolidation	2
Dividends paid	1
Net income	15

Workhard's comprehensive income for the year:

A is \$18 million.

B is increased by the derivatives accounted for as hedges.

C includes \$4 million in other comprehensive income.

25 When preparing an income statement, which of the following items would *most likely* be classified as other comprehensive income?

A A foreign currency translation adjustment

B An unrealized gain on a security held for trading purposes

C A realized gain on a derivative contract not accounted for as a hedge

SOLUTIONS

1 C is correct. IAS No. 1 states that expenses may be categorized by either nature or function.

2 C is correct. Cost of goods sold is a classification by function. The other two expenses represent classifications by nature.

3 C is correct. Gross margin is revenue minus cost of goods sold. Answer A represents net income and B represents operating income.

4 B is correct. Under IFRS, income includes increases in economic benefits from increases in assets, enhancement of assets, and decreases in liabilities.

5 B is correct. Net revenue is revenue for goods sold during the period less any returns and allowances, or \$1,000,000 minus \$100,000 = \$900,000.

6 A is correct. Apex is not the owner of the goods and should only report its net commission as revenue.

7 C is correct. Under the converged accounting standards, the incremental costs of obtaining a contract and certain costs incurred to fulfill a contract must be capitalized. If a company expensed these incremental costs in the years prior to adopting the converged standards, all else being equal, its profitability will appear higher under the converged standards.

8 B is correct. Under the first in, first out (FIFO) method, the first 10,000 units sold came from the October purchases at £10, and the next 2,000 units sold came from the November purchases at £11.

9 C is correct. Under the weighted average cost method:

October purchases	10,000 units	\$100,000
November purchases	5,000 units	\$55,000
Total	15,000 units	\$155,000

\$155,000/15,000 units = \$10.3333

\$10.3333 × 12,000 units = \$124,000

10 B is correct. The last in, first out (LIFO) method is not permitted under IFRS. The other two methods are permitted.

11 A is correct. Straight-line depreciation would be (\$600,000 – \$50,000)/10, or \$55,000.

12 C is correct. Double-declining balance depreciation would be \$600,000 × 20 percent (twice the straight-line rate). The residual value is not subtracted from the initial book value to calculate depreciation. However, the book value (carrying amount) of the asset will not be reduced below the estimated residual value.

13 C is correct. This would result in the highest amount of depreciation in the first year and hence the lowest amount of net income relative to the other choices.

14 A is correct. A fire may be infrequent, but it would still be part of continuing operations and reported in the profit and loss statement. Discontinued operations relate to a decision to dispose of an operating division.

15 C is correct. If a company changes an accounting policy, the financial statements for all fiscal years shown in a company's financial report are presented, if practical, as if the newly adopted accounting policy had been used throughout the entire period; this retrospective application of the change makes the

financial results of any prior years included in the report comparable. Notes to the financial statements describe the change and explain the justification for the change.

16 C is correct. The weighted average number of shares outstanding for 2009 is 1,050,000. Basic earnings per share would be $1,000,000 divided by 1,050,000, or $0.95.

17 B is correct. The formula to calculate diluted EPS is as follows:

Diluted EPS = (Net income − Preferred dividends)/[Weighted average number of shares outstanding + (New shares that would have been issued at option exercise − Shares that could have been purchased with cash received upon exercise) × (Proportion of year during which the financial instruments were outstanding)].

The underlying assumption is that outstanding options are exercised, and then the proceeds from the issuance of new shares are used to repurchase shares already outstanding:

Proceeds from option exercise = 100,000 × $20 = $2,000,000

Shares repurchased = $2,000,000/$25 = 80,000

The net increase in shares outstanding is thus 100,000 − 80,000 = 20,000. Therefore, the diluted EPS for CWC = ($12,000,000 − $800,000)/2,020,000 = $5.54.

18 A is correct. Basic and diluted EPS are equal for a company with a simple capital structure. A company that issues only common stock, with no financial instruments that are potentially convertible into common stock has a simple capital structure. Basic EPS is calculated using the weighted average number of shares outstanding.

19 B is correct. LB has warrants in its capital structure; if the exercise price is less than the weighted average market price during the year, the effect of their conversion is to increase the weighted average number of common shares outstanding, causing diluted EPS to be lower than basic EPS. If the exercise price is equal to the weighted average market price, the number of shares issued equals the number of shares repurchased. Therefore, the weighted average number of common shares outstanding is not affected and diluted EPS equals basic EPS. If the exercise price is greater than the weighted average market price, the effect of their conversion is anti-dilutive. As such, they are not included in the calculation of basic EPS. LB's basic EPS is $1.22 [= ($3,350,000 − $430,000)/2,400,000]. Stock dividends are treated as having been issued retroactively to the beginning of the period.

20 A is correct. With stock options, the treasury stock method must be used. Under that method, the company would receive $100,000 (10,000 × $10) and would repurchase 6,667 shares ($100,000/$15). The shares for the denominator would be:

Shares outstanding	1,000,000
Options exercises	10,000
Treasury shares purchased	(6,667)
Denominator	1,003,333

21 C is correct.

$$\begin{aligned}\text{Diluted EPS} &= (\text{Net income})/(\text{Weighted average number of shares outstanding} + \text{New common shares that would have been issued at conversion})\\ &= \$200{,}000{,}000/[50{,}000{,}000 + (2{,}000{,}000 \times 2)]\\ &= \$3.70\end{aligned}$$

The diluted EPS assumes that the preferred dividend is not paid and that the shares are converted at the beginning of the period.

22 A is correct. When a company has stock options outstanding, diluted EPS is calculated as if the financial instruments had been exercised and the company had used the proceeds from the exercise to repurchase as many shares possible at the weighted average market price of common stock during the period. As a result, the conversion of stock options increases the number of common shares outstanding but has no effect on net income available to common shareholders. The conversion of convertible debt increases the net income available to common shareholders by the after-tax amount of interest expense saved. The conversion of convertible preferred shares increases the net income available to common shareholders by the amount of preferred dividends paid; the numerator becomes the net income.

23 B is correct. Common size income statements facilitate comparison across time periods (time-series analysis) and across companies (cross-sectional analysis) by stating each line item of the income statement as a percentage of revenue. The relative performance of different companies can be more easily assessed because scaling the numbers removes the effect of size. A common size income statement states each line item on the income statement as a percentage of revenue. The standardization of each line item makes a common size income statement useful for identifying differences in companies' strategies.

24 C is correct. Comprehensive income includes both net income and other comprehensive income.

$$\begin{aligned}\text{Other comprehensive income} &= \text{Unrealized gain on available-for-sale securities} - \text{Unrealized loss on derivatives accounted for as hedges} + \text{Foreign currency translation gain on consolidation}\\ &= \$5\text{ million} - \$3\text{ million} + \$2\text{ million}\\ &= \$4\text{ million}\end{aligned}$$

Alternatively,

$$\text{Comprehensive income} - \text{Net income} = \text{Other comprehensive income}$$

$$\begin{aligned}\text{Comprehensive income} &= (\text{Ending shareholders equity} - \text{Beginning shareholders equity}) + \text{Dividends}\\ &= (\$493\text{ million} - \$475\text{ million}) + \$1\text{ million}\\ &= \$18\text{ million} + \$1\text{ million} = \$19\text{ million}\end{aligned}$$

Net income is $15 million so other comprehensive income is $4 million.

25 A is correct. Other comprehensive income includes items that affect shareholders' equity but are not reflected in the company's income statement. In consolidating the financial statements of foreign subsidiaries, the effects of translating the subsidiaries' balance sheet assets and liabilities at current exchange rates are included as other comprehensive income.

READING

18

Understanding Balance Sheets

by Elaine Henry, PhD, CFA, and Thomas R. Robinson, PhD, CFA

Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). Thomas R. Robinson, PhD, CFA, is at AACSB International (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the elements of the balance sheet: assets, liabilities, and equity;
☐	**b.** describe uses and limitations of the balance sheet in financial analysis;
☐	**c.** describe alternative formats of balance sheet presentation;
☐	**d.** contrast current and non-current assets and current and non-current liabilities;
☐	**e.** describe different types of assets and liabilities and the measurement bases of each;
☐	**f.** describe the components of shareholders' equity;
☐	**g.** demonstrate the conversion of balance sheets to common-size balance sheets and interpret common-size balance sheets;
☐	**h.** calculate and interpret liquidity and solvency ratios.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION AND COMPONENTS OF THE BALANCE SHEET

a describe the elements of the balance sheet: assets, liabilities, and equity

b describe uses and limitations of the balance sheet in financial analysis;

The balance sheet provides information on a company's resources (assets) and its sources of capital (equity and liabilities/debt). This information helps an analyst assess a company's ability to pay for its near-term operating needs, meet future debt obligations, and make distributions to owners. The basic equation underlying the balance sheet is Assets = Liabilities + Equity.

Analysts should be aware that different types of assets and liabilities may be measured differently. For example, some items are measured at historical cost or a variation thereof and others at fair value.[1] An understanding of the measurement issues will facilitate analysis. The balance sheet measurement issues are, of course, closely linked to the revenue and expense recognition issues affecting the income statement. Throughout this reading, we describe and illustrate some of the linkages between the measurement issues affecting the balance sheet and the revenue and expense recognition issues affecting the income statement.

This reading is organized as follows: In Sections 1–3, we describe and give examples of the elements and formats of balance sheets. Sections 4–6 discuss current assets and current liabilities. Sections 7–11 focus on assets, and Section 12 focuses on liabilities. Sections 13–14 describe the components of equity and illustrates the statement of changes in shareholders' equity. Sections 15–16 introduce balance sheet analysis. A summary of the key points and practice problems in the CFA Institute multiple-choice format conclude the reading.

1.1 Components and Format of the Balance Sheet

The **balance sheet** (also called the **statement of financial position** or **statement of financial condition**) discloses what an entity owns (or controls), what it owes, and what the owners' claims are at a specific point in time.[2]

The financial position of a company is described in terms of its basic elements (assets, liabilities, and equity):

- **Assets** (A) are what the company owns (or controls). More formally, assets are resources controlled by the company as a result of past events and from which future economic benefits are expected to flow *to* the entity.
- **Liabilities** (L) are what the company owes. More formally, liabilities represent obligations of a company arising from past events, the settlement of which is expected to result in a future outflow of economic benefits *from* the entity.
- **Equity** (E) represents the owners' residual interest in the company's assets after deducting its liabilities. Commonly known as **shareholders' equity** or **owners' equity**, equity is determined by subtracting the liabilities from the assets of a company, giving rise to the accounting equation: **A − L = E** or **A = L + E**.

---

[1] IFRS and US GAAP define "fair value" as an exit price, i.e., the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants at the measurement date (IFRS 13, FASB ASC Topic 820).

[2] IFRS uses the term "statement of financial position" (IAS 1 *Presentation of Financial Statements*), and US GAAP uses the terms "balance sheet" and "statement of financial position" interchangeably (ASC 210-10-05 [Balance Sheet–Overall–Overview and Background]).

The equation A = L + E is sometimes summarized as follows: The left side of the equation reflects the resources controlled by the company and the right side reflects how those resources were financed. For all financial statement items, an item should only be recognized in the financial statements if it is probable that any future economic benefit associated with the item will flow to or from the entity and if the item has a cost or value that can be measured with reliability.[3]

The balance sheet provides important information about a company's financial condition, but the balance sheet amounts of equity (assets, net of liabilities) should not be viewed as a measure of either the market or intrinsic value of a company's equity for several reasons. First, the balance sheet under current accounting standards is a mixed model with respect to measurement. Some assets and liabilities are measured based on historical cost, sometimes with adjustments, whereas other assets and liabilities are measured based on a fair value, which represents its current value as of the balance sheet date. The measurement bases may have a significant effect on the amount reported. Second, even the items measured at current value reflect the value that was current at the end of the reporting period. The values of those items obviously can change after the balance sheet is prepared. Third, the value of a company is a function of many factors, including future cash flows expected to be generated by the company and current market conditions. Important aspects of a company's ability to generate future cash flows—for example, its reputation and management skills—are not included in its balance sheet.

1.2 Balance Sheet Components

To illustrate the components and formats of balance sheets, we show the major subtotals from two companies' balance sheets. Exhibit 1 and Exhibit 2 are based on the balance sheets of SAP Group and Apple Inc. SAP Group is a leading business software company based in Germany and prepares its financial statements in accordance with IFRS. Apple is a technology manufacturer based in the United States and prepares its financial statements in accordance with US GAAP. For purposes of discussion, Exhibits 1 and 2 show only the main subtotals and totals of these companies' balance sheets. Additional exhibits throughout this reading will expand on these subtotals.

Exhibit 1 SAP Group Consolidated Statements of Financial Position (Excerpt) (in millions of €)

	31 December	
Assets	**2017**	**2016***
Total current assets	11,930	11,564
Total non-current assets	30,567	32,713
Total assets	42,497	44,277
Equity and liabilities		
Total current liabilities	10,210	9,675
Total non-current liabilities	6,747	8,205
Total liabilities	16,957	17,880

(continued)

---

[3] *Conceptual Framework for Financial Reporting (2018).*

Exhibit 1 (Continued)

	31 December	
Assets	**2017**	**2016***
Total equity	25,540	26,397
Equity and liabilities	42,497	44,277

Source: SAP Group 2017 annual report.
Notes: Numbers exactly from the annual report as prepared by the company, which reflects some rounding.
** Numbers are the reclassified numbers from the SAP Group 2017 annual report.*

Exhibit 2 Apple Inc. Consolidated Balance Sheets (Excerpt)* (in millions of $)

Assets	**30 September 2017**	**24 September 2016**
Total current assets	128,645	106,869
[All other assets]	*246,674*	*214,817*
Total assets	375,319	321,686
Liabilities and shareholders' equity		
Total current liabilities	100,814	79,006
[Total non-current liabilities]	*140,458*	*114,431*
Total liabilities	241,272	193,437
Total shareholders' equity	134,047	128,249
Total liabilities and shareholders' equity	375,319	321,686

**Note*: The italicized subtotals presented in this excerpt are not explicitly shown on the face of the financial statement as prepared by the company.
Source: Apple Inc. 2017 annual report (Form 10K).

SAP Group uses the title Statement of Financial Position and Apple uses the title Balance Sheet. Despite their different titles, both statements report the three basic elements: assets, liabilities, and equity. Both companies are reporting on a consolidated basis, i.e., including all their controlled subsidiaries. The numbers in SAP Group's balance sheet are in millions of euro, and the numbers in Apple's balance sheet are in millions of dollars.

Balance sheet information is as of a specific point in time. These exhibits are from the companies' annual financial statements, so the balance sheet information is as of the last day of their respective fiscal years. SAP Group's fiscal year is the same as the calendar year and the balance sheet information is as of 31 December. Apple's fiscal year ends on the last Saturday of September, so the actual date changes from year to year. About every six years, Apple's fiscal year will include 53 weeks rather than 52 weeks. This feature of Apple's fiscal year should be noted, but in general, the extra week is more relevant to evaluating statements spanning a period of time (the income and cash flow statements) rather than the balance sheet which captures information as of a specific point in time.

A company's ability to pay for its short term operating needs relates to the concept of **liquidity**. With respect to a company overall, liquidity refers to the availability of cash to meet those short-term needs. With respect to a particular asset or liability, liquidity refers to its "nearness to cash." A liquid asset is one that can be easily converted into cash in a short period of time at a price close to fair market value. For example, a small holding of an actively traded stock is much more liquid than an investment in an asset such as a commercial real estate property, particularly in a weak property market.

The separate presentation of current and non-current assets and liabilities facilitates analysis of a company's liquidity position (at least as of the end of the fiscal period). Both IFRS and US GAAP require that the balance sheet distinguish between current and non-current assets and between current and non-current liabilities and present these as separate classifications. An exception to this requirement, under IFRS, is that the current and non-current classifications are not required if a liquidity-based presentation provides reliable and more relevant information. Presentations distinguishing between current and non-current elements are shown in Exhibits 1 and 2. Exhibit 3 in Section 3 shows a liquidity-based presentation.

CURRENT AND NON-CURRENT CLASSIFICATION

2

c describe alternative formats of balance sheet presentation

d contrast current and non-current assets and current and non-current liabilities

Assets that are held primarily for the purpose of trading or that are expected to be sold, used up, or otherwise realized in cash within one year or one operating cycle of the business, whichever is greater, after the reporting period are classified as **current assets**. A company's operating cycle is the average amount of time that elapses between acquiring inventory and collecting the cash from sales to customers. (When the entity's normal operating cycle is not clearly identifiable, its duration is assumed to be one year.) For a manufacturer, the operating cycle is the average amount of time between acquiring raw materials and converting these into cash from a sale. Examples of companies that might be expected to have operating cycles longer than one year include those operating in the tobacco, distillery, and lumber industries. Even though these types of companies often hold inventories longer than one year, the inventory is classified as a current asset because it is expected to be sold within an operating cycle. Assets not expected to be sold or used up within one year or one operating cycle of the business, whichever is greater, are classified as **non-current assets** (long-term, long-lived assets).

Current assets are generally maintained for operating purposes, and these assets include—in addition to cash—items expected to be converted into cash (e.g., trade receivables), used up (e.g., office supplies, prepaid expenses), or sold (e.g., inventories) in the current operating cycle. Current assets provide information about the operating activities and the operating capability of the entity. For example, the item "trade receivables" or "accounts receivable" would indicate that a company provides credit to its customers. Non-current assets represent the infrastructure from which the entity operates and are not consumed or sold in the current period. Investments in such assets are made from a strategic and longer term perspective.

Similarly, liabilities expected to be settled within one year or within one operating cycle of the business, whichever is greater, after the reporting period are classified as **current liabilities**. The specific criteria for classification of a liability as current include the following:

- It is expected to be settled in the entity's normal operating cycle;

- It is held primarily for the purpose of being traded;[4]
- It is due to be settled within one year after the balance sheet date; or
- The entity does not have an unconditional right to defer settlement of the liability for at least one year after the balance sheet date.[5]

IFRS specify that some current liabilities, such as trade payables and some accruals for employee and other operating costs, are part of the working capital used in the entity's normal operating cycle. Such operating items are classified as current liabilities even if they will be settled more than one year after the balance sheet date. All other liabilities are classified as **non-current liabilities**. Non-current liabilities include financial liabilities that provide financing on a long-term basis.

The excess of current assets over current liabilities is called **working capital**. The level of working capital provides analysts with information about the ability of an entity to meet liabilities as they fall due. Although adequate working capital is essential, excessive working capital should be so that funds that could be used more productively elsewhere are not inappropriately tied up.

A balance sheet with separately classified current and non-current assets and liabilities is referred to as a **classified balance sheet**. Classification also refers generally to the grouping of accounts into subcategories. Both companies' balance sheets that are summarized in Exhibits 1 and 2 are classified balance sheets. Although both companies' balance sheets present current assets before non-current assets and current liabilities before non-current liabilities, this is not required. IFRS does not specify the order or format in which a company presents items on a current/non-current classified balance sheet.

3

LIQUIDITY-BASED PRESENTATION

c describe alternative formats of balance sheet presentation

A liquidity-based presentation, rather than a current/non-current presentation, is used when such a presentation provides information that is reliable and more relevant. With a liquidity-based presentation, all assets and liabilities are presented broadly in order of liquidity.

Entities such as banks are candidates to use a liquidity-based presentation. Exhibit 3 presents the assets portion of the balance sheet of HSBC Holdings plc (HSBC), a global financial services company that reports using IFRS. HSBC's balance sheet is ordered using a liquidity-based presentation. As shown, the asset section begins with cash and balances at central banks. Less liquid items such as "Interest in associates and joint ventures" appear near the bottom of the asset listing.

---

[4] Examples of these are financial liabilities classified as held for trading in accordance with IAS 39, which is replaced by IFRS 9 effective for periods beginning on or after 1 January 2018.

[5] IAS 1, *Presentation of Financial Statements*, paragraph 69.

Exhibit 3 HSBC Holdings plc Consolidated Statement of Financial Position (Excerpt: Assets Only) as of 31 December (in millions of US $)

Consolidated balance sheet - USD ($) $ in Millions	Dec. 31, 2017	Dec. 31, 2016
Assets		
Cash and balances at central banks	$180,624	$128,009
Items in the course of collection from other banks	6,628	5,003
Hong Kong Government certificates of indebtedness	34,186	31,228
Trading assets	287,995	235,125
Financial assets designated at fair value	29,464	24,756
Derivatives	219,818	290,872
Loans and advances to banks	90,393	88,126
Loans and advances to customers	962,964	861,504
Reverse repurchase agreements – non-trading	201,553	160,974
Financial investments	389,076	436,797
Prepayments, accrued income and other assets	67,191	63,909
Current tax assets	1,006	1,145
Interests in associates and joint ventures	22,744	20,029
Goodwill and intangible assets	23,453	21,346
Deferred tax assets	4,676	6,163
Total assets	2,521,771	2,374,986

Source: HSBC Holdings plc 2017 Annual Report and Accounts.

CURRENT ASSETS: CASH AND CASH EQUIVALENTS, MARKETABLE SECURITIES AND TRADE RECEIVABLES

4

e describe different types of assets and liabilities and the measurement bases of each

This section examines current assets and current liabilities in greater detail.

4.1 Current Assets

Accounting standards require that certain specific line items, if they are material, must be shown on a balance sheet. Among the current assets' required line items are cash and cash equivalents, trade and other receivables, inventories, and financial assets (with short maturities). Companies present other line items as needed, consistent with the requirements to separately present each material class of similar items. As examples, Exhibit 4 and Exhibit 5 present balance sheet excerpts for SAP Group and Apple Inc. showing the line items for the companies' current assets.

Exhibit 4 SAP Group Consolidated Statements of Financial Position (Excerpt: Current Assets Detail) (in millions of €)

	As of 31 December	
Assets	**2017**	**2016**
Cash and cash equivalents	€4,011	€3,702
Other financial assets	990	1,124
Trade and other receivables	5,899	5,924
Other non-financial assets	725	581
Tax assets	306	233
Total current assets	11,930	11,564
Total non-current assets	30,567	32,713
Total assets	**42,497**	**44,277**
Total current liabilities	10,210	9,674
Total non-current liabilities	6,747	8,205
Total liabilities	16,958	17,880
Total equity	25,540	26,397
Total equity and liabilities	**€42,497**	**€44,277**

Source: SAP Group 2017 annual report.

Exhibit 5 Apple Inc. Consolidated Balance Sheet (Excerpt: Current Assets Detail) * (in millions of $)

Assets	**30 September, 2017**	**24 September, 2016**
Cash and cash equivalents	$20,289	$20,484
Short-term marketable securities	53,892	46,671
Accounts receivable, less allowances of $58 and $53, respectively	17,874	15,754
Inventories	4,855	2,132
Vendor non-trade receivables	17,799	13,545
Other current assets	13,936	8,283
Total current assets	128,645	106,869
[All other assets]	*246,674*	*214,817*
Total assets	**375,319**	**321,686**
Total current liabilities	100,814	79,006
[Total non-current liabilities]	*140,458*	*114,431*
Total liabilities	241,272	193,437

Exhibit 5 (Continued)

Assets	**30 September, 2017**	**24 September, 2016**
Total shareholders' equity	134,047	128,249
Total liabilities and shareholders' equity	**$375,319**	**$321,686**

**Note*: The italicized subtotals presented in this excerpt are not explicitly shown on the face of the financial statement as prepared by the company.
Source: Apple Inc. 2017 annual report (Form 10K).

4.1.1 *Cash and Cash Equivalents*

Cash equivalents are highly liquid, short-term investments that are so close to maturity,[6] the risk is minimal that their value will change significantly with changes in interest rates. Cash and cash equivalents are financial assets. Financial assets, in general, are measured and reported at either **amortised cost** or **fair value**. Amortised cost is the historical cost (initially recognised cost) of the asset adjusted for amortisation and impairment. Under IFRS and US GAAP, fair value is based on an exit price, the price received to sell an asset or paid to transfer a liability in an orderly transaction between two market participants at the measurement date.

For cash and cash equivalents, amortised cost and fair value are likely to be immaterially different. Examples of cash equivalents are demand deposits with banks and highly liquid investments (such as US Treasury bills, commercial paper, and money market funds) with original maturities of three months or less. Cash and cash equivalents excludes amounts that are restricted in use for at least 12 months. For all companies, the Statement of Cash Flows presents information about the changes in cash over a period. For the fiscal year 2017, SAP Group's cash and cash equivalents increased from €3,702 million to €4,011 million, and Apple's cash and cash equivalents decreased from $20,484 million to $20,289 million.

4.1.2 *Marketable Securities*

Marketable securities are also financial assets and include investments in debt or equity securities that are traded in a public market, and whose value can be determined from price information in a public market. Examples of marketable securities include treasury bills, notes, bonds, and equity securities, such as common stocks and mutual fund shares. Companies disclose further detail in the notes to their financial statements about their holdings. For example, SAP Group discloses that its other financial assets consist of items such as time deposits, other receivables, and loans to employees and third parties. These do not fall into marketable securities and thus are more properly treated under trade receivables. Apple's short-term marketable securities, totaling $53.9 billion and $46.7 billion at the end of fiscal 2017 and 2016, respectively, include holdings of US treasuries, corporate securities, commercial paper, and time deposits. Financial assets such as investments in debt and equity securities involve a variety of measurement issues and will be addressed in Section 10.

---

[6] Generally, three months or less.

4.1.3 *Trade Receivables*

Trade receivables, also referred to as accounts receivable, are another type of financial asset. These are amounts owed to a company by its customers for products and services already delivered. They are typically reported at net realizable value, an approximation of fair value, based on estimates of collectability. Several aspects of accounts receivable are usually relevant to an analyst. First, the overall level of accounts receivable relative to sales (a topic to be addressed further in ratio analysis) is important because a significant increase in accounts receivable relative to sales could signal that the company is having problems collecting cash from its customers.

A second relevant aspect of accounts receivable is the allowance for doubtful accounts. The allowance for doubtful accounts reflects the company's estimate of the amount of receivables that will ultimately be uncollectible. Additions to the allowance in a particular period are reflected as bad debt expenses, and the balance of the allowance for doubtful accounts reduces the gross receivables amount to a net amount that is an estimate of net realizable value. When specific receivables are deemed to be uncollectible, they are written off by reducing accounts receivable and the allowance for doubtful accounts. The allowance for doubtful accounts is called a **contra account** because it is netted against (i.e., reduces) the balance of accounts receivable, which is an asset account. SAP Group's balance sheet, for example, reports current net trade and other receivables of €5,899 million as of 31 December 2017. The amount of the allowance for doubtful accounts (€74 million) is disclosed in the notes[7] to the financial statements. Apple discloses the allowance for doubtful accounts on the face of the balance sheet; as of 30 September 2017, the allowance was $58 million. The $17,874 million of accounts receivable on that date is net of the allowance. Apple's disclosures state that the allowance is based on "historical experience, the age of the accounts receivable balances, credit quality of the Company's customers, current economic conditions, and other factors that may affect customers' abilities to pay." The age of an accounts receivable balance refers to the length of time the receivable has been outstanding, including how many days past the due date.

Another relevant aspect of accounts receivable is the concentration of credit risk. For example, SAP Group's annual report discloses that concentration of credit risk is limited because they have a large customer base diversified across various industries, company sizes, and countries. Similarly, Apple's annual report notes that no single customer accounted for 10 percent or more of its revenues. However, Apple's disclosures for 2017 indicate that two customers individually represented 10% or more of its total trade receivables and its cellular network carriers accounted for 59% of trade receivables. Of its vendor non-trade receivables, three vendors represent 42%, 19%, and 10% of the total. [8]

EXAMPLE 1

Analysis of Accounts Receivable

1 Based on the balance sheet excerpt for Apple Inc. in Exhibit 5, what percentage of its total accounts receivable in 2017 and 2016 does Apple estimate will be uncollectible?

---

[7] Note 13 SAP Group 2017 Annual report

[8] Page 53, Apple Inc 2017 10-K

2 In general, how does the amount of allowance for doubtful accounts relate to bad debt expense?

3 In general, what are some factors that could cause a company's allowance for doubtful accounts to decrease?

Solution to 1:

($ millions) The percentage of 2017 accounts receivable estimated to be uncollectible is 0.32 percent, calculated as $58/($17,874 + $58). Note that the $17,874 is net of the $58 allowance, so the gross amount of accounts receivable is determined by adding the allowance to the net amount. The percentage of 2016 accounts receivable estimated to be uncollectible is 0.34 percent [$53/($15,754 + $53)].

Solution to 2:

Bad debt expense is an expense of the period, based on a company's estimate of the percentage of credit sales in the period, for which cash will ultimately not be collected. The allowance for bad debts is a contra asset account, which is netted against the asset accounts receivable.

To record the estimated bad debts, a company recognizes a bad debt expense (which affects net income) and increases the balance in the allowance for doubtful accounts by the same amount. To record the write off of a particular account receivable, a company reduces the balance in the allowance for doubtful accounts and reduces the balance in accounts receivable by the same amount.

Solution to 3:

In general, a decrease in a company's allowance for doubtful accounts in absolute terms could be caused by a decrease in the amount of credit sales.

Some factors that could cause a company's allowance for doubtful accounts to decrease as a percentage of accounts receivable include the following:

- Improvements in the credit quality of the company's existing customers (whether driven by a customer-specific improvement or by an improvement in the overall economy);
- Stricter credit policies (for example, refusing to allow less creditworthy customers to make credit purchases and instead requiring them to pay cash, to provide collateral, or to provide some additional form of financial backing); and/or
- Stricter risk management policies (for example, buying more insurance against potential defaults).

In addition to the business factors noted above, because the allowance is based on management's estimates of collectability, management can potentially bias these estimates to manipulate reported earnings. For example, a management team aiming to increase reported income could intentionally over-estimate collectability and under-estimate the bad debt expense for a period. Conversely, in a period of good earnings, management could under-estimate collectability and over-estimate the bad debt expense with the intent of reversing the bias in a period of poorer earnings.

5 CURRENT ASSETS: INVENTORIES AND OTHER CURRENT ASSETS

e describe different types of assets and liabilities and the measurement bases of each

Inventories are physical products that will eventually be sold to the company's customers, either in their current form (finished goods) or as inputs into a process to manufacture a final product (raw materials and work-in-process). Like any manufacturer, Apple holds inventories. The 2017 balance sheet of Apple Inc. shows $4,855 million of inventories. SAP Group's balance sheet does not include a line item for inventory, consistent with the fact that SAP Group is primarily a software and services provider.

Inventories are measured at the lower of cost and net realizable value (NRV) under IFRS. The cost of inventories comprises all costs of purchase, costs of conversion, and other costs incurred in bringing the inventories to their present location and condition. NRV is the estimated selling price less the estimated costs of completion and costs necessary to complete the sale. NRV is applicable for all inventories under IFRS. Under US GAAP, inventories are also measured at the lower of cost and NRV unless they are measured using the last-in, first-out (LIFO) or retail inventory methods. When using LIFO or the retail inventory methods, inventories are measured at the lower of cost or market value. US GAAP defines market value as current replacement cost but with upper and lower limits; the recorded value cannot exceed NRV and cannot be lower than NRV less a normal profit margin.

If the net realizable value or market value (under US GAAP, in certain cases) of a company's inventory falls below its carrying amount, the company must write down the value of the inventory. The loss in value is reflected in the income statement. For example, within its Management's Discussion and Analysis and notes, Apple indicates that the company reviews its inventory each quarter and records write-downs of inventory that has become obsolete, exceeds anticipated demand, or is carried at a value higher than its market value. Under IFRS, if inventory that was written down in a previous period subsequently increases in value, the amount of the original write-down is reversed. Subsequent reversal of an inventory write-down is not permitted under US GAAP.

When inventory is sold, the cost of that inventory is reported as an expense, "cost of goods sold." Accounting standards allow different valuation methods for determining the amounts that are included in cost of goods sold on the income statement and thus the amounts that are reported in inventory on the balance sheet. (Inventory valuation methods are referred to as cost formulas and cost flow assumptions under IFRS and US GAAP, respectively.) IFRS allows only the first-in, first-out (FIFO), weighted average cost, and specific identification methods. Some accounting standards (such as US GAAP) also allow last-in, first-out (LIFO) as an additional inventory valuation method. The LIFO method is not allowed under IFRS.

5.1 Other Current Assets

The amounts shown in "other current assets" reflect items that are individually not material enough to require a separate line item on the balance sheet and so are aggregated into a single amount. Companies usually disclose the components of other assets in a note to the financial statements. A typical item included in other current assets is prepaid expenses. **Prepaid expenses** are normal operating expenses that have been paid in advance. Because expenses are recognized in the period in which they are incurred—and not necessarily the period in which the payment is made—the advance payment of a future expense creates an asset. The asset (prepaid expenses) will be recognized as an expense in future periods as it is used up. For example, consider prepaid insurance. Assume a company pays its insurance premium for coverage over the next calendar year on 31 December of the current year. At the time of the payment, the company recognizes an asset (prepaid insurance expense). The expense is not incurred at that date; the expense is incurred as time passes (in this example, one-twelfth, 1/12, in each following month). Therefore, the expense is recognized and the value of the asset is reduced in the financial statements over the course of the year.

SAP's notes to the financial statements disclose components of the amount shown as other non-financial assets on the balance sheet. The largest portion pertains to prepaid expenses, primarily prepayments for operating leases, support services, and software royalties. Apple's notes do not disclose components of other current assets.

CURRENT LIABILITIES

6

e describe different types of assets and liabilities and the measurement bases of each

Current liabilities are those liabilities that are expected to be settled in the entity's normal operating cycle, held primarily for trading, or due to be settled within 12 months after the balance sheet date. Exhibit 6 and Exhibit 7 present balance sheet excerpts for SAP Group and Apple Inc. showing the line items for the companies' current liabilities. Some of the common types of current liabilities, including trade payables, financial liabilities, accrued expenses, and deferred income, are discussed below.

Exhibit 6 SAP Group Consolidated Statements of Financial Position (Excerpt: Current Liabilities Detail) (in millions of €)

	As of 31 December	
	2017	**2016**
Assets		
Total current assets	11,930	11,564
Total non-current assets	30,567	32,713
Total assets	42,497	44,277
Equity and liabilities		
Trade and other payables	1,151	1,281
Tax liabilities	597	316
Financial liabilities	1,561	1,813
Other non-financial liabilities	3,946	3,699
Provisions	184	183
Deferred income	2,771	2,383
Total current liabilities	10,210	9,674
Total non-current liabilities	6,747	8,205
Total liabilities	16,958	17,880
Total equity	25,540	26,397
Total equity and liabilities	€42,497	€44,277

Source: SAP Group 2017 annual report.

Exhibit 7 Apple Inc. Consolidated Balance Sheet (Excerpt: Current Liabilities Detail)* (in millions of $)

Assets	30 September 2017	24 September 2016
Total current assets	128,645	106,869
[All other assets]	*246,674*	*214,817*
Total assets	375,319	321,686
Liabilities and shareholders' equity		
Accounts payable	49,049	37,294
Accrued expenses	25,744	22,027
Deferred revenue	7,548	8,080
Commercial paper	11,977	8,105
Current portion of long-term debt	6,496	3,500
Total current liabilities	100,814	79,006
[Total non-current liabilities]	*140,458*	*114,431*
Total liabilities	241,272	193,437
Total shareholders' equity	134,047	128,249
Total liabilities and shareholders' equity	375,319	321,686

**Note*: The italicized subtotals presented in this excerpt are not explicitly shown on the face of the financial statement as prepared by the company.
Source: Apple Inc. 2017 annual report (Form 10K).

Trade payables, also called **accounts payable**, are amounts that a company owes its vendors for purchases of goods and services. In other words, these represent the unpaid amount as of the balance sheet date of the company's purchases on credit. An issue relevant to analysts is the trend in overall levels of trade payables relative to purchases (a topic to be addressed further in ratio analysis). Significant changes in accounts payable relative to purchases could signal potential changes in the company's credit relationships with its suppliers. The general term "trade credit" refers to credit provided to a company by its vendors. Trade credit is a source of financing that allows the company to make purchases and then pay for those purchases at a later date.

Financial liabilities that are due within one year or the operating cycle, whichever is longer, appear in the current liability section of the balance sheet. Financial liabilities include borrowings such as bank loans, notes payable (which refers to financial liabilities owed by a company to creditors, including trade creditors and banks, through a formal loan agreement), and commercial paper. In addition, any portions of long-term liabilities that are due within one year (i.e., the current portion of long-term liabilities) are also shown in the current liability section of the balance sheet. According to its footnote disclosures, most of SAP's €1,561 million of current financial liabilities is for bonds payable due in the next year. Apple shows $11,977 million of commercial paper borrowing (short-term promissory notes issued by companies) and $6,496 million of long-term debt due within the next year.

Accrued expenses (also called accrued expenses payable, accrued liabilities, and other non-financial liabilities) are expenses that have been recognized on a company's income statement but not yet been paid as of the balance sheet date. For example, SAP's 2017 balance sheet shows €597 million of tax liabilities. In addition to income taxes payable, other common examples of accrued expenses are accrued interest payable, accrued warranty costs, and accrued employee compensation (i.e., wages payable). SAP's notes disclose that the €3,946 million line item of other non-financial liabilities in 2017, for example, includes €2,565 million of employee-related liabilities.

Deferred income (also called **deferred revenue** or **unearned revenue**) arises when a company receives payment in advance of delivery of the goods and services associated with the payment. The company has an obligation either to provide the goods or services or to return the cash received. Examples include lease payments received at the beginning of a lease, fees for servicing office equipment received at the beginning of the service period, and payments for magazine subscriptions received at the beginning of the subscription period. SAP's balance sheet shows deferred income of €2,771 million at the end of 2017, up slightly from €2,383 million at the end of 2016. Apple's balance sheet shows deferred revenue of $7,548 million at the end of fiscal 2017, down slightly from $8,080 million at the end of fiscal 2016. Example 3 presents each company's disclosures about deferred revenue and discusses some of the implications.

EXAMPLE 2

Analysis of Deferred Revenue

In the notes to its 2017 financial statements, SAP describes its deferred income as follows:

> Deferred income consists mainly of prepayments made by our customers for cloud subscriptions and support; software support and services; fees from multiple-element arrangements allocated to undelivered elements; and amounts … for obligations to perform under acquired customer contracts in connection with acquisitions.

Apple's deferred revenue also arises from sales involving multiple elements, some delivered at the time of sale and others to be delivered in the future. In addition, Apple recognizes deferred revenue in connection with sales of gifts cards as well as service contracts. In the notes to its 2017 financial statements, Apple describes its deferred revenue as follows:

> The Company records deferred revenue when it receives payments in advance of the delivery of products or the performance of services. This includes amounts that have been deferred for unspecified and specified software upgrade rights and non-software services that are attached to hardware and software products. The Company sells gift cards redeemable at its retail and online stores ... The Company records deferred revenue upon the sale of the card, which is relieved upon redemption of the card by the customer. Revenue from AppleCare service and support contracts is deferred and recognized over the service coverage periods. AppleCare service and support contracts typically include extended phone support, repair services, web-based support resources and diagnostic tools offered under the Company's standard limited warranty.

1 In general, in the period a transaction occurs, how would a company's balance sheet reflect $100 of deferred revenue resulting from a sale? (Assume, for simplicity, that the company receives cash for all sales, the company's income tax payable is 30 percent based on cash receipts, and the company pays cash for all relevant income tax obligations as they arise. Ignore any associated deferred costs.)

2 In general, how does deferred revenue impact a company's financial statements in the periods following its initial recognition?

3 Interpret the amounts shown by SAP as deferred income and by Apple as deferred revenue.

4 Both accounts payable and deferred revenue are classified as current liabilities. Discuss the following statements:

  A When assessing a company's liquidity, the implication of amounts in accounts payable differs from the implication of amounts in deferred revenue.

  B Some investors monitor amounts in deferred revenue as an indicator of future revenue growth.

Solution to 1:

In the period that deferred revenue arises, the company would record a $100 increase in the asset Cash and a $100 increase in the liability Deferred Revenues. In addition, because the company's income tax payable is based on cash receipts and is paid in the current period, the company would record a $30 decrease in the asset Cash and a $30 increase in the asset Deferred Tax Assets. Deferred tax assets increase because the company has paid taxes on revenue it has not yet recognized for accounting purposes. In effect, the company has prepaid taxes from an accounting perspective.

Solution to 2:

In subsequent periods, the company will recognize the deferred revenue as it is earned. When the revenue is recognized, the liability Deferred Revenue will decrease. In addition, the tax expense is recognized on the income statement as the revenue is recognized and thus the associated amounts of Deferred Tax Assets will decrease.

Solution to 3:

The deferred income on SAP's balance sheet and deferred revenue on Apple's balance sheet at the end of their respective 2017 fiscal years will be recognized as revenue, sales, or a similar item in income statements subsequent to the 2017 fiscal year, as the goods or services are provided or the obligation is reduced. The costs of delivering the goods or services will also be recognised.

Solution to 4A:

The amount of accounts payable represents a future obligation to pay cash to suppliers. In contrast, the amount of deferred revenue represents payments that the company has already received from its customers, and the future obligation is to deliver the related services. With respect to liquidity, settling accounts payable will require cash outflows whereas settling deferred revenue obligations will not.

Solution to 4B:

Some investors monitor amounts in deferred revenue as an indicator of future growth because the amounts in deferred revenue will be recognized as revenue in the future. Thus, growth in the amount of deferred revenue implies future growth of that component of a company's revenue.

NON-CURRENT ASSETS: PROPERTY, PLANT AND EQUIPMENT AND INVESTMENT PROPERTY

7

e describe different types of assets and liabilities and the measurement bases of each

This section provides an overview of assets other than current assets, sometimes collectively referred to as non-current, long-term, or long-lived assets. The categories discussed are property, plant, and equipment; investment property; intangible assets; goodwill; financial assets; and deferred tax assets. Exhibit 8 and Exhibit 9 present balance sheet excerpts for SAP Group and Apple Inc. showing the line items for the companies' non-current assets.

Exhibit 8 SAP Group Consolidated Statements of Financial Position (Excerpt: Non-Current Assets Detail) (in millions of €)

	As of 31 December	
Assets	**2017**	**2016**
Total current assets	11,930	11,564
Goodwill	21,274	23,311
Intangible assets	2,967	3,786
Property, plant and equipment	2,967	2,580
Other financial assets	1,155	1,358
Trade and other receivables	118	126
Other non-financial assets	621	532
Tax assets	443	450
Deferred tax assets	1,022	571
Total non-current assets	30,567	32,713
Total assets	**42,497**	**44,277**
Total current liabilities	10,210	9,674
Total non-current liabilities	6,747	8,205
Total liabilities	**16,958**	**17,880**
Total equity	**25,540**	**26,397**
Total equity and liabilities	**€42,497**	**€44,277**

Source: SAP Group 2017 annual report.

Exhibit 9 Apple Inc. Consolidated Balance Sheet (Excerpt: Non-Current Assets Detail)* (in millions of $)

Assets	**30 September 2017**	**24 September 2016**
Total current assets	128,645	106,869
Long-term marketable securities	194,714	170,430
Property, plant and equipment, net	33,783	27,010

(continued)

Exhibit 9 (Continued)

Assets	30 September 2017	24 September 2016
Goodwill	5,717	5,414
Acquired intangible assets, net	2,298	3,206
Other non-current assets	10,162	8,757
[All other assets]	*246,674*	*214,817*
Total assets	375,319	321,686
Liabilities and shareholders' equity		
Total current liabilities	100,814	79,006
[Total non-current liabilities]	*140,458*	*114,431*
Total liabilities	241,272	193,437
Total shareholders' equity	134,047	128,249
Total liabilities and shareholders' equity	375,319	321,686

**Note*: The italicized subtotals presented in this excerpt are not explicitly shown on the face of the financial statement as prepared by the company.
Source: Apple Inc. 2017 annual report (Form 10K).

7.1 Property, Plant, and Equipment

Property, plant, and equipment (PPE) are tangible assets that are used in company operations and expected to be used (provide economic benefits) over more than one fiscal period. Examples of tangible assets treated as property, plant, and equipment, include land, buildings, equipment, machinery, furniture, and natural resources such as mineral and petroleum resources. IFRS permits companies to report PPE using either a cost model or a revaluation model.[9] While IFRS permits companies to use the cost model for some classes of assets and the revaluation model for others, the company must apply the same model to all assets within a particular class of assets. US GAAP permits only the cost model for reporting PPE.

Under the cost model, PPE is carried at amortised cost (historical cost less any accumulated depreciation or accumulated depletion, and less any impairment losses). Historical cost generally consists of an asset's purchase price, plus its delivery cost, and any other additional costs incurred to make the asset operable (such as costs to install a machine). Depreciation and depletion refer to the process of allocating (recognizing as an expense) the cost of a long-lived asset over its useful life. Land is not depreciated. Because PPE is presented on the balance sheet net of depreciation and depreciation expense is recognised in the income statement, the choice of depreciation method and the related estimates of useful life and salvage value impact both a company's balance sheet and income statement.

Whereas depreciation is the systematic allocation of cost over an asset's useful life, impairment losses reflect an unanticipated decline in value. Impairment occurs when the asset's recoverable amount is less than its carrying amount, with terms defined as follows under IFRS:[10]

- Recoverable amount: The higher of an asset's fair value less cost to sell, and its value in use.

---

[9] IAS 16, *Property, Plant and Equipment*, paragraphs 29-31.
[10] IAS 36, *Impairment of Assets*, paragraph 6. US GAAP uses a different approach to impairment.

- Fair value less cost to sell: The amount obtainable in a sale of the asset in an arms-length transaction between knowledgeable willing parties, less the costs of the sale.
- Value in use: The present value of the future cash flows expected to be derived from the asset.

When an asset is considered impaired, the company recognizes the impairment loss in the income statement in the period the impairment is identified. Reversals of impairment losses are permitted under IFRS but not under US GAAP.

Under the revaluation model, the reported and carrying value for PPE is the fair value at the date of revaluation less any subsequent accumulated depreciation. Changes in the value of PPE under the revaluation model affect equity directly or profit and loss depending upon the circumstances.

In Exhibits 8 and 9, SAP reports €2,967 million of PPE and Apple reports $33,783 million of PPE at the end of fiscal year 2017. For SAP, PPE represents approximately 7 percent of total assets and for Apple, PPE represents approximately 9 percent of total assets. Both companies disclose in the notes that PPE are generally depreciated over their expected useful lives using the straight-line method.

7.2 Investment Property

Some property is not used in the production of goods or services or for administrative purposes. Instead, it is used to earn rental income or capital appreciation (or both). Under IFRS, such property is considered to be **investment property**.[11] US GAAP does not include a specific definition for investment property. IFRS provides companies with the choice to report investment property using either a cost model or a fair value model. In general, a company must apply its chosen model (cost or fair value) to all of its investment property. The cost model for investment property is identical to the cost model for PPE: In other words, investment property is carried at cost less any accumulated depreciation and any accumulated impairment losses. Under the fair value model, investment property is carried at its fair value. When a company uses the fair value model to measure the value of its investment property, any gain or loss arising from a change in the fair value of the investment property is recognized in profit and loss, i.e., on the income statement, in the period in which it arises.[12]

Neither SAP Group nor Apple disclose ownership of investment property. The types of companies that typically hold investment property are real estate investment companies or property management companies. Entities such as life insurance companies and endowment funds may also hold investment properties as part of their investment portfolio.

NON-CURRENT ASSETS: INTANGIBLE ASSETS

8

e describe different types of assets and liabilities and the measurement bases of each

---

[11] IAS 40, *Investment Property*.

[12] IAS 40, *Investment Property*, paragraph 35.

Intangible assets are identifiable non-monetary assets without physical substance.[13] An identifiable asset can be acquired singly (can be separated from the entity) or is the result of specific contractual or legal rights or privileges. Examples include patents, licenses, and trademarks. The most common asset that is not a separately identifiable asset is accounting goodwill, which arises in business combinations and is discussed further in Section 9.

IFRS allows companies to report intangible assets using either a cost model or a revaluation model. The revaluation model can only be selected when there is an active market for an intangible asset. These measurement models are essentially the same as described for PPE. US GAAP permits only the cost model.

For each intangible asset, a company assesses whether the useful life of the asset is finite or indefinite. Amortisation and impairment principles apply as follows:

- An intangible asset with a finite useful life is amortised on a systematic basis over the best estimate of its useful life, with the amortisation method and useful life estimate reviewed at least annually.
- Impairment principles for an intangible asset with a finite useful life are the same as for PPE.
- An intangible asset with an indefinite useful life is not amortised. Instead, at least annually, the reasonableness of assuming an indefinite useful life for the asset is reviewed and the asset is tested for impairment.

Financial analysts have traditionally viewed the values assigned to intangible assets, particularly goodwill, with caution. Consequently, in assessing financial statements, analysts often exclude the book value assigned to intangibles, reducing net equity by an equal amount and increasing pretax income by any amortisation expense or impairment associated with the intangibles. An arbitrary assignment of zero value to intangibles is not advisable; instead, an analyst should examine each listed intangible and assess whether an adjustment should be made. Note disclosures about intangible assets may provide useful information to the analyst. These disclosures include information about useful lives, amortisation rates and methods, and impairment losses recognised or reversed.

Further, a company may have developed intangible assets internally that can only be recognised in certain circumstances. Companies may also have assets that are never recorded on a balance sheet because they have no physical substance and are non-identifiable. These assets might include management skill, name recognition, a good reputation, and so forth. Such assets are valuable and are, in theory, reflected in the price at which the company's equity securities trade in the market (and the price at which the entirety of the company's equity would be sold in an acquisition transaction). Such assets may be recognised as goodwill if a company is acquired, but are not recognised until an acquisition occurs.

8.1 Identifiable Intangibles

Under IFRS, identifiable intangible assets are recognised on the balance sheet if it is probable that future economic benefits will flow to the company and the cost of the asset can be measured reliably. Examples of identifiable intangible assets include patents, trademarks, copyrights, franchises, licenses, and other rights. Identifiable intangible assets may have been created internally or purchased by a company. Determining the cost of internally created intangible assets can be difficult and subjective. For these reasons, under IFRS and US GAAP, the general requirement is that internally created identifiable intangibles are expensed rather than reported on the balance sheet.

---

[13] IAS 38, *Intangible Assets*, paragraph 8.

IFRS provides that for internally created intangible assets, the company must separately identify the research phase and the development phase.[14] The research phase includes activities that seek new knowledge or products. The development phase occurs after the research phase and includes design or testing of prototypes and models. IFRS require that costs to internally generate intangible assets during the research phase must be expensed on the income statement. Costs incurred in the development stage can be capitalized as intangible assets if certain criteria are met, including technological feasibility, the ability to use or sell the resulting asset, and the ability to complete the project.

US GAAP prohibits the capitalization as an asset of most costs of internally developed intangibles and research and development. All such costs usually must be expensed. Costs related to the following categories are typically expensed under IFRS and US GAAP. They include:

- internally generated brands, mastheads, publishing titles, customer lists, etc.;
- start-up costs;
- training costs;
- administrative and other general overhead costs;
- advertising and promotion;
- relocation and reorganization expenses; and
- redundancy and other termination costs.

Generally, acquired intangible assets are reported as separately identifiable intangibles (as opposed to goodwill) if they arise from contractual rights (such as a licensing agreement), other legal rights (such as patents), or have the ability to be separated and sold (such as a customer list).

EXAMPLE 3

Measuring Intangible Assets

Alpha Inc., a motor vehicle manufacturer, has a research division that worked on the following projects during the year:

Project 1	Research aimed at finding a steering mechanism that does not operate like a conventional steering wheel but reacts to the impulses from a driver's fingers.
Project 2	The design of a prototype welding apparatus that is controlled electronically rather than mechanically. The apparatus has been determined to be technologically feasible, salable, and feasible to produce.

The following is a summary of the expenses of the research division (in thousands of €):

	General	Project 1	Project 2
Material and services	128	935	620
Labor			
• Direct labor	—	630	320

(continued)

---

[14] IAS 38, *Intangible Assets*, paragraphs 51–67.

	General	Project 1	Project 2
Administrative personnel	720	—	—
Design, construction, and testing	270	450	470

Five percent of administrative personnel costs can be attributed to each of Projects 1 and 2. Explain the accounting treatment of Alpha's costs for Projects 1 and 2 under IFRS and US GAAP.

Solution:

Under IFRS, the capitalization of development costs for Projects 1 and 2 would be as follows:

		Amount Capitalized as an Asset (€'000)
Project 1:	Classified as in the research stage, so all costs are recognized as expenses	NIL
Project 2:	Classified as in the development stage, so costs may be capitalized. Note that administrative costs are not capitalized.	(620 + 320 + 470) = 1,410

Under US GAAP, the costs of Projects 1 and 2 are expensed.

As presented in Exhibits 8 and 9, SAP's 2017 balance sheet shows €2,967 million of intangible assets, and Apple's 2017 balance sheet shows acquired intangible assets, net of $2,298 million. SAP's notes disclose the types of intangible assets (software and database licenses, purchased software to be incorporated into its products, customer contracts, and acquired trademark licenses) and notes that all of its purchased intangible assets other than goodwill have finite useful lives and are amortised either based on expected consumption of economic benefits or on a straight-line basis over their estimated useful lives which range from two to 20 years. Apple's notes disclose that its acquired intangible assets consist primarily of patents and licenses, and almost the entire amount represents definite-lived and amortisable assets for which the remaining weighted-average amortisation period is 3.4 years as of 2017.

9

NON-CURRENT ASSETS: GOODWILL

e describe different types of assets and liabilities and the measurement bases of each

When one company acquires another, the purchase price is allocated to all the identifiable assets (tangible and intangible) and liabilities acquired, based on fair value. If the purchase price is greater than the acquirer's interest in the fair value of the identifiable assets and liabilities acquired, the excess amount is recognized as an asset, described as **goodwill**. To understand why an acquirer would pay more to purchase a company than the fair value of the target company's identifiable assets net of liabilities, consider the following three observations. First, as noted, certain items not recognized in a company's own financial statements (e.g., its reputation, established distribution system, trained employees) have value. Second, a target company's expenditures in research and development may not have resulted in a separately identifiable asset that meets the criteria for recognition but nonetheless may have created some value. Third, part of the value of an acquisition may arise from strategic positioning versus a competitor or from perceived synergies. The purchase price might not pertain solely to the separately identifiable assets and liabilities acquired and thus may exceed the value of those net assets due to the acquisition's role in protecting the value of all of the acquirer's existing assets or to cost savings and benefits from combining the companies.

The subject of recognizing goodwill in financial statements has found both proponents and opponents among professionals. The proponents of goodwill recognition assert that goodwill is the present value of excess returns that a company is expected to earn. This group claims that determining the present value of these excess returns is analogous to determining the present value of future cash flows associated with other assets and projects. Opponents of goodwill recognition claim that the prices paid for acquisitions often turn out to be based on unrealistic expectations, thereby leading to future write-offs of goodwill.

Analysts should distinguish between accounting goodwill and economic goodwill. Economic goodwill is based on the economic performance of the entity, whereas accounting goodwill is based on accounting standards and is reported only in the case of acquisitions. Economic goodwill is important to analysts and investors, and it is not necessarily reflected on the balance sheet. Instead, economic goodwill is reflected in the stock price (at least in theory). Some financial statement users believe that goodwill should not be listed on the balance sheet, because it cannot be sold separately from the entity. These financial statement users believe that only assets that can be separately identified and sold should be reflected on the balance sheet. Other financial statement users analyze goodwill and any subsequent impairment charges to assess management's performance on prior acquisitions.

Under both IFRS and US GAAP, accounting goodwill arising from acquisitions is capitalized. Goodwill is not amortised but is tested for impairment annually. If goodwill is deemed to be impaired, an impairment loss is charged against income in the current period. An impairment loss reduces current earnings. An impairment loss also reduces total assets, so some performance measures, such as return on assets (net income divided by average total assets), may actually increase in future periods. An impairment loss is a non-cash item.

Accounting standards' requirements for recognizing goodwill can be summarized by the following steps:

A The total cost to purchase the target company (the acquiree) is determined.

B The acquiree's identifiable assets are measured at fair value. The acquiree's liabilities and contingent liabilities are measured at fair value. The difference between the fair value of identifiable assets and the fair value of the liabilities and contingent liabilities equals the net identifiable assets acquired.

C Goodwill arising from the purchase is the excess of a) the cost to purchase the target company over b) the net identifiable assets acquired. Occasionally, a transaction will involve the purchase of net identifiable assets with a value greater than the cost to purchase. Such a transaction is called a "bargain purchase." Any gain from a bargain purchase is recognized in profit and loss in the period in which it arises.[15]

---

[15] IFRS 3 *Business Combinations* and FASB ASC 805 [Business Combinations].

Companies are also required to disclose information that enables users to evaluate the nature and financial effect of business combinations. The required disclosures include, for example, the acquisition date fair value of the total cost to purchase the target company, the acquisition date amount recognized for each major class of assets and liabilities, and a qualitative description of the factors that make up the goodwill recognized.

Despite the guidance incorporated in accounting standards, analysts should be aware that the estimations of fair value involve considerable management judgment. Values for intangible assets, such as computer software, might not be easily validated when analyzing acquisitions. Management judgment about valuation in turn impacts current and future financial statements because identifiable intangible assets with definite lives are amortised over time. In contrast, neither goodwill nor identifiable intangible assets with indefinite lives are amortised; instead, as noted, both are tested annually for impairment.

The recognition and impairment of goodwill can significantly affect the comparability of financial statements between companies. Therefore, analysts often adjust the companies' financial statements by removing the impact of goodwill. Such adjustments include:

- excluding goodwill from balance sheet data used to compute financial ratios, and
- excluding goodwill impairment losses from income data used to examine operating trends.

In addition, analysts can develop expectations about a company's performance following an acquisition by taking into account the purchase price paid relative to the net assets and earnings prospects of the acquired company. Example 4 provides an historical example of goodwill impairment.

EXAMPLE 4

Goodwill Impairment

Safeway, Inc., is a North American food and drug retailer. On 25 February 2010, Safeway issued a press release that included the following information:

> Safeway Inc. today reported a net loss of $1,609.1 million ($4.06 per diluted share) for the 16-week fourth quarter of 2009. Excluding a non-cash goodwill impairment charge of $1,818.2 million, net of tax ($4.59 per diluted share), net income would have been $209.1 million ($0.53 per diluted share). Net income was $338.0 million ($0.79 per diluted share) for the 17-week fourth quarter of 2008.
>
> In the fourth quarter of 2009, Safeway recorded a non-cash goodwill impairment charge of $1,974.2 million ($1,818.2 million, net of tax). The impairment was due primarily to Safeway's reduced market capitalization and a weak economy.…The goodwill originated from previous acquisitions.
>
> Safeway's balance sheet as of 2 January 2010 showed goodwill of $426.6 million and total assets of $14,963.6 million. The company's balance sheet as of 3 January 2009 showed goodwill of $2,390.2 million and total assets of $17,484.7 million.

1 How significant is this goodwill impairment charge?

2 With reference to acquisition prices, what might this goodwill impairment indicate?

Solution to 1:

The goodwill impairment was more than 80 percent of the total value of goodwill and 11 percent of total assets, so it was clearly significant. (The charge of $1,974.2 million equals 82.6 percent of the $2,390.2 million of goodwill at the beginning of the year and 11.3 percent of the $17,484.7 million total assets at the beginning of the year.)

Solution to 2:

The goodwill had originated from previous acquisitions. The impairment charge implies that the acquired operations are now worth less than the price that was paid for their acquisition.

As presented in Exhibits 8 and 9, SAP's 2017 balance sheet shows €21,274 million of goodwill, and Apple's 2017 balance sheet shows goodwill of $5,717 million. Goodwill represents 50.1 percent of SAP's total assets and only 1.5 percent of Apple's total assets. An analyst may be concerned that goodwill represents such a high proportion of SAP's total assets.

NON-CURRENT ASSETS: FINANCIAL ASSETS

10

e describe different types of assets and liabilities and the measurement bases of each

IFRS define a financial instrument as a contract that gives rise to a financial asset of one entity, and a financial liability or equity instrument of another entity.[16] This section will focus on financial assets such as a company's investments in stocks issued by another company or its investments in the notes, bonds, or other fixed-income instruments issued by another company (or issued by a governmental entity). Financial liabilities such as notes payable and bonds payable issued by the company itself will be discussed in the liability portion of this reading. Some financial instruments may be classified as either an asset or a liability depending on the contractual terms and current market conditions. One example of such a financial instrument is a derivative. **Derivatives** are financial instruments for which the value is derived based on some underlying factor (interest rate, exchange rate, commodity price, security price, or credit rating) and for which little or no initial investment is required.

Financial instruments are generally recognized when the entity becomes a party to the contractual provisions of the instrument. In general, there are two basic alternative ways that financial instruments are measured subsequent to initial acquisition: fair value or amortised cost. Recall that fair value is the price that would be received to sell an asset or paid to transfer a liability in an orderly market transaction.[17] The amortised cost of a financial asset (or liability) is the amount at which it was initially recognized, minus any principal repayments, plus or minus any amortisation of discount or premium, and minus any reduction for impairment.

Under IFRS, financial assets are subsequently measured at amortised cost if the asset's cash flows occur on specified dates and consist solely of principal and interest, and if the business model is to hold the asset to maturity. The concept is similar in US GAAP, where this category of asset is referred to as **held-to-maturity**. An example is an investment in a long-term bond issued by another company or by a government; the value of the bond will fluctuate, for example with interest rate movements, but if the bond is classified as a held-to-maturity investment, it will be measured at amortised cost on the balance sheet of the investing company. Other types of financial assets measured at amortised cost are loans to other companies.

---

[16] IAS 32, *Financial Instruments: Presentation*, paragraph 11.

[17] IFRS 13 *Fair Value Measurement* and US GAAP ASC 820 *Fair Value Measurement*.

Financial assets not measured at amortised cost subsequent to acquisition are measured at fair value as of the reporting date. For financial instruments measured at fair value, there are two basic alternatives in how net changes in fair value are recognized: as profit or loss on the income statement, or as other comprehensive income (loss) which bypasses the income statement. Note that these alternatives refer to *un*realized changes in fair value, i.e., changes in the value of a financial asset that has not been sold and is still owned at the end of the period. Unrealized gains and losses are also referred to as holding period gains and losses. If a financial asset is sold within the period, a gain is realized if the selling price is greater than the carrying value and a loss is realized if the selling price is less than the carrying value. When a financial asset is sold, any realized gain or loss is reported on the income statement.

Under IFRS, financial assets are subsequently measured at fair value through other comprehensive income (i.e., any unrealized holding gains or losses are recognized in other comprehensive income) if the business model's objective involves both collecting contractual cash flows and selling the financial assets. This IFRS category applies specifically to debt investments, namely assets with cash flows occurring on specified dates and consisting solely of principal and interest. However, IFRS also permits equity investments to be measured at fair value through other comprehensive income if, at the time a company buys an equity investment, the company decides to make an irrevocable election to measure the asset in this manner.[18] The concept is similar to the US GAAP investment category **available-for-sale** in which assets are measured at fair value, with any unrealized holding gains or losses recognized in other comprehensive income. However, unlike IFRS, the US GAAP category available-for-sale applies only to debt securities and is not permitted for investments in equity securities.[19]

Under IFRS, financial assets are subsequently measured at fair value through profit or loss (i.e., any unrealized holding gains or losses are recognized in the income statement) if they are not assigned to either of the other two measurement categories described above. In addition, IFRS allows a company to make an irrevocable election at acquisition to measure a financial asset in this category. Under US GAAP, all investments in equity securities (other than investments giving rise to ownership positions that confer significant influence over the investee) are measured at fair value with unrealized holding gains or losses recognized in the income statement. Under US GAAP, debt securities designated as trading securities are also measured at fair value with unrealized holding gains or losses recognized in the income statement. The trading securities category pertains to a debt security that is acquired with the intent of selling it rather than holding it to collect the interest and principal payments.

Exhibit 10 summarizes how various financial assets are classified and measured subsequent to acquisition.

---

[18] IFRS 7 *Financial Instruments: Disclosures*, paragraph 8(h) and IFRS 9 *Financial Instruments*, paragraph 5.7.5.

[19] US GAAP ASU 2016-01 and ASC 32X *Investments.*

Exhibit 10 Measurement of Financial Assets

Measured at Cost or Amortised Cost	Measured at Fair Value through Other Comprehensive Income	Measured at Fair Value through Profit and Loss
■ Debt securities that are to be held to maturity. ■ Loans and notes receivable ■ Unquoted equity instruments (in limited circumstances where the fair value is not reliably measurable, cost may serve as a proxy (estimate) for fair value)	■ "Available-for-sale" debt securities (US GAAP); Debt securities where the business model involves both collecting interest and principal and selling the security (IFRS); ■ Equity investments for which the company irrevocably elects this measurement at acquisition (IFRS only)	■ All equity securities unless the investment gives the investor significant influence (US GAAP only) ■ "Trading" debt securities (US GAAP) ■ Securities not assigned to either of the other two categories, or investments for which the company irrevocably elects this measurement at acquisition (IFRS only)

To illustrate the different accounting treatments of the gains and losses on financial assets, consider an entity that invests €100,000,000 on 1 January 200X in a fixed-income security investment, with a 5 percent coupon paid semi-annually. After six months, the company receives the first coupon payment of €2,500,000. Additionally, market interest rates have declined such that the value of the fixed-income investment has increased by €2,000,000 as of 30 June 200X. Exhibit 11 illustrates how this situation will be portrayed in the balance sheet and income statement (ignoring taxes) of the entity concerned, under each of the following three measurement categories of financial assets: assets held for trading purposes, assets available for sale, and held-to-maturity assets.

Exhibit 11 Accounting for Gains and Losses on Marketable Securities

IFRS Categories	**Measured at Cost or Amortised Cost**	**Measured at Fair Value through Other Comprehensive Income**	**Measured at Fair Value through Profit and Loss**
US GAAP Comparable Categories	***Held to Maturity***	***Available-for-Sale Debt Securities***	***Trading Debt Securities***
Income Statement For period 1 January–30 June 200X			
Interest income	2,500,000	2,500,000	2,500,000
Unrealized gains	—	—	2,000,000
Impact on profit and loss	2,500,000	2,500,000	4,500,000
Balance Sheet As of 30 June 200X			
Assets			
Cash and cash equivalents	2,500,000	2,500,000	2,500,000
Cost of securities	100,000,000	100,000,000	100,000,000

(continued)

Exhibit 11 (Continued)

Balance Sheet As of 30 June 200X			
Unrealized gains on securities	—	2,000,000	2,000,000
	102,500,000	104,500,000	104,500,000
Liabilities			
Equity			
Paid-in capital	100,000,000	100,000,000	100,000,000
Retained earnings	2,500,000	2,500,000	4,500,000
Accumulated other comprehensive income	—	2,000,000	—
	102,500,000	104,500,000	104,500,000

In the case of held-to-maturity securities, the income statement shows only the interest income (which is then reflected in retained earnings of the ending balance sheet). Because the securities are measured at cost rather than fair value, no unrealized gain is recognized. On the balance sheet, the investment asset is shown at its amortised cost of €100,000,000. In the case of securities classified as Measured at Fair Value through Other Comprehensive Income (IFRS) or equivalently as Available-for-sale debt securities (US GAAP), the income statement shows only the interest income (which is then reflected in retained earnings of the balance sheet). The unrealized gain does not appear on the income statement; instead, it would appear on a Statement of Comprehensive Income as Other Comprehensive Income. On the balance sheet, the investment asset is shown at its fair value of €102,000,000. (Exhibit 11 shows the unrealized gain on a separate line solely to highlight the impact of the change in value. In practice, the investments would be shown at their fair value on a single line.) In the case of securities classified as Measured at Fair Value through Profit and Loss (IFRS) or equivalently as trading debt securities (US GAAP), both the interest income and the unrealized gain are included on the income statement and thus reflected in retained earnings on the balance sheet.

In Exhibits 4 and 8, SAP's 2017 balance sheet shows other financial assets of €990 million (current) and €1,155 million (non-current). The company's notes disclose that the largest component of the current financial assets are loans and other financial receivables (€793 million) and the largest component of the non-current financial assets is €827 million of available-for-sale equity investments.

In Exhibits 5 and 9, Apple's 2017 balance sheet shows $53,892 million of short-term marketable securities and $194,714 million of long-term marketable securities. In total, marketable securities represent more than 66 percent of Apple's $375.3 billion in total assets. Marketable securities plus cash and cash equivalents represent around 72 percent of the company's total assets. Apple's notes disclose that most of the company's marketable securities are fixed-income securities issued by the US government or its agencies ($60,237 million) and by other companies including commercial paper ($153,451 million). In accordance with its investment policy, Apple invests in highly rated securities (which the company defines as investment grade) and limits its credit exposure to any one issuer. The company classifies its marketable securities as available for sale and reports them on the balance sheet at fair value. Unrealized gains and losses are reported in other comprehensive income.

NON-CURRENT ASSETS: DEFERRED TAX ASSETS

11

e describe different types of assets and liabilities and the measurement bases of each

Portions of the amounts shown as **deferred tax assets** on SAP's balance sheet represent income taxes incurred prior to the time that the income tax expense will be recognized on the income statement. Deferred tax assets may result when the actual **income tax payable** based on income for tax purposes in a period exceeds the amount of income tax expense based on the reported financial statement income due to temporary timing differences. For example, a company may be required to report certain income for tax purposes in the current period but to defer recognition of that income for financial statement purposes to subsequent periods. In this case, the company will pay income tax as required by tax laws, and the difference between the taxes payable and the tax expense related to the income for which recognition was deferred on the financial statements will be reported as a deferred tax asset. When the income is subsequently recognized on the income statement, the related tax expense is also recognized which will reduce the deferred tax asset.

Also, a company may claim certain expenses for financial statement purposes that it is only allowed to claim in subsequent periods for tax purposes. In this case, as in the previous example, the financial statement income before taxes is less than taxable income. Thus, income taxes payable based on taxable income exceeds income tax expense based on accounting net income before taxes. The difference is expected to reverse in the future when the income reported on the financial statements exceeds the taxable income as a deduction for the expense becomes allowed for tax purposes. Deferred tax assets may also result from carrying forward unused tax losses and credits (these are not temporary timing differences). Deferred tax assets are only to be recognized if there is an expectation that there will be taxable income in the future, against which the temporary difference or carried forward tax losses or credits can be applied to reduce taxes payable.

NON-CURRENT LIABILITIES

12

e describe different types of assets and liabilities and the measurement bases of each

All liabilities that are not classified as current are considered to be non-current or long-term. Exhibits 12 and 13 present balance sheet excerpts for SAP Group and Apple Inc. showing the line items for the companies' non-current liabilities.

Both companies' balance sheets show non-current unearned revenue (deferred income for SAP Group and deferred revenue for Apple). These amounts represent the amounts of unearned revenue relating to goods and services expected to be delivered in periods beyond twelve months following the reporting period. The sections that follow focus on two common types of non-current (long-term) liabilities: long-term financial liabilities and deferred tax liabilities.

Exhibit 12 SAP Group Consolidated Statements of Financial Position (Excerpt: Non-Current Liabilities Detail) (in millions of €)

	as of 31 December	
	2017	**2016**
Assets		
Total current assets	11,930	11,564
Total non-current assets	30,567	32,713
Total assets	42,497	44,277
Total current liabilities	10,210	9,674
Trade and other payables	119	127
Tax liabilities	470	365
Financial liabilities	5,034	6,481
Other non-financial liabilities	503	461
Provisions	303	217
Deferred tax liabilities	240	411
Deferred income	79	143
Total non-current liabilities	6,747	8,205
Total liabilities	16,958	17,880
Total equity	25,540	26,397
Total equity and liabilities	€42,497	€44,277

Source: SAP Group 2017 annual report.

Exhibit 13 Apple Inc. Consolidated Balance Sheet (Excerpt: Non-Current Liabilities Detail)* (in millions of $)

Assets	**30 September 2017**	**24 September 2016**
Total current assets	128,645	106,869
[All other assets]	*246,674*	*214,817*
Total assets	375,319	321,686
Liabilities and shareholders' equity		
Total current liabilities	100,814	79,006
Deferred revenue, non-current	2,836	2,930
Long-term debt	97,207	75,427
Other non-current liabilities	40,415	36,074
[Total non-current liabilities]	*140,458*	*114,431*
Total liabilities	241,272	193,437
Total shareholders' equity	134,047	128,249
Total liabilities and shareholders' equity	375,319	321,686

*Note: The italicized subtotals presented in this excerpt are not explicitly shown on the face of the financial statement as prepared by the company.
Source: Apple Inc. 2017 annual report (Form 10K).

12.1 Long-term Financial Liabilities

Typical long-term financial liabilities include loans (i.e., borrowings from banks) and notes or bonds payable (i.e., fixed-income securities issued to investors). Liabilities such as loans payable and bonds payable are usually reported at amortised cost on the balance sheet. At maturity, the amortised cost of the bond (carrying amount) will be equal to the face value of the bond. For example, if a company issues $10,000,000 of bonds at par, the bonds are reported as a long-term liability of $10 million. The carrying amount (amortised cost) from the date of issue to the date of maturity remains at $10 million. As another example, if a company issues $10,000,000 of bonds at a price of 97.50 (a discount to par), the bonds are reported as a liability of $9,750,000 at issue date. Over the bond's life, the discount of $250,000 is amortised so that the bond will be reported as a liability of $10,000,000 at maturity. Similarly, any bond premium would be amortised for bonds issued at a price in excess of face or par value.

In certain cases, liabilities such as bonds issued by a company are reported at fair value. Those cases include financial liabilities held for trading, derivatives that are a liability to the company, and some non-derivative instruments such as those which are hedged by derivatives.

SAP's balance sheet in Exhibit 12 shows €5,034 million of financial liabilities, and the notes disclose that these liabilities are mostly for bonds payable. Apple's balance sheet shows $97,207 million of long-term debt, and the notes disclose that this debt includes floating- and fixed-rate notes with varying maturities.

12.2 Deferred Tax Liabilities

Deferred tax liabilities result from temporary timing differences between a company's income as reported for tax purposes (taxable income) and income as reported for financial statement purposes (reported income). Deferred tax liabilities result when taxable income and the actual income tax payable in a period based on it is less than the reported financial statement income before taxes and the income tax expense based on it. Deferred tax liabilities are defined as the amounts of income taxes payable in future periods in respect of taxable temporary differences.[20] In contrast, in the previous discussion of unearned revenue, inclusion of revenue in taxable income in an earlier period created a deferred tax asset (essentially prepaid tax).

Deferred tax liabilities typically arise when items of expense are included in taxable income in earlier periods than for financial statement net income. This results in taxable income being less than income before taxes in the earlier periods. As a result, taxes payable based on taxable income are less than income tax expense based on accounting income before taxes. The difference between taxes payable and income tax expense results in a deferred tax liability—for example, when companies use accelerated depreciation methods for tax purposes and straight-line depreciation methods for financial statement purposes. Deferred tax liabilities also arise when items of income are included in taxable income in later periods—for example, when a company's subsidiary has profits that have not yet been distributed and thus have not yet been taxed.

SAP's balance sheet in Exhibit 12 shows €240 million of deferred tax liabilities. Apple's balance sheet in Exhibit 13 does not show a separate line item for deferred tax liabilities, however, note disclosures indicate that most of the $40,415 million of other non-current liabilities reported on Apple's balance sheet represents deferred tax liabilities, which totaled $31,504 million.

---

[20] IAS 12, *Income Taxes*, paragraph 5.

13 COMPONENTS OF EQUITY

f describe the components of shareholders' equity

Equity is the owners' residual claim on a company's assets after subtracting its liabilities.[21] It represents the claim of the owner against the company. Equity includes funds directly invested in the company by the owners, as well as earnings that have been reinvested over time. Equity can also include items of gain or loss that are not recognized on the company's income statement.

13.1 Components of Equity

Six main components typically comprise total owners' equity. The first five components listed below comprise equity attributable to owners of the parent company. The sixth component is the equity attributable to non-controlling interests.

1. *Capital contributed by owners* (or common stock, or issued capital). The amount contributed to the company by owners. Ownership of a corporation is evidenced through the issuance of common shares. Common shares may have a par value (or stated value) or may be issued as no par shares (depending on regulations governing the incorporation). Where par or stated value requirements exist, it must be disclosed in the equity section of the balance sheet. In addition, the number of shares authorized, issued, and outstanding must be disclosed for each class of share issued by the company. The number of authorized shares is the number of shares that may be sold by the company under its articles of incorporation. The number of issued shares refers to those shares that have been sold to investors. The number of outstanding shares consists of the issued shares less treasury shares.
2. *Preferred shares.* Classified as equity or financial liabilities based upon their characteristics rather than legal form. For example, perpetual, non-redeemable preferred shares are classified as equity. In contrast, preferred shares with mandatory redemption at a fixed amount at a future date are classified as financial liabilities. Preferred shares have rights that take precedence over the rights of common shareholders—rights that generally pertain to receipt of dividends and receipt of assets if the company is liquidated.
3. *Treasury shares* (or treasury stock or own shares repurchased). Shares in the company that have been repurchased by the company and are held as treasury shares, rather than being cancelled. The company is able to sell (reissue) these shares. A company may repurchase its shares when management considers the shares undervalued, needs shares to fulfill employees' stock options, or wants to limit the effects of dilution from various employee stock compensation plans. A repurchase of previously issued shares reduces shareholders' equity by the amount of the cost of repurchasing the shares and reduces the number of total shares outstanding. If treasury shares are subsequently reissued, a company does not recognize any gain or loss from the reissuance on the income statement. Treasury shares are non-voting and do not receive any dividends declared by the company.

---

[21] IASB *Conceptual Framework (2018)*, paragraph 4.4 (c) and FASB ASC 505-10-05-3 [Equity–Overview and Background].

4 *Retained earnings.* The cumulative amount of earnings recognized in the company's income statements which have not been paid to the owners of the company as dividends.

5 *Accumulated other comprehensive income* (or other reserves). The cumulative amount of *other* comprehensive income or loss. The term comprehensive income includes both a) net income, which is recognized on the income statement and is reflected in retained earnings, and b) other comprehensive income which is not recognized as part of net income and is reflected in accumulated other comprehensive income.[22]

6 *Noncontrolling interest* (or minority interest). The equity interests of minority shareholders in the subsidiary companies that have been consolidated by the parent (controlling) company but that are not wholly owned by the parent company.

Exhibits 14 and 15 present excerpts of the balance sheets of SAP Group and Apple Inc., respectively, with detailed line items for each company's equity section. SAP's balance sheet indicates that the company has €1,229 million issued capital, and the notes to the financial statements disclose that the company has issued 1,229 million no-par common stock with a nominal value of €1 per share. SAP's balance sheet also indicates that the company has €1,591 million of treasury shares, and the notes to the financial statements disclose that the company holds 35 million of its shares as treasury shares. The line item share premium of €570 million includes amounts from treasury share transactions (and certain other transactions). The amount of retained earnings, €24,794 million, represents the cumulative amount of earnings that the company has recognized in its income statements, net of dividends. SAP's €508 million of "Other components of equity" includes the company's accumulated other comprehensive income. The notes disclose that this is composed of €330 million gains on exchange differences in translation, €157 million gains on remeasuring available-for-sale financial assets, and €21 million gains on cash flow hedges. The balance sheet next presents a subtotal for the amount of equity attributable to the parent company €25,509 million followed by the amount of equity attributable to non-controlling interests €31 million. Total equity includes both equity attributable to the parent company and equity attributable to non-controlling interests.

The equity section of Apple's balance sheet consists of only three line items: common stock, retained earnings, and accumulated other comprehensive income/(loss). Although Apple's balance sheet shows no treasury stock, the company does repurchase its own shares but cancels the repurchased shares rather than holding the shares in treasury. Apple's balance sheet shows that 5,126,201 thousand shares were issued and outstanding at the end of fiscal 2017 and 5,336,166 thousand shares were issued and outstanding at the end of fiscal 2016. Details on the change in shares outstanding is presented on the Statement of Shareholders' Equity in Exhibit 16, which shows that in 2017 Apple repurchased 246,496 thousand shares of its previously issued common stock and issued 36,531 thousand shares to employees.

---

[22] IFRS defines Total comprehensive income as "the change in equity during a period resulting from transactions and other events, other than those changes resulting from transactions with owners in their capacity as owners. (IAS 1, Presentation of Financial Statements, paragraph 7. Similarly, US GAAP defines comprehensive income as "the change in equity [net assets] of a business entity during a period from transactions and other events and circumstances from nonowner sources. It includes all changes in equity during a period except those resulting from investments by owners and distributions to owners." (FASB ASC *Master Glossary*.)

Exhibit 14 SAP Group Consolidated Statements of Financial Position (Excerpt: Equity Detail) (in millions of €)

	as of 31 December	
Assets	**2017**	**2016**
Total current assets	11,930	11,564
Total non-current assets	30,567	32,713
Total assets	42,497	44,277
Total current liabilities	10,210	9,674
Total non-current liabilities	6,747	8,205
Total liabilities	16,958	17,880
Issued capital	1,229	1,229
Share premium	570	599
Retained earnings	24,794	22,302
Other components of equity	508	3,346
Treasury shares	(1,591)	(1,099)
Equity attributable to owners of parent	25,509	26,376
Non-controlling interests	31	21
Total equity	25,540	26,397
Total equity and liabilities	€42,497	€44,277

Source: SAP Group 2017 annual report.

Exhibit 15 Apple Inc. Consolidated Balance Sheet (Excerpt: Equity Detail) (in millions of $) (Number of shares are reflected in thousands)

Assets	**30 September 2017**	**24 September 2016**
Total current assets	128,645	106,869
[All other assets]	*246,674*	*214,817*
Total assets	375,319	321,686
Liabilities and shareholders’ equity		
Total current liabilities	100,814	79,006
[Total non-current liabilities]	*140,458*	*114,431*
Total liabilities	241,272	193,437
Common stock and additional paid-in capital, $0.00001 par value: 12,600,000 shares authorized; 5,126,201 and 5,336,166 shares issued and outstanding, respectively	35,867	31,251
Retained earnings	98,330	96,364
Accumulated other comprehensive income/(loss)	(150)	634

Exhibit 15 (Continued)

Assets	30 September 2017	24 September 2016
Total shareholders' equity	134,047	128,249
Total liabilities and shareholders' equity	375,319	321,686

Source: Apple Inc. 2017 annual report (10K).

STATEMENT OF CHANGES IN EQUITY

14

f describe the components of shareholders' equity

The **statement of changes in equity** (or statement of shareholders' equity) presents information about the increases or decreases in a company's equity over a period. IFRS requires the following information in the statement of changes in equity:

- total comprehensive income for the period;
- the effects of any accounting changes that have been retrospectively applied to previous periods;
- capital transactions with owners and distributions to owners; and
- reconciliation of the carrying amounts of each component of equity at the beginning and end of the year.[23]

Under US GAAP, the requirement as specified by the SEC is for companies to provide an analysis of changes in each component of stockholders' equity that is shown in the balance sheet.[24]

Exhibit 16 presents an excerpt from Apple's Consolidated Statements of Changes in Shareholders' Equity. The excerpt shows only one of the years presented on the actual statement. It begins with the balance as of 24 September 2016 (i.e., the beginning of fiscal 2017) and presents the analysis of changes to 30 September 2017 in each component of equity that is shown on Apple's balance sheet. As noted above, the number of shares outstanding decreased from 5,336,166 thousand to 5,126,201 thousand as the company repurchased 246,496 thousand shares of its common stock and issued 36,531 thousand new shares which reduced the dollar balance of Paid-in Capital and Retained earnings by $913 million and $581 million, respectively. The dollar balance in common stock also increased by $ 4,909 million in connection with share-based compensation. Retained earnings increased by $48,351 million net income, minus $12,803 million dividends, $33,001 million for the share repurchase and $581 million adjustment in connection with the stock issuance. For companies that pay dividends, the amount of dividends are shown separately as a deduction from retained earnings. The statement also provides details on the $784 million change in Apple's Accumulated other comprehensive income. Note that the statement provides a subtotal for total comprehensive income that includes net income and each of the components of other comprehensive income.

---

[23] IAS 1, *Presentation of Financial Statements*, paragraph 106.

[24] FASB ASC 505-10-S99 [Equity–Overall–SEC materials] indicates that a company can present the analysis of changes in stockholders' equity either in the notes or in a separate statement.

Exhibit 16 Excerpt from Apple Inc.'s Consolidated Statements of Changes in Shareholders' Equity (in millions, except share amounts which are reflected in thousands)

	Common Stock and Additional Paid-In Capital		Retained Earnings	Accumulated Other Comprehensive Income/(Loss)	Total Shareholders' Equity
	Shares	Amount			
Balances as of September 24, 2016	5,336,166	31,251	96,364	634	128,249
Net income	—	—	48,351	—	48,351
Other comprehensive income/(loss)	—	—	—	(784)	(784)
Dividends and dividend equivalents declared	—	—	(12,803)	—	(12,803)
Repurchase of common stock	(246,496)	—	(33,001)	—	(33,001)
Share-based compensation	—	4,909	—	—	4,909
Common stock issued, net of shares withheld for employee taxes	36,531	(913)	(581)	—	(1,494)
Tax benefit from equity awards, including transfer pricing adjustments	—	620	—	—	620
Balances as of September 30, 2017	5,126,201	35,867	98,330	(150)	134,047

15 COMMON SIZE ANALYSIS OF BALANCE SHEET

g demonstrate the conversion of balance sheets to common-size balance sheets and interpret common-size balance sheets

This section describes two tools for analyzing the balance sheet: common-size analysis and balance sheet ratios. Analysis of a company's balance sheet can provide insight into the company's liquidity and solvency—as of the balance sheet date—as well as the economic resources the company controls. **Liquidity** refers to a company's ability to meet its short-term financial commitments. Assessments of liquidity focus a company's ability to convert assets to cash and to pay for operating needs. **Solvency** refers to a company's ability to meet its financial obligations over the longer term. Assessments of solvency focus on the company's financial structure and its ability to pay long-term financing obligations.

15.1 Common-Size Analysis of the Balance Sheet

The first technique, vertical common-size analysis, involves stating each balance sheet item as a percentage of total assets.[25] Common-size statements are useful in comparing a company's balance sheet composition over time (time-series analysis) and across companies in the same industry. To illustrate, Panel A of Exhibit 17 presents a balance sheet for three hypothetical companies. Company C, with assets of $9.75 million is much larger than Company A and Company B, each with only $3.25 million in assets. The common-size balance sheet presented in Panel B facilitates a comparison of these different sized companies.

---

[25] As discussed in the curriculum reading on financial statement analysis, another type of common-size analysis, known as "horizontal common-size analysis," states quantities in terms of a selected base-year value. Unless otherwise indicated, text references to "common-size analysis" refer to vertical analysis.

Exhibit 17

Panel A: Balance Sheets for Companies A, B, and C

($ Thousands)	**A**	**B**	**C**
ASSETS			
Current assets			
Cash and cash equivalents	1,000	200	3,000
Short-term marketable securities	900	—	300
Accounts receivable	500	1,050	1,500
Inventory	100	950	300
Total current assets	2,500	2,200	5,100
Property, plant, and equipment, net	750	750	4,650
Intangible assets	—	200	—
Goodwill	—	100	—
Total assets	3,250	3,250	9,750
LIABILITIES AND SHAREHOLDERS' EQUITY			
Current liabilities			
Accounts payable	—	2,500	600
Total current liabilities	—	2,500	600
Long term bonds payable	10	10	9,000
Total liabilities	10	2,510	9,600
Total shareholders' equity	3,240	740	150
Total liabilities and shareholders' equity	3,250	3,250	9,750

Panel B: Common-Size Balance Sheets for Companies A, B, and C

(Percent)	**A**	**B**	**C**
ASSETS			
Current assets			
Cash and cash equivalents	30.8	6.2	30.8
Short-term marketable securities	27.7	0.0	3.1
Accounts receivable	15.4	32.3	15.4
Inventory	3.1	29.2	3.1
Total current assets	76.9	67.7	52.3
Property, plant and equipment, net	23.1	23.1	47.7
Intangible assets	0.0	6.2	0.0
Goodwill	0.0	3.1	0.0
Total assets	100.0	100.0	100.0

(continued)

Exhibit 17 (Continued)

Panel B: Common-Size Balance Sheets for Companies A, B, and C

(Percent)	A	B	C
LIABILITIES AND SHAREHOLDERS' EQUITY			
Current liabilities			
Accounts payable	0.0	76.9	6.2
Total current liabilities	0.0	76.9	6.2
Long term bonds payable	0.3	0.3	92.3
Total liabilities	0.3	77.2	98.5
Total shareholders' equity	99.7	22.8	1.5
Total liabilities and shareholders' equity	100.0	100.0	100.0

Most of the assets of Company A and B are current assets; however, Company A has nearly 60 percent of its total assets in cash and short-term marketable securities while Company B has only 6 percent of its assets in cash. Company A is more liquid than Company B. Company A shows no current liabilities (its current liabilities round to less than $10 thousand), and it has cash on hand of $1.0 million to meet any near-term financial obligations it might have. In contrast, Company B has $2.5 million of current liabilities which exceed its available cash of only $200 thousand. To pay those near-term obligations, Company B will need to collect some of its accounts receivables, sell more inventory, borrow from a bank, and/or raise more long-term capital (e.g., by issuing more bonds or more equity). Company C also appears more liquid than Company B. It holds over 30 percent of its total assets in cash and short-term marketable securities, and its current liabilities are only 6.2 percent of the amount of total assets.

Company C's $3.3 million in cash and short-term marketable securities is substantially more than its current liabilities of $600 thousand. Turning to the question of solvency, however, note that 98.5 percent of Company C's assets are financed with liabilities. If Company C experiences significant fluctuations in cash flows, it may be unable to pay the interest and principal on its long-term bonds. Company A is far more solvent than Company C, with less than one percent of its assets financed with liabilities.

Note that these examples are hypothetical only. Other than general comparisons, little more can be said without further detail. In practice, a wide range of factors affect a company's liquidity management and capital structure. The study of optimal **capital structure** is a fundamental issue addressed in corporate finance. Capital refers to a company's long-term debt and equity financing; capital structure refers to the proportion of debt versus equity financing.

Common-size balance sheets can also highlight differences in companies' strategies. Comparing the asset composition of the companies, Company C has made a greater proportional investment in property, plant, and equipment—possibly because it manufactures more of its products in-house. The presence of goodwill on Company B's balance sheet signifies that it has made one or more acquisitions in the past. In contrast, the lack of goodwill on the balance sheets of Company A and Company C suggests that these two companies may have pursued a strategy of internal growth rather than growth by acquisition. Company A may be in either a start-up or liquidation stage of operations as evidenced by the composition of its balance sheet. It has relatively little inventory and no accounts payable. It either has not yet established trade credit or it is in the process of paying off its obligations in the process of liquidating.

EXAMPLE 5

Common-Size Analysis

Applying common-size analysis to the excerpts of SAP Group's balance sheets presented in Exhibits 4, 6, 8, and 12, answer the following: In 2017 relative to 2016, which of the following line items increased as a percentage of assets?

A Cash and cash equivalents.

B Total current assets.

C Total financial liabilities

D Total deferred income.

Solution:

A, B, and D are correct. The following items increased as a percentage of total assets:

- Cash and cash equivalents increased from 8.4 percent of total assets in 2016 (€3,702 ÷ €44,277) to 9.4 percent in 2017 (€4,011 ÷ €42,497).
- Total current assets increased from 26.1 percent of total assets in 2016 (€11,564 ÷ €44,277) to 28.1 percent in 2017 (€11,930 ÷ €42,497).
- Total deferred income increased from 5.7 percent of total assets in 2016 ((€ 2,383 +€143) ÷ €44,277) to 6.7 percent in 2017 ((€ 2,771 +€79) ÷ €42,497).

Total financial liabilities decreased both in absolute Euro amounts and as a percentage of total assets when compared with the previous year.

Note that some amounts of the company's deferred income and financial liabilities are classified as current liabilities (shown in Exhibit 6) and some amounts are classified as non-current liabilities (shown in Exhibit 12). The total amounts—current and non-current—of deferred income and financial liabilities, therefore, are obtained by summing the amounts in Exhibits 6 and 12.

Overall, aspects of the company's liquidity position are somewhat stronger in 2017 compared to 2016. The company's cash balances as a percentage of total assets increased. While current liabilities increased as a percentage of total assets and total liabilities remained approximately the same percentage, the mix of liabilities shifted. Financial liabilities, which represent future cash outlays, decreased as a percentage of total assets, while deferred revenues, which represent cash received in advance of revenue recognition, increased.

Common-size analysis of the balance sheet is particularly useful in cross-sectional analysis—comparing companies to each other for a particular time period or comparing a company with industry or sector data. The analyst could select individual peer companies for comparison, use industry data from published sources, or compile data from databases. When analyzing a company, many analysts prefer to select the peer companies for comparison or to compile their own industry statistics.

Exhibit 18 presents common-size balance sheet data compiled for the 10 sectors of the S&P 500 using 2017 data. The sector classification follows the S&P/MSCI Global Industrial Classification System (GICS). The exhibit presents mean and median common-size balance sheet data for those companies in the S&P 500 for which 2017 data was available in the Compustat database.[26]

Some interesting general observations can be made from these data:

- Energy and utility companies have the largest amounts of property, plant, and equipment (PPE). Telecommunication services, followed by utilities, have the highest level of long-term debt. Utilities also use some preferred stock.
- Financial companies have the greatest percentage of total liabilities. Financial companies typically have relatively high financial leverage.
- Telecommunications services and utility companies have the lowest level of receivables.
- Inventory levels are highest for consumer discretionary. Materials and consumer staples have the next highest inventories.
- Information technology companies use the least amount of leverage as evidenced by the lowest percentages for long-term debt and total liabilities and highest percentages for common and total equity.

Example 6 discusses an analyst using cross-sectional common-size balance sheet data.

EXAMPLE 6

Cross-Sectional Common-Size Analysis

Jason Lu is comparing two companies in the computer industry to evaluate their relative financial position as reflected on their balance sheets. He has compiled the following vertical common-size data for Apple and Microsoft.

Cross-Sectional Analysis: Consolidated Balance Sheets (as Percent of Total Assets)

	Apple	Microsoft
ASSETS:	**30 September 2017**	**30 June 2017**
Current assets:		
Cash and cash equivalents	5.4	3.2
Short-term marketable securities	14.4	52.0
Accounts receivable	4.8	8.2
Inventories	1.3	0.9
Vendor non-trade receivables	4.7	0.0
Other current assets	3.7	2.0
Total current assets	34.3	66.3
Long-term marketable securities	51.9	2.5
Property, plant and equipment, net	9.0	9.8
Goodwill	1.5	14.6

---

[26] An entry of zero for an item (e.g., current assets) was excluded from the data, except in the case of preferred stock. Note that most financial institutions did not provide current asset or current liability data, so these are reported as not available in the database.

Exhibit 18 Common-Size Balance Sheet Statistics for the S&P 500 Grouped by S&P/MSCI GICS Sector (in percent except No. of Observations; data for 2017)

Panel A. Median Data

	10 Energy	15 Materials	20 Industrials	25 Consumer Discretionary	30 Consumer Staples	35 Health Care	40 Financials	45 Information Technology	50 Telecommunication Services	55 Utilities	60 Real Estate
Number of observation	34	27	68	81	33	59	64	64	4	29	30
Cash and short-term investments	6.8%	6.3%	8.1%	8.3%	4.1%	11.2%	6.2%	22.7%	1.2%	0.7%	1.4%
Receivables	5.8%	8.8%	12.9%	6.8%	6.5%	9.7%	20.4%	9.6%	3.7%	3.6%	2.0%
Inventories	1.6%	8.9%	6.9%	14.9%	9.6%	4.3%	0.0%	1.3%	0.3%	1.7%	0.0%
Total current assets	16.1%	26.0%	30.5%	41.5%	29.1%	31.4%	N.A.	48.7%	8.6%	7.3%	10.8%
PPE	73.3%	36.3%	12.5%	19.8%	17.2%	8.1%	0.9%	6.2%	35.0%	72.0%	33.4%
Intangibles	1.6%	27.9%	33.3%	16.8%	41.9%	37.6%	2.8%	26.4%	49.6%	6.2%	1.0%
Goodwill	*0.7%*	*20.0%*	*28.3%*	*11.3%*	*26.2%*	*22.8%*	*2.2%*	*22.3%*	*26.0%*	*4.8%*	*0.0%*
Accounts payable	5.7%	7.3%	6.2%	8.0%	8.0%	3.1%	27.0%	2.7%	2.5%	3.0%	1.3%
Current liabilities	10.9%	16.5%	22.5%	25.8%	25.0%	16.5%	N.A.	21.2%	11.5%	11.5%	7.1%
LT debt	27.3%	31.4%	28.0%	28.7%	32.3%	24.3%	6.4%	22.9%	46.8%	32.5%	43.4%
Total liabilities	49.3%	64.2%	65.5%	64.9%	63.8%	59.2%	86.7%	59.9%	75.8%	71.8%	53.3%
Common equity	47.3%	33.8%	34.5%	34.7%	36.2%	39.4%	12.6%	39.3%	23.9%	27.7%	40.4%
Preferred stock	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Total equity	47.3%	33.8%	34.5%	34.7%	36.2%	39.4%	13.2%	39.3%	23.9%	28.0%	41.8%

Exhibit 18 (Continued)

Panel B. Mean Data

	10 Energy	15 Materials	20 Industrials	25 Consumer Discretionary	30 Consumer Staples	35 Health Care	40 Financials	45 Information Technology	50 Telecommunication Services	55 Utilities	60 Real Estate
Number of observations	34	27	68	81	33	59	64	64	4	29	30
Cash and short-term investments	6.9%	7.4%	9.2%	12.9%	7.3%	15.4%	11.2%	28.3%	3.6%	1.3%	2.9%
Receivables	6.6%	10.5%	15.2%	9.0%	7.7%	11.2%	31.5%	11.8%	5.0%	3.8%	3.8%
Inventories	3.4%	9.3%	7.8%	18.3%	10.6%	6.3%	3.8%	4.1%	0.3%	1.6%	0.1%
Total current assets	17.7%	28.8%	32.9%	40.6%	27.8%	36.4%	N.A.	49.4%	10.1%	8.6%	16.1%
PPE	68.0%	36.9%	24.5%	25.1%	21.6%	11.2%	2.1%	10.3%	39.0%	69.9%	34.9%
Intangibles	7.8%	26.6%	35.6%	23.0%	43.6%	43.9%	11.4%	31.1%	48.2%	6.8%	10.3%
Goodwill	5.4%	18.4%	26.8%	14.6%	24.6%	27.3%	7.7%	24.5%	25.9%	5.7%	5.7%
Accounts payable	5.9%	8.1%	7.1%	11.8%	9.8%	8.1%	35.9%	5.1%	3.1%	2.9%	2.0%
Current liabilities	11.8%	17.0%	23.0%	26.8%	24.6%	21.2%	N.A.	26.1%	11.9%	11.8%	12.8%
LT debt	28.3%	31.2%	29.4%	31.3%	32.4%	28.5%	10.3%	24.8%	47.5%	35.0%	44.8%
Total liabilities	50.3%	63.4%	67.1%	67.5%	68.3%	60.1%	80.1%	61.8%	77.6%	73.9%	54.5%
Common equity	46.4%	34.2%	32.3%	32.3%	30.9%	38.9%	18.2%	37.5%	22.2%	24.7%	40.2%
Preferred stock	0.0%	0.0%	0.1%	0.0%	0.0%	0.1%	0.4%	0.3%	0.0%	0.3%	2.2%
Total equity	46.4%	34.2%	32.4%	32.3%	30.9%	39.0%	18.5%	37.8%	22.2%	25.0%	42.3%

PPE = Property, plant, and equipment, LT = Long term.
Source: Based on data from Compustat.

(Continued)

	Apple	Microsoft
ASSETS:	**30 September 2017**	**30 June 2017**
Acquired intangible assets, net	0.6	4.2
Other assets	2.7	2.6
Total assets	100.0	100.0
LIABILITIES AND SHAREHOLDERS' EQUITY:		
Current liabilities:		
Accounts payable	13.1	3.1
Short-term debt	3.2	3.8
Current portion of long-term debt	1.7	0.4
Accrued expenses	6.9	2.7
Deferred revenue	2.0	14.1
Other current liabilities	0.0	2.6
Total current liabilities	26.9	26.8
Long-term debt	25.9	31.6
Deferred revenue non-current	0.8	4.3
Other non-current liabilities	10.8	7.3
Total liabilities	64.3	70.0
Commitments and contingencies		
Total shareholders' equity	35.7	30.0
Total liabilities and shareholders' equity	100.0	100.0

Source: Based on data from companies' annual reports.

From this data, Lu learns the following:

- Apple and Microsoft have high levels of cash and short-term marketable securities, consistent with the information technology sector as reported in Exhibit 18. Apple also has a high balance in long-term marketable securities. This may reflect the success of the company's business model, which has generated large operating cash flows in recent years.
- Apple's level of accounts receivable is lower than Microsoft's and lower than the industry average. Further research is necessary to learn the extent to which this is related to Apple's cash sales through its own retail stores. An alternative explanation would be that the company has been selling/factoring receivables to a greater degree than the other companies; however, that explanation is unlikely given Apple's cash position. Additionally, Apple shows vendor non-trade receivables, reflecting arrangements with its contract manufacturers.
- Apple and Microsoft both have low levels of inventory, similar to industry medians as reported in Exhibit 18. Apple uses contract manufacturers and can rely on suppliers to hold inventory until needed. Additionally, in the Management Discussion and Analysis section of their annual report, Apple discloses $38 billion of noncancelable manufacturing purchase obligations, $33 billion of which is due within twelve months. These amounts are not currently recorded as inventory and reflect the use of contract manufacturers to assemble and test some finished products. The use of

purchase commitments and contract manufacturers implies that inventory may be "understated." Microsoft's low level of inventory is consistent with its business mix which is more heavily weighted to software than to hardware.

- Apple and Microsoft have a level of property, plant, and equipment that is relatively close to the sector median as reported in Exhibit 18.
- Apple has a very low amount of goodwill, reflecting its strategy to grow organically rather than through acquisition. Microsoft's level of goodwill, while higher than Apple's, is lower than the industry median and mean. Microsoft made a number of major acquisitions (for example, Nokia in 2014) but subsequently (in 2015) wrote off significant amounts of goodwill as an impairment charge.
- Apple's level of accounts payable is higher than the industry, but given the company's high level of cash and investments, it is unlikely that this is a problem.
- Apple's and Microsoft's levels of long-term debt are slightly higher than industry averages. Again, given the companies' high level of cash and investments, it is unlikely that this is a problem.

16 BALANCE SHEET RATIOS

h calculate and interpret liquidity and solvency ratios

Ratios facilitate time-series and cross-sectional analysis of a company's financial position. **Balance sheet ratios** are those involving balance sheet items only. Each of the line items on a vertical common-size balance sheet is a ratio in that it expresses a balance sheet amount in relation to total assets. Other balance sheet ratios compare one balance sheet item to another. For example, the current ratio expresses current assets in relation to current liabilities as an indicator of a company's liquidity. Balance sheet ratios include **liquidity ratios** (measuring the company's ability to meet its short-term obligations) and **solvency ratios** (measuring the company's ability to meet long-term and other obligations). These ratios and others are discussed in a later reading. Exhibit 19 summarizes the calculation and interpretation of selected balance sheet ratios.

Exhibit 19 Balance Sheet Ratios

Liquidity Ratios	Calculation	Indicates
Current	Current assets ÷ Current liabilities	Ability to meet current liabilities
Quick (acid test)	(Cash + Marketable securities + Receivables) ÷ Current liabilities	Ability to meet current liabilities
Cash	(Cash + Marketable securities) ÷ Current liabilities	Ability to meet current liabilities

Exhibit 19 (Continued)

Solvency Ratios		
Long-term debt-to-equity	Total long-term debt ÷ Total equity	Financial risk and financial leverage
Debt-to-equity	Total debt ÷ Total equity	Financial risk and financial leverage
Total debt	Total debt ÷ Total assets	Financial risk and financial leverage
Financial leverage	Total assets ÷ Total equity	Financial risk and financial leverage

EXAMPLE 7

Ratio Analysis

For the following ratio questions, refer to the balance sheet information for the SAP Group presented in Exhibits 1, 4, 6, 8, and 12.

1 The current ratio for SAP Group at 31 December 2017 is *closest* to:
   A 1.17.
   B 1.20.
   C 2.00.

2 Which of the following liquidity ratios decreased in 2017 relative to 2016?
   A Cash.
   B Quick.
   C Current.

3 Which of the following leverage ratios decreased in 2017 relative to 2016?
   A Debt-to-equity.
   B Financial leverage.
   C Long-term debt-to-equity.

Solution to 1:

A is correct. SAP Group's current ratio (Current assets ÷ Current liabilities) at 31 December 2017 is 1.17 (€11,930 million ÷ €10,210 million).

Solution to 2:

B and C are correct. The ratios are shown in the table below. The quick ratio and current ratio are lower in 2017 than in 2016. The cash ratio is slightly higher.

Liquidity Ratios	Calculation	2017 € in millions	2016 € in millions
Current	Current assets ÷ Current liabilities	€11,930 ÷ €10,210 = **1.17**	€11,564 ÷ €9,674 = **1.20**
Quick (acid test)	(Cash + Marketable securities + Receivables) ÷ Current liabilities	(€4,011 + €990 + €5,899) ÷ €10,210 = **1.07**	(€3,702 + €1,124 + €5,924) ÷ €9,674 = **1.11**
Cash	(Cash + Marketable securities) ÷ Current liabilities	€4,011 ÷ €10,210 = **0.39**	€3,702 ÷ €9,674 = **0.38**

Solution to 3:

A, B, and C are correct. The ratios are shown in the table below. All three leverage ratios decreased in 2017 relative to 2016.

Solvency Ratios			
Long-term debt-to-equity	Total long-term debt ÷ Total equity	€5,034 ÷ €25,540 =**19.7%**	€6,481 ÷ €26,397 = **24.6%**
Debt-to-equity	Total debt ÷ Total equity	(€1,561 + €5,034) ÷ €25,540 = **25.8%**	(€ 1,813 + €6,481) ÷ €26,397 = **31.4%**
Financial Leverage	Total assets ÷ Total equity	€42,497 ÷ €25,540 = **1.66**	€44,277 ÷ €26,397 = **1.68**

Cross-sectional financial ratio analysis can be limited by differences in accounting methods. In addition, lack of homogeneity of a company's operating activities can limit comparability. For diversified companies operating in different industries, using industry-specific ratios for different lines of business can provide better comparisons. Companies disclose information on operating segments. The financial position and performance of the operating segments can be compared to the relevant industry.

Ratio analysis requires a significant amount of judgment. One key area requiring judgment is understanding the limitations of any ratio. The current ratio, for example, is only a rough measure of liquidity at a specific point in time. The ratio captures only the amount of current assets, but the components of current assets differ significantly in their nearness to cash (e.g., marketable securities versus inventory). Another limitation of the current ratio is its sensitivity to end-of-period financing and operating decisions that can potentially impact current asset and current liability amounts. Another overall area requiring judgment is determining whether a ratio for a company is within a reasonable range for an industry. Yet another area requiring judgment is evaluating whether a ratio signifies a persistent condition or reflects only a temporary condition. Overall, evaluating specific ratios requires an examination of the entire operations of a company, its competitors, and the external economic and industry setting in which it is operating.

SUMMARY

The balance sheet (also referred to as the statement of financial position) discloses what an entity owns (assets) and what it owes (liabilities) at a specific point in time. Equity is the owners' residual interest in the assets of a company, net of its liabilities. The amount of equity is increased by income earned during the year, or by the issuance of new equity. The amount of equity is decreased by losses, by dividend payments, or by share repurchases.

An understanding of the balance sheet enables an analyst to evaluate the liquidity, solvency, and overall financial position of a company.

- The balance sheet distinguishes between current and non-current assets and between current and non-current liabilities unless a presentation based on liquidity provides more relevant and reliable information.

- The concept of liquidity relates to a company's ability to pay for its near-term operating needs. With respect to a company overall, liquidity refers to the availability of cash to pay those near-term needs. With respect to a particular asset or liability, liquidity refers to its "nearness to cash."
- Some assets and liabilities are measured on the basis of fair value and some are measured at historical cost. Notes to financial statements provide information that is helpful in assessing the comparability of measurement bases across companies.
- Assets expected to be liquidated or used up within one year or one operating cycle of the business, whichever is greater, are classified as current assets. Assets not expected to be liquidated or used up within one year or one operating cycle of the business, whichever is greater, are classified as non-current assets.
- Liabilities expected to be settled or paid within one year or one operating cycle of the business, whichever is greater, are classified as current liabilities. Liabilities not expected to be settled or paid within one year or one operating cycle of the business, whichever is greater, are classified as non-current liabilities.
- Trade receivables, also referred to as accounts receivable, are amounts owed to a company by its customers for products and services already delivered. Receivables are reported net of the allowance for doubtful accounts.
- Inventories are physical products that will eventually be sold to the company's customers, either in their current form (finished goods) or as inputs into a process to manufacture a final product (raw materials and work-in-process). Inventories are reported at the lower of cost or net realizable value. If the net realizable value of a company's inventory falls below its carrying amount, the company must write down the value of the inventory and record an expense.
- Inventory cost is based on specific identification or estimated using the first-in, first-out or weighted average cost methods. Some accounting standards (including US GAAP but not IFRS) also allow last-in, first-out as an additional inventory valuation method.
- Accounts payable, also called trade payables, are amounts that a business owes its vendors for purchases of goods and services.
- Deferred revenue (also known as unearned revenue) arises when a company receives payment in advance of delivery of the goods and services associated with the payment received.
- Property, plant, and equipment (PPE) are tangible assets that are used in company operations and expected to be used over more than one fiscal period. Examples of tangible assets include land, buildings, equipment, machinery, furniture, and natural resources such as mineral and petroleum resources.
- IFRS provide companies with the choice to report PPE using either a historical cost model or a revaluation model. US GAAP permit only the historical cost model for reporting PPE.
- Depreciation is the process of recognizing the cost of a long-lived asset over its useful life. (Land is not depreciated.)
- Under IFRS, property used to earn rental income or capital appreciation is considered to be an investment property. IFRS provide companies with the choice to report an investment property using either a historical cost model or a fair value model.

- Intangible assets refer to identifiable non-monetary assets without physical substance. Examples include patents, licenses, and trademarks. For each intangible asset, a company assesses whether the useful life is finite or indefinite.
- An intangible asset with a finite useful life is amortised on a systematic basis over the best estimate of its useful life, with the amortisation method and useful-life estimate reviewed at least annually. Impairment principles for an intangible asset with a finite useful life are the same as for PPE.
- An intangible asset with an indefinite useful life is not amortised. Instead, it is tested for impairment at least annually.
- For internally generated intangible assets, IFRS require that costs incurred during the research phase must be expensed. Costs incurred in the development stage can be capitalized as intangible assets if certain criteria are met, including technological feasibility, the ability to use or sell the resulting asset, and the ability to complete the project.
- The most common intangible asset that is not a separately identifiable asset is goodwill, which arises in business combinations. Goodwill is not amortised; instead it is tested for impairment at least annually.
- Financial instruments are contracts that give rise to both a financial asset of one entity and a financial liability or equity instrument of another entity. In general, there are two basic alternative ways that financial instruments are measured: fair value or amortised cost. For financial instruments measured at fair value, there are two basic alternatives in how net changes in fair value are recognized: as profit or loss on the income statement, or as other comprehensive income (loss) which bypasses the income statement.
- Typical long-term financial liabilities include loans (i.e., borrowings from banks) and notes or bonds payable (i.e., fixed-income securities issued to investors). Liabilities such as bonds issued by a company are usually reported at amortised cost on the balance sheet.
- Deferred tax liabilities arise from temporary timing differences between a company's income as reported for tax purposes and income as reported for financial statement purposes.
- Six potential components that comprise the owners' equity section of the balance sheet include: contributed capital, preferred shares, treasury shares, retained earnings, accumulated other comprehensive income, and non-controlling interest.
- The statement of changes in equity reflects information about the increases or decreases in each component of a company's equity over a period.
- Vertical common-size analysis of the balance sheet involves stating each balance sheet item as a percentage of total assets.
- Balance sheet ratios include liquidity ratios (measuring the company's ability to meet its short-term obligations) and solvency ratios (measuring the company's ability to meet long-term and other obligations).

PRACTICE PROBLEMS

1 Resources controlled by a company as a result of past events are:

A equity.

B assets.

C liabilities.

2 Equity equals:

A Assets − Liabilities.

B Liabilities − Assets.

C Assets + Liabilities.

3 Distinguishing between current and non-current items on the balance sheet and presenting a subtotal for current assets and liabilities is referred to as:

A a classified balance sheet.

B an unclassified balance sheet.

C a liquidity-based balance sheet.

4 Shareholders' equity reported on the balance sheet is *most likely* to differ from the market value of shareholders' equity because:

A historical cost basis is used for all assets and liabilities.

B some factors that affect the generation of future cash flows are excluded.

C shareholders' equity reported on the balance sheet is updated continuously.

5 The information provided by a balance sheet item is limited because of uncertainty regarding:

A measurement of its cost or value with reliability.

B the change in current value following the end of the reporting period.

C the probability that any future economic benefit will flow to or from the entity.

6 Which of the following is *most likely* classified as a current liability?

A Payment received for a product due to be delivered at least one year after the balance sheet date

B Payments for merchandise due at least one year after the balance sheet date but still within a normal operating cycle

C Payment on debt due in six months for which the company has the unconditional right to defer settlement for at least one year after the balance sheet date

7 The *most likely* company to use a liquidity-based balance sheet presentation is a:

A bank.

B computer manufacturer holding inventories.

C software company with trade receivables and payables.

8 All of the following are current assets *except*:

A cash.

B goodwill.

C inventories.

© 2019 CFA Institute. All rights reserved.

9 The *most* likely costs included in both the cost of inventory and property, plant, and equipment are:

A selling costs.

B storage costs.

C delivery costs.

10 Debt due within one year is considered:

A current.

B preferred.

C convertible.

11 Money received from customers for products to be delivered in the future is recorded as:

A revenue and an asset.

B an asset and a liability.

C revenue and a liability.

12 An example of a contra asset account is:

A depreciation expense.

B sales returns and allowances.

C allowance for doubtful accounts.

13 The carrying value of inventories reflects:

A their historical cost.

B their current value.

C the lower of historical cost or net realizable value.

14 When a company pays its rent in advance, its balance sheet will reflect a reduction in:

A assets and liabilities.

B assets and shareholders' equity.

C one category of assets and an increase in another.

15 Accrued expenses (accrued liabilities) are:

A expenses that have been paid.

B created when another liability is reduced.

C expenses that have been reported on the income statement but not yet paid.

16 The initial measurement of goodwill is *most likely* affected by:

A an acquisition's purchase price.

B the acquired company's book value.

C the fair value of the acquirer's assets and liabilities.

17 Defining total asset turnover as revenue divided by average total assets, all else equal, impairment write-downs of long-lived assets owned by a company will *most likely* result in an increase for that company in:

A the debt-to-equity ratio but not the total asset turnover.

B the total asset turnover but not the debt-to-equity ratio.

C both the debt-to-equity ratio and the total asset turnover.

18 A company has total liabilities of £35 million and total stockholders' equity of £55 million. Total liabilities are represented on a vertical common-size balance sheet by a percentage *closest* to:

A 35%.

B 39%.

C 64%.

19 For financial assets classified as trading securities, how are unrealized gains and losses reflected in shareholders' equity?

A They are not recognized.

B They flow through income into retained earnings.

C They are a component of accumulated other comprehensive income.

20 For financial assets classified as available for sale, how are unrealized gains and losses reflected in shareholders' equity?

A They are not recognized.

B They flow through retained earnings.

C They are a component of accumulated other comprehensive income.

21 For financial assets classified as held to maturity, how are unrealized gains and losses reflected in shareholders' equity?

A They are not recognized.

B They flow through retained earnings.

C They are a component of accumulated other comprehensive income.

22 The non-controlling (minority) interest in consolidated subsidiaries is presented on the balance sheet:

A as a long-term liability.

B separately, but as a part of shareholders' equity.

C as a mezzanine item between liabilities and shareholders' equity.

23 The item "retained earnings" is a component of:

A assets.

B liabilities.

C shareholders' equity.

24 When a company buys shares of its own stock to be held in treasury, it records a reduction in:

A both assets and liabilities.

B both assets and shareholders' equity.

C assets and an increase in shareholders' equity.

25 Which of the following would an analyst *most likely* be able to determine from a common-size analysis of a company's balance sheet over several periods?

A An increase or decrease in sales.

B An increase or decrease in financial leverage.

C A more efficient or less efficient use of assets.

26 An investor concerned whether a company can meet its near-term obligations is *most likely* to calculate the:

A current ratio.

B return on total capital.

C financial leverage ratio.

27 The most stringent test of a company's liquidity is its:

A cash ratio.

B quick ratio.

C current ratio.

28 An investor worried about a company's long-term solvency would *most likely* examine its:

A current ratio.

B return on equity.

C debt-to-equity ratio.

29 Using the information presented in Exhibit 4 of the reading, the quick ratio for SAP Group at 31 December 2017 is *closest* to:

A 1.00.

B 1.07.

C 1.17.

30 Using the information presented in Exhibit 14 of the reading, the financial leverage ratio for SAP Group at 31 December 2017 is *closest* to:

A 1.50.

B 1.66.

C 2.00.

Questions 31 through 34 refer to Exhibit 1

Exhibit 1 Common-Size Balance Sheets for Company A, Company B, and Sector Average

	Company A	Company B	Sector Average
ASSETS			
Current assets			
Cash and cash equivalents	5	5	7
Marketable securities	5	0	2
Accounts receivable, net	5	15	12
Inventories	15	20	16
Prepaid expenses	5	15	11
Total current assets	35	55	48
Property, plant, and equipment, net	40	35	37
Goodwill	25	0	8
Other assets	0	10	7
Total assets	100	100	100
LIABILITIES AND SHAREHOLDERS' EQUITY			
Current liabilities			
Accounts payable	10	10	10
Short-term debt	25	10	15
Accrued expenses	0	5	3
Total current liabilities	35	25	28

Exhibit 1 (Continued)

LIABILITIES AND SHAREHOLDERS' EQUITY			
Long-term debt	45	20	28
Other non-current liabilities	0	10	7
Total liabilities	80	55	63
Total shareholders' equity	20	45	37
Total liabilities and shareholders' equity	100	100	100

31 Based on Exhibit 1, which statement is *most likely* correct?

A Company A has below-average liquidity risk.
B Company B has above-average solvency risk.
C Company A has made one or more acquisitions.

32 The quick ratio for Company A is *closest* to:

A 0.43.
B 0.57.
C 1.00.

33 Based on Exhibit 1, the financial leverage ratio for Company B is *closest* to:

A 0.55.
B 1.22.
C 2.22.

34 Based on Exhibit 1, which ratio indicates lower liquidity risk for Company A compared with Company B?

A Cash ratio
B Quick ratio
C Current ratio

SOLUTIONS

1 B is correct. Assets are resources controlled by a company as a result of past events.

2 A is correct. Assets = Liabilities + Equity and, therefore, Assets – Liabilities = Equity.

3 A is correct. A classified balance sheet is one that classifies assets and liabilities as current or non-current and provides a subtotal for current assets and current liabilities. A liquidity-based balance sheet broadly presents assets and liabilities in order of liquidity.

4 B is correct. The balance sheet omits important aspects of a company's ability to generate future cash flows, such as its reputation and management skills. The balance sheet measures some assets and liabilities based on historical cost and measures others based on current value. Market value of shareholders' equity is updated continuously. Shareholders' equity reported on the balance sheet is updated for reporting purposes and represents the value that was current at the end of the reporting period.

5 B is correct. Balance sheet information is as of a specific point in time, and items measured at current value reflect the value that was current at the end of the reporting period. For all financial statement items, an item should be recognized in the financial statements only if it is probable that any future economic benefit associated with the item will flow to or from the entity and if the item has a cost or value that can be measured with reliability.

6 B is correct. Payments due within one operating cycle of the business, even if they will be settled more than one year after the balance sheet date, are classified as current liabilities. Payment received in advance of the delivery of a good or service creates an obligation or liability. If the obligation is to be fulfilled at least one year after the balance sheet date, it is recorded as a non-current liability, such as deferred revenue or deferred income. Payments that the company has the unconditional right to defer for at least one year after the balance sheet may be classified as non-current liabilities.

7 A is correct. A liquidity-based presentation, rather than a current/non-current presentation, may be used by such entities as banks if broadly presenting assets and liabilities in order of liquidity is reliable and more relevant.

8 B is correct. Goodwill is a long-term asset, and the others are all current assets.

9 C is correct. Both the cost of inventory and property, plant, and equipment include delivery costs, or costs incurred in bringing them to the location for use or resale.

10 A is correct. Current liabilities are those liabilities, including debt, due within one year. Preferred refers to a class of stock. Convertible refers to a feature of bonds (or preferred stock) allowing the holder to convert the instrument into common stock.

11 B is correct. The cash received from customers represents an asset. The obligation to provide a product in the future is a liability called "unearned income" or "unearned revenue." As the product is delivered, revenue will be recognized and the liability will be reduced.

12 C is correct. A contra asset account is netted against (i.e., reduces) the balance of an asset account. The allowance for doubtful accounts reduces the balance of accounts receivable. Accumulated depreciation, not depreciation expense, is a contra asset account. Sales returns and allowances create a contra account that reduce sales, not an asset.

13 C is correct. Under IFRS, inventories are carried at historical cost, unless net realizable value of the inventory is less. Under US GAAP, inventories are carried at the lower of cost or market.

14 C is correct. Paying rent in advance will reduce cash and increase prepaid expenses, both of which are assets.

15 C is correct. Accrued liabilities are expenses that have been reported on a company's income statement but have not yet been paid.

16 A is correct. Initially, goodwill is measured as the difference between the purchase price paid for an acquisition and the fair value of the acquired, not acquiring, company's net assets (identifiable assets less liabilities).

17 C is correct. Impairment write-downs reduce equity in the denominator of the debt-to-equity ratio but do not affect debt, so the debt-to-equity ratio is expected to increase. Impairment write-downs reduce total assets but do not affect revenue. Thus, total asset turnover is expected to increase.

18 B is correct. Vertical common-size analysis involves stating each balance sheet item as a percentage of total assets. Total assets are the sum of total liabilities (£35 million) and total stockholders' equity (£55 million), or £90 million. Total liabilities are shown on a vertical common-size balance sheet as (£35 million/£90 million) ≈ 39%.

19 B is correct. For financial assets classified as trading securities, unrealized gains and losses are reported on the income statement and flow to shareholders' equity as part of retained earnings.

20 C is correct. For financial assets classified as available for sale, unrealized gains and losses are not recorded on the income statement and instead are part of *other* comprehensive income. Accumulated other comprehensive income is a component of Shareholders' equity

21 A is correct. Financial assets classified as held to maturity are measured at amortised cost. Gains and losses are recognized only when realized.

22 B is correct. The non-controlling interest in consolidated subsidiaries is shown separately as part of shareholders' equity.

23 C is correct. The item "retained earnings" is a component of shareholders' equity.

24 B is correct. Share repurchases reduce the company's cash (an asset). Shareholders' equity is reduced because there are fewer shares outstanding and treasury stock is an offset to owners' equity.

25 B is correct. Common-size analysis (as presented in the reading) provides information about composition of the balance sheet and changes over time. As a result, it can provide information about an increase or decrease in a company's financial leverage.

26 A is correct. The current ratio provides a comparison of assets that can be turned into cash relatively quickly and liabilities that must be paid within one year. The other ratios are more suited to longer-term concerns.

27 A is correct. The cash ratio determines how much of a company's near-term obligations can be settled with existing amounts of cash and marketable securities.

28 C is correct. The debt-to-equity ratio, a solvency ratio, is an indicator of financial risk.

29 B is correct. The quick ratio ([Cash + Marketable securities + Receivables] ÷ Current liabilities) is 1.07 ([= €4,011 + €990 + €5,899] ÷ €10,210). As noted in the text, the largest component of the current financial assets are loans and other financial receivables. Thus, financial assets are included in the quick ratio but not the cash ratio.

30 B is correct. The financial leverage ratio (Total assets ÷ Total equity) is 1.66 (= €42,497 ÷ €25,540).

31 C is correct. The presence of goodwill on Company A's balance sheet signifies that it has made one or more acquisitions in the past. The current, cash, and quick ratios are lower for Company A than for the sector average. These lower liquidity ratios imply above-average liquidity risk. The total debt, long-term debt-to-equity, debt-to-equity, and financial leverage ratios are lower for Company B than for the sector average. These lower solvency ratios imply below-average solvency risk.

Current ratio is (35/35) = 1.00 for Company A, versus (48/28) = 1.71 for the sector average.

Cash ratio is (5 + 5)/35 = 0.29 for Company A, versus (7 + 2)/28 = 0.32 for the sector average.

Quick ratio is (5 + 5 + 5)/35 = 0.43 for Company A, versus (7 + 2 + 12)/28 = 0.75 for the sector average.

Total debt ratio is (55/100) = 0.55 for Company B, versus (63/100) = 0.63 for the sector average.

Long-term debt-to-equity ratio is (20/45) = 0.44 for Company B, versus (28/37) = 0.76 for the sector average.

Debt-to-equity ratio is (55/45) = 1.22 for Company B, versus (63/37) = 1.70 for the sector average.

Financial leverage ratio is (100/45) = 2.22 for Company B, versus (100/37) = 2.70 for the sector average.

32 A is correct. The quick ratio is defined as (Cash and cash equivalents + Marketable securities + receivables) ÷ Current liabilities. For Company A, this calculation is (5 + 5 + 5)/35 = 0.43.

33 C is correct. The financial leverage ratio is defined as Total assets ÷ Total equity. For Company B, total assets are 100 and total equity is 45; hence, the financial leverage ratio is 100/45 = 2.22.

34 A is correct. The cash ratio is defined as (Cash + Marketable securities)/Current liabilities. Company A's cash ratio, (5 + 5)/35 = 0.29, is higher than (5 + 0)/25 = 0.20 for Company B.

READING

19

Understanding Cash Flow Statements

**by Elaine Henry, PhD, CFA, Thomas R. Robinson, PhD, CFA,
J. Hennie van Greuning, DCom, CFA, and
Michael A. Broihahn, CPA, CIA, CFA**

Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). Thomas R. Robinson, PhD, CFA, is at AACSB International (USA). J. Hennie van Greuning, DCom, CFA, is at BIBD (Brunei). Michael A. Broihahn, CPA, CIA, CFA, is at Barry University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** compare cash flows from operating, investing, and financing activities and classify cash flow items as relating to one of those three categories given a description of the items;
☐	**b.** describe how non-cash investing and financing activities are reported;
☐	**c.** contrast cash flow statements prepared under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP);
☐	**d.** compare and contrast the direct and indirect methods of presenting cash from operating activities and describe arguments in favor of each method;
☐	**e.** describe how the cash flow statement is linked to the income statement and the balance sheet;
☐	**f.** describe the steps in the preparation of direct and indirect cash flow statements, including how cash flows can be computed using income statement and balance sheet data;
☐	**g.** demonstrate the conversion of cash flows from the indirect to direct method;
☐	**h.** analyze and interpret both reported and common-size cash flow statements;
☐	**i.** calculate and interpret free cash flow to the firm, free cash flow to equity, and performance and coverage cash flow ratios.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION

The cash flow statement provides information about a company's *cash receipts* and *cash payments* during an accounting period. The cash-based information provided by the cash flow statement contrasts with the accrual-based information from the income statement. For example, the income statement reflects revenues when earned rather than when cash is collected; in contrast, the cash flow statement reflects cash receipts when collected as opposed to when the revenue was earned. A reconciliation between reported income and cash flows from operating activities provides useful information about when, whether, and how a company is able to generate cash from its operating activities. Although income is an important measure of the results of a company's activities, cash flow is also essential. As an extreme illustration, a hypothetical company that makes all sales on account, without regard to whether it will ever collect its accounts receivable, would report healthy sales on its income statement and might well report significant income; however, with zero cash inflow, the company would not survive. The cash flow statement also provides a reconciliation of the beginning and ending cash on the balance sheet.

In addition to information about cash generated (or, alternatively, cash used) in operating activities, the cash flow statement provides information about cash provided (or used) in a company's investing and financing activities. This information allows the analyst to answer such questions as:

- Does the company generate enough cash from its operations to pay for its new investments, or is the company relying on new debt issuance to finance them?
- Does the company pay its dividends to common stockholders using cash generated from operations, from selling assets, or from issuing debt?

Answers to these questions are important because, in theory, generating cash from operations can continue indefinitely, but generating cash from selling assets, for example, is possible only as long as there are assets to sell. Similarly, generating cash from debt financing is possible only as long as lenders are willing to lend, and the lending decision depends on expectations that the company will ultimately have adequate cash to repay its obligations. In summary, information about the sources and uses of cash helps creditors, investors, and other statement users evaluate the company's liquidity, solvency, and financial flexibility.

This reading explains how cash flow activities are reflected in a company's cash flow statement. The reading is organized as follows. Sections 2–8 describe the components and format of the cash flow statement, including the classification of cash flows under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (GAAP) and the direct and indirect formats for presenting the cash flow statement. Sections 9–15 discuss the linkages of the cash flow statement with the income statement and balance sheet and the steps in the preparation of the cash flow statement. Sections 16–19 demonstrate the analysis of cash flow statements, including the conversion of an indirect cash flow statement to the direct method and how to use common-size cash flow analysis, free cash flow measures, and cash flow ratios used in security analysis. A summary of the key points and practice problems in the CFA Institute multiple-choice format conclude the reading.

CLASSIFICATION OF CASH FLOWS AND NON-CASH ACTIVITIES

2

a compare cash flows from operating, investing, and financing activities and classify cash flow items as relating to one of those three categories given a description of the items

b describe how non-cash investing and financing activities are reported

The analyst needs to be able to extract and interpret information on cash flows from financial statements. The basic components and allowable formats of the cash flow statement are well established.

- The cash flow statement has subsections relating specific items to the operating, investing, and financing activities of the company.
- Two presentation formats for the operating section are allowable: direct and indirect.

The following discussion presents these topics in greater detail.

2.1 Classification of Cash Flows and Non-Cash Activities

All companies engage in operating, investing, and financing activities. These activities are the classifications used in the cash flow statement under both IFRS and US GAAP and are described as follows:[1]

- **Operating activities** include the company's day-to-day activities that create revenues, such as selling inventory and providing services, and other activities not classified as investing or financing. Cash inflows result from cash sales and from collection of accounts receivable. Examples include cash receipts from the provision of services and royalties, commissions, and other revenue. To generate revenue, companies undertake such activities as manufacturing inventory, purchasing inventory from suppliers, and paying employees. Cash outflows result from cash payments for inventory, salaries, taxes, and other operating-related expenses and from paying accounts payable. Additionally, operating activities include cash receipts and payments related to **dealing securities** or **trading securities** (as opposed to buying or selling securities as investments, as discussed below).
- **Investing activities** include purchasing and selling long-term assets and other investments. These long-term assets and other investments include property, plant, and equipment; intangible assets; other long-term assets; and both long-term and short-term investments in the equity and debt (bonds and loans) issued by other companies. For this purpose, investments in equity and debt securities exclude securities held for dealing or trading purposes, the purchase and sale of which are considered operating activities even for companies where this is not a primary business activity. Cash inflows in the investing category include cash receipts from the sale of non-trading securities; property, plant, and equipment; intangibles; and other long-term assets. Cash outflows include cash payments for the purchase of these assets.

---

[1] IAS 7 *Statement of Cash Flows*.

- **Financing activities** include obtaining or repaying capital, such as equity and long-term debt. The two primary sources of capital are shareholders and creditors. Cash inflows in this category include cash receipts from issuing stock (common or preferred) or bonds and cash receipts from borrowing. Cash outflows include cash payments to repurchase stock (e.g., treasury stock) and to repay bonds and other borrowings. Note that indirect borrowing using accounts payable is not considered a financing activity—such borrowing is classified as an operating activity. The new IFRS standard relating to lease accounting (IFRS 16) affects how operating leases are represented in the cash flow statement.[2] Under IFRS 16, operating leases are treated similarly to finance leases—that is, the interest component of lease payments will be reflected in either the operating or financing section, and the principal component of lease payments is included in the financing section.

EXAMPLE 1

Net Cash Flow from Investing Activities

A company recorded the following in Year 1:

Proceeds from issuance of long-term debt	€300,000
Purchase of equipment	€200,000
Loss on sale of equipment	€70,000
Proceeds from sale of equipment	€120,000
Equity in earnings of affiliate	€10,000

On the Year 1 statement of cash flows, the company would report net cash flow from investing activities *closest* to:

A (€150,000).

B (€80,000).

C €200,000.

Solution:

B is correct. The only two items that would affect the investing section are the purchase of equipment and the proceeds from sale of equipment: (€200,000) + €120,000 = (€80,000). The loss on sale of equipment and the equity in earnings of affiliate affect net income but are not cash flows. The issuance of debt is a financing cash flow.

IFRS provide companies with choices in reporting some items of cash flow, particularly interest and dividends. IFRS explain that although for a financial institution interest paid and received would normally be classified as operating activities, for other entities, alternative classifications may be appropriate. For this reason, under IFRS, interest received may be classified either as an operating activity or as an investing activity. Under IFRS, interest paid may be classified as either an operating activity or a financing activity. Furthermore, under IFRS, dividends received may be classified as either an operating activity or an investing activity and dividends paid may be classified as either an operating activity or a financing activity. Companies must use a consistent classification from year to year and disclose separately the amounts of interest and dividends received and paid and where the amounts are reported.

---

[2] IFRS 16 is effective for fiscal years beginning 1 January 2019, with earlier voluntary adoption allowed.

Under US GAAP, discretion is not permitted in classifying interest and dividends. Interest received and interest paid are reported as operating activities for all companies.[3] Under US GAAP, dividends received are always reported as operating activities and dividends paid are always reported as financing activities.

EXAMPLE 2

Operating versus Financing Cash Flows

On 31 December 2018, a company issued a £30,000 180-day note at 8 percent and used the cash received to pay for inventory and issued £110,000 long-term debt at 11 percent annually and used the cash received to pay for new equipment. Which of the following *most* accurately reflects the combined effect of both transactions on the company's cash flows for the year ended 31 December 2018 under IFRS? Cash flows from:

A operations are unchanged.

B financing increase £110,000.

C operations decrease £30,000.

Solution:

C is correct. The payment for inventory would decrease cash flows from operations. The issuance of debt (both short-term and long-term debt) is part of financing activities and would increase cash flows from financing activities by £140,000. The purchase of equipment is an investing activity. Note that the treatment under US GAAP would be the same for these transactions.

Companies may also engage in non-cash investing and financing transactions. A non-cash transaction is any transaction that does not involve an inflow or outflow of cash. For example, if a company exchanges one non-monetary asset for another non-monetary asset, no cash is involved. Similarly, no cash is involved when a company issues common stock either for dividends or in connection with conversion of a convertible bond or convertible preferred stock. Because no cash is involved in non-cash transactions (by definition), these transactions are not incorporated in the cash flow statement. However, because such transactions may affect a company's capital or asset structures, any significant non-cash transaction is required to be disclosed, either in a separate note or a supplementary schedule to the cash flow statement.

CASH FLOW STATEMENT: DIFFERENCES BETWEEN IFRS AND US GAAP

3

c contrast cash flow statements prepared under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP)

---

[3] FASB ASC Topic 230 [Statement of Cash Flows].

As highlighted in the previous section, there are some differences in cash flow statements prepared under IFRS and US GAAP that the analyst should be aware of when comparing the cash flow statements of companies prepared in accordance with different sets of standards. The key differences are summarized in Exhibit 1. Most significantly, IFRS allow more flexibility in the reporting of such items as interest paid or received and dividends paid or received and in how income tax expense is classified.

US GAAP classify interest and dividends received from investments as operating activities, whereas IFRS allow companies to classify those items as either operating or investing cash flows. Likewise, US GAAP classify interest expense as an operating activity, even though the principal amount of the debt issued is classified as a financing activity. IFRS allow companies to classify interest expense as either an operating activity or a financing activity. US GAAP classify dividends paid to stockholders as a financing activity, whereas IFRS allow companies to classify dividends paid as either an operating activity or a financing activity.

US GAAP classify all income tax expenses as an operating activity. IFRS also classify income tax expense as an operating activity, unless the tax expense can be specifically identified with an investing or financing activity (e.g., the tax effect of the sale of a discontinued operation could be classified under investing activities).

Exhibit 1 Cash Flow Statements: Differences between IFRS and US GAAP

Topic	IFRS	US GAAP
Classification of cash flows:		
■ Interest received	Operating or investing	Operating
■ Interest paid	Operating or financing	Operating
■ Dividends received	Operating or investing	Operating
■ Dividends paid	Operating or financing	Financing
■ Bank overdrafts	Considered part of cash equivalents	Not considered part of cash and cash equivalents and classified as financing
■ Taxes paid	Generally operating, but a portion can be allocated to investing or financing if it can be specifically identified with these categories	Operating
Format of statement	Direct or indirect; direct is encouraged	Direct or indirect; direct is encouraged. A reconciliation of net income to cash flow from operating activities must be provided regardless of method used

Sources: IAS 7; FASB ASC Topic 230; and "IFRS and US GAAP: Similarities and Differences," PricewaterhouseCoopers (November 2017), available at www.pwc.com.

Under either set of standards, companies currently have a choice of formats for presenting cash flow statements, as discussed in the next section.

CASH FLOW STATEMENT: DIRECT AND INDIRECT METHODS FOR REPORTING CASH FLOW FROM OPERATING ACTIVITIES

4

d compare and contrast the direct and indirect methods of presenting cash from operating activities and describe arguments in favor of each method

There are two acceptable formats for reporting **cash flow from operating activities** (also known as **cash flow from operations** or **operating cash flow**), defined as the net amount of cash provided from operating activities: the direct and the indirect methods. The *amount* of operating cash flow is identical under both methods; only the *presentation format* of the operating cash flow section differs. The presentation format of the cash flows from investing and financing is exactly the same, regardless of which method is used to present operating cash flows.

The **direct method** shows the specific cash inflows and outflows that result in reported cash flow from operating activities. It shows each cash inflow and outflow related to a company's cash receipts and disbursements. In other words, the direct method eliminates any impact of accruals and shows only cash receipts and cash payments. The primary argument in favor of the direct method is that it provides information on the specific sources of operating cash receipts and payments. This is in contrast to the indirect method, which shows only the net result of these receipts and payments. Just as information on the specific sources of revenues and expenses is more useful than knowing only the net result—net income—the analyst gets additional information from a direct-format cash flow statement. The additional information is useful in understanding historical performance and in predicting future operating cash flows.

The **indirect method** shows how cash flow from operations can be obtained from reported net income as the result of a series of adjustments. The **indirect format** begins with net income. To reconcile net income with operating cash flow, adjustments are made for non-cash items, for non-operating items, and for the net changes in operating accruals. The main argument for the indirect approach is that it shows the reasons for differences between net income and operating cash flows. (However, the differences between net income and operating cash flows are equally visible on an indirect-format cash flow statement and in the supplementary reconciliation required under US GAAP if the company uses the direct method.) Another argument for the indirect method is that it mirrors a forecasting approach that begins by forecasting future income and then derives cash flows by adjusting for changes in balance sheet accounts that occur because of the timing differences between accrual and cash accounting.

IFRS and US GAAP both encourage the use of the direct method but permit either method. US GAAP encourage the use of the direct method but also require companies to present a reconciliation between net income and cash flow (which is equivalent to the indirect method).[4] If the indirect method is chosen, no direct-format disclosures are required. The majority of companies, reporting under IFRS or US GAAP, present using the indirect method for operating cash flows.

Many users of financial statements prefer the **direct format**, particularly analysts and commercial lenders, because of the importance of information about operating receipts and payments in assessing a company's financing needs and capacity to repay existing obligations. Preparers argue that adjusting net income to operating cash flow, as in the indirect format, is easier and less costly than reporting gross operating cash receipts and payments, as in the direct format. With advances in accounting systems and technology, it is not clear that gathering the information required to use the direct method is difficult or costly. CFA Institute has advocated that standard setters require the use of the direct format for the main presentation of the cash flow statement, with indirect cash flows as supplementary disclosure.[5]

---

[4] FASB ASC Section 230-10-45 [Statement of Cash Flows–Overall–Other Presentation Matters].

5 CASH FLOW STATEMENT: INDIRECT METHOD UNDER IFRS

c contrast cash flow statements prepared under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP)

d compare and contrast the direct and indirect methods of presenting cash from operating activities and describe arguments in favor of each method

Exhibit 2 presents the consolidated cash flow statement prepared under IFRS from Unilever Group's 2017 annual report. The statement, covering the fiscal years ended 31 December 2017, 2016, and 2015, shows the use of the indirect method. Unilever is an Anglo-Dutch consumer products company with headquarters in the United Kingdom and the Netherlands.[6]

Exhibit 2 Unilever Group Consolidated Cash Flow Statement (€ millions)

	For the year ended 31 December		
	2017	2016	2015
Cash flow from operating activities			
Net profit	6,486	5,547	5,259
Taxation	1,667	1,922	1,961
Share of net profit of joint ventures/associates and other income (loss) from non-current investments and associates	(173)	(231)	(198)
Net finance costs:	877	563	493
Operating profit	8,857	7,801	7,515
Depreciation, amortisation and impairment	1,538	1,464	1,370
Changes in working capital:	(68)	51	720
Inventories	(104)	190	(129)
Trade and other current receivables	(506)	142	2
Trade payables and other liabilities	542	(281)	847
Pensions and similar obligations less payments	(904)	(327)	(385)
Provisions less payments	200	65	(94)
Elimination of (profits)/losses on disposals	(298)	127	26
Non-cash charge for share-based compensation	284	198	150
Other adjustments	(153)	(81)	49

---

[5] *A Comprehensive Business Reporting Model: Financial Reporting for Investors*, CFA Institute Centre for Financial Market Integrity (July 2007), p. 13.

[6] Unilever NV and Unilever PLC have independent legal structures, but a series of agreements enable the companies to operate as a single economic entity.

Exhibit 2 (Continued)

	For the year ended 31 December		
	2017	**2016**	**2015**
Cash flow from operating activities	**9,456**	**9,298**	**9,351**
Income tax paid	(2,164)	(2,251)	(2,021)
Net cash flow from operating activities	**7,292**	**7,047**	**7,330**
Interest received	154	105	119
Purchase of intangible assets	(158)	(232)	(334)
Purchase of property, plant and equipment	(1,509)	(1,804)	(1,867)
Disposal of property, plant and equipment	46	158	127
Acquisition of group companies, joint ventures and associates	(4,896)	(1,731)	(1,897)
Disposal of group companies, joint ventures and associates	561	30	199
Acquisition of other non-current investments	(317)	(208)	(78)
Disposal of other non-current investments	251	173	127
Dividends from joint ventures, associates and other non-current investments	138	186	176
(Purchase)/sale of financial assets	(149)	135	(111)
Net cash flow (used in)/from investing activities	**(5,879)**	**(3,188)**	**(3,539)**
Dividends paid on ordinary share capital	(3,916)	(3,609)	(3,331)
Interest and preference dividends paid	(470)	(472)	(579)
Net change in short-term borrowings	2,695	258	245
Additional financial liabilities	8,851	6,761	7,566
Repayment of financial liabilities	(2,604)	(5,213)	(6,270)
Capital element of finance lease rental payments	(14)	(35)	(14)
Buy back of preference shares	(448)	—	—
Repurchase of shares	(5,014)	—	—
Other movements on treasury stock	(204)	(257)	(276)
Other financing activities	(309)	(506)	(373)
Net cash flow (used in)/from financing activities	**(1,433)**	**(3,073)**	**(3,032)**
Net increase/(decrease) in cash and cash equivalents	**(20)**	**786**	**759**
Cash and cash equivalents at the beginning of the year	**3,198**	**2,128**	**1,910**
Effect of foreign exchange rate changes	(9)	284	(541)
Cash and cash equivalents at the end of the year	**3,169**	**3,198**	**2,128**

Beginning first at the bottom of the statement, we note that cash increased from €1,910 million at the beginning of 2015 to €3,169 million at the end of 2017, with the largest increase occurring in 2016. To understand the changes, we next examine the sections of the statement. In each year, the primary cash inflow derived from operating activities, as would be expected for a mature company in a relatively stable industry. In each year, the operating cash flow was more than the reported net profit, again, as would be expected from a mature company, with the largest differences primarily arising from the add-back of depreciation. Also, in each year, the operating cash flow was more than enough to cover the company's capital expenditures. For example, in 2017, the company generated €7,292 million in net cash from operating activities and—as shown in the investing section—spent €1,509 million on property, plant, and equipment. The operating cash flow was also sufficient to cover acquisitions of other companies.

The financing section of the statement shows that each year the company returned more than €3.3 billion to its common shareholders through dividends and around €500 million to its debt holders and preferred shareholders via interest and dividends. In 2017, the company used cash to repurchase about €5 billion in common stock in and generated cash from increased borrowing. The increase in short-term borrowings (€2,695 million) and additional financial liabilities (€8,851 million) exceeded the cash repayment of liabilities (€2,604 million).

Having examined each section of the statement, we return to the operating activities section of Unilever's cash flow statement, which presents a reconciliation of net profit to net cash flow from operating activities (i.e., uses the indirect method). The following discussion of certain adjustments to reconcile net profit to operating cash flows explains some of the main reconciliation adjustments and refers to the amounts in 2017. The first adjustment adds back the €1,667 million income tax expense (labeled "Taxation") that had been recognized as an expense in the computation of net profit. A €2,164 million deduction for the (cash) income taxes paid is then shown separately, as the last item in the operating activities section, consistent with the IFRS requirement that cash flows arising from income taxes be separately disclosed. The classification of taxes on income paid should be indicated. The classification is in operating activities unless the taxes can be specifically identified with financing or investing activities.

The next adjustment "removes" from the operating cash flow section the €173 million representing Unilever's share of joint ventures' income that had been included in the computation of net profit. A €138 million inflow of (cash) dividends received from those joint ventures is then shown in the investing activities section. Similarly, a €877 million adjustment removes the net finance costs from the operating activities section. Unilever then reports its €154 million (cash) interest received in the investing activities section and its €470 million (cash) interest paid (and preference dividends paid) in the financing activities section. The next adjustment in the operating section of this indirect-method statement adds back €1,538 million depreciation, amortisation, and impairment, all of which are expenses that had been deducted in the computation of net income but which did not involve any outflow of cash in the period. The €68 million adjustment for changes in working capital is necessary because these changes result from applying accrual accounting and thus do not necessarily correspond to the actual cash movement. These adjustments are described in greater detail in a later section.

In summary, some observations from an analysis of Unilever's cash flow statement include:

- Total cash increased from €1,910 million at the beginning of 2015 to €3,169 million at the end of 2017, with the largest increase occurring in 2016.
- In each year, the operating cash flow was more than the reported net profit, as would generally be expected from a mature company.
- In each year, the operating cash flow was more than enough to cover the company's capital expenditures.
- The company returned cash to its equity investors through dividends in each year and through share buybacks in 2017.

CASH FLOW STATEMENT: DIRECT METHOD UNDER IFRS

6

c contrast cash flow statements prepared under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP)

d compare and contrast the direct and indirect methods of presenting cash from operating activities and describe arguments in favor of each method

In the direct format of the cash flow statement, the cash received from customers, as well as other operating items, is clearly shown.

Exhibit 3 presents a direct-method format cash flow statement prepared under IFRS for Telefónica Group, a diversified telecommunications company based in Madrid.[7]

Exhibit 3 Telefónica Group Consolidated Statement of Cash Flows (€ millions)

for the years ended 31 December	2017	2016	2015
Cash flows from operating activities			
Cash received from operations	63,456	63,514	67,582
Cash paid from operations	(46,929)	(47,384)	(50,833)
Net interest and other financial expenses net of dividends received	(1,726)	(2,143)	(2,445)
Taxes paid	(1,005)	(649)	(689)
Net cash flow provided by operating activities	**13,796**	**13,338**	**13,615**
Cash flows from investing activities			
(Payments on investments)/proceeds from the sale in property, plant and equipment and intangible assets, net	(8,992)	(9,187)	(10,256)
Proceeds on disposals of companies, net of cash and cash equivalents disposed	40	767	354
Payments on investments in companies, net of cash and cash equivalents acquired	(128)	(54)	(3,181)
Proceeds on financial investments not included under cash equivalents	296	489	1,142
Payments made on financial investments not included under cash equivalents	(1,106)	(265)	(426)
(Payments)/proceeds on placements of cash surpluses not included under cash equivalents	(357)	42	(557)
Government grants received	2	—	7
Net cash used in investing activities	**(10,245)**	**(8,208)**	**(12,917)**
Cash flows from financing activities			
Dividends paid	(2,459)	(2,906)	(2,775)
Proceeds from share capital increase	2	—	4,255
Proceeds/(payments) of treasury shares and other operations with shareholders and with minority interests	1,269	(660)	(1,772)

(continued)

---

[7] This statement excludes the supplemental cash flow reconciliation provided at the bottom of the original cash flow statement by the company.

Exhibit 3 (Continued)

Operations with other equity holders	646	656	83
Proceeds on issue of debentures and bonds, and other debts	8,390	5,693	1,602
Proceeds on loans, borrowings and promissory notes	4,844	10,332	8,784
Repayments of debentures and bonds and other debts	(6,687)	(6,873)	(3,805)
Repayments of loans, borrowings and promissory notes	(6,711)	(8,506)	(9,858)
Financed operating payments and investments in property, plant and equipment and intangible assets payments	(1,046)	(1,956)	(126)
Net cash flow used in financing activities	**(1,752)**	**(4,220)**	**(3,612)**
Effect of changes in exchange rates	(341)	185	(1,000)
Effect of changes in consolidation methods and others	(2)	26	—
Net increase (decrease) in cash and cash equivalents during the period	**1,456**	**1,121**	**(3,914)**
Cash and cash equivalents at 1 January	3,736	2,615	6,529
Cash and cash equivalents at 31 December	5,192	3,736	2,615

As shown at the bottom of the statement, cash and cash equivalents decreased from €6,529 million at the beginning of 2015 to €5,192 million at the end of 2017. The largest decrease in cash occurred in 2015. Cash from operations was the primary source of cash, consistent with the profile of a mature company in a relatively stable industry. Each year, the company generated significantly more cash from operations than it required for its capital expenditures. For example, in 2017, the company generated €13.8 billion cash from operations and spent—as shown in the investing section—only €9 billion on property, plant, and equipment, net of proceeds from sales. Another notable item from the investing section is the company's limited acquisition activity in 2017 and 2016 compared with 2015. In 2015, the company made over €3 billion of acquisitions. As shown in the financing section, cash flows from financing was negative in all three years, although the components of the negative cash flows differed. In 2015, for example, the company generated cash with an equity issuance of €4.2 billion but made significant net repayments of debts resulting in negative cash from financing activities.

In summary, some observations from an analysis of Telefónica's cash flow statement include

- Total cash and cash equivalents decreased over the three-year period, with 2015 showing the biggest decrease.
- Cash from operating activities was large enough in each year to cover the company's capital expenditures.
- The amount paid for property, plant, and equipment and intangible assets was the largest investing expenditure each year.
- The company had a significant amount of acquisition activity in 2015.
- The company paid dividends each year although the amount in 2017 is somewhat lower than in prior years.

CASH FLOW STATEMENT: DIRECT METHOD UNDER US GAAP

7

c contrast cash flow statements prepared under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP)

d compare and contrast the direct and indirect methods of presenting cash from operating activities and describe arguments in favor of each method

Previously, we presented cash flow statements prepared under IFRS. In this section, we illustrate cash flow statements prepared under US GAAP. This section presents the cash flow statements of two companies, Tech Data Corporation and Walmart. Tech Data reports its operating activities using the direct method, whereas Walmart reports its operating activities using the more common indirect method.

Tech Data Corporation is a leading distributor of information technology products. Exhibit 4 presents comparative cash flow statements from the company's annual report for the fiscal years ended 31 January 2016 through 2018.

Exhibit 4 Tech Data Corporation and Subsidiaries Consolidated Cash Flow Statements (in Thousands)

Years Ended 31 January	2018	2017	2016
Cash flows from operating activities:			
Cash received from customers	$42,981,601	$29,427,357	$28,119,687
Cash paid to vendors and employees	(41,666,356)	(28,664,222)	(27,819,886)
Interest paid, net	(86,544)	(22,020)	(20,264)
Income taxes paid	(131,632)	(84,272)	(85,645)
Net cash provided by operating activities	**1,097,069**	**656,843**	**193,892**
Cash flows from investing activities:			
Acquisition of business, net of cash acquired	(2,249,849)	(2,916)	(27,848)
Expenditures for property and equipment	(192,235)	(24,971)	(20,917)
Software and software development costs	(39,702)	(14,364)	(13,055)
Proceeds from sale of subsidiaries	0	0	20,020
Net cash used in investing activities	**(2,481,786)**	**(42,251)**	**(41,800)**
Cash flows from financing activities:			
Borrowings on long-term debt	1,008,148	998,405	—
Principal payments on long-term debt	(861,394)	—	(319)
Cash paid for debt issuance costs	(6,348)	(21,581)	—
Net borrowings on revolving credit loans	(16,028)	3,417	5,912
Cash paid for purchase of treasury stock	—	—	(147,003)
Payments for employee withholdings on equity awards	(6,027)	(4,479)	(4,662)
Proceeds from the reissuance of treasury stock	1,543	733	561
Acquisition of earn-out payments	—	—	(2,736)
Net cash provided by (used in) financing activities	**119,894**	**976,495**	**(148,247)**

(continued)

Exhibit 4 (Continued)

Years Ended 31 January	**2018**	**2017**	**2016**
Effect of exchange rate changes on cash and cash equivalents	94,860	3,335	(15,671)
Net (decrease) increase in cash and cash equivalents	(1,169,963)	1,594,422	(11,826)
Cash and cash equivalents at beginning of year	2,125,591	531,169	542,995
Cash and cash equivalents at end of year	$955,628	$2,125,591	$531,169
Reconciliation of net income to net cash provided by operating activities:			
Net income	$116,641	$195,095	$265,736
Adjustments to reconcile net income to net cash provided by operating activities:			
Depreciation and amortization	150,046	54,437	57,253
Provision for losses on accounts receivable	21,022	5,026	6,061
Stock-based compensation expense	29,381	13,947	14,890
Loss on disposal of subsidiaries	—	—	699
Accretion of debt discount and debt issuance costs	3,326	835	839
Deferred income taxes	(4,261)	(11,002)	2,387
Changes in operating assets and liabilities:			
Accounts receivable	(554,627)	(91,961)	(297,637)
Inventories	(502,352)	(20,838)	(219,482)
Prepaid expenses and other assets	32,963	66,027	(44,384)
Accounts payable	1,704,307	459,146	426,412
Accrued expenses and other liabilities	100,623	(13,869)	(18,882)
Total adjustments	980,428	461,748	(71,844)
Net cash provided by operating activities	$1,097,069	$656,843	$193,892

Tech Data Corporation prepares its cash flow statements under the direct method. The company's cash increased from $543 million at the beginning of 2016 to $956 million at the end of January 2018, with the biggest increase occurring in 2017. The 2017 increase was driven by changes in both operating cash flow and financing cash flow. In the cash flows from operating activities section of Tech Data's cash flow statements, the company identifies the amount of cash it received from customers, $43 billion for 2018, and the amount of cash that it paid to suppliers and employees, $41.7 billion for 2018. Cash receipts increased from $29.4 billion in the prior year and cash paid also increased substantially. Net cash provided by operating activities was adequate to cover the company's investing activities in 2016 and 2017 but not in 2018, primarily because of increased amounts of cash used for acquisition of business. Related to this investing cash outflow for an acquisition, footnotes disclose that the major acquisition in 2018 accounted for the large increase in cash receipts and cash payments in the operating section. Also related to the 2018 acquisition, the financing section shows that the company borrowed more debt that it repaid in both 2017 and 2018. In 2017, borrowings on long-term debt were $998.4 million, and net borrowings on revolving credit loans were $3.4 million. In 2018, the company generated cash by borrowing more long-term debt than it repaid but used cash to pay down its revolving credit loans. There are no dividend payments, although in 2016, the company paid $147 million to repurchase its common stock.

Whenever the direct method is used, US GAAP require a disclosure note and a schedule that reconciles net income with the net cash flow from operating activities. Tech Data shows this reconciliation at the bottom of its consolidated statements of cash flows. The disclosure note and reconciliation schedule are exactly the information that would have been presented in the body of the cash flow statement if the company had elected to use the indirect method rather than the direct method. For 2018, the reconciliation highlights an increase in the company's accounts receivable, inventory, and payables.

In summary, some observations from an analysis of Tech Data's cash flow statement include:

- The company's cash increased by over $412 million over the three years ending in January 2018, with the biggest increase occurring in 2017.
- The company's operating cash was adequate to cover the company's investments in 2016 and 2017, but not in 2018 primarily because of a major acquisition.
- Related to the 2018 acquisition, the financing section shows an increase in long-term borrowings in 2017 and 2018, including a $998 million increase in 2017.
- The company has not paid dividends in the past three years, but the financing section shows that in 2016 the company repurchased stock.

CASH FLOW STATEMENT: INDIRECT METHOD UNDER US GAAP

8

c contrast cash flow statements prepared under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (US GAAP)

d compare and contrast the direct and indirect methods of presenting cash from operating activities and describe arguments in favor of each method

Walmart is a global retailer that conducts business under the names of Walmart and Sam's Club. Exhibit 5 presents the comparative cash flow statements from the company's annual report for the fiscal years ended 31 January 2018, 2017, and 2016.

Exhibit 5 Walmart Cash Flow Statements Fiscal Years Ended 31 January ($ millions)

Fiscal Year Ended 31 January	2018	2017	2016
Cash flows from operating activities:			
Consolidated net income	10,523	14,293	15,080
Adjustments to reconcile income from continuing operations to net cash provided by operating activities:			
Depreciation and amortization	10,529	10,080	9,454
Deferred income taxes	(304)	761	(672)

(continued)

Exhibit 5 (Continued)

Fiscal Year Ended 31 January	**2018**	**2017**	**2016**
Loss on extinguishment of debt	3,136	—	—
Other operating activities	1,210	206	1,410
Changes in certain assets and liabilities, net of effects of acquisitions:			
Receivables, net	(1.074)	(402)	(19)
Inventories	(140)	1,021	(703)
Accounts payable	4,086	3,942	2,008
Accrued liabilities	928	1,280	1,466
Accrued income taxes	(557)	492	(472)
Net cash provided by operating activities	28,337	31,673	27,552
Cash flows from investing activities:			
Payments for property and equipment	(10,051)	(10,619)	(11,477)
Proceeds from disposal of property and equipment	378	456	635
Proceeds from the disposal of certain operations	1,046	662	246
Purchase of available for sale securities	—	(1,901)	—
Investment and business acquisitions, net of cash acquired	(375)	(2,463)	—
Other investing activities	(58)	(122)	(79)
Net cash used in investing activities	(9,060)	(13,987)	(10,675)
Cash flows from financing activities:			
Net change in short-term borrowings	4,148	(1,673)	1,235
Proceeds from issuance of long-term debt	7,476	137	39
Payments of long-term debt	(13,061)	(2,055)	(4,432)
Payment for debt extinguishment or debt prepayment cost	(3,059)	—	—
Dividends paid	(6,124)	(6,216)	(6,294)
Purchase of Company stock	(8,296)	(8,298)	(4,112)
Dividends paid to noncontrolling interest	(690)	(479)	(719)
Purchase of noncontrolling interest	(8)	(90)	(1,326)
Other financing activities	(261)	(398)	(676)
Net cash used in financing activities	(19,875)	(19,072)	(16,285)
Effect of exchange rates on cash and cash equivalents	487	(452)	(1,022)
Net increase (decrease) in cash and cash equivalents	(111)	(1,838)	(430)
Cash and cash equivalents at beginning of yea	6,867	8,705	9,135
Cash and cash equivalents at end of yea	**6,756**	**6,867**	**8,705**
Supplemental disclosure of cash flow information			
Income taxes paid	6,179	4,507	8,111
Interest paid	2,450	2,351	2,540

Walmart's cash flow statement indicates the following:

- Cash and cash equivalents declined over the three years, from $9.1 billion at the beginning of fiscal 2016 to $6.8 billion at the end of fiscal 2018.

- Operating cash flow was relatively steady at $27.6 billion, $31.7 billion, and $28.3 billion in fiscal 2016, 2017, and 2018, respectively. Further, operating cash flow was significantly greater than the company's expenditures on property and equipment in every year.
- Over the three years, the company used significant amounts of cash to pay dividends and to repurchase its common stock. The company also repaid borrowing, particularly in fiscal 2018.

Walmart prepares its cash flow statements under the indirect method. In the cash flows from operating activities section of Walmart's cash flow statement, the company reconciles its net income for 2018 of $10.5 billion to net cash provided by operating activities of $28.3 billion. The largest adjustment is for depreciation and amortization of $10.5 billion. Depreciation and amortization expense requires an adjustment because it was a non-cash expense on the income statement. As illustrated in previous examples, depreciation is the largest or one of the largest adjustments made by many companies in the reconciliation of net income to operating cash flow.

Whenever the indirect method is used, US GAAP mandate disclosure of how much cash was paid for interest and income taxes. Note that these are line items in cash flow statements using the direct method, so disclosure does not have to be mandated. Walmart discloses the amount of cash paid for income tax ($6.2 billion) and interest ($2.5 billion) at the bottom of its cash flow statements.

LINKAGES OF CASH FLOW STATEMENT WITH THE INCOME STATEMENT AND BALANCE SHEET

9

e describe how the cash flow statement is linked to the income statement and the balance sheet

The indirect format of the cash flow statement demonstrates that changes in balance sheet accounts are an important factor in determining cash flows. The next section addresses the linkages between the cash flow statement and other financial statements.

9.1 Linkages of the Cash Flow Statement with the Income Statement and Balance Sheet

Recall the accounting equation that summarizes the balance sheet:

Assets = Liabilities + Equity

Cash is an asset. The statement of cash flows ultimately shows the change in cash during an accounting period. The beginning and ending balances of cash are shown on the company's balance sheets for the previous and current years, and the bottom of the cash flow statement reconciles beginning cash with ending cash. The relationship, stated in general terms, is as shown below.

Beginning Balance Sheet at 31 December 20X8	**Statement of Cash Flows for Year Ended 31 December 20X9**		**Ending Balance Sheet at 31 December 20X9**
Beginning cash	Plus: Cash receipts (from operating, investing, and financing activities)	Less: Cash payments (for operating, investing, and financing activities)	Ending cash

In the case of cash held in foreign currencies, there would also be an impact from changes in exchange rates. For example, Walmart's cash flow statement for 2018, presented in Exhibit 5, shows overall cash flows from operating, investing, and financing activities that total $(111) million during the year, including $487 million net effect of exchange rates on cash and cash equivalents.

The body of Walmart's cash flow statement shows why the change in cash occurred; in other words, it shows the company's operating, investing, and financing activities (as well as the impact of foreign currency translation). The beginning and ending balance sheet values of cash and cash equivalents are linked through the cash flow statement.

The current assets and current liabilities sections of the balance sheet typically reflect a company's operating decisions and activities. Because a company's operating activities are reported on an accrual basis in the income statement, any differences between the accrual basis and the cash basis of accounting for an operating transaction result in an increase or decrease in some (usually) short-term asset or liability on the balance sheet. For example, if revenue reported using accrual accounting is higher than the cash actually collected, the result will typically be an increase in accounts receivable. If expenses reported using accrual accounting are lower than cash actually paid, the result will typically be a decrease in accounts payable or another accrued liability account[8]. As an example of how items on the balance sheet are related to the income statement and/or cash flow statement through the change in the beginning and ending balances, consider accounts receivable:

Beginning Balance Sheet at 31 December 20X8	Income Statement for Year Ended 31 December 20X9	Statement of Cash Flows for Year Ended 31 December 20X9	Ending Balance Sheet at 31 December 20X9
Beginning accounts receivable	Plus: Revenues	Minus: Cash collected from customers	Equals: Ending accounts receivable

Knowing any three of these four items makes it easy to compute the fourth. For example, if you know beginning accounts receivable, revenues, and cash collected from customers, you can compute ending accounts receivable. Understanding the interrelationships among the balance sheet, income statement, and cash flow statement is useful not only in evaluating the company's financial health but also in detecting accounting irregularities. Recall the extreme illustration of a hypothetical company that makes sales on account without regard to future collections and thus reports healthy sales and significant income on its income statement yet lacks cash inflow. Such a pattern would occur if a company improperly recognized revenue.

A company's investing activities typically relate to the long-term asset section of the balance sheet, and its financing activities typically relate to the equity and long-term debt sections of the balance sheet. The next section demonstrates the preparation of cash flow information based on income statement and balance sheet information.

---

[8] There are other less typical explanations of the differences. For example, if revenue reported using accrual accounting is higher than the cash actually collected, it is possible that it is the result of a decrease in an unearned revenue liability account. If expenses reported using accrual accounting are lower than cash actually paid, it is possible that it is the result of an increase in prepaid expenses, inventory, or another asset account.

PREPARING THE CASH FLOW STATEMENT: THE DIRECT METHOD FOR OPERATING ACTIVITIES

10

f describe the steps in the preparation of direct and indirect cash flow statements, including how cash flows can be computed using income statement and balance sheet data

The preparation of the cash flow statement uses data from both the income statement and the comparative balance sheets.

As noted earlier, companies often only disclose indirect operating cash flow information, whereas analysts prefer direct-format information. Understanding how cash flow information is put together will enable you to take an indirect statement apart and reconfigure it in a more useful manner. The result is an approximation of a direct cash flow statement, which—while not perfectly accurate—can be helpful to an analyst. The following demonstration of how an approximation of a direct cash flow statement is prepared uses the income statement and the comparative balance sheets for Acme Corporation (a fictitious retail company) shown in Exhibits 6 and 7.

Exhibit 6 Acme Corporation Income Statement Year Ended 31 December 2018

Revenue (net)		$23,598
Cost of goods sold		11,456
Gross profit		12,142
Salary and wage expense	$4,123	
Depreciation expense	1,052	
Other operating expenses	3,577	
Total operating expenses		8,752
Operating profit		3,390
Other revenues (expenses):		
Gain on sale of equipment	205	
Interest expense	(246)	(41)
Income before tax		3,349
Income tax expense		1,139
Net income		$2,210

Exhibit 7 Acme Corporation Comparative Balance Sheets 31 December 2018 and 2017

	2018	2017	Net Change
Cash	$1,011	$1,163	$(152)
Accounts receivable	1,012	957	55
Inventory	3,984	3,277	707
Prepaid expenses	155	178	(23)

(continued)

Exhibit 7 (Continued)

	2018	2017	Net Change
Total current assets	6,162	5,575	587
Land	510	510	—
Buildings	3,680	3,680	—
Equipment*	8,798	8,555	243
Less: accumulated depreciation	(3,443)	(2,891)	(552)
Total long-term assets	9,545	9,854	(309)
Total assets	$15,707	$15,429	$278
Accounts payable	$3,588	$3,325	$263
Salary and wage payable	85	75	10
Interest payable	62	74	(12)
Income tax payable	55	50	5
Other accrued liabilities	1,126	1,104	22
Total current liabilities	4,916	4,628	288
Long-term debt	3,075	3,575	(500)
Common stock	3,750	4,350	(600)
Retained earnings	3,966	2,876	1,090
Total liabilities and equity	$15,707	$15,429	$278

* During 2018, Acme purchased new equipment for a total cost of $1,300. No items impacted retained earnings other than net income and dividends.

The first step in preparing the cash flow statement is to determine the total cash flows from operating activities. The direct method of presenting cash from operating activities is illustrated in sections 10–13. Section 14 illustrates the indirect method of presenting cash flows from operating activities. Cash flows from investing activities and from financing activities are identical regardless of whether the direct or indirect method is used to present operating cash flows.

10.1 Operating Activities: Direct Method

We first determine how much cash Acme received from its customers, followed by how much cash was paid to suppliers and to employees as well as how much cash was paid for other operating expenses, interest, and income taxes.

10.1.1 *Cash Received from Customers*

The income statement for Acme reported revenue of $23,598 for the year ended 31 December 2018. To determine the approximate cash receipts from its customers, it is necessary to adjust this revenue amount by the net change in accounts receivable for the year. If accounts receivable increase during the year, revenue on an accrual basis is higher than cash receipts from customers, and vice versa. For Acme Corporation, accounts receivable increased by $55, so cash received from customers was $23,543, as follows:

Revenue	$23,598
Less: Increase in accounts receivable	(55)
Cash received from customers	**$23,543**

Cash received from customers affects the accounts receivable account as follows:

Beginning accounts receivable	957
Plus revenue	23,598
Minus cash collected from customers	**(23,543)**
Ending accounts receivable	$1,012

The accounts receivable account information can also be presented as follows:

Beginning accounts receivable	$957
Plus revenue	23,598
Minus ending accounts receivable	(1,012)
Cash collected from customers	**$23,543**

EXAMPLE 3

Computing Cash Received from Customers

Blue Bayou, a fictitious advertising company, reported revenues of $50 million, total expenses of $35 million, and net income of $15 million in the most recent year. If accounts receivable decreased by $12 million, how much cash did the company receive from customers?

A $38 million.

B $50 million.

C $62 million.

Solution:

C is correct. Revenues of $50 million plus the decrease in accounts receivable of $12 million equals $62 million cash received from customers. The decrease in accounts receivable means that the company received more in cash than the amount of revenue it reported.

"Cash received from customers" is sometimes referred to as "cash collections from customers" or "cash collections."

10.1.2 *Cash Paid to Suppliers*

For Acme, the cash paid to suppliers was $11,900, determined as follows:

Cost of goods sold	$11,456
Plus: Increase in inventory	707
Equals purchases from suppliers	$12,163
Less: Increase in accounts payable	(263)
Cash paid to suppliers	**$11,900**

There are two pieces to this calculation: the amount of inventory purchased and the amount paid for it. To determine purchases from suppliers, cost of goods sold is adjusted for the change in inventory. If inventory increased during the year, then purchases during the year exceeded cost of goods sold, and vice versa. Acme reported cost of goods sold of $11,456 for the year ended 31 December 2018. For Acme Corporation, inventory increased by $707, so purchases from suppliers was $12,163. Purchases from suppliers affect the inventory account, as shown below:

Beginning inventory	$3,277
Plus purchases	12,163
Minus cost of goods sold	(11,456)
Ending inventory	$3,984

Acme purchased $12,163 of inventory from suppliers in 2018, but is this the amount of cash that Acme paid to its suppliers during the year? Not necessarily. Acme may not have yet paid for all of these purchases and may yet owe for some of the purchases made this year. In other words, Acme may have paid less cash to its suppliers than the amount of this year's purchases, in which case Acme's liability (accounts payable) will have increased by the difference. Alternatively, Acme may have paid even more to its suppliers than the amount of this year's purchases, in which case Acme's accounts payable will have decreased.

Therefore, once purchases have been determined, cash paid to suppliers can be calculated by adjusting purchases for the change in accounts payable. If the company made all purchases with cash, then accounts payable would not change and cash outflows would equal purchases. If accounts payable increased during the year, then purchases on an accrual basis would be higher than they would be on a cash basis, and vice versa. In this example, Acme made more purchases than it paid in cash, so the balance in accounts payable increased. For Acme, the cash paid to suppliers was $11,900, determined as follows:

Purchases from suppliers	$12,163
Less: Increase in accounts payable	(263)
Cash paid to suppliers	**$11,900**

The amount of cash paid to suppliers is reflected in the accounts payable account, as shown below:

Beginning accounts payable	$3,325
Plus purchases	12,163
Minus cash paid to suppliers	**(11,900)**
Ending accounts payable	$3,588

EXAMPLE 4

Computing Cash Paid to Suppliers

Orange Beverages Plc., a fictitious manufacturer of tropical drinks, reported cost of goods sold for the year of $100 million. Total assets increased by $55 million, but inventory declined by $6 million. Total liabilities increased by $45 million, but accounts payable decreased by $2 million. How much cash did the company pay to its suppliers during the year?

A $96 million.

B $104 million.

C $108 million.

Solution:

A is correct. Cost of goods sold of $100 million less the decrease in inventory of $6 million equals purchases from suppliers of $94 million. The decrease in accounts payable of $2 million means that the company paid $96 million in cash ($94 million plus $2 million).

10.1.3 *Cash Paid to Employees*

To determine the cash paid to employees, it is necessary to adjust salary and wages expense by the net change in salary and wages payable for the year. If salary and wages payable increased during the year, then salary and wages expense on an accrual basis would be higher than the amount of cash paid for this expense, and vice versa. For Acme, salary and wages payable increased by $10, so cash paid for salary and wages was $4,113, as follows:

Salary and wages expense	$4,123
Less: Increase in salary and wages payable	(10)
Cash paid to employees	**$4,113**

The amount of cash paid to employees is reflected in the salary and wages payable account, as shown below:

Beginning salary and wages payable	$75
Plus salary and wages expense	4,123
Minus cash paid to employees	**(4,113)**
Ending salary and wages payable	$85

10.1.4 *Cash Paid for Other Operating Expenses*

To determine the cash paid for other operating expenses, it is necessary to adjust the other operating expenses amount on the income statement by the net changes in prepaid expenses and accrued expense liabilities for the year. If prepaid expenses increased during the year, other operating expenses on a cash basis would be higher than on an accrual basis, and vice versa. Likewise, if accrued expense liabilities increased during the year, other operating expenses on a cash basis would be lower than on an accrual basis, and vice versa. For Acme Corporation, the amount of cash paid for operating expenses in 2018 was $3,532, as follows:

Other operating expenses	$3,577
Less: Decrease in prepaid expenses	(23)
Less: Increase in other accrued liabilities	(22)
Cash paid for other operating expenses	**$3,532**

EXAMPLE 5

Computing Cash Paid for Other Operating Expenses

Black Ice, a fictitious sportswear manufacturer, reported other operating expenses of $30 million. Prepaid insurance expense increased by $4 million, and accrued utilities payable decreased by $7 million. Insurance and utilities are the only two components of other operating expenses. How much cash did the company pay in other operating expenses?

A $19 million.

B $33 million.

C $41 million.

Solution:

C is correct. Other operating expenses of $30 million plus the increase in prepaid insurance expense of $4 million plus the decrease in accrued utilities payable of $7 million equals $41 million.

10.1.5 *Cash Paid for Interest*

The cash paid for interest is included in operating cash flows under US GAAP and may be included in operating or financing cash flows under IFRS. To determine the cash paid for interest, it is necessary to adjust interest expense by the net change in interest payable for the year. If interest payable increases during the year, then interest expense on an accrual basis will be higher than the amount of cash paid for interest, and vice versa. For Acme Corporation, interest payable decreased by $12, and cash paid for interest was $258, as follows:

Interest expense	$246
Plus: Decrease in interest payable	12
Cash paid for interest	**$258**

Alternatively, cash paid for interest may also be determined by an analysis of the interest payable account, as shown below:

Beginning interest payable	$74
Plus interest expense	246
Minus cash paid for interest	**(258)**
Ending interest payable	$62

10.1.6 *Cash Paid for Income Taxes*

To determine the cash paid for income taxes, it is necessary to adjust the income tax expense amount on the income statement by the net changes in taxes receivable, taxes payable, and deferred income taxes for the year. If taxes receivable or deferred tax assets increase during the year, income taxes on a cash basis will be higher than on an accrual basis, and vice versa. Likewise, if taxes payable or deferred tax liabilities increase during the year, income tax expense on a cash basis will be lower than on an accrual basis, and vice versa. For Acme Corporation, the amount of cash paid for income taxes in 2018 was $1,134, as follows:

Income tax expense	$1,139
Less: Increase in income tax payable	(5)
Cash paid for income taxes	**$1,134**

11 PREPARING THE CASH FLOW STATEMENT: INVESTING ACTIVITIES

f describe the steps in the preparation of direct and indirect cash flow statements, including how cash flows can be computed using income statement and balance sheet data

The second and third steps in preparing the cash flow statement are to determine the total cash flows from investing activities and from financing activities. The presentation of this information is identical, regardless of whether the direct or indirect method is used for operating cash flows.

Purchases and sales of equipment were the only investing activities undertaken by Acme in 2018, as evidenced by the fact that the amounts reported for land and buildings were unchanged during the year. An informational note in Exhibit 7 tells us that Acme *purchased* new equipment in 2018 for a total cost of $1,300. However, the amount of equipment shown on Acme's balance sheet increased by only $243 (ending balance of $8,798 minus beginning balance of $8,555); therefore, Acme must have also *sold or otherwise disposed of* some equipment during the year. To determine the cash inflow from the sale of equipment, we analyze the equipment and accumulated depreciation accounts as well as the gain on the sale of equipment from Exhibits 6 and 7. Assuming that the entire accumulated depreciation is related to equipment, the cash received from sale of equipment is determined as follows.

The historical cost of the equipment sold was $1,057. This amount is determined as follows:

Beginning balance equipment (from balance sheet)	$8,555
Plus equipment purchased (from informational note)	1,300
Minus ending balance equipment (from balance sheet)	(8,798)
Equals historical cost of equipment sold	$1,057

The accumulated depreciation on the equipment sold was $500, determined as follows:

Beginning balance accumulated depreciation (from balance sheet)	$2,891
Plus depreciation expense (from income statement)	1,052
Minus ending balance accumulated depreciation (from balance sheet)	(3,443)
Equals accumulated depreciation on equipment sold	$500

The historical cost information, accumulated depreciation information, and information from the income statement about the gain on the sale of equipment can be used to determine the cash received from the sale.

Historical cost of equipment sold (calculated above)	$1,057
Less accumulated depreciation on equipment sold (calculated above)	(500)
Equals book value of equipment sold	$557
Plus gain on sale of equipment (from the income statement)	205
Equals cash received from sale of equipment	**$762**

EXAMPLE 6

Computing Cash Received from the Sale of Equipment

Copper, Inc., a fictitious brewery and restaurant chain, reported a gain on the sale of equipment of $12 million. In addition, the company's income statement shows depreciation expense of $8 million and the cash flow statement shows capital expenditure of $15 million, all of which was for the purchase of new equipment.

Balance sheet item	12/31/2017	12/31/2018	Change
Equipment	$100 million	$109 million	$9 million
Accumulated depreciation—equipment	$30 million	$36 million	$6 million

Using the above information from the comparative balance sheets, how much cash did the company receive from the equipment sale?

A $12 million.

B $16 million.

C $18 million.

Solution:

B is correct. Selling price (cash inflow) minus book value equals gain or loss on sale; therefore, gain or loss on sale plus book value equals selling price (cash inflow). The amount of gain is given, $12 million. To calculate the book value of the equipment sold, find the historical cost of the equipment and the accumulated depreciation on the equipment.

- Beginning balance of equipment of $100 million plus equipment purchased of $15 million minus ending balance of equipment of $109 million equals historical cost of equipment sold, or $6 million.
- Beginning accumulated depreciation on equipment of $30 million plus depreciation expense for the year of $8 million minus ending balance of accumulated depreciation of $36 million equals accumulated depreciation on the equipment sold, or $2 million.
- Therefore, the book value of the equipment sold was $6 million minus $2 million, or $4 million.
- Because the gain on the sale of equipment was $12 million, the amount of cash received must have been $16 million.

12

PREPARING THE CASH FLOW STATEMENT: FINANCING ACTIVITIES

f describe the steps in the preparation of direct and indirect cash flow statements, including how cash flows can be computed using income statement and balance sheet data

As with investing activities, the presentation of financing activities is identical, regardless of whether the direct or indirect method is used for operating cash flows.

12.1 Long-Term Debt and Common Stock

The change in long-term debt, based on the beginning 2018 (ending 2017) and ending 2018 balances in Exhibit 7, was a decrease of $500. Absent other information, this indicates that Acme retired $500 of long-term debt. Retiring long-term debt is a cash outflow relating to financing activities.

Similarly, the change in common stock during 2018 was a decrease of $600. Absent other information, this indicates that Acme repurchased $600 of its common stock. Repurchase of common stock is also a cash outflow related to financing activity.

12.2 Dividends

Recall the following relationship:

> Beginning retained earnings + Net income − Dividends = Ending retained earnings

Based on this relationship, the amount of cash dividends paid in 2018 can be determined from an analysis of retained earnings, as follows:

Beginning balance of retained earnings (from the balance sheet)	$2,876
Plus net income (from the income statement)	2,210
Minus ending balance of retained earnings (from the balance sheet)	(3,966)
Equals dividends paid	**$1,120**

Note that dividends paid are presented in the statement of changes in equity.

PREPARING THE CASH FLOW STATEMENT: OVERALL STATEMENT OF CASH FLOWS UNDER THE DIRECT METHOD

13

f describe the steps in the preparation of direct and indirect cash flow statements, including how cash flows can be computed using income statement and balance sheet data

Exhibit 8 summarizes the information about Acme's operating, investing, and financing cash flows in the statement of cash flows. At the bottom of the statement, the total net change in cash is shown to be a decrease of $152 (from $1,163 to $1,011). This decrease can also be seen on the comparative balance sheet in Exhibit 7. The cash provided by operating activities of $2,606 was adequate to cover the net cash used in investing activities of $538; however, the company's debt repayments, cash payments for dividends, and repurchase of common stock (i.e., its financing activities) of $2,220 resulted in an overall decrease in cash of $152.

Exhibit 8 Acme Corporation Cash Flow Statement (Direct Method) for Year Ended 31 December 2018

Cash flow from operating activities:	
Cash received from customers	$23,543
Cash paid to suppliers	(11,900)
Cash paid to employees	(4,113)
Cash paid for other operating expenses	(3,532)
Cash paid for interest	(258)
Cash paid for income tax	(1,134)
Net cash provided by operating activities	2,606
Cash flow from investing activities:	
Cash received from sale of equipment	762
Cash paid for purchase of equipment	(1,300)
Net cash used for investing activities	(538)

(continued)

Exhibit 8 (Continued)

Cash flow from financing activities:	
Cash paid to retire long-term debt	(500)
Cash paid to retire common stock	(600)
Cash paid for dividends	(1,120)
Net cash used for financing activities	(2,220)
Net increase (decrease) in cash	(152)
Cash balance, 31 December 2017	1,163
Cash balance, 31 December 2018	$1,011

14

PREPARING THE CASH FLOW STATEMENT: OVERALL STATEMENT OF CASH FLOWS UNDER THE INDIRECT METHOD

f describe the steps in the preparation of direct and indirect cash flow statements, including how cash flows can be computed using income statement and balance sheet data

Using the alternative approach to reporting cash from operating activities, the indirect method, we will present the same amount of cash provided by operating activities. Under this approach, we reconcile Acme's net income of $2,210 to its operating cash flow of $2,606.

To perform this reconciliation, net income is adjusted for the following: a) any non-operating activities, b) any non-cash expenses, and c) changes in operating working capital items.

The only non-operating activity in Acme's income statement, the sale of equipment, resulted in a gain of $205. This amount is removed from the operating cash flow section; the cash effects of the sale are shown in the investing section.

Acme's only non-cash expense was depreciation expense of $1,052. Under the indirect method, depreciation expense must be added back to net income because it was a non-cash deduction in the calculation of net income.

Changes in working capital accounts include increases and decreases in the current operating asset and liability accounts. The changes in these accounts arise from applying accrual accounting; that is, recognizing revenues when they are earned and expenses when they are incurred instead of when the cash is received or paid. To make the working capital adjustments under the indirect method, any increase in a current operating asset account is subtracted from net income and a net decrease is added to net income. As described above, the increase in accounts receivable, for example, resulted from Acme recording income statement revenue higher than the amount of cash received from customers; therefore, to reconcile back to operating cash flow, that increase in accounts receivable must be deducted from net income. For current operating liabilities, a net increase is added to net income and a net decrease is subtracted from net income. As described above, the increase in wages payable, for example, resulted from Acme recording income statement expenses higher than the amount of cash paid to employees.

Exhibit 9 presents a tabulation of the most common types of adjustments that are made to net income when using the indirect method to determine net cash flow from operating activities.

Exhibit 9 Adjustments to Net Income Using the Indirect Method

Additions

- Non-cash items
  - Depreciation expense of tangible assets
  - Amortisation expense of intangible assets
  - Depletion expense of natural resources
  - Amortisation of bond discount
- Non-operating losses
  - Loss on sale or write-down of assets
  - Loss on retirement of debt
  - Loss on investments accounted for under the equity method
- Increase in deferred income tax liability
- Changes in working capital resulting from accruing higher amounts for expenses than the amounts of cash payments or lower amounts for revenues than the amounts of cash receipts
  - Decrease in current operating assets (e.g., accounts receivable, inventory, and prepaid expenses)
  - Increase in current operating liabilities (e.g., accounts payable and accrued expense liabilities)

Subtractions

- Non-cash items (e.g., amortisation of bond premium)
- Non-operating items
  - Gain on sale of assets
  - Gain on retirement of debt
  - Income on investments accounted for under the equity method
- Decrease in deferred income tax liability
- Changes in working capital resulting from accruing lower amounts for expenses than for cash payments or higher amounts for revenues than for cash receipts
  - Increase in current operating assets (e.g., accounts receivable, inventory, and prepaid expenses)
  - Decrease in current operating liabilities (e.g., accounts payable and accrued expense liabilities)

Accordingly, for Acme Corporation, the $55 increase in accounts receivable and the $707 increase in inventory are subtracted from net income and the $23 decrease in prepaid expenses is added to net income. For Acme's current liabilities, the increases in accounts payable, salary and wage payable, income tax payable, and other accrued liabilities ($263, $10, $5, and $22, respectively) are added to net income and the $12 decrease in interest payable is subtracted from net income. Exhibit 10 presents the cash flow statement for Acme Corporation under the indirect method by using the information that we have determined from our analysis of the income statement and the comparative balance sheets. Note that the investing and financing sections are identical to the statement of cash flows prepared using the direct method.

Exhibit 10 **Acme Corporation Cash Flow Statement (Indirect Method) Year Ended 31 December 2018**

Cash flow from operating activities:	
Net income	$2,210
Depreciation expense	1,052
Gain on sale of equipment	(205)
Increase in accounts receivable	(55)
Increase in inventory	(707)
Decrease in prepaid expenses	23
Increase in accounts payable	263
Increase in salary and wage payable	10
Decrease in interest payable	(12)
Increase in income tax payable	5
Increase in other accrued liabilities	22
Net cash provided by operating activities	2,606
Cash flow from investing activities:	
Cash received from sale of equipment	762
Cash paid for purchase of equipment	(1,300)
Net cash used for investing activities	(538)
Cash flow from financing activities:	
Cash paid to retire long-term debt	(500)
Cash paid to retire common stock	(600)
Cash paid for dividends	(1,120)
Net cash used for financing activities	(2,220)
Net decrease in cash	(152)
Cash balance, 31 December 2017	1,163
Cash balance, 31 December 2018	$1,011

EXAMPLE 7

Adjusting Net Income to Compute Operating Cash Flow

Based on the following information for Pinkerly Inc., a fictitious company, what are the total adjustments that the company would make to net income in order to derive operating cash flow?

	Year Ended		
Income statement item		12/31/2018	
Net income		$30 million	
Depreciation		$7 million	
Balance sheet item	12/31/2017	12/31/2018	Change
Accounts receivable	$15 million	$30 million	$15 million

Inventory	$16 million	$13 million	($3 million)
Accounts payable	$10 million	$20 million	$10 million

A Add $5 million.

B Add $21 million.

C Subtract $9 million.

Solution:

A is correct. To derive operating cash flow, the company would make the following adjustments to net income: add depreciation (a non-cash expense) of $7 million; add the decrease in inventory of $3 million; add the increase in accounts payable of $10 million; and subtract the increase in accounts receivable of $15 million. Total additions of $20 million and total subtractions of $15 million result in net total additions of $5 million.

CONVERSION OF CASH FLOWS FROM THE INDIRECT TO DIRECT METHOD

15

g demonstrate the conversion of cash flows from the indirect to direct method

An analyst may desire to review direct-format operating cash flow to review trends in cash receipts and payments (such as cash received from customers or cash paid to suppliers). If a direct-format statement is not available, cash flows from operating activities reported under the indirect method can be converted to the direct method. Accuracy of conversion depends on adjustments using data available in published financial reports. The method described here is sufficiently accurate for most analytical purposes.

The three-step conversion process is demonstrated for Acme Corporation in Exhibit 11. Referring again to Exhibits 6 and 7 for Acme Corporation's income statement and balance sheet information, begin by disaggregating net income of $2,210 into total revenues and total expenses (Step 1). Next, remove any non-operating and non-cash items (Step 2). For Acme, we therefore remove the non-operating gain on the sale of equipment of $205 and the non-cash depreciation expense of $1,052. Then, convert accrual amounts of revenues and expenses to cash flow amounts of receipts and payments by adjusting for changes in working capital accounts (Step 3). The results of these adjustments are the items of information for the direct format of operating cash flows. These line items are shown as the results of Step 3.

Exhibit 11 Conversion from the Indirect to the Direct Method

Step 1	Total revenues	$23,803
Aggregate all revenue and all expenses	Total expenses	21,593
	Net income	$2,210
Step 2	Total revenue less noncash item revenues:	

(continued)

Exhibit 11 (Continued)

Remove all noncash items from aggregated revenues and expenses and break out remaining items into relevant cash flow items	($23,803 – $205) =	$23,598
	Revenue	$23,598
	Total expenses less noncash item expenses:	
	($21,593 – $1,052) =	$20,541
	Cost of goods sold	$11,456
	Salary and wage expenses	4,123
	Other operating expenses	3,577
	Interest expense	246
	Income tax expense	1,139
	Total	$20,541
Step 3 Convert accrual amounts to cash flow amounts by adjusting for working capital changes	Cash received from customers[a]	$23,543
	Cash paid to suppliers[b]	(11,900)
	Cash paid to employees[c]	(4,113)
	Cash paid for other operating expenses[d]	(3,532)
	Cash paid for interest[e]	(258)
	Cash paid for income tax[f]	(1,134)
	Net cash provided by operating activities	$2,606

Calculations for Step 3:

[a] Revenue of $23,598 less increase in accounts receivable of $55.

[b] Cost of goods sold of $11,456 plus increase in inventory of $707 less increase in accounts payable of $263.

[c] Salary and wage expense of $4,123 less increase in salary and wage payable of $10.

[d] Other operating expenses of $3,577 less decrease in prepaid expenses of $23 less increase in other accrued liabilities of $22.

[e] Interest expense of $246 plus decrease in interest payable of $12.

[f] Income tax expense of $1,139 less increase in income tax payable of $5.

16

CASH FLOW STATEMENT ANALYSIS: EVALUATION OF SOURCES AND USES OF CASH

h analyze and interpret both reported and common-size cash flow statements

The analysis of a company's cash flows can provide useful information for understanding a company's business and earnings and for predicting its future cash flows. This section describes tools and techniques for analyzing the statement of cash flows, including the analysis of sources and uses of cash and cash flow, common-size analysis, and calculation of free cash flow measures and cash flow ratios.

16.1 Evaluation of the Sources and Uses of Cash

Evaluation of the cash flow statement should involve an overall assessment of the sources and uses of cash between the three main categories as well as an assessment of the main drivers of cash flow within each category, as follows:

1 Evaluate where the major sources and uses of cash flow are between operating, investing, and financing activities.

2. Evaluate the primary determinants of operating cash flow.
3. Evaluate the primary determinants of investing cash flow.
4. Evaluate the primary determinants of financing cash flow.

Step 1

The major sources of cash for a company can vary with its stage of growth. For a mature company, it is expected and desirable that operating activities are the primary source of cash flows. Over the long term, a company must generate cash from its operating activities. If operating cash flow were consistently negative, a company would need to borrow money or issue stock (financing activities) to fund the shortfall. Eventually, these providers of capital need to be repaid from operations or they will no longer be willing to provide capital. Cash generated from operating activities can be used in either investing or financing activities. If the company has good opportunities to grow the business or other investment opportunities, it is desirable to use the cash in investing activities. If the company does not have profitable investment opportunities, the cash should be returned to capital providers, a financing activity. For a new or growth stage company, operating cash flow may be negative for some period of time as it invests in such assets as inventory and receivables (extending credit to new customers) in order to grow the business. This situation is not sustainable over the long term, so eventually the cash must start to come primarily from operating activities so that capital can be returned to the providers of capital. Lastly, it is desirable that operating cash flows are sufficient to cover capital expenditures (in other words, the company has free cash flow as discussed further in Section 18). In summary, major points to consider at this step are:

- What are the major sources and uses of cash flow?
- Is operating cash flow positive and sufficient to cover capital expenditures?

Step 2

Turning to the operating section, the analysts should examine the most significant determinants of operating cash flow. Companies need cash for use in operations (for example, to hold receivables and inventory and to pay employees and suppliers) and receive cash from operating activities (for example, payments from customers). Under the indirect method, the increases and decreases in receivables, inventory, payables, and so on can be examined to determine whether the company is using or generating cash in operations and why. It is also useful to compare operating cash flow with net income. For a mature company, because net income includes non-cash expenses (depreciation and amortisation), it is expected and desirable that operating cash flow exceeds net income. The relationship between net income and operating cash flow is also an indicator of earnings quality. If a company has large net income but poor operating cash flow, it may be a sign of poor earnings quality. The company may be making aggressive accounting choices to increase net income but not be generating cash for its business. You should also examine the variability of both earnings and cash flow and consider the impact of this variability on the company's risk as well as the ability to forecast future cash flows for valuation purposes. In summary:

- What are the major determinants of operating cash flow?
- Is operating cash flow higher or lower than net income? Why?
- How consistent are operating cash flows?

Step 3

Within the investing section, you should evaluate each line item. Each line item represents either a source or use of cash. This enables you to understand where the cash is being spent (or received). This section will tell you how much cash is being invested for the future in property, plant, and equipment; how much is used to acquire entire companies; and how much is put aside in liquid investments, such as stocks and bonds. It will also tell you how much cash is being raised by selling these types of assets. If the company is making major capital investments, you should consider where the cash is coming from to cover these investments (e.g., is the cash coming from excess operating cash flow or from the financing activities described in Step 4). If assets are being sold, it is important to determine why and to assess the effects on the company.

Step 4

Within the financing section, you should examine each line item to understand whether the company is raising capital or repaying capital and what the nature of its capital sources are. If the company is borrowing each year, you should consider when repayment may be required. This section will also present dividend payments and repurchases of stock that are alternative means of returning capital to owners. It is important to assess why capital is being raised or repaid.

We now provide an example of a cash flow statement evaluation.

EXAMPLE 8

Analysis of the Cash Flow Statement

Derek Yee, CFA, is preparing to forecast cash flow for Groupe Danone as an input into his valuation model. He has asked you to evaluate the historical cash flow statement of Groupe Danone, which is presented in Exhibit 12. Groupe Danone prepares its financial statements in conformity with IFRS. Note that Groupe Danone presents the most recent period on the right. Exhibit 13 presents excerpts from Danone's 2017 Registration Document.

Yee would like answers to the following questions:

- What are the major sources of cash for Groupe Danone?
- What are the major uses of cash for Groupe Danone?
- Is cash flow from operating activities sufficient to cover capital expenditures?
- What is the relationship between net income and cash flow from operating activities?
- What types of financing cash flows does Groupe Danone have?

Exhibit 12 Groupe Danone Consolidated Financial Statements Consolidated Statements of Cash Flows (in € Millions)

Years Ended 31 December	2016	2017
Net income	1,827	2,563
Share of profits of associates net of dividends received	52	(54)
Depreciation, amortization and impairment of tangible and intangible assets	786	974
Increases in (reversals of) provisions	51	153
Change in deferred taxes	(65)	(353)

Exhibit 12 (Continued)

Years Ended 31 December	**2016**	**2017**
(Gains) losses on disposal of property, plant and equipment and financial investments	(74)	(284)
Expense related to Group performance shares	24	22
Cost of net financial debt	149	265
Net interest paid	(148)	(186)
Net change in interest income (expense)	—	80
Other components with no cash impact	13	(15)
Cash flows provided by operating activities, before changes in net working capital	**2,615**	**3,085**
(Increase) decrease in inventories	(24)	(122)
(Increase) decrease in trade receivables	(110)	(190)
Increase (decrease) in trade payables	298	145
Changes in other receivables and payables	(127)	40
Change in other working capital requirements	37	(127)
Cash flows provided by (used in) operating activities	**2,652**	**2,958**
Capital expenditure	(925)	(969)
Proceeds from the disposal of property, plant and equipment	27	45
Net cash outflows on purchases of subsidiaries and financial investments	(66)	(10,949)
Net cash inflows on disposal of subsidiaries and financial investments	110	441
(Increase) decrease in long-term loans and other long-term financial assets	6	(4)
Cash flows provided by (used in) investing activities	**(848)**	**(11,437)**
Increase in capital and additional paid-in capital	46	47
Purchases of treasury stock (net of disposals) and DANONE call options	32	13
Issue of perpetual subordinated debt securities	—	1,245
Interest on perpetual subordinated debt securities	—	—
Dividends paid to Danone shareholders	(985)	(279)
Buyout of non-controlling interests	(295)	(107)
Dividends paid	(94)	(86)
Contribution from non-controlling interests to capital increases	6	1
Transactions with non-controlling interests	(383)	(193)
Net cash flows on hedging derivatives	50	(52)
Bonds issued during the period	11,237	—
Bonds repaid during the period	(638)	(1,487)
Net cash flows from other current and non-current financial debt	(442)	(564)
Net cash flows from short-term investments	(10,531)	9,559
Cash flows provided by (used in) financing activities	**(1,616)**	**8,289**
Effect of exchange rate and other changes	(151)	272
Increase (decrease) in cash and cash equivalents	**38**	**81**
Cash and cash equivalents at beginning of period	519	557
Cash and cash equivalents at end of period	557	638
Supplemental disclosures		

(continued)

Exhibit 12 (Continued)

Years Ended 31 December	2016	2017
Income tax payments during the year	(891)	(1,116)

Note: the numbers in the consolidated statement of cash flows were derived straight from company filings; some sub-totals may not sum exactly due to rounding by the company.

Exhibit 13 Groupe Danone Excerpt from 2017 Registration Statement

Excerpt from Footnote 2 to the financial statements:

… On July 7, 2016, Danone announced the signing of an agreement to acquire The WhiteWave Foods Company ("WhiteWave"), the global leader in plant-based foods and beverages and organic produce. The acquisition in cash, for USD 56.25 per share, represented, as of the date of the agreement, a total enterprise value of approximately USD 12.5 billion, including debt and certain other WhiteWave liabilities. …

"Acquisition expenses recognized in Danone's consolidated financial statements totaled €51 million before tax, of which €48 million was recognized in 2016 in Other operating income (expense), with the balance recognized in 2017.

"WhiteWave's contribution to 2017 consolidated sales totaled €2.7 billion. Had the transaction been completed on January 1, 2017, the Group's 2017 consolidated sales would have been €25.7 billion, with recurring operating income of €3.6 billion.

"Meanwhile, integration expenses for the period totaled €91 million, recognized under Other operating income (expense)…

Excerpt from Overview of Activities:

"… As part of its transformation plan aimed at ensuring a safe journey to deliver strong, profitable and sustainable growth, Danone set objectives for 2020 that include like-for-like sales growth between 4% and 5% …. a recurring operating margin of over 16% in 2020 … Finally, Danone will continue to focus on growing its free cash flow, which will contribute to financial deleverage with an objective of a ratio of Net debt/EBITDA below 3x in 2020. Danone is committed to reaching a ROIC level around 12% in 2020."

Solution:

The major categories of cash flows can be summarized as follows (in € millions):

	2016	2017
Cash flows provided by operating activities	2,652	2,958
Cash flows provided by (used in) investing activities	(848)	(11,437)
Cash flows provided by (used in) financing activities	(1,616)	8,289

Exchange rate effects on cash	(151)	272
Increase in cash	38	81

The primary source of cash for Groupe Danone in 2016 is operating activities. In both 2016 and 2017, there was sufficient operating cash flow to cover usual capital expenditures, and operating cash flow exceeded net income. Evaluating the five prior years [not shown in this Example], you confirm that Danone typically derives most of its cash from operating activities, reports operating cash flow greater than net income, and generates sufficient operating cash flow to cover capital expenditures.

The fact that the primary source of cash is from operations is positive and desirable for a mature company. Additionally, the fact that operating cash flow exceeds net income in both years is a positive sign. Finally, operating cash flows exceed normal capital expenditures, indicating that the company can fund capital expenditures from operations.

In 2017, however, the primary source of cash was financing activities, and the investing section shows significant use of cash for purchase of subsidiaries within investing activities. Footnotes disclose a major acquisition with an aggregate value of €12.5 billion, some of which was funded through proceeds from an earlier bond issuance, which appears as a financing cash flow in the financing section for 2016.

For purposes of Yee's cash flow forecast, the company's targets for free cash flow and debt reduction—as well as disclosures concerning the acquisition's impact on 2017 operating results—are potentially helpful.

CASH FLOW STATEMENT ANALYSIS: COMMON SIZE ANALYSIS

17

h analyze and interpret both reported and common-size cash flow statements

In common-size analysis of a company's income statement, each income and expense line item is expressed as a percentage of net revenues (net sales). For the common-size balance sheet, each asset, liability, and equity line item is expressed as a percentage of total assets. For the common-size cash flow statement, there are two alternative approaches. The first approach is to express each line item of cash inflow (outflow) as a percentage of total inflows (outflows) of cash, and the second approach is to express each line item as a percentage of net revenue.

Exhibit 14 demonstrates the total cash inflows/total cash outflows method for Acme Corporation. Under this approach, each of the cash inflows is expressed as a percentage of the total cash inflows, whereas each of the cash outflows is expressed as a percentage of the total cash outflows. In Panel A, Acme's common-size statement is based on a cash flow statement using the direct method of presenting operating cash flows. Operating cash inflows and outflows are separately presented on the cash flow statement, and therefore, the common-size cash flow statement shows each of these operating inflows (outflows) as a percentage of total inflows (outflows). In Panel B, Acme's common-size statement is based on a cash flow statement using the indirect method of presenting operating cash flows. When a cash flow statement has been presented using the indirect method, operating cash inflows and outflows are not separately presented; therefore, the common-size cash flow statement shows only the net operating cash flow (net cash provided by or used in operating activities) as a percentage of total inflows or outflows, depending on whether the net amount was a cash inflow or outflow. Because Acme's net operating cash flow is positive, it is shown as a percentage of total inflows.

Exhibit 14 Acme Corporation Common-Size Cash Flow Statement Year Ended 31 December 2018

Panel A. Direct Format for Cash Flow

Inflows		Percentage of Total Inflows
Receipts from customers	$23,543	96.86%
Sale of equipment	762	3.14
Total	$24,305	100.00%

Outflows		Percentage of Total Outflows
Payments to suppliers	$11,900	48.66%
Payments to employees	4,113	16.82
Payments for other operating expenses	3,532	14.44
Payments for interest	258	1.05
Payments for income tax	1,134	4.64
Purchase of equipment	1,300	5.32
Retirement of long-term debt	500	2.04
Retirement of common stock	600	2.45
Dividend payments	1,120	4.58
Total	$24,457	100.00%
Net increase (decrease) in cash	($152)	

Panel B. Indirect Format for Cash Flow

Inflows		Percentage of Total Inflows
Net cash provided by operating activities	$2,606	77.38%
Sale of equipment	762	22.62
Total	$3,368	100.00%

Outflows		Percentage of Total Outflows
Purchase of equipment	$1,300	36.93%
Retirement of long-term debt	500	14.20
Retirement of common stock	600	17.05
Dividend payments	1,120	31.82

Exhibit 14 (Continued)

Outflows		Percentage of Total Outflows
Total	$3,520	100.00%
Net increase (decrease) in cash	($152)	

Exhibit 15 demonstrates the net revenue common-size cash flow statement for Acme Corporation. Under the net revenue approach, each line item in the cash flow statement is shown as a percentage of net revenue. The common-size statement in this exhibit has been developed based on Acme's cash flow statement using the indirect method for operating cash flows and using net revenue of $23,598 as shown in Exhibit 6. Each line item of the reconciliation between net income and net operating cash flows is expressed as a percentage of net revenue. The common-size format makes it easier to see trends in cash flow rather than just looking at the total amount. This method is also useful to the analyst in forecasting future cash flows because individual items in the common-size statement (e.g., depreciation, fixed capital expenditures, debt borrowing, and repayment) are expressed as a percentage of net revenue. Thus, once the analyst has forecast revenue, the common-size statement provides a basis for forecasting cash flows for those items with an expected relation to net revenue.

Exhibit 15 Acme Corporation Common-Size Cash Flow Statement: Indirect Format Year Ended 31 December 2018

		Percentage of Net Revenue
Cash flow from operating activities:		
Net income	$2,210	9.37%
Depreciation expense	1,052	4.46
Gain on sale of equipment	(205)	(0.87)
Increase in accounts receivable	(55)	(0.23)
Increase in inventory	(707)	(3.00)
Decrease in prepaid expenses	23	0.10
Increase in accounts payable	263	1.11
Increase in salary and wage payable	10	0.04
Decrease in interest payable	(12)	(0.05)
Increase in income tax payable	5	0.02
Increase in other accrued liabilities	22	0.09
Net cash provided by operating activities	$2,606	11.04%
Cash flow from investing activities:		
Cash received from sale of equipment	$762	3.23%
Cash paid for purchase of equipment	(1,300)	(5.51)
Net cash used for investing activities	$(538)	(2.28)%
Cash flow from financing activities:		

(continued)

Exhibit 15 (Continued)

		Percentage of Net Revenue
Cash paid to retire long-term debt	$(500)	(2.12)%
Cash paid to retire common stock	(600)	(2.54)
Cash paid for dividends	(1,120)	(4.75)
Net cash used for financing activities	$(2,220)	(9.41)%
Net decrease in cash	$(152)	(0.64)%

EXAMPLE 9

Analysis of a Common-Size Cash Flow Statement

Andrew Potter is examining an abbreviated common-size cash flow statement for Apple Inc., a multinational technology company. The common-size cash flow statement was prepared by dividing each line item by total net sales for the same year.

Apple Inc. Common Size Statements OF Cash Flows as Percentage of Total Net Sales

	12 Months Ended		
	30 Sep. 2017	24 Sep. 2016	26 Sep. 2015
Statement of Cash Flows [Abstract]			
Operating activities:			
Net income	21.1%	21.2%	22.8%
Adjustments to reconcile net income to cash generated by operating activities:			
Depreciation and amortization	4.4%	4.9%	4.8%
Share-based compensation expense	2.1%	2.0%	1.5%
Deferred income tax expense	2.6%	2.3%	0.6%
Other	–0.1%	0.2%	0.2%
Changes in operating assets and liabilities:			
Accounts receivable, net	–0.9%	0.2%	0.2%
Inventories	–1.2%	0.1%	–0.1%
Vendor non-trade receivables	–1.9%	0.0%	–1.6%
Other current and non-current assets	–2.3%	0.5%	–0.1%
Accounts payable	4.2%	0.9%	2.1%
Deferred revenue	–0.3%	–0.7%	0.4%

(Continued)

	12 Months Ended		
	30 Sep. 2017	**24 Sep. 2016**	**26 Sep. 2015**
Other current and non-current liabilities	–0.1%	–0.9%	3.9%
Cash generated by operating activities	27.7%	30.5%	34.8%
Investing activities:			
Purchases of marketable securities	–69.6%	–66.0%	–71.2%
Proceeds from maturities of marketable securities	13.9%	9.9%	6.2%
Proceeds from sales of marketable securities	41.3%	42.0%	46.0%
Payments made in connection with business acquisitions, net	–0.1%	–0.1%	–0.1%
Payments for acquisition of property, plant and equipment	–5.4%	–5.9%	–4.8%
Payments for acquisition of intangible assets	–0.2%	–0.4%	–0.1%
Payments for strategic investments, net	–0.2%	–0.6%	0.0%
Other	0.1%	–0.1%	0.0%
Cash used in investing activities	–20.3%	–21.3%	–24.1%
Financing activities:			
Proceeds from issuance of common stock	0.2%	0.2%	0.2%
Excess tax benefits from equity awards	0.3%	0.2%	0.3%
Payments for taxes related to net share settlement of equity awards	–0.8%	–0.7%	–0.6%
Payments for dividends and dividend equivalents	–5.6%	–5.6%	–4.9%
Repurchases of common stock	–14.4%	–13.8%	–15.1%
Proceeds from issuance of term debt, net	12.5%	11.6%	—
Repayments of term debt	–1.5%	–1.2%	0.0%
Change in commercial paper, net	1.7%	–0.2%	0.9%
Cash used in financing activities	–7.6%	–9.5%	–7.6%
Increase/(Decrease) in cash and cash equivalents	–0.1%	–0.3%	3.1%

Based on the information in the above exhibit:

1 Discuss the significance of

- **A** depreciation and amortization.
- **B** capital expenditures.

2 Compare Apple's operating cash flow as a percentage of revenue with Apple's net profit margin.

3 Discuss Apple's use of its positive operating cash flow.

Solution to 1:

A Apple's depreciation and amortization expense was consistently just less than 5% of total net revenue in 2015 and 2016, declining to 4.4% in 2017.

B Apple's level of capital expenditures is greater than depreciation and amortization in 2016 and 2017 whereas it was at about the same level as depreciation and amortization in 2015. In 2017 capital expenditures approached 6%. This is an indication that Apple is doing more than replacing property, plant, and equipment, and is expanding those investments. With cash generated from operating activities exceeding 27% of sales in every year, however, Apple has more than enough cash flow from operations to fund these expenditures.

Solution to 2:

Apple's operating cash flow as a percentage of sales is much higher than net profit margin in every year. This gap appears to be declining however over the three year period. In 2015 net profit margin was 22.8% while operating cash flow as a percentage of sales was 34.8%. By 2017 the net profit margin declined slightly to 21.1% while the operating cash flow as a percentage of sales declined more to 27.7%. The primary difference appears to have been an increase in the level of receivables and inventory purchases, somewhat offset by an increase in accounts payable.

Solution to 3:

Apple has a very strong cash flow statement. Apple generates a large amount of operating cash flow in every year, exceeding net income. This cash flow is used for relatively modest purchases of property, plant and equipment, substantial purchases of marketable securities (investments), dividend payments and repurchases of its own stock.

18

CASH FLOW STATEMENT ANALYSIS: FREE CASH FLOW TO FIRM AND FREE CASH FLOW TO EQUITY

i. calculate and interpret free cash flow to the firm, free cash flow to equity, and performance and coverage cash flow ratios

It was mentioned earlier that it is desirable that operating cash flows are sufficient to cover capital expenditures. The excess of operating cash flow over capital expenditures is known generically as **free cash flow**. For purposes of valuing a company or its equity securities, an analyst may want to determine and use other cash flow measures, such as free cash flow to the firm (FCFF) or free cash flow to equity (FCFE).

FCFF is the cash flow available to the company's suppliers of debt and equity capital after all operating expenses (including income taxes) have been paid and necessary investments in working capital and fixed capital have been made. FCFF can be computed starting with net income as

$$\text{FCFF} = \text{NI} + \text{NCC} + \text{Int}(1 - \text{Tax rate}) - \text{FCInv} - \text{WCInv}$$

where

NI = Net income
NCC = Non-cash charges (such as depreciation and amortisation)
Int = Interest expense
FCInv = Capital expenditures (fixed capital, such as equipment)
WCInv = Working capital expenditures

The reason for adding back interest is that FCFF is the cash flow available to the suppliers of debt capital as well as equity capital. Conveniently, FCFF can also be computed from cash flow from operating activities as

$$\text{FCFF} = \text{CFO} + \text{Int}(1 - \text{Tax rate}) - \text{FCInv}$$

CFO represents cash flow from operating activities under US GAAP or under IFRS where the company has included interest paid in operating activities. If interest paid was included in financing activities, then CFO does not have to be adjusted for Int(1 – Tax rate). Under IFRS, if the company has placed interest and dividends received in investing activities, these should be added back to CFO to determine FCFF. Additionally, if dividends paid were subtracted in the operating section, these should be added back in to compute FCFF.

The computation of FCFF for Acme Corporation (based on the data from Exhibits 6, 7, and 8) is as follows:

CFO	$2,606
Plus: Interest paid times (1 – income tax rate)	
{$258 [1 – 0.34[a]]}	170
Less: Net investments in fixed capital	
($1,300 – $762)	(538)
FCFF	$2,238

---

[a] Income tax rate of 0.34 = (Tax expense ÷ Pretax income) = ($1,139 ÷ $3,349).

FCFE is the cash flow available to the company's common stockholders after all operating expenses and borrowing costs (principal and interest) have been paid and necessary investments in working capital and fixed capital have been made. FCFE can be computed as

$$\text{FCFE} = \text{CFO} - \text{FCInv} + \text{Net borrowing}$$

When net borrowing is negative, debt repayments exceed receipts of borrowed funds. In this case, FCFE can be expressed as

$$\text{FCFE} = \text{CFO} - \text{FCInv} - \text{Net debt repayment}$$

The computation of FCFE for Acme Corporation (based on the data from Exhibits 6, 7, and 8) is as follows:

CFO	$2,606
Less: Net investments in fixed capital ($1,300 – $762)	(538)
Less: Debt repayment	(500)
FCFE	$1,568

Positive FCFE means that the company has an excess of operating cash flow over amounts needed for capital expenditures and repayment of debt. This cash would be available for distribution to owners.

19

CASH FLOW STATEMENT ANALYSIS: CASH FLOW RATIOS

i. calculate and interpret free cash flow to the firm, free cash flow to equity, and performance and coverage cash flow ratios

The statement of cash flows provides information that can be analyzed over time to obtain a better understanding of the past performance of a company and its future prospects. This information can also be effectively used to compare the performance and prospects of different companies in an industry and of different industries. There are several ratios based on cash flow from operating activities that are useful in this analysis. These ratios generally fall into cash flow performance (profitability) ratios and cash flow coverage (solvency) ratios. Exhibit 15 summarizes the calculation and interpretation of some of these ratios.

Exhibit 16 Cash Flow Ratios

Performance Ratios	Calculation	What It Measures
Cash flow to revenue	CFO ÷ Net revenue	Operating cash generated per dollar of revenue
Cash return on assets	CFO ÷ Average total assets	Operating cash generated per dollar of asset investment
Cash return on equity	CFO ÷ Average shareholders' equity	Operating cash generated per dollar of owner investment
Cash to income	CFO ÷ Operating income	Cash generating ability of operations
Cash flow per share[a]	(CFO − Preferred dividends) ÷ Number of common shares outstanding	Operating cash flow on a per-share basis

Coverage Ratios	Calculation	What It Measures
Debt coverage	CFO ÷ Total debt	Financial risk and financial leverage
Interest coverage[b]	(CFO + Interest paid + Taxes paid) ÷ Interest paid	Ability to meet interest obligations
Reinvestment	CFO ÷ Cash paid for long-term assets	Ability to acquire assets with operating cash flows
Debt payment	CFO ÷ Cash paid for long-term debt repayment	Ability to pay debts with operating cash flows

Exhibit 16 (Continued)

Coverage Ratios	Calculation	What It Measures
Dividend payment	CFO ÷ Dividends paid	Ability to pay dividends with operating cash flows
Investing and financing	CFO ÷ Cash outflows for investing and financing activities	Ability to acquire assets, pay debts, and make distributions to owners

Notes:

[a] If the company reports under IFRS and includes total dividends paid as a use of cash in the operating section, total dividends should be added back to CFO as reported and then preferred dividends should be subtracted. Recall that CFO reported under US GAAP and IFRS may differ depending on the treatment of interest and dividends, received and paid.

[b] If the company reports under IFRS and included interest paid as a use of cash in the financing section, then interest paid should not be added back to the numerator.

EXAMPLE 10

A Cash Flow Analysis of Comparables

Andrew Potter is comparing the cash-flow-generating ability of Microsoft with that of Apple Inc. He collects information from the companies' annual reports and prepares the following table.

Cash Flow from Operating Activities as a Percentage of Total Net Revenue

	2017 (%)	2016 (%)	2015 (%)
Microsoft	43.9	39.1	31.7
Apple Inc.	27.7	30.5	34.8

As a Percentage of Average Total Assets

	2017 (%)	2016 (%)	2015 (%)
Microsoft	18.2	18.1	17.1
Apple Inc.	18.2	21.5	31.1

What is Potter likely to conclude about the relative cash-flow-generating ability of these two companies?

Solution:

On both measures—operating cash flow divided by revenue and operating cash flow divided by assets—both companies have overall strong results. However, Microsoft has higher cash flow from operating activities as a percentage of revenues in both 2016 and 2017. Further, Microsoft has an increasing trend. While Apple had a higher operating cash flow as a percent of revenue in 2015 compared to Microsoft, it has had a declining trend and was below Microsoft in the two more recent years. Microsoft's operating cash flow relative to assets is the same as Apple's in 2017 and relatively stable with a slight increase since 2015. Apple started the three years with a much stronger ratio but saw a declining trend such that its ratio is now at the same level as Microsoft. We should note that this ratio is heavily influenced by substantial investments in financial instruments that Apple has made over the years due to its strong historic cash flow.

SUMMARY

The cash flow statement provides important information about a company's cash receipts and cash payments during an accounting period as well as information about a company's operating, investing, and financing activities. Although the income statement provides a measure of a company's success, cash and cash flow are also vital to a company's long-term success. Information on the sources and uses of cash helps creditors, investors, and other statement users evaluate the company's liquidity, solvency, and financial flexibility. Key concepts are as follows:

- Cash flow activities are classified into three categories: operating activities, investing activities, and financing activities. Significant non-cash transaction activities (if present) are reported by using a supplemental disclosure note to the cash flow statement.
- Cash flow statements under IFRS and US GAAP are similar; however, IFRS provide companies with more choices in classifying some cash flow items as operating, investing, or financing activities.
- Companies can use either the direct or the indirect method for reporting their operating cash flow:
  - The direct method discloses operating cash inflows by source (e.g., cash received from customers, cash received from investment income) and operating cash outflows by use (e.g., cash paid to suppliers, cash paid for interest) in the operating activities section of the cash flow statement.
  - The indirect method reconciles net income to operating cash flow by adjusting net income for all non-cash items and the net changes in the operating working capital accounts.
- The cash flow statement is linked to a company's income statement and comparative balance sheets and to data on those statements.
- Although the indirect method is most commonly used by companies, an analyst can generally convert it to an approximation of the direct format by following a simple three-step process.
- An evaluation of a cash flow statement should involve an assessment of the sources and uses of cash and the main drivers of cash flow within each category of activities.
- The analyst can use common-size statement analysis for the cash flow statement. Two approaches to developing the common-size statements are the total cash inflows/total cash outflows method and the percentage of net revenues method.
- The cash flow statement can be used to determine free cash flow to the firm (FCFF) and free cash flow to equity (FCFE).
- The cash flow statement may also be used in financial ratios that measure a company's profitability, performance, and financial strength.

PRACTICE PROBLEMS

1 The three major classifications of activities in a cash flow statement are:

- **A** inflows, outflows, and net flows.
- **B** operating, investing, and financing.
- **C** revenues, expenses, and net income.

2 The sale of a building for cash would be classified as what type of activity on the cash flow statement?

- **A** Operating.
- **B** Investing.
- **C** Financing.

3 Under which section of a manufacturing company's cash flow statement are the following activities reported?

Item 1: Purchases of securities held for trading

Item 2: Purchases of securities held for investment

- **A** Both items are investing activities.
- **B** Only Item 1 is an operating activity.
- **C** Only Item 2 is an operating activity.

4 Which of the following is an example of a financing activity on the cash flow statement under US GAAP?

- **A** Payment of interest.
- **B** Receipt of dividends.
- **C** Payment of dividends.

5 A conversion of a face value $1 million convertible bond for $1 million of common stock would most likely be:

- **A** reported as a $1 million investing cash inflow and outflow.
- **B** reported as a $1 million financing cash outflow and inflow.
- **C** reported as supplementary information to the cash flow statement.

6 A company recently engaged in a non-cash transaction that significantly affected its property, plant, and equipment. The transaction is:

- **A** reported under the investing section of the cash flow statement.
- **B** reported differently in cash flow from operations under the direct and indirect methods.
- **C** disclosed as a separate note or in a supplementary schedule to the cash flow statement.

7 Interest paid is classified as an operating cash flow under:

- **A** US GAAP but may be classified as either operating or investing cash flows under IFRS.
- **B** IFRS but may be classified as either operating or investing cash flows under US GAAP.
- **C** US GAAP but may be classified as either operating or financing cash flows under IFRS.

8 Cash flows from taxes on income must be separately disclosed under:

© 2019 CFA Institute. All rights reserved.

A IFRS only.

B US GAAP only.

C both IFRS and US GAAP.

9 Which of the following components of the cash flow statement may be prepared under the indirect method under both IFRS and US GAAP?

A Operating.

B Investing.

C Financing.

10 Which of the following is *most likely* to appear in the operating section of a cash flow statement under the indirect method?

A Net income.

B Cash paid to suppliers.

C Cash received from customers.

11 A benefit of using the direct method rather than the indirect method when reporting operating cash flows is that the direct method:

A mirrors a forecasting approach.

B is easier and less costly.

C provides specific information on the sources of operating cash flows.

12 Mabel Corporation (MC) reported accounts receivable of $66 million at the end of its second fiscal quarter. MC had revenues of $72 million for its third fiscal quarter and reported accounts receivable of $55 million at the end of its third fiscal quarter. Based on this information, the amount of cash MC collected from customers during the third fiscal quarter is:

A $61 million.

B $72 million.

C $83 million.

13 When computing net cash flow from operating activities using the indirect method, an addition to net income is *most likely* to occur when there is a:

A gain on the sale of an asset.

B loss on the retirement of debt.

C decrease in a deferred tax liability.

14 Red Road Company, a consulting company, reported total revenues of $100 million, total expenses of $80 million, and net income of $20 million in the most recent year. If accounts receivable increased by $10 million, how much cash did the company receive from customers?

A $90 million.

B $100 million.

C $110 million.

15 In 2018, a company using US GAAP made cash payments of $6 million for salaries, $2 million for interest expense, and $4 million for income taxes. Additional information for the company is provided in the table:

($ millions)	2017	2018
Revenue	42	37
Cost of goods sold	18	16
Inventory	36	40

($ millions)	2017	2018
Accounts receivable	22	19
Accounts payable	14	12

Based only on the information given, the company's operating cash flow for 2018 is *closest to*:

A $6 million.

B $10 million.

C $14 million.

16 Green Glory Corp., a garden supply wholesaler, reported cost of goods sold for the year of $80 million. Total assets increased by $55 million, including an increase of $5 million in inventory. Total liabilities increased by $45 million, including an increase of $2 million in accounts payable. The cash paid by the company to its suppliers is most likely *closest* to:

A $73 million.

B $77 million.

C $83 million.

17 Purple Fleur S.A., a retailer of floral products, reported cost of goods sold for the year of $75 million. Total assets increased by $55 million, but inventory declined by $6 million. Total liabilities increased by $45 million, and accounts payable increased by $2 million. The cash paid by the company to its suppliers is most likely *closest* to:

A $67 million.

B $79 million.

C $83 million.

18 White Flag, a women's clothing manufacturer, reported salaries expense of $20 million. The beginning balance of salaries payable was $3 million, and the ending balance of salaries payable was $1 million. How much cash did the company pay in salaries?

A $18 million.

B $21 million.

C $22 million.

19 An analyst gathered the following information from a company's 2018 financial statements (in $ millions):

Year ended 31 December	2017	2018
Net sales	245.8	254.6
Cost of goods sold	168.3	175.9
Accounts receivable	73.2	68.3
Inventory	39.0	47.8
Accounts payable	20.3	22.9

Based only on the information above, the company's 2018 statement of cash flows in the direct format would include amounts (in $ millions) for cash received from customers and cash paid to suppliers, respectively, that are *closest* to:

	cash received from customers	cash paid to suppliers
A	249.7	169.7
B	259.5	174.5
C	259.5	182.1

20 Golden Cumulus Corp., a commodities trading company, reported interest expense of $19 million and taxes of $6 million. Interest payable increased by $3 million, and taxes payable decreased by $4 million over the period. How much cash did the company pay for interest and taxes?

A $22 million for interest and $10 million for taxes.

B $16 million for interest and $2 million for taxes.

C $16 million for interest and $10 million for taxes.

21 An analyst gathered the following information from a company's 2018 financial statements (in $ millions):

Balances as of Year Ended 31 December	2017	2018
Retained earnings	120	145
Accounts receivable	38	43
Inventory	45	48
Accounts payable	36	29

In 2018, the company declared and paid cash dividends of $10 million and recorded depreciation expense in the amount of $25 million. The company considers dividends paid a financing activity. The company's 2018 cash flow from operations (in $ millions) was *closest* to

A 25.

B 45.

C 75.

22 Silverago Incorporated, an international metals company, reported a loss on the sale of equipment of $2 million in 2018. In addition, the company's income statement shows depreciation expense of $8 million and the cash flow statement shows capital expenditure of $10 million, all of which was for the purchase of new equipment. Using the following information from the comparative balance sheets, how much cash did the company receive from the equipment sale?

Balance Sheet Item	12/31/2017	12/31/2018	Change
Equipment	$100 million	$105 million	$5 million
Accumulated depreciation—equipment	$40 million	$46 million	$6 million

A $1 million.

B $2 million.

C $3 million.

23 Jaderong Plinkett Stores reported net income of $25 million. The company has no outstanding debt. Using the following information from the comparative balance sheets (in millions), what should the company report in the financing section of the statement of cash flows in 2018?

Balance Sheet Item	12/31/2017	12/31/2018	Change
Common stock	$100	$102	$2
Additional paid-in capital common stock	$100	$140	$40
Retained earnings	$100	$115	$15
Total stockholders' equity	$300	$357	$57

- **A** Issuance of common stock of $42 million; dividends paid of $10 million.
- **B** Issuance of common stock of $38 million; dividends paid of $10 million.
- **C** Issuance of common stock of $42 million; dividends paid of $40 million.

24 Based on the following information for Star Inc., what are the total net adjustments that the company would make to net income in order to derive operating cash flow?

	Year Ended
Income Statement Item	12/31/2018
Net income	$20 million
Depreciation	$2 million

Balance Sheet Item	12/31/2017	12/31/2018	Change
Accounts receivable	$25 million	$22 million	($3 million)
Inventory	$10 million	$14 million	$4 million
Accounts payable	$8 million	$13 million	$5 million

- **A** Add $2 million.
- **B** Add $6 million.
- **C** Subtract $6 million.

25 The first step in cash flow statement analysis should be to:

- **A** evaluate consistency of cash flows.
- **B** determine operating cash flow drivers.
- **C** identify the major sources and uses of cash.

26 Which of the following would be valid conclusions from an analysis of the cash flow statement for Telefónica Group presented in Exhibit 3?

- **A** The primary use of cash is financing activities.
- **B** The primary source of cash is operating activities.
- **C** Telefónica classifies dividends paid as an operating activity.

27 The following information is extracted from Sweetfall Incorporated's financial statements.

Income Statement		Balance Sheet Changes	
Revenue	$56,800	Decrease in accounts receivable	$1,324
Cost of goods sold	27,264	Decrease in inventory	501
Other operating expense	562	Increase in prepaid expense	6
Depreciation expense	2,500	Increase in accounts payable	1,063

The amount of cash Sweetfall Inc. paid to suppliers is:

A $25,700.

B $26,702.

C $27,826.

28 Which is an appropriate method of preparing a common-size cash flow statement?

A Show each item of revenue and expense as a percentage of net revenue.

B Show each line item on the cash flow statement as a percentage of net revenue.

C Show each line item on the cash flow statement as a percentage of total cash outflows.

29 Which of the following is an appropriate method of computing free cash flow to the firm?

A Add operating cash flows to capital expenditures and deduct after-tax interest payments.

B Add operating cash flows to after-tax interest payments and deduct capital expenditures.

C Deduct both after-tax interest payments and capital expenditures from operating cash flows.

30 An analyst has calculated a ratio using as the numerator the sum of operating cash flow, interest, and taxes and as the denominator the amount of interest. What is this ratio, what does it measure, and what does it indicate?

A This ratio is an interest coverage ratio, measuring a company's ability to meet its interest obligations and indicating a company's solvency.

B This ratio is an effective tax ratio, measuring the amount of a company's operating cash flow used for taxes and indicating a company's efficiency in tax management.

C This ratio is an operating profitability ratio, measuring the operating cash flow generated accounting for taxes and interest and indicating a company's liquidity.

SOLUTIONS

1 B is correct. Operating, investing, and financing are the three major classifications of activities in a cash flow statement. Revenues, expenses, and net income are elements of the income statement. Inflows, outflows, and net flows are items of information in the statement of cash flows.

2 B is correct. Purchases and sales of long-term assets are considered investing activities. Note that if the transaction had involved the exchange of a building for other than cash (for example, for another building, common stock of another company, or a long-term note receivable), it would have been considered a significant non-cash activity.

3 B is correct. The purchase and sale of securities held for trading are considered operating activities even for companies in which this activity is not a primary business activity.

4 C is correct. Payment of dividends is a financing activity under US GAAP. Payment of interest and receipt of dividends are included in operating cash flows under US GAAP. Note that IFRS allow companies to include receipt of interest and dividends as either operating or investing cash flows and to include payment of interest and dividends as either operating or financing cash flows.

5 C is correct. Non-cash transactions, if significant, are reported as supplementary information, not in the investing or financing sections of the cash flow statement.

6 C is correct. Because no cash is involved in non-cash transactions, these transactions are not incorporated in the cash flow statement. However, non-cash transactions that significantly affect capital or asset structures are required to be disclosed either in a separate note or a supplementary schedule to the cash flow statement.

7 C is correct. Interest expense is always classified as an operating cash flow under US GAAP but may be classified as either an operating or financing cash flow under IFRS.

8 C is correct. Taxes on income are required to be separately disclosed under IFRS and US GAAP. The disclosure may be in the cash flow statement or elsewhere.

9 A is correct. The operating section may be prepared under the indirect method. The other sections are always prepared under the direct method.

10 A is correct. Under the indirect method, the operating section would begin with net income and adjust it to arrive at operating cash flow. The other two items would appear in the operating section under the direct method.

11 C is correct. The primary argument in favor of the direct method is that it provides information on the specific sources of operating cash receipts and payments. Arguments for the indirect method include that it mirrors a forecasting approach and it is easier and less costly

12 C is correct. The amount of cash collected from customers during the quarter is equal to beginning accounts receivable plus revenues minus ending accounts receivable: \$66 million + \$72 million – \$55 million = \$83 million. A reduction in accounts receivable indicates that cash collected during the quarter was greater than revenue on an accrual basis.

13 B is correct. An addition to net income is made when there is a loss on the retirement of debt, which is a non-operating loss. A gain on the sale of an asset and a decrease in deferred tax liability are both subtracted from net-income.

14 A is correct. Revenues of $100 million minus the increase in accounts receivable of $10 million equal $90 million cash received from customers. The increase in accounts receivable means that the company received less in cash than it reported as revenue.

15 A is correct.

$$\begin{aligned}\text{Operating cash flows} &= \text{Cash received from customers} - (\text{Cash paid to suppliers} + \text{Cash paid to employees} \\ &\quad + \text{Cash paid for other operating expenses} + \text{Cash paid for interest} + \text{Cash paid for income taxes}) \\ \text{Cash received from customers} &= \text{Revenue} + \text{Decrease in accounts receivable} \\ &= \$37 + \$3 = \$40 \text{ million} \\ \text{Cash paid to suppliers} &= \text{Cost of goods sold} + \text{Increase in inventory} + \text{Decrease in accounts payable} \\ &= \$16 + \$4 + \$2 = \$22 \text{ million}\end{aligned}$$

Therefore, the company's operating cash flow = $40 – $22 – Cash paid for salaries – Cash paid for interest – Cash paid for taxes = $40 – $22 – $6 – $2 – $4 = $6 million.

16 C is correct. Cost of goods sold of $80 million plus the increase in inventory of $5 million equals purchases from suppliers of $85 million. The increase in accounts payable of $2 million means that the company paid $83 million in cash ($85 million minus $2 million) to its suppliers.

17 A is correct. Cost of goods sold of $75 million less the decrease in inventory of $6 million equals purchases from suppliers of $69 million. The increase in accounts payable of $2 million means that the company paid $67 million in cash ($69 million minus $2 million).

18 C is correct. Beginning salaries payable of $3 million plus salaries expense of $20 million minus ending salaries payable of $1 million equals $22 million. Alternatively, the expense of $20 million plus the $2 million decrease in salaries payable equals $22 million.

19 C is correct. Cash received from customers = Sales + Decrease in accounts receivable = 254.6 + 4.9 = 259.5. Cash paid to suppliers = Cost of goods sold + Increase in inventory – Increase in accounts payable = 175.9 + 8.8 – 2.6 = 182.1.

20 C is correct. Interest expense of $19 million less the increase in interest payable of $3 million equals interest paid of $16 million. Tax expense of $6 million plus the decrease in taxes payable of $4 million equals taxes paid of $10 million.

21 B is correct. All dollar amounts are in millions. Net income (NI) for 2018 is $35. This amount is the increase in retained earnings, $25, plus the dividends paid, $10. Depreciation of $25 is added back to net income, and the increases in accounts receivable, $5, and in inventory, $3, are subtracted from net income because they are uses of cash. The decrease in accounts payable is also a use of cash and, therefore, a subtraction from net income. Thus, cash flow from operations is $25 + $10 + $25 – $5 – $3 – $7 = $45.

22 A is correct. Selling price (cash inflow) minus book value equals gain or loss on sale; therefore, gain or loss on sale plus book value equals selling price (cash inflow). The amount of loss is given—$2 million. To calculate the book value of the equipment sold, find the historical cost of the equipment and the accumulated depreciation on the equipment.

- Beginning balance of equipment of $100 million plus equipment purchased of $10 million minus ending balance of equipment of $105 million equals the historical cost of equipment sold, or $5 million.
- Beginning accumulated depreciation of $40 million plus depreciation expense for the year of $8 million minus ending balance of accumulated depreciation of $46 million equals accumulated depreciation on the equipment sold, or $2 million.
- Therefore, the book value of the equipment sold was $5 million minus $2 million, or $3 million.
- Because the loss on the sale of equipment was $2 million, the amount of cash received must have been $1 million.

23 A is correct. The increase of $42 million in common stock and additional paid-in capital indicates that the company issued stock during the year. The increase in retained earnings of $15 million indicates that the company paid $10 million in cash dividends during the year, determined as beginning retained earnings of $100 million plus net income of $25 million minus ending retained earnings of $115 million, which equals $10 million in cash dividends.

24 B is correct. To derive operating cash flow, the company would make the following adjustments to net income: Add depreciation (a non-cash expense) of $2 million; add the decrease in accounts receivable of $3 million; add the increase in accounts payable of $5 million; and subtract the increase in inventory of $4 million. Total additions would be $10 million, and total subtractions would be $4 million, which gives net additions of $6 million.

25 C is correct. An overall assessment of the major sources and uses of cash should be the first step in evaluating a cash flow statement.

26 B is correct. The primary source of cash is operating activities. Cash flow provided by operating activity totaled €13,796 million in the most recent year. The primary use of cash is investing activities (total of €10,245 million). Dividends paid are classified as a financing activity.

27 A is correct. The amount of cash paid to suppliers is calculated as follows:

= Cost of goods sold – Decrease in inventory – Increase in accounts payable
= $27,264 – $501 – $1,063
= $25,700.

28 B is correct. An appropriate method to prepare a common-size cash flow statement is to show each line item on the cash flow statement as a percentage of net revenue. An alternative way to prepare a statement of cash flows is to show each item of cash inflow as a percentage of total inflows and each item of cash outflows as a percentage of total outflows.

29 B is correct. Free cash flow to the firm can be computed as operating cash flows plus after-tax interest expense less capital expenditures.

30 A is correct. This ratio is an interest coverage ratio, measuring a company's ability to meet its interest obligations and indicating a company's solvency. This coverage ratio is based on cash flow information; another common coverage ratio uses a measure based on the income statement (earnings before interest, taxes, depreciation, and amortisation).

READING

20

Financial Analysis Techniques

by Elaine Henry, PhD, CFA, Thomas R. Robinson, PhD, CFA, and
J. Hennie van Greuning, DCom, CFA

Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). Thomas R. Robinson, PhD, CFA, is at AACSB International (USA). J. Hennie van Greuning, DCom, CFA, is at BIBD (Brunei).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe tools and techniques used in financial analysis, including their uses and limitations;
☐	**b.** identify, calculate, and interpret activity, liquidity, solvency, profitability, and valuation ratios;
☐	**c.** describe relationships among ratios and evaluate a company using ratio analysis;
☐	**d.** demonstrate the application of DuPont analysis of return on equity and calculate and interpret effects of changes in its components;
☐	**e.** calculate and interpret ratios used in equity analysis and credit analysis;
☐	**f.** explain the requirements for segment reporting and calculate and interpret segment ratios;
☐	**g.** describe how ratio analysis and other techniques can be used to model and forecast earnings.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1

INTRODUCTION

a describe tools and techniques used in financial analysis, including their uses and limitations

Financial analysis tools can be useful in assessing a company's performance and trends in that performance. In essence, an analyst converts data into financial metrics that assist in decision making. Analysts seek to answer such questions as: How successfully has the company performed, relative to its own past performance and relative to its competitors? How is the company likely to perform in the future? Based on expectations about future performance, what is the value of this company or the securities it issues?

A primary source of data is a company's annual report, including the financial statements and notes, and management commentary (operating and financial review or management's discussion and analysis). This reading focuses on data presented in financial reports prepared under International Financial Reporting Standards (IFRS) and United States generally accepted accounting principles (US GAAP). However, financial reports do not contain all the information needed to perform effective financial analysis. Although financial statements do contain data about the *past* performance of a company (its income and cash flows) as well as its *current* financial condition (assets, liabilities, and owners' equity), such statements do not necessarily provide all the information useful for analysis nor do they forecast *future* results. The financial analyst must be capable of using financial statements in conjunction with other information to make projections and reach valid conclusions. Accordingly, an analyst typically needs to supplement the information found in a company's financial reports with other information, including information on the economy, industry, comparable companies, and the company itself.

This reading describes various techniques used to analyze a company's financial statements. Financial analysis of a company may be performed for a variety of reasons, such as valuing equity securities, assessing credit risk, conducting due diligence related to an acquisition, or assessing a subsidiary's performance. This reading will describe techniques common to any financial analysis and then discuss more specific aspects for the two most common categories: equity analysis and credit analysis.

Equity analysis incorporates an owner's perspective, either for valuation or performance evaluation. Credit analysis incorporates a creditor's (such as a banker or bondholder) perspective. In either case, there is a need to gather and analyze information to make a decision (ownership or credit); the focus of analysis varies because of the differing interest of owners and creditors. Both equity and credit analyses assess the entity's ability to generate and grow earnings, and cash flow, as well as any associated risks. Equity analysis usually places a greater emphasis on growth, whereas credit analysis usually places a greater emphasis on risks. The difference in emphasis reflects the different fundamentals of these types of investments: The value of a company's equity generally increases as the company's earnings and cash flow increase, whereas the value of a company's debt has an upper limit.[1]

The balance of this reading is organized as follows: Section 1 recaps the framework for financial statements and the place of financial analysis techniques within the framework. Sections 2–6 provide a description of analytical tools and techniques. Sections 7–13 explain how to compute, analyze, and interpret common financial ratios. Sections 14–19 explain the use of ratios and other analytical data in equity

---

[1] The upper limit is equal to the undiscounted sum of the principal and remaining interest payments (i.e., the present value of these contractual payments at a zero percent discount rate).

analysis, credit analysis, segment analysis, and forecasting, respectively. A summary of the key points and practice problems in the CFA Institute multiple-choice format conclude the reading.

1.1 The Financial Analysis Process

In financial analysis, it is essential to clearly identify and understand the final objective and the steps required to reach that objective. In addition, the analyst needs to know where to find relevant data, how to process and analyze the data (in other words, know the typical questions to address when interpreting data), and how to communicate the analysis and conclusions.

1.1.1 *The Objectives of the Financial Analysis Process*

Because of the variety of reasons for performing financial analysis, the numerous available techniques, and the often substantial amount of data, it is important that the analytical approach be tailored to the specific situation. Prior to beginning any financial analysis, the analyst should clarify the purpose and context, and clearly understand the following:

- What is the purpose of the analysis? What questions will this analysis answer?
- What level of detail will be needed to accomplish this purpose?
- What data are available for the analysis?
- What are the factors or relationships that will influence the analysis?
- What are the analytical limitations, and will these limitations potentially impair the analysis?

Having clarified the purpose and context of the analysis, the analyst can select the set of techniques (e.g., ratios) that will best assist in making a decision. Although there is no single approach to structuring the analysis process, a general framework is set forth in Exhibit 1.[2] The steps in this process were discussed in more detail in an earlier reading; the primary focus of this reading is on Phases 3 and 4, processing and analyzing data.

---

[2] Components of this framework have been adapted from van Greuning and Bratanovic (2003, p. 300) and Benninga and Sarig (1997, pp. 134–156).

Exhibit 1 A Financial Statement Analysis Framework

Phase	Sources of Information	Output
1 Articulate the purpose and context of the analysis.	■ The nature of the analyst's function, such as evaluating an equity or debt investment or issuing a credit rating. ■ Communication with client or supervisor on needs and concerns. ■ Institutional guidelines related to developing specific work product.	■ Statement of the purpose or objective of analysis. ■ A list (written or unwritten) of specific questions to be answered by the analysis. ■ Nature and content of report to be provided. ■ Timetable and budgeted resources for completion.
2 Collect input data.	■ Financial statements, other financial data, questionnaires, and industry/economic data. ■ Discussions with management, suppliers, customers, and competitors. ■ Company site visits (e.g., to production facilities or retail stores).	■ Organized financial statements. ■ Financial data tables. ■ Completed questionnaires, if applicable.
3 Process data.	■ Data from the previous phase.	■ Adjusted financial statements. ■ Common-size statements. ■ Ratios and graphs. ■ Forecasts.
4 Analyze/interpret the processed data.	■ Input data as well as processed data.	■ Analytical results.
5 Develop and communicate conclusions and recommendations (e.g., with an analysis report).	■ Analytical results and previous reports. ■ Institutional guidelines for published reports.	■ Analytical report answering questions posed in Phase 1. ■ Recommendation regarding the purpose of the analysis, such as whether to make an investment or grant credit.
6 Follow-up.	■ Information gathered by periodically repeating above steps as necessary to determine whether changes to holdings or recommendations are necessary.	■ Updated reports and recommendations.

1.1.2 *Distinguishing between Computations and Analysis*

An effective analysis encompasses both computations and interpretations. A well-reasoned analysis differs from a mere compilation of various pieces of information, computations, tables, and graphs by integrating the data collected into a cohesive whole. Analysis of past performance, for example, should address not only what happened but also why it happened and whether it advanced the company's strategy. Some of the key questions to address include:

- What aspects of performance are critical for this company to successfully compete in this industry?

- How well did the company's performance meet these critical aspects? (Established through computation and comparison with appropriate benchmarks, such as the company's own historical performance or competitors' performance.)
- What were the key causes of this performance, and how does this performance reflect the company's strategy? (Established through analysis.)

If the analysis is forward looking, additional questions include:

- What is the likely impact of an event or trend? (Established through interpretation of analysis.)
- What is the likely response of management to this trend? (Established through evaluation of quality of management and corporate governance.)
- What is the likely impact of trends in the company, industry, and economy on future cash flows? (Established through assessment of corporate strategy and through forecasts.)
- What are the recommendations of the analyst? (Established through interpretation and forecasting of results of analysis.)
- What risks should be highlighted? (Established by an evaluation of major uncertainties in the forecast and in the environment within which the company operates.)

Example 1 demonstrates how a company's financial data can be analyzed in the context of its business strategy and changes in that strategy. An analyst must be able to understand the "why" behind the numbers and ratios, not just what the numbers and ratios are.

EXAMPLE 1

Strategy Reflected in Financial Performance

Apple Inc. engages in the design, manufacture, and sale of computer hardware, mobile devices, operating systems and related products, and services. It also operates retail and online stores. Microsoft develops, licenses, and supports software products, services, and technology devices through a variety of channels including retail stores in recent years. Selected financial data for 2015 through 2017 for these two companies are given below. Apple's fiscal year (FY) ends on the final Saturday in September (for example, FY2017 ended on 30 September 2017). Microsoft's fiscal year ends on 30 June (for example, FY2017 ended on 30 June 2017).

Selected Financial Data for Apple (Dollars in Millions)

Fiscal year	2017	2016	2015
Net sales (or Revenue)	229,234	215,639	233,715
Gross margin	88,186	84,263	93,626
Operating income	61,344	60,024	71,230

Selected Financial Data for Microsoft (Dollars in Millions)*

Fiscal year	2017	2016	2015
Net sales (or Revenue)	89,950	85,320	93,580
Gross margin	55,689	52,540	60,542
Operating income	22,326	20,182	18,161

* Microsoft revenue for 2017 and 2016 were subsequently revised in the company's 2018 10-K report due to changes in revenue recognition and lease accounting standards.
Source: 10-K reports for Apple and Microsoft.

Apple reported a 7.7 percent decrease in net sales from FY2015 to FY2016 and an increase of 6.3 percent from FY2016 to FY2017 for an overall slight decline over the three-year period. Gross margin decreased 10.0 percent from FY2015 to FY2016 and increased 4.7 percent from FY2016 to FY2017. This also represented an overall decline in gross margin over the three-year period. The company's operating income exhibited similar trends.

Microsoft reported an 8.8 percent decrease in net sales from FY2015 to FY2016 and an increase of 5.4 percent from FY2016 to FY2017 for an overall slight decline over the three-year period. Gross margin decreased 13.2 percent from FY2015 to FY2016 and increased 6.0 percent from FY2016 to FY2017. Similar to Apple, this represented an overall decline in gross margin over the three-year period. Microsoft's operating income on the other hand exhibited growth each year and for the three-year period. Overall growth in operating income was 23%.

What caused Microsoft's growth in operating income while Apple and Microsoft had similar negative trends in sales and gross margin? Apple's decline in sales, gross margin, and operating income from FY2015 to FY2016 was caused by declines in iPhone sales and weakness in foreign currencies relative to the US dollar. FY2017 saw a rebound in sales of iPhones, Mac computers, and services offset somewhat by continued weaknesses in foreign currencies. Microsoft similarly had declines in revenue and gross margin from sales of its devices and Windows software in FY2016, as well as negative impacts from foreign currency weakness. Microsoft's increase in revenue and gross margin in FY2017 was driven by the acquisition of LinkedIn, higher sales of Microsoft Office software, and higher sales of cloud services. The driver in the continuous increase in operating income for Microsoft was a large decline over the three-year period in impairment, integration, and restructuring charges. Microsoft recorded a $10 billion charge in FY2015 related to its phone business, and there were further charges of $1.1 billion in FY2016 and $306 million in FY2017. Absent these large write-offs, Microsoft would have had a trend similar to Apple's in operating income over the three-year period.

Analysts often need to communicate the findings of their analysis in a written report. Their reports should communicate how conclusions were reached and why recommendations were made. For example, a report might present the following:

- the purpose of the report, unless it is readily apparent;
- relevant aspects of the business context:
  - economic environment (country/region, macro economy, sector);
  - financial and other infrastructure (accounting, auditing, rating agencies);

  - legal and regulatory environment (and any other material limitations on the company being analyzed);
- evaluation of corporate governance and assessment of management strategy, including the company's competitive advantage(s);
- assessment of financial and operational data, including key assumptions in the analysis; and
- conclusions and recommendations, including limitations of the analysis and risks.

An effective narrative and well supported conclusions and recommendations are normally enhanced by using 3–10 years of data, as well as analytic techniques appropriate to the purpose of the report.

ANALYTICAL TOOLS AND TECHNIQUES

2

a describe tools and techniques used in financial analysis, including their uses and limitations

The tools and techniques presented in this section facilitate evaluations of company data. Evaluations require comparisons. It is difficult to say that a company's financial performance was "good" without clarifying the basis for comparison. In assessing a company's ability to generate and grow earnings and cash flow, and the risks related to those earnings and cash flows, the analyst draws comparisons to other companies (cross-sectional analysis) and over time (trend or time-series analysis).

For example, an analyst may wish to compare the profitability of companies competing in a global industry. If the companies differ significantly in size and/or report their financial data in different currencies, comparing net income as reported is not useful. Ratios (which express one number in relation to another) and common-size financial statements can remove size as a factor and enable a more relevant comparison. To achieve comparability across companies reporting in different currencies, one approach is to translate all reported numbers into a common currency using exchange rates at the end of a period. Others may prefer to translate reported numbers using the average exchange rates during the period. Alternatively, if the focus is primarily on ratios, comparability can be achieved without translating the currencies.

The analyst may also want to examine comparable performance over time. Again, the nominal currency amounts of sales or net income may not highlight significant changes. To address this challenge, horizontal financial statements (whereby quantities are stated in terms of a selected base year value) can make such changes more apparent. Another obstacle to comparison is differences in fiscal year end. To achieve comparability, one approach is to develop trailing twelve months data, which will be described in a section below. Finally, it should be noted that differences in accounting standards can limit comparability.

EXAMPLE 2

Ratio Analysis

An analyst is examining the profitability of two international companies with large shares of the global personal computer market: Acer Inc. and Lenovo Group Limited. Acer has pursued a strategy of selling its products at affordable prices. In contrast, Lenovo aims to achieve higher selling prices by stressing the high engineering quality of its personal computers for business use. Acer reports in TWD,[3] and Lenovo reports in USD. For Acer, fiscal year end is 31 December. For Lenovo, fiscal year end is 31 March; thus, FY2017 ended 31 March 2018.

The analyst collects the data shown in Exhibit 2 below. Use this information to answer the following questions:

1 Which company is larger based on the amount of revenue, in US$, reported in fiscal year 2017? For FY2017, assume the relevant, average exchange rate was 30.95 TWD/USD.

2 Which company had the higher revenue growth from FY2016 to FY2017? FY2013 to FY2017?

3 How do the companies compare, based on profitability?

Exhibit 2

Acer

TWD Millions	**FY2013**	**FY2014**	**FY2015**	**FY2016**	**FY2017**
Revenue	360,132	329,684	263,775	232,724	237,275
Gross profit	22,550	28,942	24,884	23,212	25,361
Net income	(20,519)	1,791	604	(4,901)	2,797

Lenovo

USD Millions	**FY2013***	**FY2014***	**FY2015***	**FY2016***	**FY2017***
Revenue	38,707	46,296	44,912	43,035	45,350
Gross profit	5,064	6,682	6,624	6,105	6,272
Net income (Loss)	817	837	(145)	530	(127)

* Fiscal years for Lenovo end 31 March. Thus FY2017 represents the fiscal year ended 31 March 2018; the same applies respectively for prior years.

Solution to 1:

Lenovo is much larger than Acer based on FY2017 revenues in USD terms. Lenovo's FY2017 revenues of $USD45.35 billion are considerably higher than Acer's USD7.67 billion (= TWD237.275 million/30.95).

Acer: At the assumed average exchange rate of 30.95 TWD/USD, Acer's FY2017 revenues are equivalent to USD7.67 billion (= TWD237.275 million ÷ 30.95 TWD/USD).

Lenovo: Lenovo's FY2017 revenues totaled USD45.35 billion.

Note: Comparing the size of companies reporting in different currencies requires translating reported numbers into a common currency using exchange rates at some point in time. This solution converts the revenues of Acer to billions of USD using the average exchange rate of the fiscal period. It would be equally informative (and would yield the same conclusion) to convert the revenues of Lenovo to TWD.

Solution to 2:

The growth in Lenovo's revenue was much higher than Acer's in the most recent fiscal year and for the five-year period.

---

[3] TWD is the three-letter ISO 4217 currency code for Taiwan New Dollar.

	Change in Revenue FY2016 versus FY2017 (%)	Change in Revenue FY2013 to FY2017 (%)
Acer	1.96	(34.11)
Lenovo	5.38	17.16

The table shows two growth metrics. Calculations are illustrated using the revenue data for Acer:

The change in Acer's revenue for FY2016 versus FY2017 is 1.96% percent calculated as (237,275 – 232,724) ÷ 232,724 or equivalently (237,275 ÷ 232,724) – 1. The change in Acer's revenue from FY2013 to FY2017 is a decline of 34.11%.

Solution to 3:

Profitability can be assessed by comparing the amount of gross profit to revenue and the amount of net income to revenue. The following table presents these two profitability ratios—**gross profit margin** (gross profit divided by revenue) and **net profit margin** (net income divided by revenue)—for each year.

Acer	FY2013 (%)	FY2014 (%)	FY2015 (%)	FY2016 (%)	FY2017 (%)
Gross profit margin	6.26	8.78	9.43	9.97	10.69
Net profit margin	(5.70)	0.54	0.23	(2.11)	1.18

Lenovo	FY2013 (%)	FY2014 (%)	FY2015 (%)	FY2016 (%)	FY2017 (%)
Gross profit margin	13.08	14.43	14.75	14.19	13.83
Net profit margin	2.11	1.81	(0.32)	1.23	(0.28)

The net profit margins indicate that both companies' profitability is relatively low. Acer's net profit margin is lower than Lenovo's in three out of the five years. Acer's gross profit margin increased each year but remains significantly below that of Lenovo. Lenovo's gross profit margin grew from FY2013 to FY2015 and then declined in FY2016 and FY2017. Overall, Lenovo is the more profitable company, likely attributable to its larger size and commensurate economies of scale. (Lenovo has the largest share of the personal computer market relative to other personal computer companies.)

Section 3 describes the tools and techniques of ratio analysis in more detail. Sections 4 to 6 describe other tools and techniques.

FINANCIAL RATIO ANALYSIS

3

a describe tools and techniques used in financial analysis, including their uses and limitations

There are many relationships among financial accounts and various expected relationships from one point in time to another. Ratios are a useful way of expressing these relationships. Ratios express one quantity in relation to another (usually as a quotient).

Extensive academic research has examined the importance of ratios in predicting stock returns (Ou and Penman, 1989; Abarbanell and Bushee, 1998) or credit failure (Altman, 1968; Ohlson, 1980; Hopwood et al., 1994). This research has found that financial statement ratios are effective in selecting investments and in predicting financial distress. Practitioners routinely use ratios to derive and communicate the value of companies and securities.

Several aspects of ratio analysis are important to understand. First, the computed ratio is not "the answer." The ratio is an *indicator* of some aspect of a company's performance, telling what happened but not why it happened. For example, an analyst might want to answer the question: Which of two companies was more profitable? As demonstrated in the previous example, the net profit margin, which expresses profit relative to revenue, can provide insight into this question. Net profit margin is calculated by dividing net income by revenue:[4]

$$\frac{\text{Net income}}{\text{Revenue}}$$

Assume Company A has €100,000 of net income and Company B has €200,000 of net income. Company B generated twice as much income as Company A, but was it more profitable? Assume further that Company A has €2,000,000 of revenue, and thus a net profit margin of 5 percent, and Company B has €6,000,000 of revenue, and thus a net profit margin of 3.33 percent. Expressing net income as a percentage of revenue clarifies the relationship: For each €100 of revenue, Company A earns €5 in net income, whereas Company B earns only €3.33 for each €100 of revenue. So, we can now answer the question of which company was more profitable in percentage terms: Company A was more profitable, as indicated by its higher net profit margin of 5 percent. Note that Company A was more *profitable* despite the fact that Company B reported higher absolute amounts of net income and revenue. However, this ratio by itself does not tell us *why* Company A has a higher profit margin. Further analysis is required to determine the reason (perhaps higher relative sales prices or better cost control or lower effective tax rates).

Company size sometimes confers economies of scale, so the absolute amounts of net income and revenue are useful in financial analysis. However, ratios control for the effect of size, which enhances comparisons between companies and over time.

A second important aspect of ratio analysis is that differences in accounting policies (across companies and across time) can distort ratios, and a meaningful comparison may, therefore, involve adjustments to the financial data. Third, not all ratios are necessarily relevant to a particular analysis. The ability to select a relevant ratio or ratios to answer the research question is an analytical skill. Finally, as with financial analysis in general, ratio analysis does not stop with computation; interpretation of the result is essential. In practice, differences in ratios across time and across companies can be subtle, and interpretation is situation specific.

3.1 The Universe of Ratios

There are no authoritative bodies specifying exact formulas for computing ratios or providing a standard, comprehensive list of ratios. Formulas and even names of ratios often differ from analyst to analyst or from database to database. The number of different ratios that can be created is practically limitless. There are, however, widely accepted ratios that have been found to be useful. Sections 7–13 of this reading focus primarily on these broad classes and commonly accepted definitions of key ratios. However, the analyst should be aware that different ratios may be used in practice and that certain industries have unique ratios tailored to the characteristics of that industry. When faced with an unfamiliar ratio, the analyst can examine the underlying formula to gain insight into what the ratio is measuring. For example, consider the following ratio formula:

---

[4] The term "sales" is often used interchangeably with the term "revenues." Other times it is used to refer to revenues derived from sales of products versus services. The income statement usually reflects "revenues" or "sales" after returns and allowances (e.g., returns of products or discounts offered after a sale to induce the customer to not return a product). Additionally, in some countries, including the United Kingdom and South Africa, the term "turnover" is used in the sense of "revenue."

$$\frac{\text{Operating income}}{\text{Average total assets}}$$

Never having seen this ratio, an analyst might question whether a result of 12 percent is better than 8 percent. The answer can be found in the ratio itself. The numerator is operating income and the denominator is average total assets, so the ratio can be interpreted as the amount of operating income generated per unit of assets. For every €100 of average total assets, generating €12 of operating income is better than generating €8 of operating income. Furthermore, it is apparent that this particular ratio is an indicator of profitability (and, to a lesser extent, efficiency in use of assets in generating operating profits). When encountering a ratio for the first time, the analyst should evaluate the numerator and denominator to assess what the ratio is attempting to measure and how it should be interpreted. This is demonstrated in Example 3.

EXAMPLE 3

Interpreting a Financial Ratio

A US insurance company reports that its "combined ratio" is determined by dividing losses and expenses incurred by net premiums earned. It reports the following combined ratios:

Fiscal Year	5	4	3	2	1
Combined ratio	90.1%	104.0%	98.5%	104.1%	101.1%

Explain what this ratio is measuring and compare the results reported for each of the years shown in the chart. What other information might an analyst want to review before making any conclusions on this information?

Solution:

The combined ratio is a profitability measure. The ratio is explaining how much costs (losses and expenses) were incurred for every dollar of revenue (net premiums earned). The underlying formula indicates that a *lower* value for this ratio is better. The Year 5 ratio of 90.1 percent means that for every dollar of net premiums earned, the costs were $0.901, yielding a gross profit of $0.099. Ratios greater than 100 percent indicate an overall loss. A review of the data indicates that there does not seem to be a consistent trend in this ratio. Profits were achieved in Years 5 and 3. The results for Years 4 and 2 show the most significant costs at approximately 104 percent.

The analyst would want to discuss this data further with management and understand the characteristics of the underlying business. He or she would want to understand why the results are so volatile. The analyst would also want to determine what should be used as a benchmark for this ratio.

The Operating income/Average total assets ratio shown above is one of many versions of the **return on assets (ROA)** ratio. Note that there are other ways of specifying this formula based on how assets are defined. Some financial ratio databases compute ROA using the ending value of assets rather than average assets. In limited cases, one may also see beginning assets in the denominator. Which one is right? It depends on what you are trying to measure and the underlying company trends. If the company has a stable level of assets, the answer will not differ greatly under the three measures of assets (beginning, average, and ending). However, if the assets are growing (or shrinking), the results will differ among the three measures. When assets are growing, operating income divided by ending assets may not make sense because some of the income would have been generated before some assets were purchased, and this would understate the company's performance. Similarly, if beginning assets are used, some of the operating income later in the year may have been generated only because of the addition of assets; therefore, the ratio would overstate the company's performance. Because operating income occurs throughout the period, it generally makes sense to use some average measure of assets. A good general rule is that when an income statement or cash flow statement number is in the numerator of a ratio and a balance sheet number is in the denominator, then an average should be used for the denominator. It is generally not necessary to use averages when only balance sheet numbers are used in both the numerator and denominator because both are determined as of the same date. However, in some instances, even ratios that only use balance sheet data may use averages. For example, **return on equity (ROE)**, which is defined as net income divided by average shareholders' equity, can be decomposed into other ratios, some of which only use balance sheet data. In decomposing ROE into component ratios, if an average is used in one of the component ratios then it should be used in the other component ratios. The decomposition of ROE is discussed further in Section 13.

If an average is used, judgment is also required about what average should be used. For simplicity, most ratio databases use a simple average of the beginning and end-of-year balance sheet amounts. If the company's business is seasonal so that levels of assets vary by interim period (semiannual or quarterly), then it may be beneficial to take an average over all interim periods, if available. (If the analyst is working within a company and has access to monthly data, this can also be used.)

3.2 Value, Purposes, and Limitations of Ratio Analysis

The value of ratio analysis is that it enables a financial analyst to evaluate past performance, assess the current financial position of the company, and gain insights useful for projecting future results. As noted previously, the ratio itself is not "the answer" but is an indicator of some aspect of a company's performance. Financial ratios provide insights into:

- economic relationships within a company that help analysts project earnings and free cash flow;
- a company's financial flexibility, or ability to obtain the cash required to grow and meet its obligations, even if unexpected circumstances develop;
- management's ability;
- changes in the company and/or industry over time; and
- comparability with peer companies or the relevant industry(ies).

There are also limitations to ratio analysis. Factors to consider include:

- *The heterogeneity or homogeneity of a company's operating activities.* Companies may have divisions operating in many different industries. This can make it difficult to find comparable industry ratios to use for comparison purposes.
- *The need to determine whether the results of the ratio analysis are consistent.* One set of ratios may indicate a problem, whereas another set may indicate that the potential problem is only short term in nature.

- *The need to use judgment.* A key issue is whether a ratio for a company is within a reasonable range. Although financial ratios are used to help assess the growth potential and risk of a company, they cannot be used alone to directly value a company or its securities, or to determine its creditworthiness. The entire operation of the company must be examined, and the external economic and industry setting in which it is operating must be considered when interpreting financial ratios.
- *The use of alternative accounting methods.* Companies frequently have latitude when choosing certain accounting methods. Ratios taken from financial statements that employ different accounting choices may not be comparable unless adjustments are made. Some important accounting considerations include the following:
  - FIFO (first in, first out), LIFO (last in, first out), or average cost inventory valuation methods (IFRS does not allow LIFO);
  - Cost or equity methods of accounting for unconsolidated affiliates;
  - Straight line or accelerated methods of depreciation; and
  - Operating or finance lease treatment for lessors (under US GAAP, the type of lease affects classifications of expenses; under IFRS, operating lease treatment for lessors is not applicable).

The expanding use of IFRS and past convergence efforts between IFRS and US GAAP make the financial statements of different companies more comparable and may overcome some of these difficulties. Nonetheless, there will remain accounting choices that the analyst must consider.

3.3 Sources of Ratios

Ratios may be computed using data obtained directly from companies' financial statements or from a database such as Bloomberg, Compustat, FactSet, or Thomson Reuters. The information provided by the database may include information as reported in companies' financial statements and ratios calculated based on the information. These databases are popular because they provide easy access to many years of historical data so that trends over time can be examined. They also allow for ratio calculations based on periods other than the company's fiscal year, such as for the trailing 12 months (TTM) or most recent quarter (MRQ).

EXAMPLE 4

Trailing Twelve Months

On 15 July, an analyst is examining a company with a fiscal year ending on 31 December. Use the following data to calculate the company's trailing 12 month earnings (for the period ended 30 June 2018):

- Earnings for the year ended 31 December, 2017: $1,200;
- Earnings for the six months ended 30 June 2017: $550; and
- Earnings for the six months ended 30 June 2018: $750.

Solution:

The company's trailing 12 months earnings is $1,400, calculated as $1,200 – $550 + $750.

Analysts should be aware that the underlying formulas for ratios may differ by vendor. The formula used should be obtained from the vendor, and the analyst should determine whether any adjustments are necessary. Furthermore, database providers often exercise judgment when classifying items. For example, operating income may not appear directly on a company's income statement, and the vendor may use judgment to classify income statement items as "operating" or "non-operating." Variation in such judgments would affect any computation involving operating income. It is therefore a good practice to use the same source for data when comparing different companies or when evaluating the historical record of a single company. Analysts should verify the consistency of formulas and data classifications by the data source. Analysts should also be mindful of the judgments made by a vendor in data classifications and refer back to the source financial statements until they are comfortable that the classifications are appropriate.

Collection of financial data from regulatory filings and calculation of ratios can be automated. The eXtensible Business Reporting Language (XBRL) is a mechanism that attaches "smart tags" to financial information (e.g., total assets), so that software can automatically collect the data and perform desired computations. The organization developing XBRL (www.xbrl.org) is an international nonprofit consortium of over 600 members from companies, associations, and agencies, including the International Accounting Standards Board. Many stock exchanges and regulatory agencies around the world now use XBRL for receiving and distributing public financial reports from listed companies.

Analysts can compare a subject company to similar (peer) companies in these databases or use aggregate industry data. For non-public companies, aggregate industry data can be obtained from such sources as Annual Statement Studies by the Risk Management Association or Dun & Bradstreet. These publications typically provide industry data with companies sorted into quartiles. By definition, twenty-five percent of companies' ratios fall within the lowest quartile, 25 percent have ratios between the lower quartile and median value, and so on. Analysts can then determine a company's relative standing in the industry.

4

COMMON SIZE BALANCE SHEETS AND INCOME STATEMENTS

a describe tools and techniques used in financial analysis, including their uses and limitations

Common-size analysis involves expressing financial data, including entire financial statements, in relation to a single financial statement item, or base. Items used most frequently as the bases are total assets or revenue. In essence, common-size analysis creates a ratio between every financial statement item and the base item.

Common-size analysis was demonstrated in readings for the income statement, balance sheet, and cash flow statement. In this section, we present common-size analysis of financial statements in greater detail and include further discussion of their interpretation.

4.1 Common-Size Analysis of the Balance Sheet

A vertical[5] common-size balance sheet, prepared by dividing each item on the balance sheet by the same period's total assets and expressing the results as percentages, highlights the composition of the balance sheet. What is the mix of assets being used? How is the company financing itself? How does one company's balance sheet composition compare with that of peer companies, and what are the reasons for any differences?

A horizontal common-size balance sheet, prepared by computing the increase or decrease in percentage terms of each balance sheet item from the prior year or prepared by dividing the quantity of each item by a base year quantity of the item, highlights changes in items. These changes can be compared to expectations. The section on trend analysis below will illustrate a horizontal common-size balance sheet.

Exhibit 3 presents a vertical common-size (partial) balance sheet for a hypothetical company in two time periods. In this example, receivables have increased from 35 percent to 57 percent of total assets and the ratio has increased by 63 percent from Period 1 to Period 2. What are possible reasons for such an increase? The increase might indicate that the company is making more of its sales on a credit basis rather than a cash basis, perhaps in response to some action taken by a competitor. Alternatively, the increase in receivables as a percentage of assets may have occurred because of a change in another current asset category, for example, a decrease in the level of inventory; the analyst would then need to investigate why that asset category has changed. Another possible reason for the increase in receivables as a percentage of assets is that the company has lowered its credit standards, relaxed its collection procedures, or adopted more aggressive revenue recognition policies. The analyst can turn to other comparisons and ratios (e.g., comparing the rate of growth in accounts receivable with the rate of growth in sales) to help determine which explanation is most likely.

Exhibit 3 Vertical Common-Size (Partial) Balance Sheet for a Hypothetical Company

	Period 1 Percent of Total Assets	Period 2 Percent of Total Assets
Cash	25	15
Receivables	35	57
Inventory	35	20
Fixed assets, net of depreciation	5	8
Total assets	100	100

4.2 Common-Size Analysis of the Income Statement

A vertical common-size income statement divides each income statement item by revenue, or sometimes by total assets (especially in the case of financial institutions). If there are multiple revenue sources, a decomposition of revenue in percentage terms is useful. Exhibit 4 presents a hypothetical company's vertical common-size income statement in two time periods. Revenue is separated into the company's four services, each shown as a percentage of total revenue.

---

[5] The term **vertical analysis** is used to denote a common-size analysis using only one reporting period or one base financial statement, whereas **horizontal analysis** refers to an analysis comparing a specific financial statement with prior or future time periods or to a cross-sectional analysis of one company with another.

In this example, revenues from Service A have become a far greater percentage of the company's total revenue (30 percent in Period 1 and 45 percent in Period 2). What are possible reasons for and implications of this change in business mix? Did the company make a strategic decision to sell more of Service A, perhaps because it is more profitable? Apparently not, because the company's earnings before interest, taxes, depreciation, and amortisation (EBITDA) declined from 53 percent of sales to 45 percent, so other possible explanations should be examined. In addition, we note from the composition of operating expenses that the main reason for this decline in profitability is that salaries and employee benefits have increased from 15 percent to 25 percent of total revenue. Are more highly compensated employees required for Service A? Were higher training costs incurred in order to increase revenues from Service A? If the analyst wants to predict future performance, the causes of these changes must be understood.

In addition, Exhibit 4 shows that the company's income tax as a percentage of sales has declined dramatically (from 15 percent to 8 percent). Furthermore, taxes as a percentage of earnings before tax (EBT) (the effective tax rate, which is usually the more relevant comparison), have decreased from 36 percent (= 15/42) to 24 percent (= 8/34). Is Service A, which in Period 2 is a greater percentage of total revenue, provided in a jurisdiction with lower tax rates? If not, what is the explanation for the change in effective tax rate?

The observations based on Exhibit 4 summarize the issues that can be raised through analysis of the vertical common-size income statement.

Exhibit 4 Vertical Common-Size Income Statement for Hypothetical Company

	Period 1 Percent of Total Revenue	Period 2 Percent of Total Revenue
Revenue source: Service A	30	45
Revenue source: Service B	23	20
Revenue source: Service C	30	30
Revenue source: Service D	17	5
Total revenue	**100**	**100**
Operating expenses (excluding depreciation)		
Salaries and employee benefits	15	25
Administrative expenses	22	20
Rent expense	10	10
EBITDA	**53**	**45**
Depreciation and amortisation	4	4
EBIT	**49**	**41**
Interest paid	7	7
EBT	**42**	**34**
Income tax provision	15	8
Net income	**27**	**26**

EBIT = earnings before interest and tax.

CROSS-SECTIONAL, TREND ANALYSIS & RELATIONSHIPS IN FINANCIAL STATEMENTS

5

a describe tools and techniques used in financial analysis, including their uses and limitations

As noted previously, ratios and common-size statements derive part of their meaning through comparison to some benchmark. **Cross-sectional analysis** (sometimes called "relative analysis") compares a specific metric for one company with the same metric for another company or group of companies, allowing comparisons even though the companies might be of significantly different sizes and/or operate in different currencies. This is illustrated in Exhibit 5.

Exhibit 5 Vertical Common-Size (Partial) Balance Sheet for Two Hypothetical Companies

Assets	Company 1 Percent of Total Assets	Company 2 Percent of Total Assets
Cash	38	12
Receivables	33	55
Inventory	27	24
Fixed assets net of depreciation	1	2
Investments	1	7
Total Assets	**100**	**100**

Exhibit 5 presents a vertical common-size (partial) balance sheet for two hypothetical companies at the same point in time. Company 1 is clearly more liquid (liquidity is a function of how quickly assets can be converted into cash) than Company 2, which has only 12 percent of assets available as cash, compared with the highly liquid Company 1, which has 38 percent of assets available as cash. Given that cash is generally a relatively low-yielding asset and thus not a particularly efficient use of excess funds, why does Company 1 hold such a large percentage of total assets in cash? Perhaps the company is preparing for an acquisition, or maintains a large cash position as insulation from a particularly volatile operating environment. Another issue highlighted by the comparison in this example is the relatively high percentage of receivables in Company 2's assets, which may indicate a greater proportion of credit sales, overall changes in asset composition, lower credit or collection standards, or aggressive accounting policies.

5.1 Trend Analysis[6]

When looking at financial statements and ratios, trends in the data, whether they are improving or deteriorating, are as important as the current absolute or relative levels. Trend analysis provides important information regarding historical performance and growth and, given a sufficiently long history of accurate seasonal information, can be of great assistance as a planning and forecasting tool for management and analysts.

Exhibit 6A presents a partial balance sheet for a hypothetical company over five periods. The last two columns of the table show the changes for Period 5 compared with P+eriod 4, expressed both in absolute currency (in this case, dollars) and in percentages. A small percentage change could hide a significant currency change and vice versa, prompting the analyst to investigate the reasons despite one of the changes being relatively small. In this example, the largest percentage change was in investments, which decreased by 33.3 percent.[7] However, an examination of the absolute currency amount of changes shows that investments changed by only $2 million, and the more significant change was the $12 million increase in receivables.

Another way to present data covering a period of time is to show each item in relation to the same item in a base year (i.e., a horizontal common-size balance sheet). Exhibits 6B and 6C illustrate alternative presentations of horizontal common-size balance sheets. Exhibit 6B presents the information from the same partial balance sheet as in Exhibit 6A, but indexes each item relative to the same item in Period 1. For example, in Period 2, the company had $29 million cash, which is 74 percent or 0.74 of the amount of cash it had in Period 1. Expressed as an index relative to Period 1, where each item in Period 1 is given a value of 1.00, the value in Period 2 would be 0.74 ($29/$39 = 0.74). In Period 3, the company had $27 million cash, which is 69 percent of the amount of cash it had in Period 1 ($27/$39 = 0.69).

Exhibit 6C presents the percentage change in each item, relative to the previous year. For example, the change in cash from Period 1 to Period 2 was –25.6 percent ($29/$39 – 1 = –0.256), and the change in cash from Period 2 to Period 3 was –6.9 percent ($27/$29 – 1 = –0.069). An analyst will select the horizontal common-size balance that addresses the particular period of interest. Exhibit 6B clearly highlights that in Period 5 compared to Period 1, the company has less than half the amount of cash, four times the amount of investments, and eight times the amount of property, plant, and equipment. Exhibit 6C highlights year-to-year changes: For example, cash has declined in each period. Presenting data this way highlights significant changes. Again, note that a mathematically big change is not necessarily an important change. For example, fixed assets increased 100 percent, i.e., doubled between Period 1 and 2; however, as a proportion of total assets, fixed assets increased from 1 percent of total assets to 2 percent of total assets. The company's working capital assets (receivables and inventory) are a far higher proportion of total assets and would likely warrant more attention from an analyst.

An analysis of horizontal common-size balance sheets highlights structural changes that have occurred in a business. Past trends are obviously not necessarily an accurate predictor of the future, especially when the economic or competitive environment changes. An examination of past trends is more valuable when the macroeconomic and competitive environments are relatively stable and when the analyst is reviewing a stable or mature business. However, even in less stable contexts, historical analysis can serve as a basis for developing expectations. Understanding of past trends is helpful in assessing whether these trends are likely to continue or if the trend is likely to change direction.

---

[6] In financial statement analysis, the term "trend analysis" usually refers to comparisons across time periods of 3–10 years not involving statistical tools. This differs from the use of the term in the quantitative methods portion of the CFA curriculum, where "trend analysis" refers to statistical methods of measuring patterns in time-series data.

[7] Percentage change is calculated as (Ending value – Beginning value)/Beginning value, or equivalently, (Ending value/Beginning value) – 1.

Exhibit 6A Partial Balance Sheet for a Hypothetical Company over Five Periods

Assets ($ Millions)	Period 1	2	3	4	5	Change 4 to 5 ($ Million)	Change 4 to 5 (Percent)
Cash	39	29	27	19	16	–3	–15.8
Investments	1	7	7	6	4	–2	–33.3
Receivables	44	41	37	67	79	12	17.9
Inventory	15	25	36	25	27	2	8.0
Fixed assets net of depreciation	1	2	6	9	8	–1	–11.1
Total assets	100	104	113	126	134	8	6.3

Exhibit 6B Horizontal Common-Size (Partial) Balance Sheet for a Hypothetical Company over Five Periods, with Each Item Expressed Relative to the Same Item in Period One

Assets	Period 1	2	3	4	5
Cash	1.00	0.74	0.69	0.49	0.41
Investments	1.00	7.00	7.00	6.00	4.00
Receivables	1.00	0.93	0.84	1.52	1.80
Inventory	1.00	1.67	2.40	1.67	1.80
Fixed assets net of depreciation	1.00	2.00	6.00	9.00	8.00
Total assets	1.00	1.04	1.13	1.26	1.34

Exhibit 6C Horizontal Common-Size (Partial) Balance Sheet for a Hypothetical Company over Five Periods, with Percent Change in Each Item Relative to the Prior Period

Assets	Period 2 (%)	3 (%)	4 (%)	5 (%)
Cash	–25.6	–6.9	–29.6	–15.8
Investments	600.0	0.0	–14.3	–33.3
Receivables	–6.8	–9.8	81.1	17.9
Inventory	66.7	44.0	–30.6	8.0

(continued)

Exhibit 6C (Continued)

	Period			
Assets	**2 (%)**	**3 (%)**	**4 (%)**	**5 (%)**
Fixed assets net of depreciation	100.0	200.0	50.0	–11.1
Total assets	4.0	8.7	11.5	6.3

One measure of success is for a company to grow at a rate greater than the rate of the overall market in which it operates. Companies that grow slowly may find themselves unable to attract equity capital. Conversely, companies that grow too quickly may find that their administrative and management information systems cannot keep up with the rate of expansion.

5.2 Relationships among Financial Statements

Trend data generated by a horizontal common-size analysis can be compared across financial statements. For example, the growth rate of assets for the hypothetical company in Exhibit 6 can be compared with the company's growth in revenue over the same period of time. If revenue is growing more quickly than assets, the company may be increasing its efficiency (i.e., generating more revenue for every dollar invested in assets).

As another example, consider the following year-over-year percentage changes for a hypothetical company:

Revenue	+20%
Net income	+25%
Operating cash flow	–10%
Total assets	+30%

Net income is growing faster than revenue, which indicates increasing profitability. However, the analyst would need to determine whether the faster growth in net income resulted from continuing operations or from non-operating, non-recurring items. In addition, the 10 percent decline in operating cash flow despite increasing revenue and net income clearly warrants further investigation because it could indicate a problem with earnings quality (perhaps aggressive reporting of revenue). Lastly, the fact that assets have grown faster than revenue indicates the company's efficiency may be declining. The analyst should examine the composition of the increase in assets and the reasons for the changes. Example 5 illustrates a historical example of a company where comparisons of trend data from different financial statements were actually indicative of aggressive accounting policies.

EXAMPLE 5

Use of Comparative Growth Information[8]

In July 1996, Sunbeam, a US company, brought in new management to turn the company around. In the following year, 1997, using 1996 as the base, the following was observed based on reported numbers:

---

[8] Adapted from Robinson and Munter (2004, pp. 2–15).

Revenue	+19%
Inventory	+58%
Receivables	+38%

It is generally more desirable to observe inventory and receivables growing at a slower (or similar) rate compared to revenue growth. Receivables growing faster than revenue can indicate operational issues, such as lower credit standards or aggressive accounting policies for revenue recognition. Similarly, inventory growing faster than revenue can indicate an operational problem with obsolescence or aggressive accounting policies, such as an improper overstatement of inventory to increase profits.

In this case, the explanation lay in aggressive accounting policies. Sunbeam was later charged by the US Securities and Exchange Commission with improperly accelerating the recognition of revenue and engaging in other practices, such as billing customers for inventory prior to shipment.

THE USE OF GRAPHS AND REGRESSION ANALYSIS

6

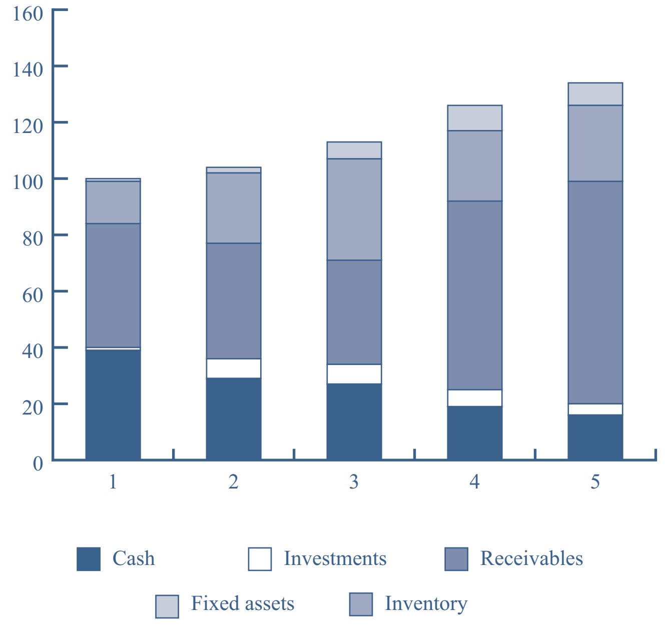
a describe tools and techniques used in financial analysis, including their uses and limitations

Graphs facilitate comparison of performance and financial structure over time, highlighting changes in significant aspects of business operations. In addition, graphs provide the analyst (and management) with a visual overview of risk trends in a business. Graphs may also be used effectively to communicate the analyst's conclusions regarding financial condition and risk management aspects.

Exhibit 7 presents the information from Exhibit 6A in a stacked column format. The graph makes the significant decline in cash and growth in receivables (both in absolute terms and as a percentage of assets) readily apparent. In Exhibit 7, the vertical axis shows US$ millions and the horizontal axis denotes the period.

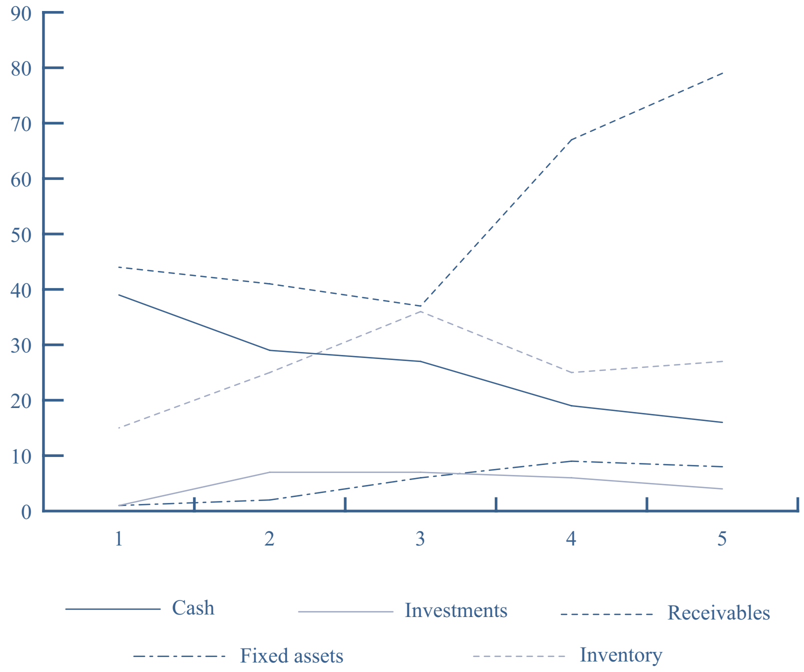
Choosing the appropriate graph to communicate the most significant conclusions of a financial analysis is a skill. In general, pie graphs are most useful to communicate the composition of a total value (e.g., assets over a limited amount of time, say one or two periods). Line graphs are useful when the focus is on the change in amount for a limited number of items over a relatively longer time period. When the composition and amounts, as well as their change over time, are all important, a stacked column graph can be useful.

Exhibit 7 Stacked Column Graph of Asset Composition of Hypothetical Company over Five Periods



When comparing Period 5 with Period 4, the growth in receivables appears to be within normal bounds; but when comparing Period 5 with earlier periods, the dramatic growth becomes apparent. In the same manner, a simple line graph will also illustrate the growth trends in key financial variables. Exhibit 8 presents the information from Exhibit 6A as a line graph, illustrating the growth of assets of a hypothetical company over five periods. The steady decline in cash, volatile movements of inventory, and dramatic growth of receivables is clearly illustrated. Again, the vertical axis is shown in US$ millions and the horizontal axis denotes periods.

Exhibit 8 Line Graph of Growth of Assets of Hypothetical Company over Five Periods



6.1 Regression Analysis

When analyzing the trend in a specific line item or ratio, frequently it is possible simply to visually evaluate the changes. For more complex situations, regression analysis can help identify relationships (or correlation) between variables. For example, a regression analysis could relate a company's sales to GDP over time, providing insight into whether the company is cyclical. In addition, the statistical relationship between sales and GDP could be used as a basis for forecasting sales.

Other examples where regression analysis may be useful include the relationship between a company's sales and inventory over time, or the relationship between hotel occupancy and a company's hotel revenues. In addition to providing a basis for forecasting, regression analysis facilitates identification of items or ratios that are not behaving as expected, given historical statistical relationships.

COMMON RATIO CATEGORIES & INTERPRETATION AND CONTEXT

7

b identify, calculate, and interpret activity, liquidity, solvency, profitability, and valuation ratios

In the previous section, we focused on ratios resulting from common-size analysis. In this section, we expand the discussion to include other commonly used financial ratios and the broad classes into which they are categorized. There is some overlap with common-size financial statement ratios. For example, a common indicator of profitability is the net profit margin, which is calculated as net income divided by sales. This ratio appears on a vertical common-size income statement. Other ratios involve information from multiple financial statements or even data from outside the financial statements.

Because of the large number of ratios, it is helpful to think about ratios in terms of broad categories based on what aspects of performance a ratio is intended to detect. Financial analysts and data vendors use a variety of categories to classify ratios. The category names and the ratios included in each category can differ. Common ratio categories include activity, liquidity, solvency, profitability, and valuation. These categories are summarized in Exhibit 9. Each category measures a different aspect of the company's business, but all are useful in evaluating a company's overall ability to generate cash flows from operating its business and the associated risks.

Exhibit 9 Categories of Financial Ratios

Category	Description
Activity	**Activity ratios** measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.
Liquidity	**Liquidity ratios** measure the company's ability to meet its short-term obligations.
Solvency	**Solvency ratios** measure a company's ability to meet long-term obligations. Subsets of these ratios are also known as "leverage" and "long-term debt" ratios.
Profitability	**Profitability ratios** measure the company's ability to generate profits from its resources (assets).
Valuation	**Valuation ratios** measure the quantity of an asset or flow (e.g., earnings) associated with ownership of a specified claim (e.g., a share or ownership of the enterprise).

These categories are not mutually exclusive; some ratios are useful in measuring multiple aspects of the business. For example, an activity ratio measuring how quickly a company collects accounts receivable is also useful in assessing the company's liquidity because collection of revenues increases cash. Some profitability ratios also reflect the operating efficiency of the business. In summary, analysts appropriately use certain ratios to evaluate multiple aspects of the business. Analysts also need to be aware of variations in industry practice in the calculation of financial ratios. In the text that follows, alternative views on ratio calculations are often provided.

7.1 Interpretation and Context

Financial ratios can only be interpreted in the context of other information, including benchmarks. In general, the financial ratios of a company are compared with those of its major competitors (cross-sectional and trend analysis) and to the company's prior periods (trend analysis). The goal is to understand the underlying causes of divergence between a company's ratios and those of the industry. Even ratios that remain consistent require understanding because consistency can sometimes indicate accounting policies selected to smooth earnings. An analyst should evaluate financial ratios based on the following:

1 *Company goals and strategy.* Actual ratios can be compared with company objectives to determine whether objectives are being attained and whether the results are consistent with the company's strategy.

2 *Industry norms* (*cross-sectional analysis*). A company can be compared with others in its industry by relating its financial ratios to industry norms or to a subset of the companies in an industry. When industry norms are used to make judgments, care must be taken because:

- Many ratios are industry specific, and not all ratios are important to all industries.
- Companies may have several different lines of business. This will cause aggregate financial ratios to be distorted. It is better to examine industry-specific ratios by lines of business.
- Differences in accounting methods used by companies can distort financial ratios.
- Differences in corporate strategies can affect certain financial ratios.

3 *Economic conditions*. For cyclical companies, financial ratios tend to improve when the economy is strong and weaken during recessions. Therefore, financial ratios should be examined in light of the current phase of the business cycle.

The following sections discuss activity, liquidity, solvency, and profitability ratios in turn. Selected valuation ratios are presented later in the section on equity analysis.

8

ACTIVITY RATIOS

Activity ratios are also known as **asset utilization ratios** or **operating efficiency ratios**. This category is intended to measure how well a company manages various activities, particularly how efficiently it manages its various assets. Activity ratios are analyzed as indicators of ongoing operational performance—how effectively assets are used by a company. These ratios reflect the efficient management of both working capital and longer term assets. As noted, efficiency has a direct impact on liquidity (the ability of a company to meet its short-term obligations), so some activity ratios are also useful in assessing liquidity.

8.1 Calculation of Activity Ratios

Exhibit 10 presents the most commonly used activity ratios. The exhibit shows the numerator and denominator of each ratio.

Exhibit 10 Definitions of Commonly Used Activity Ratios

Activity Ratios	Numerator	Denominator
Inventory turnover	Cost of sales or cost of goods sold	Average inventory
Days of inventory on hand (DOH)	Number of days in period	Inventory turnover
Receivables turnover	Revenue	Average receivables
Days of sales outstanding (DSO)	Number of days in period	Receivables turnover
Payables turnover	Purchases	Average trade payables
Number of days of payables	Number of days in period	Payables turnover
Working capital turnover	Revenue	Average working capital

(continued)

Exhibit 10 (Continued)

Activity Ratios	Numerator	Denominator
Fixed asset turnover	Revenue	Average net fixed assets
Total asset turnover	Revenue	Average total assets

Activity ratios measure how efficiently the company utilizes assets. They generally combine information from the income statement in the numerator with balance sheet items in the denominator. Because the income statement measures what happened *during* a period whereas the balance sheet shows the condition only at the end of the period, average balance sheet data are normally used for consistency. For example, to measure inventory management efficiency, cost of sales or cost of goods sold (from the income statement) is divided by average inventory (from the balance sheet). Most databases, such as Bloomberg and Baseline, use this averaging convention when income statement and balance sheet data are combined. These databases typically average only two points: the beginning of the year and the end of the year. The examples that follow based on annual financial statements illustrate that practice. However, some analysts prefer to average more observations if they are available, especially if the business is seasonal. If a semiannual report is prepared, an average can be taken over three data points (beginning, middle, and end of year). If quarterly data are available, a five-point average can be computed (beginning of year and end of each quarterly period) or a four-point average using the end of each quarterly period. Note that if the company's year ends at a low or high point for inventory for the year, there can still be bias in using three or five data points, because the beginning and end of year occur at the same time of the year and are effectively double counted.

Because cost of goods sold measures the cost of inventory that has been sold, this ratio measures how many times per year the entire inventory was theoretically turned over, or sold. (We say that the entire inventory was "theoretically" sold because in practice companies do not generally sell out their entire inventory.) If, for example, a company's cost of goods sold for a recent year was €120,000 and its average inventory was €10,000, the inventory turnover ratio would be 12. The company theoretically turns over (i.e., sells) its entire inventory 12 times per year (i.e., once a month). (Again, we say "theoretically" because in practice the company likely carries some inventory from one month into another.) Turnover can then be converted to days of inventory on hand (DOH) by dividing inventory turnover into the number of days in the accounting period. In this example, the result is a DOH of 30.42 (365/12), meaning that, on average, the company's inventory was on hand for about 30 days, or, equivalently, the company kept on hand about 30 days' worth of inventory, on average, during the period.

Activity ratios can be computed for any annual or interim period, but care must be taken in the interpretation and comparison across periods. For example, if the same company had cost of goods sold for the first quarter (90 days) of the following year of €35,000 and average inventory of €11,000, the inventory turnover would be 3.18 times. However, this turnover rate is 3.18 times per quarter, which is not directly comparable to the 12 times per year in the preceding year. In this case, we can annualize the quarterly inventory turnover rate by multiplying the quarterly turnover by 4 (12 months/3 months; or by 4.06, using 365 days/90 days) for comparison to the annual turnover rate. So, the quarterly inventory turnover is equivalent to a 12.72 annual inventory turnover (or 12.91 if we annualize the ratio using a 90-day quarter and a 365-day year). To compute the DOH using quarterly data, we can use the quarterly turnover rate and the number of days in the quarter for the numerator—or, we can use the annualized turnover rate and 365 days; either results in DOH of around 28.3, with slight differences due to rounding (90/3.18 = 28.30 and 365/12.91 = 28.27). Another time-related computational detail is that for companies using a 52/53-week annual period and for leap years, the actual days in the year should be used rather than 365.

In some cases, an analyst may want to know how many days of inventory are on hand at the end of the year rather than the average for the year. In this case, it would be appropriate to use the year-end inventory balance in the computation rather than the average. If the company is growing rapidly or if costs are increasing rapidly, analysts should consider using cost of goods sold just for the fourth quarter in this computation because the cost of goods sold of earlier quarters may not be relevant. Example 6 further demonstrates computation of activity ratios using Hong Kong Stock Exchange(HKEX)-listed Lenovo Group Limited.

EXAMPLE 6

Computation of Activity Ratios

An analyst would like to evaluate Lenovo Group's efficiency in collecting its trade accounts receivable during the fiscal year ended 31 March 2018 (FY2017). The analyst gathers the following information from Lenovo's annual and interim reports:

	US$ in Thousands
Trade receivables as of 31 March 2017	4,468,392
Trade receivables as of 31 March 2018	4,972,722
Revenue for year ended 31 March 2018	45,349,943

Calculate Lenovo's receivables turnover and number of days of sales outstanding (DSO) for the fiscal year ended 31 March 2018.

Solution:

Receivables turnover = Revenue/Average receivables
= 45,349,943/[(4,468,392 + 4,972,722)/2]
= 45,349,943/4,720,557
= 9.6069 times, or 9.6 rounded

DSO = Number of days in period/Receivables turnover
= 365/9.6
= 38.0 days

On average, it took Lenovo 38 days to collect receivables during the fiscal year ended 31 March 2018.

8.2 Interpretation of Activity Ratios

In the following section, we further discuss the activity ratios that were defined in Exhibit 10.

Inventory Turnover and DOH

Inventory turnover lies at the heart of operations for many entities. It indicates the resources tied up in inventory (i.e., the carrying costs) and can, therefore, be used to indicate inventory management effectiveness. A higher inventory turnover ratio implies a shorter period that inventory is held, and thus a lower DOH. In general, inventory turnover and DOH should be benchmarked against industry norms.

A high inventory turnover ratio relative to industry norms might indicate highly effective inventory management. Alternatively, a high inventory turnover ratio (and commensurately low DOH) could possibly indicate the company does not carry adequate inventory, so shortages could potentially hurt revenue. To assess which explanation is more likely, the analyst can compare the company's revenue growth with that of the industry. Slower growth combined with higher inventory turnover could indicate inadequate inventory levels. Revenue growth at or above the industry's growth supports the interpretation that the higher turnover reflects greater inventory management efficiency.

A low inventory turnover ratio (and commensurately high DOH) relative to the rest of the industry could be an indicator of slow-moving inventory, perhaps due to technological obsolescence or a change in fashion. Again, comparing the company's sales growth with the industry can offer insight.

Receivables Turnover and DSO.

The number of DSO represents the elapsed time between a sale and cash collection, reflecting how fast the company collects cash from customers to whom it offers credit. Although limiting the numerator to sales made on credit in the receivables turnover would be more appropriate, credit sales information is not always available to analysts; therefore, revenue as reported in the income statement is generally used as an approximation.

A relatively high receivables turnover ratio (and commensurately low DSO) might indicate highly efficient credit and collection. Alternatively, a high receivables turnover ratio could indicate that the company's credit or collection policies are too stringent, suggesting the possibility of sales being lost to competitors offering more lenient terms. A relatively low receivables turnover ratio would typically raise questions about the efficiency of the company's credit and collections procedures. As with inventory management, comparison of the company's sales growth relative to the industry can help the analyst assess whether sales are being lost due to stringent credit policies. In addition, comparing the company's estimates of uncollectible accounts receivable and actual credit losses with past experience and with peer companies can help assess whether low turnover reflects credit management issues. Companies often provide details of receivables aging (how much receivables have been outstanding by age). This can be used along with DSO to understand trends in collection, as demonstrated in Example 7.

EXAMPLE 7

Evaluation of an Activity Ratio

An analyst has computed the average DSO for Lenovo for fiscal years ended 31 March 2018 and 2017:

	FY2017	FY2016
Days of sales outstanding	38.0	37.6

Revenue increased from US$43.035 billion for fiscal year ended 31 March 2017 (FY2016) to US$45.350 billion for fiscal year ended 31 March 2018 (FY2017). The analyst would like to better understand the change in the company's DSO from FY2016 to FY2017 and whether the increase is indicative of any issues with the customers' credit quality. The analyst collects accounts receivable aging information from Lenovo's annual reports and computes the percentage of accounts receivable by days outstanding. This information is presented in Exhibit 11:

Exhibit 11

	FY2017		FY2016		FY2015	
	US$000	**Percent**	**US$000**	**Percent**	**US$000**	**Percent**
Accounts receivable						
0–30 days	3,046,240	59.95	2,923,083	63.92	3,246,600	71.99
31–60 days	1,169,286	23.01	985,251	21.55	617,199	13.69
61–90 days	320,183	6.30	283,050	6.19	240,470	5.33
Over 90 days	545,629	10.74	381,387	8.34	405,410	8.99
Total	5,081,338	100.00	4,572,771	100.00	4,509,679	100.00
Less: Provision for impairment	–108,616	–2.14	–104,379	–2.28	–106,172	–2.35
Trade receivables, net	4,972,722	97.86	4,468,392	97.72	4,403,507	97.65
Total sales	*45,349,943*		*43,034,731*		*44,912,097*	

Note: Lenovo's footnotes disclose that general trade customers are provided with credit terms ranging from 0 to 120 days.

These data indicate that total accounts receivable increased by about 11.3% in FY2017 versus FY2016, while total sales increased by only 5.4%. Further, the percentage of receivables in all categories older than 30 days has increased over the three-year period, indicating that customers are indeed taking longer to pay. On the other hand, the provision for impairment (estimate of uncollectible accounts) has declined as a percent of total receivables. Considering all this information, the company may be increasing customer financing purposely to drive its sales growth. They also may be underestimating the impairment. This should be investigated further by the analyst.

Payables Turnover and the Number of Days of Payables

The number of days of payables reflects the average number of days the company takes to pay its suppliers, and the payables turnover ratio measures how many times per year the company theoretically pays off all its creditors. For purposes of calculating these ratios, an implicit assumption is that the company makes all its purchases using credit. If the amount of purchases is not directly available, it can be computed as cost of goods sold plus ending inventory less beginning inventory. Alternatively, cost of goods sold is sometimes used as an approximation of purchases.

A payables turnover ratio that is high (low days payable) relative to the industry could indicate that the company is not making full use of available credit facilities; alternatively, it could result from a company taking advantage of early payment discounts. An excessively low turnover ratio (high days payable) could indicate trouble making payments on time, or alternatively, exploitation of lenient supplier terms. This is another example where it is useful to look simultaneously at other ratios. If liquidity ratios indicate that the company has sufficient cash and other short-term assets to pay obligations and yet the days payable ratio is relatively high, the analyst would favor the lenient supplier credit and collection policies as an explanation.

Working Capital Turnover

Working capital is defined as current assets minus current liabilities. Working capital turnover indicates how efficiently the company generates revenue with its working capital. For example, a working capital turnover ratio of 4.0 indicates that the company generates €4 of revenue for every €1 of working capital. A high working capital turnover ratio indicates greater efficiency (i.e., the company is generating a high level of revenues relative to working capital). For some companies, working capital can be near zero or negative, rendering this ratio incapable of being interpreted. The following two ratios are more useful in those circumstances.

Fixed Asset Turnover

This ratio measures how efficiently the company generates revenues from its investments in fixed assets. Generally, a higher fixed asset turnover ratio indicates more efficient use of fixed assets in generating revenue. A low ratio can indicate inefficiency, a capital-intensive business environment, or a new business not yet operating at full capacity—in which case the analyst will not be able to link the ratio directly to efficiency. In addition, asset turnover can be affected by factors other than a company's efficiency. The fixed asset turnover ratio would be lower for a company whose assets are newer (and, therefore, less depreciated and so reflected in the financial statements at a higher carrying value) than the ratio for a company with older assets (that are thus more depreciated and so reflected at a lower carrying value). The fixed asset ratio can be erratic because, although revenue may have a steady growth rate, increases in fixed assets may not follow a smooth pattern; so, every year-to-year change in the ratio does not necessarily indicate important changes in the company's efficiency.

Total Asset Turnover

The total asset turnover ratio measures the company's overall ability to generate revenues with a given level of assets. A ratio of 1.20 would indicate that the company is generating €1.20 of revenues for every €1 of average assets. A higher ratio indicates greater efficiency. Because this ratio includes both fixed and current assets, inefficient working capital management can distort overall interpretations. It is therefore helpful to analyze working capital and fixed asset turnover ratios separately.

A low asset turnover ratio can be an indicator of inefficiency or of relative capital intensity of the business. The ratio also reflects strategic decisions by management—for example, the decision whether to use a more labor-intensive (and less capital-intensive) approach to its business or a more capital-intensive (and less labor-intensive) approach.

When interpreting activity ratios, the analysts should examine not only the individual ratios but also the collection of relevant ratios to determine the overall efficiency of a company. Example 8 demonstrates the evaluation of activity ratios, both narrow (e.g., days of inventory on hand) and broad (e.g., total asset turnover) for a hypothetical manufacturer.

EXAMPLE 8

Evaluation of Activity Ratios

ZZZ Company is a hypothetical manufacturing company. As part of an analysis of management's operating efficiency, an analyst collects the following activity ratios from a data provider:

Ratio	2018	2017	2016	2015
DOH	35.68	40.70	40.47	48.51
DSO	45.07	58.28	51.27	76.98
Total asset turnover	0.36	0.28	0.23	0.22

These ratios indicate that the company has improved on all three measures of activity over the four-year period. The company appears to be managing its inventory more efficiently, is collecting receivables faster, and is generating a higher level of revenues relative to total assets. The overall trend appears good, but thus far, the analyst has only determined *what* happened. A more important question is *why* the ratios improved, because understanding good changes as well as bad ones facilitates judgments about the company's future performance. To answer this question, the analyst examines company financial reports as well as external information about the industry and economy. In examining the annual report, the analyst notes that in the fourth quarter of 2018, the company experienced an "inventory correction" and that the company recorded an allowance for the decline in market value and obsolescence of inventory of about 15 percent of year-end inventory value (compared with about a 6 percent allowance in the prior year). This reduction in the value of inventory accounts for a large portion of the decline in DOH from 40.70 in 2017 to 35.68 in 2018. Management claims that this inventory obsolescence is a short-term issue; analysts can watch DOH in future interim periods to confirm this assertion. In any event, all else being equal, the analyst would likely expect DOH to return to a level closer to 40 days going forward.

More positive interpretations can be drawn from the total asset turnover. The analyst finds that the company's revenues increased more than 35 percent while total assets only increased by about 6 percent. Based on external information about the industry and economy, the analyst attributes the increased revenues both to overall growth in the industry and to the company's increased market share. Management was able to achieve growth in revenues with a comparatively modest increase in assets, leading to an improvement in total asset turnover. Note further that part of the reason for the increase in asset turnover is lower DOH and DSO.

LIQUIDITY RATIOS

9

b identify, calculate, and interpret activity, liquidity, solvency, profitability, and valuation ratios

Liquidity analysis, which focuses on cash flows, measures a company's ability to meet its short-term obligations. Liquidity measures how quickly assets are converted into cash. Liquidity ratios also measure the ability to pay off short-term obligations. In day-to-day operations, liquidity management is typically achieved through efficient use of assets. In the medium term, liquidity in the non-financial sector is also addressed by managing the structure of liabilities. (See the discussion on financial sector below.)

The level of liquidity needed differs from one industry to another. A particular company's liquidity position may vary according to the anticipated need for funds at any given time. Judging whether a company has adequate liquidity requires analysis of its historical funding requirements, current liquidity position, anticipated future funding needs, and options for reducing funding needs or attracting additional funds (including actual and potential sources of such funding).

Larger companies are usually better able to control the level and composition of their liabilities than smaller companies. Therefore, they may have more potential funding sources, including public capital and money markets. Greater discretionary access to capital markets also reduces the size of the liquidity buffer needed relative to companies without such access.

Contingent liabilities, such as letters of credit or financial guarantees, can also be relevant when assessing liquidity. The importance of contingent liabilities varies for the non-banking and banking sector. In the non-banking sector, contingent liabilities (usually disclosed in the footnotes to the company's financial statements) represent potential cash outflows, and when appropriate, should be included in an assessment of a company's liquidity. In the banking sector, contingent liabilities represent potentially significant cash outflows that are not dependent on the bank's financial condition. Although outflows in normal market circumstances typically may be low, a general macroeconomic or market crisis can trigger a substantial increase in cash outflows related to contingent liabilities because of the increase in defaults and business bankruptcies that often accompany such events. In addition, such crises are usually characterized by diminished levels of overall liquidity, which can further exacerbate funding shortfalls. Therefore, for the banking sector, the effect of contingent liabilities on liquidity warrants particular attention.

9.1 Calculation of Liquidity Ratios

Common liquidity ratios are presented in Exhibit 12. These liquidity ratios reflect a company's position at a point in time and, therefore, typically use data from the ending balance sheet rather than averages. The current, quick, and cash ratios reflect three measures of a company's ability to pay current liabilities. Each uses a progressively stricter definition of liquid assets.

The **defensive interval ratio** measures how long a company can pay its daily cash expenditures using only its existing liquid assets, without additional cash flow coming in. This ratio is similar to the "burn rate" often computed for start-up internet companies in the late 1990s or for biotechnology companies. The numerator of this ratio includes the same liquid assets used in the quick ratio, and the denominator is an estimate of daily cash expenditures. To obtain daily cash expenditures, the total of cash expenditures for the period is divided by the number of days in the period. Total cash expenditures for a period can be approximated by summing all expenses on the income statement—such as cost of goods sold; selling, general, and administrative expenses; and research and development expenses—and then subtracting any non-cash expenses, such as depreciation and amortisation. (Typically, taxes are not included.)

The **cash conversion cycle**, a financial metric not in ratio form, measures the length of time required for a company to go from cash paid (used in its operations) to cash received (as a result of its operations). The cash conversion cycle is sometimes expressed as the length of time funds are tied up in working capital. During this period of time, the company needs to finance its investment in operations through other sources (i.e., through debt or equity).

Exhibit 12 Definitions of Commonly Used Liquidity Ratios

Liquidity Ratios	Numerator	Denominator
Current ratio	Current assets	Current liabilities
Quick ratio	Cash + Short-term marketable investments + Receivables	Current liabilities
Cash ratio	Cash + Short-term marketable investments	Current liabilities
Defensive interval ratio	Cash + Short-term marketable investments + Receivables	Daily cash expenditures
Additional Liquidity Measure		
Cash conversion cycle (net operating cycle)	DOH + DSO − Number of days of payables	

9.2 Interpretation of Liquidity Ratios

In the following, we discuss the interpretation of the five basic liquidity measures presented in Exhibit 12.

Current Ratio

This ratio expresses current assets in relation to current liabilities. A higher ratio indicates a higher level of liquidity (i.e., a greater ability to meet short-term obligations). A current ratio of 1.0 would indicate that the book value of its current assets exactly equals the book value of its current liabilities.

A lower ratio indicates less liquidity, implying a greater reliance on operating cash flow and outside financing to meet short-term obligations. Liquidity affects the company's capacity to take on debt. The current ratio implicitly assumes that inventories and accounts receivable are indeed liquid (which is presumably not the case when related turnover ratios are low).

Quick Ratio

The quick ratio is more conservative than the current ratio because it includes only the more liquid current assets (sometimes referred to as "quick assets") in relation to current liabilities. Like the current ratio, a higher quick ratio indicates greater liquidity.

The quick ratio reflects the fact that certain current assets—such as prepaid expenses, some taxes, and employee-related prepayments—represent costs of the current period that have been paid in advance and cannot usually be converted back into cash. This ratio also reflects the fact that inventory might not be easily and quickly converted into cash, and furthermore, that a company would probably not be able to sell all of its inventory for an amount equal to its carrying value, especially if it were required to sell the inventory quickly. In situations where inventories are illiquid (as indicated, for example, by low inventory turnover ratios), the quick ratio may be a better indicator of liquidity than is the current ratio.

Cash Ratio

The cash ratio normally represents a reliable measure of an entity's liquidity in a crisis situation. Only highly marketable short-term investments and cash are included. In a general market crisis, the fair value of marketable securities could decrease significantly as a result of market factors, in which case even this ratio might not provide reliable information.

Defensive Interval Ratio

This ratio measures how long the company can continue to pay its expenses from its existing liquid assets without receiving any additional cash inflow. A defensive interval ratio of 50 would indicate that the company can continue to pay its operating expenses for 50 days before running out of quick assets, assuming no additional cash inflows. A higher defensive interval ratio indicates greater liquidity. If a company's defensive interval ratio is very low relative to peer companies or to the company's own history, the analyst would want to ascertain whether there is sufficient cash inflow expected to mitigate the low defensive interval ratio.

Cash Conversion Cycle (Net Operating Cycle)

This metric indicates the amount of time that elapses from the point when a company invests in working capital until the point at which the company collects cash. In the typical course of events, a merchandising company acquires inventory on credit, incurring accounts payable. The company then sells that inventory on credit, increasing accounts receivable. Afterwards, it pays out cash to settle its accounts payable, and it collects cash in settlement of its accounts receivable. The time between the outlay of cash and the collection of cash is called the "cash conversion cycle." A shorter cash conversion cycle indicates greater liquidity. A short cash conversion cycle implies that the company only needs to finance its inventory and accounts receivable for a short period of time. A longer cash conversion cycle indicates lower liquidity; it implies that the company must finance its inventory and accounts receivable for a longer period of time, possibly indicating a need for a higher level of capital to fund current assets. Example 9 demonstrates the advantages of a short cash conversion cycle as well as how a company's business strategies are reflected in financial ratios.

EXAMPLE 9

Evaluation of Liquidity Measures

An analyst is evaluating the liquidity of Apple and calculates the number of days of receivables, inventory, and accounts payable, as well as the overall cash conversion cycle, as follows:

	FY2017	FY2016	FY2015
DSO	27	28	27
DOH	9	6	6
Less: Number of days of payables	112	101	86
Equals: Cash conversion cycle	(76)	(67)	(53)

The minimal DOH indicates that Apple maintains lean inventories, which is attributable to key aspects of the company's business model where manufacturing is outsourced. In isolation, the increase in number of days payable (from 86 days in FY2015 to 112 days in FY2017) might suggest an inability to pay suppliers; however, in Apple's case, the balance sheet indicates that the company has more than $70 billion of cash and short-term investments, which would be more than enough to pay suppliers sooner if Apple chose to do so. Instead, Apple takes advantage of the favorable credit terms granted by its suppliers. The overall effect is a negative cash cycle, a somewhat unusual result. Instead of requiring additional capital to fund working capital as is the case for most companies, Apple has excess cash to invest for over 50 days during that three-year period (reflected on the balance sheet as short-term investments) on which it is earning, rather than paying, interest.

EXAMPLE 10

Bounds and Context of Financial Measures

The previous example focused on the cash conversion cycle, which many companies identify as a key performance metric. The less positive the number of days in the cash conversion cycle, typically, the better it is considered to be. However, is this always true?

This example considers the following question: If a larger negative number of days in a cash conversion cycle is considered to be a desirable performance metric, does identifying a company with a large negative cash conversion cycle necessarily imply good performance?

Using a historical example, National Datacomputer, a technology company, had large negative number of days in its cash conversion cycle during the 2005 to 2009 period. In 2008 its cash conversion cycle was 275.5 days.

Exhibit 13 National Datacomputer Inc. ($ millions)

Fiscal year	2004	2005	2006	2007	2008	2009
Sales	3.248	2.672	2.045	1.761	1.820	1.723
Cost of goods sold	1.919	1.491	0.898	1.201	1.316	1.228
Receivables, Total	0.281	0.139	0.099	0.076	0.115	0.045
Inventories, Total	0.194	0.176	0.010	0.002	0.000	0.000
Accounts payable	0.223	0.317	0.366	1.423	0.704	0.674
DSO		28.69	21.24	18.14	19.15	16.95
DOH		45.29	37.80	1.82	0.28	0.00
*Less: Number of days of payables**		66.10	138.81	271.85	294.97	204.79
Equals: Cash conversion cycle		7.88	–79.77	–251.89	–275.54	–187.84

**Notes*: Calculated using Cost of goods sold as an approximation of purchases. Ending inventories 2008 and 2009 are reported as $0 million; therefore, inventory turnover for 2009 cannot be measured. However, given inventory and average sales per day, DOH in 2009 is 0.00.

Source: Raw data from Compustat. Ratios calculated.

The reason for the negative cash conversion cycle is that the company's accounts payable increased substantially over the period. An increase from approximately 66 days in 2005 to 295 days in 2008 to pay trade creditors is clearly a negative signal. In addition, the company's inventories disappeared, most likely because the company did not have enough cash to purchase new inventory and was unable to get additional credit from its suppliers.

Of course, an analyst would have immediately noted the negative trends in these data, as well as additional data throughout the company's financial statements. In its MD&A, the company clearly reports the risks as follows:

> Because we have historically had losses and only a limited amount of cash has been generated from operations, we have funded our operating activities to date primarily from the sale of securities and from the sale of a product line in 2009. In order to continue to fund our operations, we may need to raise additional capital, through the sale of securities. We cannot be certain that any such financing will be available on acceptable terms, or at all. Moreover, additional equity financing, if available, would likely be dilutive to the holders of our common stock, and debt financing, if available, would likely involve restrictive covenants and a security interest in all or substantially all of our assets. If we fail to obtain acceptable financing when needed, we may not have sufficient resources to fund our normal operations which would have a material adverse effect on our business.
>
> IF WE ARE UNABLE TO GENERATE ADEQUATE WORKING CAPITAL FROM OPERATIONS OR RAISE ADDITIONAL CAPITAL THERE IS SUBSTANTIAL DOUBT ABOUT THE COMPANY'S ABILITY TO CONTINUE AS A GOING CONCERN. (emphasis added by company)
>
> *Source*: National Datacomputer Inc., 2009 Form 10-K, page 7.

Subsequently, the company's 2010 Form 10K reported:

> "In January 2011, due to our inability to meet our financial obligations and the impending loss of a critical distribution agreement granting us the right to distribute certain products, our secured lenders ("Secured Parties") acting upon an event of default, sold certain of our assets (other than cash and accounts receivable) to Micronet, Ltd. ("Micronet"), an unaffiliated corporation pursuant to the terms of an asset purchase agreement between the Secured Parties and Micronet dated January 10, 2010 (the "Asset Purchase Agreement"). In order to induce Micronet to enter into the agreement, the Company also provided certain representations and warranties regarding certain business matters."

In summary, it is always necessary to consider ratios within bounds of reasonability and to understand the reasons underlying changes in ratios. Ratios must not only be calculated but must also be interpreted by an analyst.

10 SOLVENCY RATIOS

b identify, calculate, and interpret activity, liquidity, solvency, profitability, and valuation ratios

Solvency refers to a company's ability to fulfill its long-term debt obligations. Assessment of a company's ability to pay its long-term obligations (i.e., to make interest and principal payments) generally includes an in-depth analysis of the components of its financial structure. Solvency ratios provide information regarding the relative amount of debt in the company's capital structure and the adequacy of earnings and cash flow to cover interest expenses and other fixed charges (such as lease or rental payments) as they come due.

Analysts seek to understand a company's use of debt for several main reasons. One reason is that the amount of debt in a company's capital structure is important for assessing the company's risk and return characteristics, specifically its financial leverage. Leverage is a magnifying effect that results from the use of **fixed costs**—costs that stay the same within some range of activity—and can take two forms: operating leverage and financial leverage.

Operating leverage results from the use of fixed costs in conducting the company's business. Operating leverage magnifies the effect of changes in sales on operating income. Profitable companies may use operating leverage because when revenues increase, with operating leverage, their operating income increases at a faster rate. The explanation is that, although **variable costs** will rise proportionally with revenue, fixed costs will not.

When financing a company (i.e., raising capital for it), the use of debt constitutes **financial leverage** because interest payments are essentially fixed financing costs. As a result of interest payments, a given percent change in EBIT results in a larger percent change in earnings before taxes (EBT). Thus, financial leverage tends to magnify the effect of changes in EBIT on returns flowing to equity holders. Assuming that a company can earn more on funds than it pays in interest, the inclusion of some level of debt in a company's capital structure may lower a company's overall cost of capital and increase returns to equity holders. However, a higher level of debt in a company's capital structure increases the risk of default and results in higher borrowing costs for the company to compensate lenders for assuming greater credit risk. Starting with Modigliani and Miller (1958, 1963), a substantial amount of research has focused on determining a company's optimal capital structure and the subject remains an important one in corporate finance.

In analyzing financial statements, an analyst aims to understand levels and trends in a company's use of financial leverage in relation to past practices and the practices of peer companies. Analysts also need to be aware of the relationship between operating leverage (results from the use of non-current assets with fixed costs) and financial leverage (results from the use of long-term debt with fixed costs). The greater a company's operating leverage, the greater the risk of the operating income stream available to cover debt payments; operating leverage can thus limit a company's capacity to use financial leverage.

A company's relative solvency is fundamental to valuation of its debt securities and its creditworthiness. Finally, understanding a company's use of debt can provide analysts with insight into the company's future business prospects because management's decisions about financing may signal their beliefs about a company's future. For example, the issuance of long-term debt to repurchase common shares may indicate that management believes the market is underestimating the company's prospects and that the shares are undervalued.

10.1 Calculation of Solvency Ratios

Solvency ratios are primarily of two types. Debt ratios, the first type, focus on the balance sheet and measure the amount of debt capital relative to equity capital. Coverage ratios, the second type, focus on the income statement and measure the ability of a company to cover its debt payments. These ratios are useful in assessing a company's solvency and, therefore, in evaluating the quality of a company's bonds and other debt obligations.

Exhibit 14 describes commonly used solvency ratios. The first three of the debt ratios presented use total debt in the numerator. The definition of total debt used in these ratios varies among informed analysts and financial data vendors, with some using the total of interest-bearing short-term and long-term debt, excluding liabilities such as accrued expenses and accounts payable. (For calculations in this reading, we use this definition.) Other analysts use definitions that are more inclusive (e.g., all liabilities) or restrictive (e.g., long-term debt only, in which case the ratio is sometimes qualified as "long-term," as in "long-term debt-to-equity ratio"). If using different definitions of total debt materially changes conclusions about a company's solvency, the reasons for the discrepancies warrant further investigation.

Exhibit 14 Definitions of Commonly Used Solvency Ratios

Solvency Ratios	Numerator	Denominator
Debt Ratios		
Debt-to-assets ratio[a]	Total debt[b]	Total assets
Debt-to-capital ratio	Total debt[b]	Total debt[b] + Total shareholders' equity
Debt-to-equity ratio	Total debt[b]	Total shareholders' equity
Financial leverage ratio[c]	Average total assets	Average total equity
Debt-to-EBITDA	Total debt	EBITDA
Coverage Ratios		
Interest coverage	EBIT	Interest payments
Fixed charge coverage	EBIT + Lease payments	Interest payments + Lease payments

[a] "Total debt ratio" is another name sometimes used for this ratio.
[b] In this reading, total debt is the sum of interest-bearing short-term and long-term debt.
[c] *Average* total assets divided by *average* total equity is used for the purposes of this reading (in particular, Dupont analysis covered later). In practice, period-end total assets divided by period-end total equity is often used.

10.2 Interpretation of Solvency Ratios

In the following, we discuss the interpretation of the basic solvency ratios presented in Exhibit 14.

Debt-to-Assets Ratio

This ratio measures the percentage of total assets financed with debt. For example, a **debt-to-assets ratio** of 0.40 or 40 percent indicates that 40 percent of the company's assets are financed with debt. Generally, higher debt means higher financial risk and thus weaker solvency.

Debt-to-Capital Ratio

The **debt-to-capital ratio** measures the percentage of a company's capital (debt plus equity) represented by debt. As with the previous ratio, a higher ratio generally means higher financial risk and thus indicates weaker solvency.

Debt-to-Equity Ratio

The **debt-to-equity ratio** measures the amount of debt capital relative to equity capital. Interpretation is similar to the preceding two ratios (i.e., a higher ratio indicates weaker solvency). A ratio of 1.0 would indicate equal amounts of debt and equity, which is equivalent to a debt-to-capital ratio of 50 percent. Alternative definitions of this ratio use the market value of stockholders' equity rather than its book value (or use the market values of both stockholders' equity and debt).

Financial Leverage Ratio

This ratio (often called simply the "leverage ratio") measures the amount of total assets supported for each one money unit of equity. For example, a value of 3 for this ratio means that each €1 of equity supports €3 of total assets. The higher the **financial leverage ratio**, the more leveraged the company is in the sense of using debt and other liabilities to finance assets. This ratio is often defined in terms of average total assets and average total equity and plays an important role in the DuPont decomposition of return on equity that will be presented in Section 13.

Debt-to-EBITDA Ratio

This ratio estimates how many years it would take to repay total debt based on earnings before income taxes, depreciation and amortization (an approximation of operating cash flow).

Interest Coverage

This ratio measures the number of times a company's EBIT could cover its interest payments. Thus, it is sometimes referred to as "times interest earned." A higher **interest coverage** ratio indicates stronger solvency, offering greater assurance that the company can service its debt (i.e., bank debt, bonds, notes) from operating earnings.

Fixed Charge Coverage

This ratio relates fixed charges, or obligations, to the cash flow generated by the company. It measures the number of times a company's earnings (before interest, taxes, and lease payments) can cover the company's interest and lease payments.[9] Similar to the interest coverage ratio, a higher **fixed charge coverage** ratio implies stronger solvency, offering greater assurance that the company can service its debt (i.e., bank debt, bonds, notes, and leases) from normal earnings. The ratio is sometimes used as an indication of the quality of the preferred dividend, with a higher ratio indicating a more secure preferred dividend.

Example 11 demonstrates the use of solvency ratios in evaluating the creditworthiness of a company.

EXAMPLE 11

Evaluation of Solvency Ratios

A credit analyst is evaluating the solvency of Eskom, a South African public utility based on financial statements for the year ended 31 March 2017. The following data are gathered from the company's 2017 annual report:

---

[9] For computing this ratio, an assumption sometimes made is that one-third of the lease payment amount represents interest on the lease obligation and that the rest is a repayment of principal on the obligation. For this variant of the fixed charge coverage ratio, the numerator is EBIT plus one-third of lease payments and the denominator is interest payments plus one-third of lease payments.

South African Rand, millions	**2017**	**2016**	**2015**
Total Assets	710,009	663,170	559,688
Short Term Debt	18,530	15,688	19,976
Long Term Debt	336,770	306,970	277,458
Total Liabilities	534,067	480,818	441,269
Total Equity	175,942	182,352	118,419

1 **A** Calculate the company's financial leverage ratio for 2016 and 2017.

B Interpret the financial leverage ratio calculated in Part A.

2 **A** What are the company's debt-to-assets, debt-to-capital, and debt-to-equity ratios for the three years?

B Is there any discernable trend over the three years?

Solutions to 1:

(Amounts are millions of Rand.)

A For 2017, average total assets were (710,009 + 663,170)/2 = 686,590, and average total equity was (175,942 + 182,352)/2 = 179,147. Thus, financial leverage was 686,590/179,942 = 3.83. For 2016, financial leverage was 4.07.

	2017	**2016**
Average Assets	686,590	611,429
Average Equity	179,147	150,386
Financial Leverage	3.83	4.07

B For 2017, every Rand in total equity supported R3.83 in total assets, on average. Financial leverage decreased from 2016 to 2017 on this measure.

Solutions to 2:

(Amounts are millions of Rand other than ratios)

A

	2017	**2016**	**2015**
Total Debt	355,300	322,658	297,434
Total Capital	531,242	505,010	415,853
Debt/Assets	50.0%	48.7%	53.1%
Debt/Capital	66.9%	63.9%	71.5%
Debt/Equity	2.02	1.77	2.51

B On all three metrics, the company's leverage decreased from 2015 to 2016 and increased from 2016 to 2017. For 2016 the decrease in leverage resulted from a conversion of subordinated debt into equity as well as additional issuance of equity. However, in 2017 debt levels increased again relative to assets, capital and equity indicating that the company's solvency has weakened. From a creditor's perspective, lower solvency (higher debt) indicates higher risk of default on obligations.

As with all ratio analysis, it is important to consider leverage ratios in a broader context. In general, companies with lower business risk and operations that generate steady cash flows are better positioned to take on more leverage without a commensurate increase in the risk of insolvency. In other words, a higher proportion of debt financing poses less risk of non-payment of interest and debt principal to a company with steady cash flows than to a company with volatile cash flows.

PROFITABILITY RATIOS

11

b identify, calculate, and interpret activity, liquidity, solvency, profitability, and valuation ratios;

The ability to generate profit on capital invested is a key determinant of a company's overall value and the value of the securities it issues. Consequently, many equity analysts would consider profitability to be a key focus of their analytical efforts.

Profitability reflects a company's competitive position in the market, and by extension, the quality of its management. The income statement reveals the sources of earnings and the components of revenue and expenses. Earnings can be distributed to shareholders or reinvested in the company. Reinvested earnings enhance solvency and provide a cushion against short-term problems.

11.1 Calculation of Profitability Ratios

Profitability ratios measure the return earned by the company during a period. Exhibit 15 provides the definitions of a selection of commonly used profitability ratios. Return-on-sales profitability ratios express various subtotals on the income statement (e.g., gross profit, operating profit, net profit) as a percentage of revenue. Essentially, these ratios constitute part of a common-size income statement discussed earlier. Return on investment profitability ratios measure income relative to assets, equity, or total capital employed by the company. For operating ROA, returns are measured as operating income, i.e., prior to deducting interest on debt capital. For ROA and ROE, returns are measured as net income, i.e., after deducting interest paid on debt capital. For return on common equity, returns are measured as net income minus preferred dividends (because preferred dividends are a return to preferred equity).

Exhibit 15 Definitions of Commonly Used Profitability Ratios

Profitability Ratios	Numerator	Denominator
Return on Sales[a]		
Gross profit margin	Gross profit	Revenue
Operating profit margin	Operating income[b]	Revenue
Pretax margin	EBT (earnings before tax but after interest)	Revenue
Net profit margin	Net income	Revenue
Return on Investment		
Operating ROA	Operating income	Average total assets
ROA	Net income	Average total assets

(continued)

Exhibit 15 (Continued)

Return on Investment		
Return on total capital	EBIT	Average short- and long-term debt and equity
ROE	Net income	Average total equity
Return on common equity	Net income − Preferred dividends	Average common equity

[a] "Sales" is being used as a synonym for "revenue."

[b] Some analysts use EBIT as a shortcut representation of operating income. Note that EBIT, strictly speaking, includes non-operating items such as dividends received and gains and losses on investment securities. Of utmost importance is that the analyst compute ratios consistently whether comparing different companies or analyzing one company over time.

11.2 Interpretation of Profitability Ratios

In the following, we discuss the interpretation of the profitability ratios presented in Exhibit 15. For each of the profitability ratios, a higher ratio indicates greater profitability.

Gross Profit Margin

Gross profit margin indicates the percentage of revenue available to cover operating and other expenses and to generate profit. Higher gross profit margin indicates some combination of higher product pricing and lower product costs. The ability to charge a higher price is constrained by competition, so gross profits are affected by (and usually inversely related to) competition. If a product has a competitive advantage (e.g., superior branding, better quality, or exclusive technology), the company is better able to charge more for it. On the cost side, higher gross profit margin can also indicate that a company has a competitive advantage in product costs.

Operating Profit Margin

Operating profit is calculated as gross profit minus operating costs. So, an **operating profit margin** increasing faster than the gross profit margin can indicate improvements in controlling operating costs, such as administrative overheads. In contrast, a declining operating profit margin could be an indicator of deteriorating control over operating costs.

Pretax Margin

Pretax income (also called "earnings before tax" or "EBT") is calculated as operating profit minus interest, and the **pretax margin** is the ratio of pretax income to revenue. The pretax margin reflects the effects on profitability of leverage and other (non-operating) income and expenses. If a company's pretax margin is increasing primarily as a result of increasing amounts of non-operating income, the analyst should evaluate whether this increase reflects a deliberate change in a company's business focus and, therefore, the likelihood that the increase will continue.

Net Profit Margin

Net profit, or net income, is calculated as revenue minus all expenses. Net income includes both recurring and non-recurring components. Generally, the net income used in calculating the net profit margin is adjusted for non-recurring items to offer a better view of a company's potential future profitability.

ROA

ROA measures the return earned by a company on its assets. The higher the ratio, the more income is generated by a given level of assets. Most databases compute this ratio as:

$$\frac{\text{Net income}}{\text{Average total assets}}$$

An issue with this computation is that net income is the return to equity holders, whereas assets are financed by both equity holders and creditors. Interest expense (the return to creditors) has already been subtracted in the numerator. Some analysts, therefore, prefer to add back interest expense in the numerator. In such cases, interest must be adjusted for income taxes because net income is determined after taxes. With this adjustment, the ratio would be computed as:

$$\frac{\text{Net income} + \text{Interest expense}(1 - \text{Tax rate})}{\text{Average total assets}}$$

Alternatively, some analysts elect to compute ROA on a pre-interest and pre-tax basis (operating ROA in Exhibit 15) as:

$$\frac{\text{Operating income or EBIT}}{\text{Average total assets}}$$

In this ROA calculation, returns are measured prior to deducting interest on debt capital (i.e., as operating income or EBIT). This measure reflects the return on all assets invested in the company, whether financed with liabilities, debt, or equity. Whichever form of ROA is chosen, the analyst must use it consistently in comparisons to other companies or time periods.

Return on Total Capital

Return on total capital measures the profits a company earns on all of the capital that it employs (short-term debt, long-term debt, and equity). As with operating ROA, returns are measured prior to deducting interest on debt capital (i.e., as operating income or EBIT).

ROE

ROE measures the return earned by a company on its equity capital, including minority equity, preferred equity, and common equity. As noted, return is measured as net income (i.e., interest on debt capital is not included in the return on equity capital). A variation of ROE is return on common equity, which measures the return earned by a company only on its common equity.

Both ROA and ROE are important measures of profitability and will be explored in more detail in section 13. As with other ratios, profitability ratios should be evaluated individually and as a group to gain an understanding of what is driving profitability (operating versus non-operating activities). Example 12 demonstrates the evaluation of profitability ratios and the use of the management report (sometimes called management's discussion and analysis or management commentary) that accompanies financial statements to explain the trend in ratios.

EXAMPLE 12

Evaluation of Profitability Ratios

Recall from Example 1 that an analysis found that Apple's gross margin declined over the three-year period FY2015 to FY2017. An analyst would like to further explore Apple's profitability using a five-year period. He gathers the following revenue data and calculates the following profitability ratios from information in Apple's annual reports:

Dollars in millions	2017	2016	2015	2014	2013
Sales	229,234	215,639	233,715	182,795	170,910
Gross Profit	88,186	84,263	93,626	70,537	64,304
Operating Income	61,344	60,024	71,230	52,503	48,999
Pre-tax Income	64,089	61,372	72,515	53,483	50,155
Net Income	48,351	45,687	53,394	39,510	37,037
Gross profit margin	38.47%	39.08%	40.06%	38.59%	37.62%
Operating income margin	26.76%	27.84%	30.48%	28.72%	28.67%
Pre-tax income	27.96%	28.46%	31.03%	29.26%	29.35%
Net profit margin	21.09%	21.19%	22.85%	21.61%	21.67%

Evaluate the overall trend in Apple's profitability ratios for the five-year period.

Solution:

Sales had increased steadily through 2015, dropped in 2016, and rebounded somewhat in 2017. As noted in Example 1, the sales decline in 2016 was related to a decline in iPhone sales and weakness in foreign currencies. Margins also rose from 2013 to 2015 and declined in 2016. However, in spite of the increase in sales in 2017, all margins declined slightly indicating costs were rising faster than sales. In spite of the fluctuations, Apple's bottom line net profit margin was relatively stable over the five-year period.

12

INTEGRATED FINANCIAL RATIO ANALYSIS

c describe relationships among ratios and evaluate a company using ratio analysis

In prior sections, the text presented separately activity, liquidity, solvency, and profitability ratios. Prior to discussing valuation ratios, the following sections demonstrate the importance of examining a variety of financial ratios—not a single ratio or category of ratios in isolation—to ascertain the overall position and performance of a company. Experience shows that the information from one ratio category can be helpful in answering questions raised by another category and that the most accurate overall picture comes from integrating information from all sources. Section 12 provides some introductory examples of such analysis and Section 13 shows how return on equity can be analyzed into components related to profit margin, asset utilization (activity), and financial leverage.

12.1 The Overall Ratio Picture: Examples

This section presents two simple illustrations to introduce the use of a variety of ratios to address an analytical task. Example 13 shows how the analysis of a pair of activity ratios resolves an issue concerning a company's liquidity. Example 14 shows that examining the overall ratios of multiple companies can assist an analyst in drawing conclusions about their relative performances.

EXAMPLE 13

A Variety of Ratios

An analyst is evaluating the liquidity of a Canadian manufacturing company and obtains the following liquidity ratios:

Fiscal Year	10	9	8
Current ratio	2.1	1.9	1.6
Quick ratio	0.8	0.9	1.0

The ratios present a contradictory picture of the company's liquidity. Based on the increase in its current ratio from 1.6 to 2.1, the company appears to have strong and improving liquidity; however, based on the decline of the quick ratio from 1.0 to 0.8, its liquidity appears to be deteriorating. Because both ratios have exactly the same denominator, current liabilities, the difference must be the result of changes in some asset that is included in the current ratio but not in the quick ratio (e.g., inventories). The analyst collects the following activity ratios:

DOH	55	45	30
DSO	24	28	30

The company's DOH has deteriorated from 30 days to 55 days, meaning that the company is holding increasingly larger amounts of inventory relative to sales. The decrease in DSO implies that the company is collecting receivables faster. If the proceeds from these collections were held as cash, there would be no effect on either the current ratio or the quick ratio. However, if the proceeds from the collections were used to purchase inventory, there would be no effect on the current ratio and a decline in the quick ratio (i.e., the pattern shown in this example). Collectively, the ratios suggest that liquidity is declining and that the company may have an inventory problem that needs to be addressed.

EXAMPLE 14

A Comparison of Two Companies (1)

An analyst collects the information[10] shown in Exhibit 16 for two hypothetical companies:

Exhibit 16

	Fiscal Year			
Anson Industries	**5**	**4**	**3**	**2**
Inventory turnover	76.69	89.09	147.82	187.64
DOH	4.76	4.10	2.47	1.95
Receivables turnover	10.75	9.33	11.14	7.56
DSO	33.95	39.13	32.77	48.29
Accounts payable turnover	4.62	4.36	4.84	4.22
Days payable	78.97	83.77	75.49	86.56
Cash from operations/Total liabilities	31.41%	11.15%	4.04%	8.81%
ROE	5.92%	1.66%	1.62%	–0.62%
ROA	3.70%	1.05%	1.05%	–0.39%
Net profit margin (Net income/Revenue)	3.33%	1.11%	1.13%	–0.47%
Total asset turnover (Revenue/Average assets)	1.11	0.95	0.93	0.84
Leverage (Average assets/Average equity)	1.60	1.58	1.54	1.60

	Fiscal Year			
Clarence Corporation	**5**	**4**	**3**	**2**
Inventory turnover	9.19	9.08	7.52	14.84
DOH	39.73	40.20	48.51	24.59
Receivables turnover	8.35	7.01	6.09	5.16
DSO	43.73	52.03	59.92	70.79
Accounts payable turnover	6.47	6.61	7.66	6.52
Days payable	56.44	55.22	47.64	56.00
Cash from operations/Total liabilities	13.19%	16.39%	15.80%	11.79%
ROE	9.28%	6.82%	–3.63%	–6.75%
ROA	4.64%	3.48%	–1.76%	–3.23%
Net profit margin (Net income/Revenue)	4.38%	3.48%	–1.60%	–2.34%
Total asset turnover (Revenue/Average assets)	1.06	1.00	1.10	1.38
Leverage (Average assets/Average equity)	2.00	1.96	2.06	2.09

---

[10] Note that ratios are expressed in terms of two decimal places and are rounded. Therefore, expected relationships may not hold perfectly.

Which of the following choices best describes reasonable conclusions an analyst might make about the companies' efficiency?

A Over the past four years, Anson has shown greater improvement in efficiency than Clarence, as indicated by its total asset turnover ratio increasing from 0.84 to 1.11.

B In FY5, Anson's DOH of only 4.76 indicated that it was less efficient at inventory management than Clarence, which had DOH of 39.73.

C In FY5, Clarence's receivables turnover of 8.35 times indicated that it was more efficient at receivables management than Anson, which had receivables turnover of 10.75.

Solution:

A is correct. Over the past four years, Anson has shown greater improvement in efficiency than Clarence, as indicated by its total asset turnover ratio increasing from 0.84 to 1.11. Over the same period of time, Clarence's total asset turnover ratio has declined from 1.38 to 1.06. Choices B and C are incorrect because DOH and receivables turnover are misinterpreted.

DUPONT ANALYSIS: THE DECOMPOSITION OF ROE

13

d demonstrate the application of DuPont analysis of return on equity and calculate and interpret effects of changes in its components

As noted earlier, ROE measures the return a company generates on its equity capital. To understand what drives a company's ROE, a useful technique is to decompose ROE into its component parts. (Decomposition of ROE is sometimes referred to as **DuPont analysis** because it was developed originally at that company.) Decomposing ROE involves expressing the basic ratio (i.e., net income divided by average shareholders' equity) as the product of component ratios. Because each of these component ratios is an indicator of a distinct aspect of a company's performance that affects ROE, the decomposition allows us to evaluate how these different aspects of performance affected the company's profitability as measured by ROE.[11]

Decomposing ROE is useful in determining the reasons for changes in ROE over time for a given company and for differences in ROE for different companies in a given time period. The information gained can also be used by management to determine which areas they should focus on to improve ROE. This decomposition will also show why a company's overall profitability, measured by ROE, is a function of its efficiency, operating profitability, taxes, and use of financial leverage. DuPont analysis shows the relationship between the various categories of ratios discussed in this reading and how they all influence the return to the investment of the owners.

Analysts have developed several different methods of decomposing ROE. The decomposition presented here is one of the most commonly used and the one found in popular research databases, such as Bloomberg. Return on equity is calculated as:

ROE = Net income/Average shareholders' equity

---

[11] For purposes of analyzing ROE, this method usually uses average balance sheet factors; however, the math will work out if beginning or ending balances are used throughout. For certain purposes, these alternative methods may be appropriate.

The decomposition of ROE makes use of simple algebra and illustrates the relationship between ROE and ROA. Expressing ROE as a product of only two of its components, we can write:

$$\begin{aligned}\text{ROE} &= \frac{\text{Net income}}{\text{Average shareholders' equity}} \\ &= \frac{\text{Net income}}{\text{Average total assets}} \times \frac{\text{Average total assets}}{\text{Average shareholders' equity}}\end{aligned} \quad \text{(1a)}$$

which can be interpreted as:

ROE = ROA × Leverage

In other words, ROE is a function of a company's ROA and its use of financial leverage ("leverage" for short, in this discussion). A company can improve its ROE by improving ROA or making more effective use of leverage. Consistent with the definition given earlier, leverage is measured as average total assets divided by average shareholders' equity. If a company had no leverage (no liabilities), its leverage ratio would equal 1.0 and ROE would exactly equal ROA. As a company takes on liabilities, its leverage increases. As long as a company is able to borrow at a rate lower than the marginal rate it can earn investing the borrowed money in its business, the company is making an effective use of leverage and ROE would increase as leverage increases. If a company's borrowing cost exceeds the marginal rate it can earn on investing in the business, ROE would decline as leverage increased because the effect of borrowing would be to depress ROA.

Using the data from Example 14 for Anson Industries, an analyst can examine the trend in ROE and determine whether the increase from an ROE of –0.625 percent in FY2 to 5.925 percent in FY5 is a function of ROA or the use of leverage:

	ROE	=	ROA	×	Leverage
FY5	5.92%		3.70%		1.60
FY4	1.66%		1.05%		1.58
FY3	1.62%		1.05%		1.54
FY2	–0.62%		–0.39%		1.60

Over the four-year period, the company's leverage factor was relatively stable. The primary reason for the increase in ROE is the increase in profitability measured by ROA.

Just as ROE can be decomposed, the individual components such as ROA can be decomposed. Further decomposing ROA, we can express ROE as a product of three component ratios:

$$\begin{aligned}\frac{\text{Net income}}{\text{Average shareholders' equity}} &= \frac{\text{Net income}}{\text{Revenue}} \times \frac{\text{Revenue}}{\text{Average total assets}} \\ &\times \frac{\text{Average total assets}}{\text{Average shareholders' equity}}\end{aligned} \quad \text{(1b)}$$

which can be interpreted as:

ROE = Net profit margin × Total asset turnover × Leverage

The first term on the right-hand side of this equation is the net profit margin, an indicator of profitability: how much income a company derives per one monetary unit (e.g., euro or dollar) of sales. The second term on the right is the asset turnover ratio, an indicator of efficiency: how much revenue a company generates per one money unit of assets. Note that ROA is decomposed into these two components: net profit margin and total asset turnover. A company's ROA is a function of profitability (net profit margin) and efficiency (total asset turnover). The third term on the right-hand side of Equation 1b is a measure of financial leverage, an indicator of solvency: the total amount of a company's assets relative to its equity capital. This decomposition illustrates that a company's ROE is a function of its net profit margin, its efficiency, and its leverage. Again, using the data from Example 14 for Anson Industries, the analyst can evaluate in more detail the reasons behind the trend in ROE:[12]

	ROE	=	Net profit margin	×	Total asset turnover	×	Leverage
FY5	5.92%		3.33%		1.11		1.60
FY4	1.66%		1.11%		0.95		1.58
FY3	1.62%		1.13%		0.93		1.54
FY2	–0.62%		–0.47%		0.84		1.60

This further decomposition confirms that increases in profitability (measured here as net profit margin) are indeed an important contributor to the increase in ROE over the four-year period. However, Anson's asset turnover has also increased steadily. The increase in ROE is, therefore, a function of improving profitability and improving efficiency. As noted above, ROE decomposition can also be used to compare the ROEs of peer companies, as demonstrated in Example 15.

EXAMPLE 15

A Comparison of Two Companies (2)

Referring to the data for Anson Industries and Clarence Corporation in Example 14, which of the following choices best describes reasonable conclusions an analyst might make about the companies' ROE?

A Anson's inventory turnover of 76.69 indicates it is more profitable than Clarence.

B The main driver of Clarence's superior ROE in FY5 is its more efficient use of assets.

C The main drivers of Clarence's superior ROE in FY5 are its greater use of debt financing and higher net profit margin.

Solution:

C is correct. The main driver of Clarence's superior ROE (9.28 percent compared with only 5.92 percent for Anson) in FY5 is its greater use of debt financing (leverage of 2.00 compared with Anson's leverage of 1.60) and higher net profit margin (4.38 percent compared with only 3.33 percent for Anson). A is incorrect because inventory turnover is not a direct indicator of profitability. An increase in inventory turnover may indicate more efficient use of inventory which in turn could affect profitability; however, an increase in inventory turnover would also be observed if a company was selling more goods even if it was not selling those goods at a profit. B is incorrect because Clarence has less efficient use of assets than Anson, indicated by turnover of 1.06 for Clarence compared with Anson's turnover of 1.11.

---

[12] Ratios are expressed in terms of two decimal places and are rounded. Therefore, ROE may not be the exact product of the three ratios.

To separate the effects of taxes and interest, we can further decompose the net profit margin and write:

$$\frac{\text{Net income}}{\text{Average shareholders' equity}} = \frac{\text{Net income}}{\text{EBT}} \times \frac{\text{EBT}}{\text{EBIT}} \times \frac{\text{EBIT}}{\text{Revenue}} \times \frac{\text{Revenue}}{\text{Average total assets}} \times \frac{\text{Average total assets}}{\text{Average shareholders' equity}} \quad (1c)$$

which can be interpreted as:

ROE = Tax burden × Interest burden × EBIT margin × Total asset turnover × Leverage

This five-way decomposition is the one found in financial databases such as Bloomberg. The first term on the right-hand side of this equation measures the effect of taxes on ROE. Essentially, it reflects one minus the average tax rate, or how much of a company's pretax profits it gets to keep. This can be expressed in decimal or percentage form. So, a 30 percent tax rate would yield a factor of 0.70 or 70 percent. A higher value for the tax burden implies that the company can keep a higher percentage of its pretax profits, indicating a lower tax rate. A decrease in the tax burden ratio implies the opposite (i.e., a higher tax rate leaving the company with less of its pretax profits).

The second term on the right-hand side captures the effect of interest on ROE. Higher borrowing costs reduce ROE. Some analysts prefer to use operating income instead of EBIT for this term and the following term. Either operating income or EBIT is acceptable as long as it is applied consistently. In such a case, the second term would measure both the effect of interest expense and non-operating income on ROE.

The third term on the right-hand side captures the effect of operating margin (if operating income is used in the numerator) or EBIT margin (if EBIT is used) on ROE. In either case, this term primarily measures the effect of operating profitability on ROE.

The fourth term on the right-hand side is again the total asset turnover ratio, an indicator of the overall efficiency of the company (i.e., how much revenue it generates per unit of total assets). The fifth term on the right-hand side is the financial leverage ratio described above—the total amount of a company's assets relative to its equity capital.

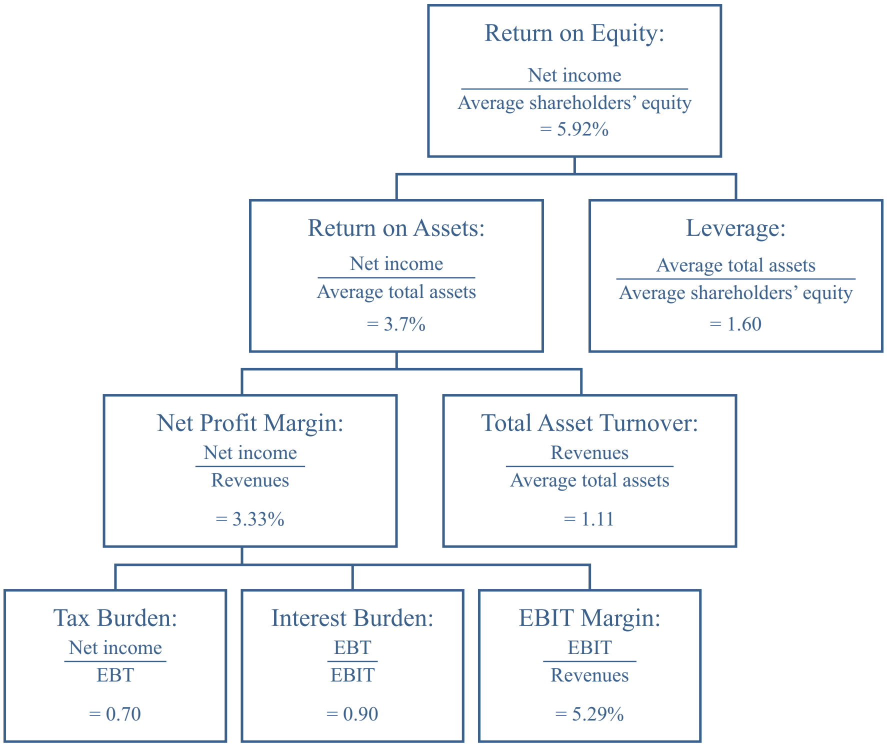
This decomposition expresses a company's ROE as a function of its tax rate, interest burden, operating profitability, efficiency, and leverage. An analyst can use this framework to determine what factors are driving a company's ROE. The decomposition of ROE can also be useful in forecasting ROE based upon expected efficiency, profitability, financing activities, and tax rates. The relationship of the individual factors, such as ROA to the overall ROE, can also be expressed in the form of an ROE tree to study the contribution of each of the five factors, as shown in Exhibit 17 for Anson Industries.[13]

Exhibit 17 shows that Anson's ROE of 5.92 percent in FY5 can be decomposed into ROA of 3.70 percent and leverage of 1.60. ROA can further be decomposed into a net profit margin of 3.33 percent and total asset turnover of 1.11. Net profit margin can be decomposed into a tax burden of 0.70 (an average tax rate of 30 percent), an interest burden of 0.90, and an EBIT margin of 5.29 percent. Overall ROE is decomposed into five components.

---

[13] Note that a breakdown of net profit margin was not provided in Example 14, but is added here.

Exhibit 17 DuPont Analysis of Anson Industries' ROE: Fiscal Year 5


Return on Equity:
Net income / Average shareholders' equity
= 5.92%
Return on Assets:
Net income / Average total assets
= 3.7%
Leverage:
Average total assets / Average shareholders' equity
= 1.60
Net Profit Margin:
Net income / Revenues
= 3.33%
Total Asset Turnover:
Revenues / Average total assets
= 1.11
Tax Burden:
Net income / EBT
= 0.70
Interest Burden:
EBT / EBIT
= 0.90
EBIT Margin:
EBIT / Revenues
= 5.29%

Example 16 demonstrates how the five-component decomposition can be used to determine reasons behind the trend in a company's ROE.

EXAMPLE 16

Five-Way Decomposition of ROE

An analyst examining Amsterdam PLC (a hypothetical company) wishes to understand the factors driving the trend in ROE over a four-year period. The analyst obtains and calculates the following data from Amsterdam's annual reports:

	2017	2016	2015	2014
ROE	9.53%	20.78%	26.50%	24.72%
Tax burden	60.50%	52.10%	63.12%	58.96%
Interest burden	97.49%	97.73%	97.86%	97.49%
EBIT margin	7.56%	11.04%	13.98%	13.98%
Asset turnover	0.99	1.71	1.47	1.44
Leverage	2.15	2.17	2.10	2.14

What might the analyst conclude?

Solution:

The tax burden measure has varied, with no obvious trend. In the most recent year, 2017, taxes declined as a percentage of pretax profit. (Because the tax burden reflects the relation of after-tax profits to pretax profits, the increase from 52.10 percent in 2016 to 60.50 percent in 2017 indicates that taxes declined as a percentage of pretax profits.) This decline in average tax rates could be a result of lower tax rates from new legislation or revenue in a lower tax jurisdiction. The interest burden has remained fairly constant over the four-year period indicating that the company maintains a fairly constant capital structure. Operating margin (EBIT margin) declined over the period, indicating the company's operations were less profitable. This decline is generally consistent with declines in oil prices in 2017 and declines in refining industry gross margins in 2016 and 2017. The company's efficiency (asset turnover) decreased in 2017. The company's leverage remained constant, consistent with the constant interest burden. Overall, the trend in ROE (declining substantially over the recent years) resulted from decreases in operating profits and a lower asset turnover. Additional research on the causes of these changes is required in order to develop expectations about the company's future performance.

The most detailed decomposition of ROE that we have presented is a five-way decomposition. Nevertheless, an analyst could further decompose individual components of a five-way analysis. For example, EBIT margin (EBIT/Revenue) could be further decomposed into a non-operating component (EBIT/Operating income) and an operating component (Operating income/Revenue). The analyst can also examine which other factors contributed to these five components. For example, an improvement in efficiency (total asset turnover) may have resulted from better management of inventory (DOH) or better collection of receivables (DSO).

14

EQUITY ANALYSIS AND VALUATION RATIOS

e calculate and interpret ratios used in equity analysis and credit analysis

One application of financial analysis is to select securities as part of the equity portfolio management process. Analysts are interested in valuing a security to assess its merits for inclusion or retention in a portfolio. The valuation process has several steps, including:

1. understanding the business and the existing financial profile
2. forecasting company performance
3. selecting the appropriate valuation model
4. converting forecasts to a valuation
5. making the investment decision

Financial analysis assists in providing the core information to complete the first two steps of this valuation process: understanding the business and forecasting performance.

Fundamental equity analysis involves evaluating a company's performance and valuing its equity in order to assess its relative attractiveness as an investment. Analysts use a variety of methods to value a company's equity, including valuation ratios (e.g., the price-to-earnings or P/E ratio), discounted cash flow approaches, and residual income approaches (ROE compared with the cost of capital), among others. The following section addresses the first of these approaches—the use of valuation ratios.

14.1 Valuation Ratios

Valuation ratios have long been used in investment decision making. A well known example is the **price to earnings ratio** (P/E ratio)—probably the most widely cited indicator in discussing the value of equity securities—which relates share price to the earnings per share (EPS). Additionally, some analysts use other market multiples, such as price to book value (P/B) and price to cash flow (P/CF). The following sections explore valuation ratios and other quantities related to valuing equities.

14.1.1 *Calculation of Valuation Ratios and Related Quantities*

Exhibit 18 describes the calculation of some common valuation ratios and related quantities.

Exhibit 18 Definitions of Selected Valuation Ratios and Related Quantities

Valuation Ratios	Numerator	Denominator
P/E	Price per share	Earnings per share
P/CF	Price per share	Cash flow per share
P/S	Price per share	Sales per share
P/BV	Price per share	Book value per share

Per-Share Quantities	Numerator	Denominator
Basic EPS	Net income minus preferred dividends	Weighted average number of ordinary shares outstanding
Diluted EPS	Adjusted income available for ordinary shares, reflecting conversion of dilutive securities	Weighted average number of ordinary and potential ordinary shares outstanding
Cash flow per share	Cash flow from operations	Weighted average number of shares outstanding
EBITDA per share	EBITDA	Weighted average number of shares outstanding
Dividends per share	Common dividends declared	Weighted average number of ordinary shares outstanding

Dividend-Related Quantities	Numerator	Denominator
Dividend payout ratio	Common share dividends	Net income attributable to common shares
Retention rate (*b*)	Net income attributable to common shares – Common share dividends	Net income attributable to common shares
Sustainable growth rate	$b \times \text{ROE}$	

The P/E ratio expresses the relationship between the price per share and the amount of earnings attributable to a single share. In other words, the P/E ratio tells us how much an investor in common stock pays per dollar of earnings.

Because P/E ratios are calculated using net income, the ratios can be sensitive to non-recurring earnings or one-time earnings events. In addition, because net income is generally considered to be more susceptible to manipulation than are cash flows, analysts may use **price to cash flow** as an alternative measure—particularly in situations where earnings quality may be an issue. EBITDA per share, because it is calculated using income before interest, taxes, and depreciation, can be used to eliminate the effect of different levels of fixed asset investment across companies. It facilitates comparison between companies in the same sector but at different stages of infrastructure maturity. **Price to sales** is calculated in a similar manner and is sometimes used as a comparative price metric when a company does not have positive net income.

Another price-based ratio that facilitates useful comparisons of companies' stock prices is **price to book value**, or P/B, which is the ratio of price to book value per share. This ratio is often interpreted as an indicator of market judgment about the relationship between a company's required rate of return and its actual rate of return. Assuming that book values reflect the fair values of the assets, a price to book ratio of one can be interpreted as an indicator that the company's future returns are expected to be exactly equal to the returns required by the market. A ratio greater than one would indicate that the future profitability of the company is expected to exceed the required rate of return, and values of this ratio less than one indicate that the company is not expected to earn excess returns.[14]

14.1.2 *Interpretation of Earnings per Share*

Exhibit 18 presented a number of per-share quantities that can be used in valuation ratios. In this section, we discuss the interpretation of one such critical quantity, earnings per share or EPS.[15]

EPS is simply the amount of earnings attributable to each share of common stock. In isolation, EPS does not provide adequate information for comparison of one company with another. For example, assume that two companies have only common stock outstanding and no dilutive securities outstanding. In addition, assume the two companies have identical net income of $10 million, identical book equity of $100 million and, therefore, identical profitability (10 percent, using ending equity in this case for simplicity). Furthermore, assume that Company A has 100 million weighted average common shares outstanding, whereas Company B has 10 million weighted average common shares outstanding. So, Company A will report EPS of $0.10 per share, and Company B will report EPS of $1 per share. The difference in EPS does not reflect a difference in profitability—the companies have identical profits and profitability. The difference reflects only a different number of common shares outstanding. Analysts should understand in detail the types of EPS information that companies report:

Basic EPS provides information regarding the earnings attributable to each share of common stock.[16] To calculate basic EPS, the weighted average number of shares outstanding during the period is first calculated. The weighted average number of shares consists of the number of ordinary shares outstanding at the beginning of the period, adjusted by those bought back or issued during the period, multiplied by a time-weighting factor.

Accounting standards generally require the disclosure of basic as well as **diluted EPS** (diluted EPS includes the effect of all the company's securities whose conversion or exercise would result in a reduction of basic EPS; dilutive securities include convertible debt, convertible preferred, warrants, and options). Basic EPS and diluted EPS must be shown with equal prominence on the face of the income statement for each class of ordinary share. Disclosure includes the amounts used as the numerators in calculating basic and diluted EPS, and a reconciliation of those amounts to the company's profit or loss for the period. Because both basic and diluted EPS are presented in a company's financial statements, an analyst does not need to calculate these measures for reported financial statements. Understanding the calculations is, however, helpful for situations requiring an analyst to calculate expected future EPS.

---

[14] For more detail on valuation ratios as used in equity analysis, see the curriculum reading "Equity Valuation: Concepts and Basic Tools."

[15] For more detail on EPS calculation, see the reading "Understanding Income Statements."

[16] IAS 33, *Earnings per Share* and FASB ASC Topic 260 [Earnings per Share].

To calculate diluted EPS, earnings are adjusted for the after-tax effects assuming conversion, and the following adjustments are made to the weighted number of shares:

- The weighted average number of shares for basic EPS, *plus* those that would be issued on conversion of all potentially dilutive ordinary shares. Potential ordinary shares are treated as dilutive when their conversion would decrease net profit per share from continuing ordinary operations.
- These shares are deemed to have been converted into ordinary shares at the beginning of the period or, if later, at the date of the issue of the shares.
- Options, warrants (and their equivalents), convertible instruments, contingently issuable shares, contracts that can be settled in ordinary shares or cash, purchased options, and written put options should be considered.

14.1.3 *Dividend-Related Quantities*

In this section, we discuss the interpretation of the dividend-related quantities presented in Exhibit 18. These quantities play a role in some present value models for valuing equities.

Dividend Payout Ratio The **dividend payout ratio** measures the percentage of earnings that the company pays out as dividends to shareholders. The amount of dividends per share tends to be relatively fixed because any reduction in dividends has been shown to result in a disproportionately large reduction in share price. Because dividend amounts are relatively fixed, the dividend payout ratio tends to fluctuate with earnings. Therefore, conclusions about a company's dividend payout policies should be based on examination of payout over a number of periods. Optimal dividend policy, similar to optimal capital structure, has been examined in academic research and continues to be a topic of significant interest in corporate finance.

Retention Rate The retention rate, or earnings retention rate, is the complement of the payout ratio or dividend payout ratio (i.e., 1 – payout ratio). Whereas the payout ratio measures the percentage of earnings that a company pays out as dividends, the retention rate is the percentage of earnings that a company retains. (Note that both the payout ratio and retention rate are both percentages of earnings. The difference in terminology—"ratio" versus "rate" versus "percentage"—reflects common usage rather than any substantive differences.)

Sustainable Growth Rate A company's **sustainable growth rate** is viewed as a function of its profitability (measured as ROE) and its ability to finance itself from internally generated funds (measured as the retention rate). The sustainable growth rate is ROE times the retention rate. A higher ROE and a higher retention rate result in a higher sustainable growth rate. This calculation can be used to estimate a company's growth rate, a factor commonly used in equity valuation.

INDUSTRY-SPECIFIC FINANCIAL RATIOS

15

e calculate and interpret ratios used in equity analysis and credit analysis

As stated earlier in this reading, a universally accepted definition and classification of ratios does not exist. The purpose of ratios is to serve as indicators of important aspects of a company's performance and value. Aspects of performance that are considered important in one industry may be irrelevant in another, and industry-specific ratios reflect these differences. For example, companies in the retail industry may report same-store sales changes because, in the retail industry, it is important to distinguish between growth that results from opening new stores and growth that results from generating more sales at existing stores. Industry-specific metrics can be especially important to the value of equity in early stage industries, where companies are not yet profitable.

In addition, regulated industries—especially in the financial sector—often are required to comply with specific regulatory ratios. For example, the banking sector's liquidity and cash reserve ratios provide an indication of banking liquidity and reflect monetary and regulatory requirements. Banking capital adequacy requirements attempt to relate banks' solvency requirements directly to their specific levels of risk exposure.

Exhibit 19 presents, for illustrative purposes only, some industry-specific and task-specific ratios.[17]

Exhibit 19 Definitions of Some Common Industry- and Task-Specific Ratios

Ratio	Numerator	Denominator
Business Risk		
Coefficient of variation of operating income	Standard deviation of operating income	Average operating income
Coefficient of variation of net income	Standard deviation of net income	Average net income
Coefficient of variation of revenues	Standard deviation of revenue	Average revenue

Financial Sector Ratios	Numerator	Denominator
Capital adequacy—banks	Various components of capital	Various measures such as risk-weighted assets, market risk exposure, or level of operational risk assumed
Monetary reserve requirement (Cash reserve ratio)	Reserves held at central bank	Specified deposit liabilities
Liquid asset requirement	Approved "readily marketable" securities	Specified deposit liabilities
Net interest margin	Net interest income	Total interest-earning assets

Retail Ratios	Numerator	Denominator
Same (or comparable) store sales	Average revenue growth year over year for stores open in both periods	Not applicable
Sales per square meter (or square foot)	Revenue	Total retail space in square meters (or square feet)

---

[17] There are many other industry- and task-specific ratios that are outside the scope of this reading. Resources such as Standard and Poor's Industry Surveys present useful ratios for each industry. Industry organizations may present useful ratios for the industry or a task specific to the industry.

Exhibit 19 (Continued)

Service Companies	Numerator	Denominator
Revenue per employee	Revenue	Total number of employees
Net income per employee	Net income	Total number of employees

Hotel	Numerator	Denominator
Average daily rate	Room revenue	Number of rooms sold
Occupancy rate	Number of rooms sold	Number of rooms available

RESEARCH ON FINANCIAL RATIOS IN CREDIT AND EQUITY ANALYSIS

16

e calculate and interpret ratios used in equity analysis and credit analysis

Some ratios may be particularly useful in equity analysis. The end product of equity analysis is often a valuation and investment recommendation. Theoretical valuation models are useful in selecting ratios that would be useful in this process. For example, a company's P/B is theoretically linked to ROE, growth, and the required return. ROE is also a primary determinant of residual income in a residual income valuation model. In both cases, higher ROE relative to the required return denotes a higher valuation. Similarly, profit margin is related to justified price-to-sales (P/S) ratios. Another common valuation method involves forecasts of future cash flows that are discounted back to the present. Trends in ratios can be useful in forecasting future earnings and cash flows (e.g., trends in operating profit margin and collection of customer receivables). Future growth expectations are a key component of all of these valuation models. Trends may be useful in assessing growth prospects (when used in conjunction with overall economic and industry trends). The variability in ratios and common-size data can be useful in assessing risk, an important component of the required rate of return in valuation models. A great deal of academic research has focused on the use of these fundamental ratios in evaluating equity investments.

A classic study, Ou and Penman (1989a and 1989b), found that ratios and common-size metrics generated from accounting data were useful in forecasting earnings and stock returns. Ou and Penman examined 68 such metrics and found that these could be reduced to a more parsimonious list of relevant variables, including percentage changes in a variety of measures such as current ratio, inventory, and sales; gross and pretax margins; and returns on assets and equity. These variables were found to be useful in forecasting earnings and stock returns.

Subsequent studies have also demonstrated the usefulness of ratios in evaluation of equity investments and valuation. Lev and Thiagarajan (1993) examined fundamental financial variables used by analysts to assess whether they are useful in security valuation. They found that fundamental variables add about 70 percent to the explanatory power of earnings alone in predicting excess returns (stock returns in excess of those expected). The fundamental variables they found useful included percentage changes in inventory and receivables relative to sales, gross margin, sales per employee, and the change in bad debts relative to the change in accounts receivable, among others. Abarbanell and Bushee (1997) found some of the same variables useful in predicting future accounting earnings. Abarbanell and Bushee (1998) devised an investment strategy using these same variables and found that they can generate excess returns under this strategy.

Piotroski (2000) used financial ratios to supplement a value investing strategy and found that he can generate significant excess returns. Variables used by Piotroski include ROA, cash flow ROA, change in ROA, change in leverage, change in liquidity, change in gross margin, and change in inventory turnover.

This research shows that in addition to being useful in evaluating the past performance of a company, ratios can be useful in predicting future earnings and equity returns.

17 CREDIT ANALYSIS

e calculate and interpret ratios used in equity analysis and credit analysis

Credit risk is the risk of loss caused by a counterparty's or debtor's failure to make a promised payment. For example, credit risk with respect to a bond is the risk that the obligor (the issuer of the bond) is not able to pay interest and principal according to the terms of the bond indenture (contract). **Credit analysis** is the evaluation of credit risk.

Approaches to credit analysis vary and, as with all financial analysis, depend on the purpose of the analysis and the context in which it is done. Credit analysis for specific types of debt (e.g., acquisition financing and other highly leveraged financing) often involves projections of period-by-period cash flows similar to projections made by equity analysts. Whereas the equity analyst may discount projected cash flows to determine the value of the company's equity, a credit analyst would use the projected cash flows to assess the likelihood of a company complying with its financial covenants in each period and paying interest and principal as due.[18] The analysis would also include expectations about asset sales and refinancing options open to the company.

Credit analysis may relate to the borrower's credit risk in a particular transaction or to its overall creditworthiness. In assessing overall creditworthiness, one general approach is credit scoring, a statistical analysis of the determinants of credit default.

Another general approach to credit analysis is the credit rating process that is used, for example, by credit rating agencies to assess and communicate the probability of default by an issuer on its debt obligations (e.g., commercial paper, notes, and bonds). A credit rating can be either long term or short term and is an indication of the rating agency's opinion of the creditworthiness of a debt issuer with respect to a specific debt security or other obligation. Where a company has no debt outstanding, a rating agency can also provide an issuer credit rating that expresses an opinion of the issuer's overall capacity and willingness to meet its financial obligations. The following sections review research on the use of ratios in credit analysis and the ratios commonly used in credit analysis.

---

[18] Financial covenants are clauses in bond indentures relating to the financial condition of the bond issuer.

17.1 The Credit Rating Process

The credit rating process involves both the analysis of a company's financial reports as well as a broad assessment of a company's operations. In assigning credit ratings, rating agencies emphasize the importance of the relationship between a company's business risk profile and its financial risk.

For corporate entities, credit ratings typically reflect a combination of qualitative and quantitative factors. Qualitative factors generally include an industry's growth prospects, volatility, technological change, and competitive environment. At the individual company level, qualitative factors may include operational effectiveness, strategy, governance, financial policies, risk management practices, and risk tolerance. In contrast, quantitative factors generally include profitability, leverage, cash flow adequacy, and liquidity.[19]

When analyzing financial ratios, rating agencies normally investigate deviations of ratios from the median ratios of the universe of companies for which such ratios have been calculated and also use the median ratings as an indicator for the ratings grade given to a specific debt issuer. This so-called universe of rated companies frequently changes, and any calculations are obviously affected by economic factors as well as by mergers and acquisitions. International ratings include the influence of country and economic risk factors. Exhibit 20 presents a few key financial ratios used by Standard & Poor's in evaluating industrial companies. Note that before calculating ratios, rating agencies make certain adjustments to reported financials such as adjusting debt to include off-balance sheet debt in a company's total debt.

Exhibit 20 Selected Credit Ratios

Credit Ratio	Numerator[a]	Denominator[a]
EBITDA interest coverage	EBITDA[b]	Interest expense, including non-cash interest on conventional debt instruments
FFO[c] (Funds from operations) to debt	FFO	Total debt
Free operating cash flow to debt	CFO[d] (adjusted) minus capital expenditures	Total debt
EBIT margin	EBIT[e]	Total revenues
EBITDA margin	EBITDA	Total revenues
Debt to EBITDA	Total debt	EBITDA
Return on capital	EBIT	Average beginning-of-year and end-of-year capital[f]

[a] Note that both the numerator and the denominator definitions are adjusted from ratio to ratio and may not correspond to the definitions used elsewhere in this reading.
[b] EBITDA = earnings before interest, taxes, depreciation, and amortization.
[c] FFO = funds from operations, defined as EBITDA minus net interest expense minus current tax expense (plus or minus all applicable adjustments).
[d] CFO = cash flow from operations.
[e] EBIT = earnings before interest and taxes.
[f] Capital = debt plus noncurrent deferred taxes plus equity (plus or minus all applicable adjustments).
Source: Based on data from Standard & Poor's *Corporate Methodology: Ratios And Adjustments* (2013). This represents the last updated version at the time of publication.

---

[19] Concepts in this paragraph are based on Standard & Poor's *General Criteria: Principles of Credit Ratings* (2011). This represents the last updated version at the time of publication.

17.2 Historical Research on Ratios in Credit Analysis

A great deal of academic and practitioner research has focused on determining which ratios are useful in assessing the credit risk of a company, including the risk of bankruptcy.

One of the earliest studies examined individual ratios to assess their ability to predict failure of a company up to five years in advance. Beaver (1967) found that six ratios could correctly predict company failure one year in advance 90 percent of the time and five years in advance at least 65 percent of the time. The ratios found effective by Beaver were cash flow to total debt, ROA, total debt to total assets, working capital to total assets, the current ratio, and the no-credit interval ratio (the length of time a company could go without borrowing). Altman (1968) and Altman, Haldeman, and Narayanan (1977) found that financial ratios could be combined in an effective model for predicting bankruptcy. Altman's initial work involved creation of a Z-score that was able to correctly predict financial distress. The Z-score was computed as

$$\begin{aligned} Z \;&= 1.2 \times (\text{Current assets} - \text{Current liabilities})/\text{Total assets} \\ &\quad + 1.4 \times (\text{Retained earnings/Total assets}) \\ &\quad + 3.3 \times (\text{EBIT/Total assets}) \\ &\quad + 0.6 \times (\text{Market value of stock/Book value of liabilities}) \\ &\quad + 1.0 \times (\text{Sales/Total assets}) \end{aligned}$$

In his initial study, a Z-score of lower than 1.81 predicted failure and the model was able to accurately classify 95 percent of companies studied into a failure group and a non-failure group. The original model was designed for manufacturing companies. Subsequent refinements to the models allow for other company types and time periods. Generally, the variables found to be useful in prediction include profitability ratios, coverage ratios, liquidity ratios, capitalization ratios, and earnings variability (Altman 2000).

Similar research has been performed on the ability of ratios to predict bond ratings and bond yields. For example, Ederington, Yawtiz, and Roberts (1987) found that a small number of variables (total assets, interest coverage, leverage, variability of coverage, and subordination status) were effective in explaining bond yields. Similarly, Ederington (1986) found that nine variables in combination could correctly classify more than 70 percent of bond ratings. These variables included ROA, long-term debt to assets, interest coverage, cash flow to debt, variability of coverage and cash flow, total assets, and subordination status. These studies have shown that ratios are effective in evaluating credit risk, bond yields, and bond ratings.

18 BUSINESS AND GEOGRAPHIC SEGMENTS

f explain the requirements for segment reporting and calculate and interpret segment ratios

Analysts often need to evaluate the performance underlying business segments (subsidiary companies, operating units, or simply operations in different geographic areas) to understand in detail the company as a whole. Although companies are not required to provide full financial statements for segments, they are required to provide segment information under both IFRS and US GAAP.[20]

---

[20] IFRS 8, *Operating Segments* and FASB ASC Topic 280 [Segment Reporting].

18.1 Segment Reporting Requirements

An operating segment is defined as a component of a company: a) that engages in activities that may generate revenue and create expenses, including a start-up segment that has yet to earn revenues, b) whose results are regularly reviewed by the company's senior management, and c) for which discrete financial information is available.[21] A company must disclose separate information about any operating segment which meets certain quantitative criteria—namely, the segment constitutes 10 percent or more of the combined operating segments' revenue, assets, or profit. (For purposes of determining whether a segment constitutes 10 percent or more of combined profits or losses, the criteria is expressed in terms of the absolute value of the segment's profit or loss as a percentage of the greater of (i) the combined profits of all profitable segments and (ii) the absolute amount of the combined losses of all loss-making segments.) If, after applying these quantitative criteria, the combined revenue from external customers for all reportable segments combined is less than 75 percent of the total company revenue, the company must identify additional reportable segments until the 75 percent level is reached. Small segments might be combined as one if they share a substantial number of factors that define a business or geographical segment, or they might be combined with a similar significant reportable segment. Information about operating segments and businesses that are not reportable is combined in an "all other segments" category.

Companies may internally report business results in a variety of ways (e.g., product segments and geographical segments). Companies identify the segments for external reporting purposes considering the definition of an operating segment and using factors such as what information is reported to the board of directors and whether a manager is responsible for each segment. Companies must disclose the factors used to identify reportable segments and the types of products and services sold by each reportable segment.

For each reportable segment, the following should also be disclosed:

- a measure of profit or loss;
- a measure of total assets and liabilities[22] (if these amounts are regularly reviewed by the company's chief decision-making officer);
- segment revenue, distinguishing between revenue to external customers and revenue from other segments;
- interest revenue and interest expense;
- cost of property, plant, and equipment, and intangible assets acquired;
- depreciation and amortisation expense;
- other non-cash expenses;
- income tax expense or income; and
- share of the net profit or loss of an investment accounted for under the equity method.

Companies also must provide a reconciliation between the information of reportable segments and the consolidated financial statements in terms of segment revenue, profit or loss, assets, and liabilities.

---

[21] IFRS 8, *Operating Segments*, paragraph 5.

[22] IFRS 8 and FASB ASC Topic 280 are largely converged. One notable difference is that US GAAP does not require disclosure of segment liabilities, while IFRS requires disclosure of segment liabilities if that information is regularly provided to the company's "chief operating decision maker."

Another disclosure required is the company's reliance on any single customer. If any single customer represents 10 percent or more of the company's total revenues, the company must disclose that fact. From an analysts' perspective, information about a concentrated customer base can be useful in assessing the risks faced by the company.

18.2 Segment Ratios

Based on the segment information that companies are required to present, a variety of useful ratios can be computed, as shown in Exhibit 21.

Exhibit 21 Definitions of Segment Ratios

Segment Ratios	Numerator	Denominator
Segment margin	Segment profit (loss)	Segment revenue
Segment turnover	Segment revenue	Segment assets
Segment ROA	Segment profit (loss)	Segment assets
Segment debt ratio	Segment liabilities	Segment assets

The segment margin measures the operating profitability of the segment relative to revenues, whereas the segment ROA measures the operating profitability relative to assets. Segment turnover measures the overall efficiency of the segment: how much revenue is generated per unit of assets. The segment debt ratio examines the level of liabilities (hence solvency) of the segment. Example 17 demonstrates the evaluation of segment ratios.

EXAMPLE 17

The Evaluation of Segment Ratios

The information contained in Exhibit 22 relates to the business segments of Groupe Danone for 2016 and 2017 in millions of euro. According to the company's 2017 annual report the company operates in four business segments which are primarily evaluated on operating income and operating margin and in two geographic segments for which they also provide information on assets deployed.

Evaluate the performance of the segments using the relative proportion of sales of each segment, the segment margins, segment ROA where available, and segment turnover where available.

Exhibit 22 Group Danone Segment Disclosures (in € millions)

	2016		2017	
Business Segments	Sales	Recurring Operating Income	Sales	Recurring Operating Income
Fresh Dairy Products – International	8,229	731	8,424	760
Fresh Dairy Products – North America	2,506	351	4,530	556
Specialized Nutrition	6,634	1,419	7,102	1,685
Waters	4,574	521	4,621	541
Group Total	21,944	3,022	24,677	3,542

Exhibit 22 (Continued)

Geographic Segments	2016 Sales	2016 Recurring Operating Income	2016 Non-Current Assets	2017 Sales	2017 Recurring Operating Income	2017 Non-Current Assets
Europe and North America	10,933	1,842	11,532	13,193	2,048	22,517
Rest of World	11,011	1,180	9,307	11,484	1,495	8,433
Group Total	21,944	3,022	20,839	24,677	3,543	30,950

Source: Company's 2017 Annual Report.

Solution:

Business Segments	2016 Segment Revenue Percent	2016 Recurring Operating Margin	2017 Segment Revenue Percent	2017 Recurring Operating Margin
Fresh Dairy Products – International	37.5%	8.9%	34.1%	9.0%
Fresh Dairy Products – North America	11.4%	14.0%	18.4%	12.3%
Specialized Nutrition	30.2%	21.4%	28.8%	23.7%
Waters	20.8%	11.4%	18.7%	11.7%
Group Total	100.0%	13.8%	100.0%	14.4%

Business Segments	2017 % change in revenue
Fresh Dairy Products – International	2.4%
Fresh Dairy Products – North America	80.8%
Specialized Nutrition	7.1%
Waters	1.0%
Group Total	12.5%

The business segment analysis shows that the largest proportion of the company's revenues occurs in the Fresh Dairy Products – International segment: 37.5% and 34.1% of the total in 2016 and 2017, respectively. The greatest increase in relative revenue, however, came from the Fresh Dairy Products – North America segment which grew by 80.8% and increased from 11.4% of total revenues in 2016 to 18.4% of total revenues in 2017. Examination of the company's full annual report reveals that Danone Group acquired a large health-oriented North American food company, Whitewave, in 2017. This caused the shift in the relative proportion of sales. The highest segment operating margin in both years comes from the Specialized Nutrition segment with operating margins of 21.4% in 2016 increasing to 23.7% in 2017. Margins increased slightly in the Fresh Dairy Products – International and Waters segments, while margins declined in Fresh Dairy Products – North America. The latter is likely due to costs associated with the Whitewave acquisition.

	2016				2017			
Geographic Segments	**Segment Revenue Percent**	**Recurring Operating Margin**	**Segment ROA**	**Segment Asset Turnover**	**Segment Revenue Percent**	**Recurring Operating Margin**	**Segment ROA**	**Segment Asset Turnover**
Europe and North America	49.8%	16.8%	16.0%	0.9	53.5%	15.5%	9.1%	0.6
Rest of World	50.2%	10.7%	12.7%	1.2	46.5%	13.0%	17.7%	1.4
Group Total	100.0%	13.8%	14.5%	1.1	100.0%	14.4%	11.4%	0.8

As used in this table, ROA refers to operating income divided by ending assets, and Asset Turnover is defined as Revenue divided by non-current assets.

The geographic segment analysis shows that the company's sales are split roughly evenly between the two geographic segments. Operating margins were higher in the Europe and North America segment in both years but declined from 16.8% in 2016 to 15.5% in 2017, likely in connection with the North American acquisition of Whitewave. Operating margins in the rest of the world, however, increased in 2017. Segment return on assets and segment asset turnover declined significantly for the Europe and North America segment in 2017, again largely due to the acquisition of Whitewave. An examination of the annual report disclosures reveals that the large increase in segment assets came from intangible assets (mainly goodwill) recorded in the Whitewave acquisition. In contrast, segment return on assets and turnover improved significantly in the Rest of World segment.

19

MODEL BUILDING AND FORECASTING

g describe how ratio analysis and other techniques can be used to model and forecast earnings

Analysts often need to forecast future financial performance. For example, analysts' EPS forecasts and related equity valuations are widely followed by Wall Street. Analysts use data about the economy, industry, and company in arriving at a company's forecast. The results of an analyst's financial analysis, including common-size and ratio analyses, are integral to this process, along with the judgment of the analysts.

Based upon forecasts of growth and expected relationships among the financial statement data, the analyst can build a model (sometimes referred to as an "earnings model") to forecast future performance. In addition to budgets, pro forma financial statements are widely used in financial forecasting within companies, especially for use by senior executives and boards of directors. Last but not least, these budgets and forecasts are also used in presentations to credit analysts and others in obtaining external financing.

For example, based on a revenue forecast, an analyst may budget expenses based on expected common-size data. Forecasts of balance sheet and cash flow statements can be derived from expected ratio data, such as DSO. Forecasts are not limited to a single point estimate but should involve a range of possibilities. This can involve several techniques:

- **Sensitivity analysis**: Also known as "what if" analysis, sensitivity analysis shows the range of possible outcomes as specific assumptions are changed; this could, in turn, influence financing needs or investment in fixed assets.
- **Scenario analysis**: This type of analysis shows the changes in key financial quantities that result from given (economic) events, such as the loss of customers, the loss of a supply source, or a catastrophic event. If the list of events is mutually exclusive and exhaustive and the events can be assigned probabilities, the analyst can evaluate not only the range of outcomes but also standard statistical measures such as the mean and median value for various quantities of interest.
- **Simulation**: This is computer-generated sensitivity or scenario analysis based on probability models for the factors that drive outcomes. Each event or possible outcome is assigned a probability. Multiple scenarios are then run using the probability factors assigned to the possible values of a variable.

SUMMARY

Financial analysis techniques, including common-size financial statements and ratio analysis, are useful in summarizing financial reporting data and evaluating the performance and financial position of a company. The results of financial analysis techniques provide important inputs into security valuation. Key facets of financial analysis include the following:

- Common-size financial statements and financial ratios remove the effect of size, allowing comparisons of a company with peer companies (cross-sectional analysis) and comparison of a company's results over time (trend or time-series analysis).
- Activity ratios measure the efficiency of a company's operations, such as collection of receivables or management of inventory. Major activity ratios include inventory turnover, days of inventory on hand, receivables turnover, days of sales outstanding, payables turnover, number of days of payables, working capital turnover, fixed asset turnover, and total asset turnover.
- Liquidity ratios measure the ability of a company to meet short-term obligations. Major liquidity ratios include the current ratio, quick ratio, cash ratio, and defensive interval ratio.
- Solvency ratios measure the ability of a company to meet long-term obligations. Major solvency ratios include debt ratios (including the debt-to-assets ratio, debt-to-capital ratio, debt-to-equity ratio, and financial leverage ratio) and coverage ratios (including interest coverage and fixed charge coverage).
- Profitability ratios measure the ability of a company to generate profits from revenue and assets. Major profitability ratios include return on sales ratios (including gross profit margin, operating profit margin, pretax margin, and net profit margin) and return on investment ratios (including operating ROA, ROA, return on total capital, ROE, and return on common equity).

- Ratios can also be combined and evaluated as a group to better understand how they fit together and how efficiency and leverage are tied to profitability.
- ROE can be analyzed as the product of the net profit margin, asset turnover, and financial leverage. This decomposition is sometimes referred to as DuPont analysis.
- Valuation ratios express the relation between the market value of a company or its equity (for example, price per share) and some fundamental financial metric (for example, earnings per share).
- Ratio analysis is useful in the selection and valuation of debt and equity securities and is a part of the credit rating process.
- Ratios can also be computed for business segments to evaluate how units within a business are performing.
- The results of financial analysis provide valuable inputs into forecasts of future earnings and cash flow.

REFERENCES

Abarbanell, J.S., and B. J. Bushee. 1997. “Fundamental Analysis, Future Earnings, and Stock Prices.” *Journal of Accounting Research*, vol. 35, no. 1:1–24.

Abarbanell, J.S., and B.J. Bushee. 1998. “Abnormal Returns to a Fundamental Analysis Strategy.” *Accounting Review*, vol. 73, no. 1:19–46.

Altman, E. 1968. “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy.” *Journal of Finance*, vol. 23, no. 4:589–609.

Altman, E.I. 2013, “Predicting Financial Distress of Companies: Revisiting the Z-Score and Zeta Models,” in Prokopczuk, Marcel; Brooks, Chris; Bell, Adrian R., *Handbook of Research Methods and Applications in Empirical Finance*: 428-456.

Altman, E., R. Haldeman, and P. Narayanan. 1977. “Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations.” *Journal of Banking & Finance*, vol. 1, no. 1.

Beaver, W. 1967. “Financial Ratios as Predictors of Failures.” *Empirical Research in Accounting*, selected studies supplement to *Journal of Accounting Research*, 4 (1).

Benninga, Simon Z., and Oded H. Sarig. 1997. *Corporate Finance: A Valuation Approach*. New York: McGraw-Hill Publishing.

Ederington, L.H. 1986. “Why Split Ratings Occur.” *Financial Management*, vol. 15, no. 1:37–47.

Ederington, L.H., J.B. Yawitz, and B.E. Robert. 1987. “The Information Content of Bond Ratings.” *Journal of Financial Research*, vol. 10, no. 3:211–226.

Lev, B., and S.R. Thiagarajan. 1993. “Fundamental Information Analysis.” *Journal of Accounting Research*, vol. 31, no. 2:190–215.

Modigliani, F., and M. Miller. 1958. “The Cost of Capital, Corporation Finance and the Theory of Investment.” *American Economic Review*, vol. 48:261–298.

Modigliani, F., and M. Miller. 1963. “Corporate Income Taxes and the Cost of Capital: A Correction.” *American Economic Review*, vol. 53:433–444.

Ou, J.A., and S.H. Penman. 1989a. “Financial Statement Analysis and the Prediction of Stock Returns.” *Journal of Accounting and Economics*, vol. 11, no. 4:295–329.

Ou, J.A., and S.H. Penman. 1989b. “Accounting Measurement, Price-Earnings Ratio, and the Information Content of Security Prices.” *Journal of Accounting Research*, vol. 27, no. Supplement:111–144.

Piotroski, J.D. 2000. “Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers.” *Journal of Accounting Research*, vol. 38, no. Supplement:1–41.

Robinson, T., and P. Munter. 2004. “Financial Reporting Quality: Red Flags and Accounting Warning Signs.” *Commercial Lending Review*, vol. 19, no. 1:2–15.

van Greuning, H., and S. Brajovic Bratanovic. 2003. *Analyzing and Managing Banking Risk: A Framework for Assessing Corporate Governance and Financial Risk*. Washington, DC: World Bank.

PRACTICE PROBLEMS

1 Comparison of a company's financial results to other peer companies for the same time period is called:

A technical analysis.

B time-series analysis.

C cross-sectional analysis.

2 In order to assess a company's ability to fulfill its long-term obligations, an analyst would *most likely* examine:

A activity ratios.

B liquidity ratios.

C solvency ratios.

3 Which ratio would a company *most likely* use to measure its ability to meet short-term obligations?

A Current ratio.

B Payables turnover.

C Gross profit margin.

4 Which of the following ratios would be *most* useful in determining a company's ability to cover its lease and interest payments?

A ROA.

B Total asset turnover.

C Fixed charge coverage.

5 An analyst is interested in assessing both the efficiency and liquidity of Spherion PLC. The analyst has collected the following data for Spherion:

	FY3	FY2	FY1
Days of inventory on hand	32	34	40
Days sales outstanding	28	25	23
Number of days of payables	40	35	35

Based on this data, what is the analyst *least likely* to conclude?

A Inventory management has contributed to improved liquidity.

B Management of payables has contributed to improved liquidity.

C Management of receivables has contributed to improved liquidity.

6 An analyst is evaluating the solvency and liquidity of Apex Manufacturing and has collected the following data (in millions of euro):

	FY5 (€)	FY4 (€)	FY3 (€)
Total debt	2,000	1,900	1,750
Total equity	4,000	4,500	5,000

Which of the following would be the analyst's *most likely* conclusion?

A The company is becoming increasingly less solvent, as evidenced by the increase in its debt-to-equity ratio from 0.35 to 0.50 from FY3 to FY5.

© 2011 CFA Institute. All rights reserved.

- **B** The company is becoming less liquid, as evidenced by the increase in its debt-to-equity ratio from 0.35 to 0.50 from FY3 to FY5.
- **C** The company is becoming increasingly more liquid, as evidenced by the increase in its debt-to-equity ratio from 0.35 to 0.50 from FY3 to FY5.

7 With regard to the data in Problem 6, what would be the *most* reasonable explanation of the financial data?

- **A** The decline in the company's equity results from a decline in the market value of this company's common shares.
- **B** The €250 increase in the company's debt from FY3 to FY5 indicates that lenders are viewing the company as increasingly creditworthy.
- **C** The decline in the company's equity indicates that the company may be incurring losses, paying dividends greater than income, and/or repurchasing shares.

8 An analyst observes a decrease in a company's inventory turnover. Which of the following would *most likely* explain this trend?

- **A** The company installed a new inventory management system, allowing more efficient inventory management.
- **B** Due to problems with obsolescent inventory last year, the company wrote off a large amount of its inventory at the beginning of the period.
- **C** The company installed a new inventory management system but experienced some operational difficulties resulting in duplicate orders being placed with suppliers.

9 Which of the following would *best* explain an increase in receivables turnover?

- **A** The company adopted new credit policies last year and began offering credit to customers with weak credit histories.
- **B** Due to problems with an error in its old credit scoring system, the company had accumulated a substantial amount of uncollectible accounts and wrote off a large amount of its receivables.
- **C** To match the terms offered by its closest competitor, the company adopted new payment terms now requiring net payment within 30 days rather than 15 days, which had been its previous requirement.

10 Brown Corporation had average days of sales outstanding of 19 days in the most recent fiscal year. Brown wants to improve its credit policies and collection practices and decrease its collection period in the next fiscal year to match the industry average of 15 days. Credit sales in the most recent fiscal year were \$300 million, and Brown expects credit sales to increase to \$390 million in the next fiscal year. To achieve Brown's goal of decreasing the collection period, the change in the average accounts receivable balance that must occur is *closest* to:

- **A** +\$0.41 million.
- **B** –\$0.41 million.
- **C** –\$1.22 million.

11 An analyst observes the following data for two companies:

	Company A (\$)	Company B (\$)
Revenue	4,500	6,000
Net income	50	1,000
Current assets	40,000	60,000
Total assets	100,000	700,000

	Company A ($)	Company B ($)
Current liabilities	10,000	50,000
Total debt	60,000	150,000
Shareholders' equity	30,000	500,000

Which of the following choices *best* describes reasonable conclusions that the analyst might make about the two companies' ability to pay their current and long-term obligations?

A Company A's current ratio of 4.0 indicates it is more liquid than Company B, whose current ratio is only 1.2, but Company B is more solvent, as indicated by its lower debt-to-equity ratio.

B Company A's current ratio of 0.25 indicates it is less liquid than Company B, whose current ratio is 0.83, and Company A is also less solvent, as indicated by a debt-to-equity ratio of 200 percent compared with Company B's debt-to-equity ratio of only 30 percent.

C Company A's current ratio of 4.0 indicates it is more liquid than Company B, whose current ratio is only 1.2, and Company A is also more solvent, as indicated by a debt-to-equity ratio of 200 percent compared with Company B's debt-to-equity ratio of only 30 percent.

The following information relates to Questions 12–15

The data in Exhibit 1 appear in the five-year summary of a major international company. A business combination with another major manufacturer took place in FY13.

Exhibit 1

	FY10	FY11	FY12	FY13	FY14
Financial statements	GBP m	GBP m	GBP m	GBP m	GBP m
Income statements					
Revenue	4,390	3,624	3,717	8,167	11,366
Profit before interest and taxation (EBIT)	844	700	704	933	1,579
Net interest payable	–80	–54	–98	–163	–188
Taxation	–186	–195	–208	–349	–579
Minorities	–94	–99	–105	–125	–167
Profit for the year	484	352	293	296	645
Balance sheets					
Fixed assets	3,510	3,667	4,758	10,431	11,483
Current asset investments, cash at bank and in hand	316	218	290	561	682
Other current assets	558	514	643	1,258	1,634
Total assets	4,384	4,399	5,691	12,250	13,799
Interest bearing debt (long term)	–602	–1,053	–1,535	–3,523	–3,707

(continued)

Exhibit 1 (Continued)

	FY10	FY11	FY12	FY13	FY14
Other creditors and provisions (current)	−1,223	−1,054	−1,102	−2,377	−3,108
Total liabilities	−1,825	−2,107	−2,637	−5,900	−6,815
Net assets	2,559	2,292	3,054	6,350	6,984
Shareholders' funds	2,161	2,006	2,309	5,572	6,165
Equity minority interests	398	286	745	778	819
Capital employed	2,559	2,292	3,054	6,350	6,984
Cash flow					
Working capital movements	−53	5	71	85	107
Net cash inflow from operating activities	864	859	975	1,568	2,292

12 The company's total assets at year-end FY9 were GBP 3,500 million. Which of the following choices *best* describes reasonable conclusions an analyst might make about the company's efficiency?

A Comparing FY14 with FY10, the company's efficiency improved, as indicated by a total asset turnover ratio of 0.86 compared with 0.64.

B Comparing FY14 with FY10, the company's efficiency deteriorated, as indicated by its current ratio.

C Comparing FY14 with FY10, the company's efficiency deteriorated due to asset growth faster than turnover revenue growth.

13 Which of the following choices *best* describes reasonable conclusions an analyst might make about the company's solvency?

A Comparing FY14 with FY10, the company's solvency improved, as indicated by an increase in its debt-to-assets ratio from 0.14 to 0.27.

B Comparing FY14 with FY10, the company's solvency deteriorated, as indicated by a decrease in interest coverage from 10.6 to 8.4.

C Comparing FY14 with FY10, the company's solvency improved, as indicated by the growth in its profits to GBP 645 million.

14 Which of the following choices *best* describes reasonable conclusions an analyst might make about the company's liquidity?

A Comparing FY14 with FY10, the company's liquidity improved, as indicated by an increase in its debt-to-assets ratio from 0.14 to 0.27.

B Comparing FY14 with FY10, the company's liquidity deteriorated, as indicated by a decrease in interest coverage from 10.6 to 8.4.

C Comparing FY14 with FY10, the company's liquidity improved, as indicated by an increase in its current ratio from 0.71 to 0.75.

15 Which of the following choices *best* describes reasonable conclusions an analyst might make about the company's profitability?

A Comparing FY14 with FY10, the company's profitability improved, as indicated by an increase in its debt-to-assets ratio from 0.14 to 0.27.

B Comparing FY14 with FY10, the company's profitability deteriorated, as indicated by a decrease in its net profit margin from 11.0 percent to 5.7 percent.

C Comparing FY14 with FY10, the company's profitability improved, as indicated by the growth in its shareholders' equity to GBP 6,165 million.

---

16 Assuming no changes in other variables, which of the following would decrease ROA?

A A decrease in the effective tax rate.

B A decrease in interest expense.

C An increase in average assets.

17 An analyst compiles the following data for a company:

	FY13	FY14	FY15
ROE	19.8%	20.0%	22.0%
Return on total assets	8.1%	8.0%	7.9%
Total asset turnover	2.0	2.0	2.1

Based only on the information above, the *most* appropriate conclusion is that, over the period FY13 to FY15, the company's:

A net profit margin and financial leverage have decreased.

B net profit margin and financial leverage have increased.

C net profit margin has decreased but its financial leverage has increased.

18 A decomposition of ROE for Integra SA is as follows:

	FY12	FY11
ROE	18.90%	18.90%
Tax burden	0.70	0.75
Interest burden	0.90	0.90
EBIT margin	10.00%	10.00%
Asset turnover	1.50	1.40
Leverage	2.00	2.00

Which of the following choices *best* describes reasonable conclusions an analyst might make based on this ROE decomposition?

A Profitability and the liquidity position both improved in FY12.

B The higher average tax rate in FY12 offset the improvement in profitability, leaving ROE unchanged.

C The higher average tax rate in FY12 offset the improvement in efficiency, leaving ROE unchanged.

19 A decomposition of ROE for Company A and Company B is as follows:

	Company A		Company B	
	FY15	FY14	FY15	FY14
ROE	26.46%	18.90%	26.33%	18.90%
Tax burden	0.7	0.75	0.75	0.75
Interest burden	0.9	0.9	0.9	0.9

(continued)

	Company A		Company B	
	FY15	**FY14**	**FY15**	**FY14**
EBIT margin	7.00%	10.00%	13.00%	10.00%
Asset turnover	1.5	1.4	1.5	1.4
Leverage	4	2	2	2

An analyst is *most likely* to conclude that:

A Company A's ROE is higher than Company B's in FY15, and one explanation consistent with the data is that Company A may have purchased new, more efficient equipment.

B Company A's ROE is higher than Company B's in FY15, and one explanation consistent with the data is that Company A has made a strategic shift to a product mix with higher profit margins.

C The difference between the two companies' ROE in FY15 is very small and Company A's ROE remains similar to Company B's ROE mainly due to Company A increasing its financial leverage.

20 What does the P/E ratio measure?

A The "multiple" that the stock market places on a company's EPS.

B The relationship between dividends and market prices.

C The earnings for one common share of stock.

21 A creditor *most likely* would consider a decrease in which of the following ratios to be positive news?

A Interest coverage (times interest earned).

B Debt-to-total assets.

C Return on assets.

22 When developing forecasts, analysts should *most likely*:

A develop possibilities relying exclusively on the results of financial analysis.

B use the results of financial analysis, analysis of other information, and judgment.

C aim to develop extremely precise forecasts using the results of financial analysis.

SOLUTIONS

1 C is correct. Cross-sectional analysis involves the comparison of companies with each other for the same time period. Technical analysis uses price and volume data as the basis for investment decisions. Time-series or trend analysis is the comparison of financial data across different time periods.

2 C is correct. Solvency ratios are used to evaluate the ability of a company to meet its long-term obligations. An analyst is more likely to use activity ratios to evaluate how efficiently a company uses its assets. An analyst is more likely to use liquidity ratios to evaluate the ability of a company to meet its short-term obligations.

3 A is correct. The current ratio is a liquidity ratio. It compares the net amount of current assets expected to be converted into cash within the year with liabilities falling due in the same period. A current ratio of 1.0 would indicate that the company would have just enough current assets to pay current liabilities.

4 C is correct. The fixed charge coverage ratio is a coverage ratio that relates known fixed charges or obligations to a measure of operating profit or cash flow generated by the company. Coverage ratios, a category of solvency ratios, measure the ability of a company to cover its payments related to debt and leases.

5 C is correct. The analyst is *unlikely* to reach the conclusion given in Statement C because days of sales outstanding increased from 23 days in FY1 to 25 days in FY2 to 28 days in FY3, indicating that the time required to collect receivables has increased over the period. This is a negative factor for Spherion's liquidity. By contrast, days of inventory on hand dropped over the period FY1 to FY3, a positive for liquidity. The company's increase in days payable, from 35 days to 40 days, shortened its cash conversion cycle, thus also contributing to improved liquidity.

6 A is correct. The company is becoming increasingly less solvent, as evidenced by its debt-to-equity ratio increasing from 0.35 to 0.50 from FY3 to FY5. The amount of a company's debt and equity do not provide direct information about the company's liquidity position.

Debt to equity:

FY5: 2,000/4,000 = 0.5000

FY4: 1,900/4,500 = 0.4222

FY3: 1,750/5,000 = 0.3500

7 C is correct. The decline in the company's equity indicates that the company may be incurring losses, paying dividends greater than income, or repurchasing shares. Recall that Beginning equity + New shares issuance − Shares repurchased + Comprehensive income − Dividends = Ending equity. The book value of a company's equity is not affected by changes in the market value of its common stock. An increased amount of lending does not necessarily indicate that lenders view a company as increasingly creditworthy. Creditworthiness is not evaluated based on how much a company has increased its debt but rather on its willingness and ability to pay its obligations. (Its financial strength is indicated by its solvency, liquidity, profitability, efficiency, and other aspects of credit analysis.)

8 C is correct. The company's problems with its inventory management system causing duplicate orders would likely result in a higher amount of inventory and would, therefore, result in a decrease in inventory turnover. A more efficient inventory management system and a write off of inventory at the beginning of

the period would both likely decrease the average inventory for the period (the denominator of the inventory turnover ratio), thus increasing the ratio rather than decreasing it.

9 B is correct. A write off of receivables would decrease the average amount of accounts receivable (the denominator of the receivables turnover ratio), thus increasing this ratio. Customers with weaker credit are more likely to make payments more slowly or to pose collection difficulties, which would likely increase the average amount of accounts receivable and thus decrease receivables turnover. Longer payment terms would likely increase the average amount of accounts receivable and thus decrease receivables turnover.

10 A is correct. The average accounts receivable balances (actual and desired) must be calculated to determine the desired change. The average accounts receivable balance can be calculated as an average day's credit sales times the DSO. For the most recent fiscal year, the average accounts receivable balance is \$15.62 million [= (\$300,000,000/365) × 19]. The desired average accounts receivable balance for the next fiscal year is \$16.03 million (= (\$390,000,000/365) × 15). This is an increase of \$0.41 million (= 16.03 million – 15.62 million). An alternative approach is to calculate the turnover and divide sales by turnover to determine the average accounts receivable balance. Turnover equals 365 divided by DSO. Turnover is 19.21 (= 365/19) for the most recent fiscal year and is targeted to be 24.33 (= 365/15) for the next fiscal year. The average accounts receivable balances are \$15.62 million (= \$300,000,000/19.21), and \$16.03 million (= \$390,000,000/24.33). The change is an increase in receivables of \$0.41 million

11 A is correct. Company A's current ratio of 4.0 (= \$40,000/\$10,000) indicates it is more liquid than Company B, whose current ratio is only 1.2 (= \$60,000/\$50,000). Company B is more solvent, as indicated by its lower debt-to-equity ratio of 30 percent (= \$150,000/\$500,000) compared with Company A's debt-to-equity ratio of 200 percent (= \$60,000/\$30,000).

12 C is correct. The company's efficiency deteriorated, as indicated by the decline in its total asset turnover ratio from 1.11 {= 4,390/[(4,384 + 3,500)/2]} for FY10 to 0.87 {= 11,366/[(12,250 + 13,799)/2]} for FY14. The decline in the total asset turnover ratio resulted from an increase in average total assets from GBP3,942 [= (4,384 + 3,500)/2] for FY10 to GBP13,024.5 for FY14, an increase of 230 percent, compared with an increase in revenue from GBP4,390 in FY10 to GBP11,366 in FY14, an increase of only 159 percent. The current ratio is not an indicator of efficiency.

13 B is correct. Comparing FY14 with FY10, the company's solvency deteriorated, as indicated by a decrease in interest coverage from 10.6 (= 844/80) in FY10 to 8.4 (= 1,579/188) in FY14. The debt-to-asset ratio increased from 0.14 (= 602/4,384) in FY10 to 0.27 (= 3,707/13,799) in FY14. This is also indicative of deteriorating solvency. In isolation, the amount of profits does not provide enough information to assess solvency.

14 C is correct. Comparing FY14 with FY10, the company's liquidity improved, as indicated by an increase in its current ratio from 0.71 [= (316 + 558)/1,223] in FY10 to 0.75 [= (682 + 1,634)/3,108] in FY14. Note, however, comparing only current investments with the level of current liabilities shows a decline in liquidity from 0.26 (= 316/1,223) in FY10 to 0.22 (= 682/3,108) in FY14. Debt-to-assets ratio and interest coverage are measures of solvency not liquidity.

15 B is correct. Comparing FY14 with FY10, the company's profitability deteriorated, as indicated by a decrease in its net profit margin from 11.0 percent (= 484/4,390) to 5.7 percent (= 645/11,366). Debt-to-assets ratio is a measure of solvency not an indicator of profitability. Growth in shareholders' equity, in isolation, does not provide enough information to assess profitability.

16 C is correct. Assuming no changes in other variables, an increase in average assets (an increase in the denominator) would decrease ROA. A decrease in either the effective tax rate or interest expense, assuming no changes in other variables, would increase ROA.

17 C is correct. The company's net profit margin has decreased and its financial leverage has increased. ROA = Net profit margin × Total asset turnover. ROA decreased over the period despite the increase in total asset turnover; therefore, the net profit margin must have decreased.

ROE = Return on assets × Financial leverage. ROE increased over the period despite the drop in ROA; therefore, financial leverage must have increased.

18 C is correct. The increase in the average tax rate in FY12, as indicated by the decrease in the value of the tax burden (the tax burden equals one minus the average tax rate), offset the improvement in efficiency indicated by higher asset turnover) leaving ROE unchanged. The EBIT margin, measuring profitability, was unchanged in FY12 and no information is given on liquidity.

19 C is correct. The difference between the two companies' ROE in 2010 is very small and is mainly the result of Company A's increase in its financial leverage, indicated by the increase in its Assets/Equity ratio from 2 to 4. The impact of efficiency on ROE is identical for the two companies, as indicated by both companies' asset turnover ratios of 1.5. Furthermore, if Company A had purchased newer equipment to replace older, depreciated equipment, then the company's asset turnover ratio (computed as sales/assets) would have declined, assuming constant sales. Company A has experienced a significant decline in its operating margin, from 10 percent to 7 percent which, all else equal, would not suggest that it is selling more products with higher profit margins.

20 A is correct. The P/E ratio measures the "multiple" that the stock market places on a company's EPS.

21 B is correct. In general, a creditor would consider a decrease in debt to total assets as positive news. A higher level of debt in a company's capital structure increases the risk of default and will, in general, result in higher borrowing costs for the company to compensate lenders for assuming greater credit risk. A decrease in either interest coverage or return on assets is likely to be considered negative news.

22 B is correct. The results of an analyst's financial analysis are integral to the process of developing forecasts, along with the analysis of other information and judgment of the analysts. Forecasts are not limited to a single point estimate but should involve a range of possibilities.

FINANCIAL STATEMENT ANALYSIS

STUDY SESSION 7

Financial Statement Analysis (3)

This study session examines financial reporting for specific categories of assets and liabilities. Inventories, long-lived assets, income taxes, and non-current liabilities are examined in greater detail because of their effect on financial statements and reported measures of profitability, liquidity, and solvency. For these items in particular, the analyst should be attentive to chosen accounting treatment, corresponding effect on reported performance, and the potential for financial statement manipulation.

READING ASSIGNMENTS

Reading 21	Inventories by Michael Broihahn, CPA, CIA, CFA
Reading 22	Long-lived Assets by Elaine Henry, PhD, CFA, and Elizabeth A. Gordon, PhD, MBA, CPA
Reading 23	Income Taxes by Elbie Louw, PhD, CFA, CIPM, and Michael A. Broihahn, CPA, CIA, CFA
Reading 24	Non-current (Long-term) Liabilities by Elizabeth A. Gordon, PhD, MBA, CPA, and Elaine Henry, PhD, CFA

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2021 CFA Institute. All rights reserved.

READING

21

Inventories

by Michael A. Broihahn, CPA, CIA, CFA

Michael A. Broihahn, CPA, CIA, CFA, is at Barry University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** contrast costs included in inventories and costs recognised as expenses in the period in which they are incurred;
☐	**b.** describe different inventory valuation methods (cost formulas);
☐	**c.** calculate and compare cost of sales, gross profit, and ending inventory using different inventory valuation methods and using perpetual and periodic inventory systems;
☐	**d.** calculate and explain how inflation and deflation of inventory costs affect the financial statements and ratios of companies that use different inventory valuation methods;
☐	**e.** explain LIFO reserve and LIFO liquidation and their effects on financial statements and ratios;
☐	**f.** demonstrate the conversion of a company's reported financial statements from LIFO to FIFO for purposes of comparison;
☐	**g.** describe the measurement of inventory at the lower of cost and net realisable value;
☐	**h.** describe implications of valuing inventory at net realisable value for financial statements and ratios;
☐	**i.** describe the financial statement presentation of and disclosures relating to inventories;
☐	**j.** explain issues that analysts should consider when examining a company's inventory disclosures and other sources of information;
☐	**k.** calculate and compare ratios of companies, including companies that use different inventory methods;
☐	**l.** analyze and compare the financial statements of companies, including companies that use different inventory methods.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION

Merchandising and manufacturing companies generate revenues and profits through the sale of inventory. Further, inventory may represent a significant asset on these companies' balance sheets. Merchandisers (wholesalers and retailers) purchase inventory, ready for sale, from manufacturers and thus account for only one type of inventory—finished goods inventory. Manufacturers, however, purchase raw materials from suppliers and then add value by transforming the raw materials into finished goods. They typically classify inventory into three different categories:[1] raw materials, work in progress,[2] and finished goods. Work-in-progress inventories have started the conversion process from raw materials but are not yet finished goods ready for sale. Manufacturers may report either the separate carrying amounts of their raw materials, work-in-progress, and finished goods inventories on the balance sheet or simply the total inventory amount. If the latter approach is used, the company must then disclose the carrying amounts of its raw materials, work-in-progress, and finished goods inventories in a footnote to the financial statements.

Inventories and cost of sales (cost of goods sold)[3] are significant items in the financial statements of many companies. Comparing the performance of these companies is challenging because of the allowable choices for valuing inventories: Differences in the choice of inventory valuation method can result in significantly different amounts being assigned to inventory and cost of sales. Financial statement analysis would be much easier if all companies used the same inventory valuation method or if inventory price levels remained constant over time. If there was no inflation or deflation with respect to inventory costs and thus unit costs were unchanged, the choice of inventory valuation method would be irrelevant. However, inventory price levels typically do change over time.

International Financial Reporting Standards (IFRS) permit the assignment of inventory costs (costs of goods available for sale) to inventories and cost of sales by three cost formulas: specific identification, first-in, first-out (FIFO), and weighted average cost.[4] US generally accepted accounting principles (US GAAP) allow the same three inventory valuation methods, referred to as cost flow assumptions in US GAAP, but also include a fourth method called last-in, first-out (LIFO).[5] The choice of inventory valuation method affects the allocation of the cost of goods available for sale to ending inventory and cost of sales. Analysts must understand the various inventory valuation methods and the related impact on financial statements and financial ratios in order to evaluate a company's performance over time and relative to industry peers. The company's financial statements and related notes provide important information that the analyst can use in assessing the impact of the choice of inventory valuation method on financial statements and financial ratios.

This reading is organized as follows: Section 2 discusses the costs that are included in inventory and the costs that are recognised as expenses in the period in which they are incurred. Sections 3–6 describe inventory valuation methods and compare the measurement of ending inventory, cost of sales and gross profit under each method, and when using periodic versus perpetual inventory systems. Sections 7 and 8 describe the LIFO method, LIFO reserve, and effects of LIFO liquidations, and demonstrate the adjustments required to compare a company that uses LIFO with one that uses

---

[1] Other classifications are possible. Inventory classifications should be appropriate to the entity.

[2] This category is commonly referred to as *work in process* under US GAAP.

[3] Typically, *cost of sales* is IFRS terminology and *cost of goods sold* is US GAAP terminology.

[4] International Accounting Standard (IAS) 2 [Inventories].

[5] Financial Accounting Standards Board *Accounting Standards Codification* (FASB ASC) Topic 330 [Inventory].

FIFO. Section 9 describes the financial statement effects of a change in inventory valuation method. Section 10 discusses the measurement and reporting of inventory when its value changes. Sections 11–13 describe the presentation of inventories on the financial statements and related disclosures, discuss inventory ratios and their interpretation, and show examples of financial analysis with respect to inventories. A summary and practice problems conclude the reading.

COST OF INVENTORIES

2

a contrast costs included in inventories and costs recognised as expenses in the period in which they are incurred

Under IFRS, the costs to include in inventories are "all costs of purchase, costs of conversion, and other costs incurred in bringing the inventories to their present location and condition."[6] The costs of purchase include the purchase price, import and tax-related duties, transport, insurance during transport, handling, and other costs directly attributable to the acquisition of finished goods, materials, and services. Trade discounts, rebates, and similar items reduce the price paid and the costs of purchase. The costs of conversion include costs directly related to the units produced, such as direct labour, and fixed and variable overhead costs.[7] Including these product-related costs in inventory (i.e., as an asset) means that they will not be recognised as an expense (i.e., as cost of sales) on the income statement until the inventory is sold. US GAAP provide a similar description of the costs to be included in inventory.[8]

Both IFRS and US GAAP exclude the following costs from inventory: abnormal costs incurred as a result of waste of materials, labour or other production conversion inputs, any storage costs (unless required as part of the production process), and all administrative overhead and selling costs. These excluded costs are treated as expenses and recognised on the income statement in the period in which they are incurred. Including costs in inventory defers their recognition as an expense on the income statement until the inventory is sold. Therefore, including costs in inventory that should be expensed will overstate profitability on the income statement (because of the inappropriate deferral of cost recognition) and create an overstated inventory value on the balance sheet.

EXAMPLE 1

Treatment of Inventory-Related Costs

Acme Enterprises, a hypothetical company that prepares its financial statements in accordance with IFRS, manufactures tables. In 2018, the factory produced 900,000 finished tables and scrapped 1,000 tables. For the finished tables, raw material costs were €9 million, direct labour conversion costs were €18 million, and production overhead costs were €1.8 million. The 1,000 scrapped tables (attributable to abnormal waste) had a total production cost of €30,000 (€10,000 raw material costs and €20,000 conversion costs; these costs are not included in the €9 million raw material and €19.8 million total conversion costs of the finished tables). During the year, Acme spent €1 million for freight delivery charges on raw materials and €500,000 for storing finished goods inventory. Acme does not have any work-in-progress inventory at the end of the year.

---

[6] International Accounting Standard (IAS) 2 [Inventories].

[7] Fixed production overhead costs (depreciation, factory maintenance, and factory management and administration) represent indirect costs of production that remain relatively constant regardless of the volume of production. Variable production overhead costs are indirect production costs (indirect labour and materials) that vary with the volume of production.

[8] FASB Accounting Standards Codification™ (ASC) Topic 330 [Inventory].

1 What costs should be included in inventory in 2018?

2 What costs should be expensed in 2018?

Solution to 1:

Total inventory costs for 2018 are as follows:

Raw materials	€9,000,000
Direct labour	18,000,000
Production overhead	1,800,000
Transportation for raw materials	1,000,000
Total inventory costs	€29,800,000

Solution to 2:

Total costs that should be expensed (not included in inventory) are as follows:

Abnormal waste	€30,000
Storage of finished goods inventory	500,000
Total	€530,000

3

INVENTORY VALUATION METHODS

b describe different inventory valuation methods (cost formulas)

Generally, inventory purchase costs and manufacturing conversion costs change over time. As a result, the allocation of total inventory costs (i.e., cost of goods available for sale) between cost of sales on the income statement and inventory on the balance sheet will vary depending on the inventory valuation method used by the company. As mentioned in the introduction, inventory valuation methods are referred to as cost formulas and cost flow assumptions under IFRS and US GAAP, respectively. If the choice of method results in more cost being allocated to cost of sales and less cost being allocated to inventory than would be the case with other methods, the chosen method will cause, in the current year, reported gross profit, net income, and inventory carrying amount to be lower than if alternative methods had been used. Accounting for inventory, and consequently the allocation of costs, thus has a direct impact on financial statements and their comparability.

Both IFRS and US GAAP allow companies to use the following inventory valuation methods: specific identification; first-in, first-out (FIFO); and weighted average cost. US GAAP allow companies to use an additional method: last-in, first-out (LIFO). A company must use the same inventory valuation method for all items that have a similar nature and use. For items with a different nature or use, a different inventory valuation method can be used.[9] When items are sold, the carrying amount of the inventory is recognised as an expense (cost of sales) according to the cost formula (cost flow assumption) in use.

Specific identification is used for inventory items that are not ordinarily interchangeable, whereas FIFO, weighted average cost, and LIFO are typically used when there are large numbers of interchangeable items in inventory. Specific identification matches the actual historical costs of the specific inventory items to their physical flow; the costs remain in inventory until the actual identifiable inventory is sold. FIFO, weighted average cost, and LIFO are based on cost flow assumptions. Under these methods, companies must make certain assumptions about which goods are sold and which goods remain in ending inventory. As a result, the allocation of costs to the units sold and to the units in ending inventory can be different from the physical movement of the items.

The choice of inventory valuation method would be largely irrelevant if inventory costs remained constant or relatively constant over time. Given relatively constant prices, the allocation of costs between cost of goods sold and ending inventory would be very similar under each of the four methods. Given changing price levels, however, the choice of inventory valuation method can have a significant impact on the amount of reported cost of sales and inventory. And the reported cost of sales and inventory balances affect other items, such as gross profit, net income, current assets, and total assets.

3.1 Specific Identification

The specific identification method is used for inventory items that are not ordinarily interchangeable and for goods that have been produced and segregated for specific projects. This method is also commonly used for expensive goods that are uniquely identifiable, such as precious gemstones. Under this method, the cost of sales and the cost of ending inventory reflect the actual costs incurred to purchase (or manufacture) the items specifically identified as sold and the items specifically identified as remaining in inventory. Therefore, this method matches the physical flow of the specific items sold and remaining in inventory to their actual cost.

3.2 First-In, First-Out (FIFO)

FIFO assumes that the oldest goods purchased (or manufactured) are sold first and the newest goods purchased (or manufactured) remain in ending inventory. In other words, the first units included in inventory are assumed to be the first units sold from inventory. Therefore, cost of sales reflects the cost of goods in beginning inventory plus the cost of items purchased (or manufactured) earliest in the accounting period, and the value of ending inventory reflects the costs of goods purchased (or manufactured) more recently. In periods of rising prices, the costs assigned to the units in ending inventory are higher than the costs assigned to the units sold. Conversely, in periods of declining prices, the costs assigned to the units in ending inventory are lower than the costs assigned to the units sold.

---

[9] For example, if a clothing manufacturer produces both a retail line and one-of-a-kind designer garments, the retail line might be valued using FIFO and the designer garments using specific identification.

3.3 Weighted Average Cost

Weighted average cost assigns the average cost of the goods available for sale (beginning inventory plus purchase, conversion, and other costs) during the accounting period to the units that are sold as well as to the units in ending inventory. In an accounting period, the weighted average cost per unit is calculated as the total cost of the units available for sale divided by the total number of units available for sale in the period (Total cost of goods available for sale/Total units available for sale).

3.4 Last-In, First-Out (LIFO)

LIFO is permitted only under US GAAP. This method assumes that the newest goods purchased (or manufactured) are sold first and the oldest goods purchased (or manufactured), including beginning inventory, remain in ending inventory. In other words, the last units included in inventory are assumed to be the first units sold from inventory. Therefore, cost of sales reflects the cost of goods purchased (or manufactured) more recently, and the value of ending inventory reflects the cost of older goods. In periods of rising prices, the costs assigned to the units in ending inventory are lower than the costs assigned to the units sold. Conversely, in periods of declining prices, the costs assigned to the units in ending inventory are higher than the costs assigned to the units sold.

4

CALCULATIONS OF COST OF SALES, GROSS PROFIT, AND ENDING INVENTORY

c calculate and compare cost of sales, gross profit, and ending inventory using different inventory valuation methods and using perpetual and periodic inventory systems

In periods of changing prices, the allocation of total inventory costs (i.e., cost of goods available for sale) between cost of sales on the income statement and inventory on the balance sheet will vary depending on the inventory valuation method used by the company. The following example illustrates how cost of sales, gross profit, and ending inventory differ based on the choice of inventory valuation method.

EXAMPLE 2

Inventory Cost Flow Illustration for the Specific Identification, Weighted Average Cost, FIFO, and LIFO Methods

Global Sales, Inc. (GSI) is a hypothetical Dubai-based distributor of consumer products, including bars of luxury soap. The soap is sold by the kilogram. GSI began operations in 2018, during which it purchased and received initially 100,000 kg of soap at 110 dirham (AED)/kg, then 200,000 kg of soap at 100 AED/kg, and finally 300,000 kg of soap at 90 AED/kg. GSI sold 520,000 kg of soap at 240 AED/kg. GSI stores its soap in its warehouse so that soap from each shipment received is readily identifiable. During 2018, the entire 100,000 kg from the first shipment received, 180,000 kg of the second shipment received, and 240,000 kg of the final shipment received was sent to customers. Answers to the following questions should be rounded to the nearest 1,000 AED.

1 What are the reported cost of sales, gross profit, and ending inventory balances for 2018 under the specific identification method?

2 What are the reported cost of sales, gross profit, and ending inventory balances for 2018 under the weighted average cost method?

3 What are the reported cost of sales, gross profit, and ending inventory balances for 2018 under the FIFO method?

4 What are the reported cost of sales, gross profit, and ending inventory balances for 2018 under the LIFO method?

Solution to 1:

Under the specific identification method, the physical flow of the specific inventory items sold is matched to their actual cost.

Sales = 520,000 × 240 = 124,800,000 AED

Cost of sales = (100,000 × 110) + (180,000 × 100) + (240,000 × 90) = 50,600,000 AED

Gross profit = 124,800,000 – 50,600,000 = 74,200,000 AED

Ending inventory = (20,000 × 100) + (60,000 × 90) = 7,400,000 AED

Note that in spite of the segregation of inventory within the warehouse, it would be inappropriate to use specific identification for this inventory of interchangeable items. The use of specific identification could potentially result in earnings manipulation through the shipment decision.

Solution to 2:

Under the weighted average cost method, costs are allocated to cost of sales and ending inventory by using a weighted average mix of the actual costs incurred for all inventory items. The weighted average cost per unit is determined by dividing the total cost of goods available for sale by the number of units available for sale.

Weighted average cost = [(100,000 × 110) + (200,000 × 100) + (300,000 × 90)]/600,000 = 96.667 AED/kg

Sales = 520,000 × 240 = 124,800,000 AED

Cost of sales = 520,000 × 96.667 = 50,267,000 AED

Gross profit = 124,800,000 – 50,267,000 = 74,533,000 AED

Ending inventory = 80,000 × 96.667 = 7,733,360 AED

Solution to 3:

Under the FIFO method, the oldest inventory units acquired are assumed to be the first units sold. Ending inventory, therefore, is assumed to consist of those inventory units most recently acquired.

Sales = 520,000 × 240 = 124,800,000 AED

Cost of sales = (100,000 × 110) + (200,000 × 100) + (220,000 × 90) = 50,800,000 AED

Gross profit = 124,800,000 – 50,800,000 = 74,000,000 AED

Ending inventory = 80,000 × 90 = 7,200,000 AED

Solution to 4:

Under the LIFO method, the newest inventory units acquired are assumed to be the first units sold. Ending inventory, therefore, is assumed to consist of the oldest inventory units.

Sales = 520,000 × 240 = 124,800,000 AED

Cost of sales = (20,000 × 110) + (200,000 × 100) + (300,000 × 90) = 49,200,000 AED

Gross profit = 124,800,000 – 49,200,000 = 75,600,000 AED

Ending inventory = 80,000 × 110 = 8,800,000 AED

The following table (in thousands of AED) summarizes the cost of sales, the ending inventory, and the cost of goods available for sale that were calculated for each of the four inventory valuation methods. Note that in the first year of operation, the total cost of goods available for sale is the same under all four methods. Subsequently, the cost of goods available for sale will typically differ because beginning inventories will differ. Also shown is the gross profit figure for each of the four methods. Because the cost of a kg of soap declined over the period, LIFO had the highest ending inventory amount, the lowest cost of sales, and the highest gross profit. FIFO had the lowest ending inventory amount, the highest cost of sales, and the lowest gross profit.

Inventory Valuation Method	Specific ID	Weighted Average Cost	FIFO	LIFO
Cost of sales	50,600	50,267	50,800	49,200
Ending inventory	7,400	7,733	7,200	8,800
Total cost of goods available for sale	58,000	58,000	58,000	58,000
Gross profit	74,200	74,533	74,000	75,600

5 PERIODIC VERSUS PERPETUAL INVENTORY SYSTEMS

c calculate and compare cost of sales, gross profit, and ending inventory using different inventory valuation methods and using perpetual and periodic inventory systems

Companies typically record changes to inventory using either a periodic inventory system or a perpetual inventory system. Under a periodic inventory system, inventory values and costs of sales are determined at the end of an accounting period. Purchases are recorded in a purchases account. The total of purchases and beginning inventory is the amount of goods available for sale during the period. The ending inventory amount is subtracted from the goods available for sale to arrive at the cost of sales. The quantity of goods in ending inventory is usually obtained or verified through a physical count of the units in inventory. Under a perpetual inventory system, inventory values and cost of sales are continuously updated to reflect purchases and sales.

Under either system, the allocation of goods available for sale to cost of sales and ending inventory is the same if the inventory valuation method used is either specific identification or FIFO. This is not generally true for the weighted average cost method.

Under a periodic inventory system, the amount of cost of goods available for sale allocated to cost of sales and ending inventory may be quite different using the FIFO method compared to the weighted average cost method. Under a perpetual inventory system, inventory values and cost of sales are continuously updated to reflect purchases and sales. As a result, the amount of cost of goods available for sale allocated to cost of sales and ending inventory is similar under the FIFO and weighted average cost methods. Because of lack of disclosure and the dominance of perpetual inventory systems, analysts typically do not make adjustments when comparing a company using the weighted average cost method with a company using the FIFO method.

Using the LIFO method, the periodic and perpetual inventory systems will generally result in different allocations to cost of sales and ending inventory. Under either a perpetual or periodic inventory system, the use of the LIFO method will generally result in significantly different allocations to cost of sales and ending inventory compared to other inventory valuation methods. When inventory costs are increasing and inventory unit levels are stable or increasing, using the LIFO method will result in higher cost of sales and lower inventory carrying amounts than using the FIFO method. The higher cost of sales under LIFO will result in lower gross profit, operating income, income before taxes, and net income. Income tax expense will be lower under LIFO, causing the company's net operating cash flow to be higher. On the balance sheet, the lower inventory carrying amount will result in lower reported current assets, working capital, and total assets. Analysts must carefully assess the financial statement implications of the choice of inventory valuation method when comparing companies that use the LIFO method with companies that use the FIFO method.

Example 3 illustrates the impact of the choice of system under LIFO.

EXAMPLE 3

Perpetual versus Periodic Inventory Systems

If GSI (the company in Example 2) had used a perpetual inventory system, the timing of purchases and sales would affect the amounts of cost of sales and inventory. Below is a record of the purchases, sales, and quantity of inventory on hand after the transaction in 2018.

Date	Purchased	Sold	Inventory on Hand
5 January	100,000 kg at 110 AED/kg		100,000 kg
1 February		80,000 kg at 240 AED/kg	20,000 kg
8 March	200,000 kg at 100 AED/kg		220,000 kg
6 April		100,000 kg at 240 AED/kg	120,000 kg
23 May		60,000 kg at 240 AED/kg	60,000 kg
7 July		40,000 kg at 240 AED/kg	20,000 kg
2 August	300,000 kg at 90 AED/kg		320,000 kg
5 September		70,000 kg at 240 AED/kg	250,000 kg
17 November		90,000 kg at 240 AED/kg	160,000 kg
8 December		80,000 kg at 240 AED/kg	80,000 kg
	Total goods available for sale = 58,000,000 AED	Total sales = 124,800,000 AED	

The amounts for total goods available for sale and sales are the same under either the perpetual or periodic system in this first year of operation. The carrying amount of the ending inventory, however, may differ because the perpetual system will apply LIFO continuously throughout the year. Under the periodic system, it was assumed that the ending inventory was composed of 80,000 units of the oldest inventory, which cost 110 AED/kg.

What are the ending inventory, cost of sales, and gross profit amounts using the perpetual system and the LIFO method? How do these compare with the amounts using the periodic system and the LIFO method, as in Example 2?

Solution:

The carrying amounts of the inventory at the different time points using the perpetual inventory system are as follows:

Date	Quantity on Hand	Quantities and Cost	Carrying Amount
5 January	100,000 kg	100,000 kg at 110 AED/kg	11,000,000 AED
1 February	20,000 kg	20,000 kg at 110 AED/kg	2,200,000 AED
8 March	220,000 kg	20,000 kg at 110 AED/kg + 200,000 kg at 100 AED/kg	22,200,000 AED
6 April	120,000 kg	20,000 kg at 110 AED/kg + 100,000 kg at 100 AED/kg	12,200,000 AED
23 May	60,000 kg	20,000 kg at 110 AED/kg + 40,000 kg at 100 AED/kg	6,200,000 AED
7 July	20,000 kg	20,000 kg at 110 AED/kg	2,200,000 AED
2 August	320,000 kg	20,000 kg at 110 AED/kg + 300,000 kg at 90 AED/kg	29,200,000 AED
5 September	250,000 kg	20,000 kg at 110 AED/kg + 230,000 kg at 90 AED/kg	22,900,000 AED
17 November	160,000 kg	20,000 kg at 110 AED/kg + 140,000 kg at 90 AED/kg	14,800,000 AED
8 December	80,000 kg	20,000 kg at 110 AED/kg + 60,000 kg at 90 AED/kg	7,600,000 AED

Perpetual system

Sales = 520,000 × 240 = 124,800,000 AED

Cost of sales = 58,000,000 – 7,600,000 = 50,400,000 AED

Gross profit = 124,800,000 – 50,400,000 = 74,400,000 AED

Ending inventory = 7,600,000 AED
Periodic system from Example 2

Sales = 520,000 × 240 = 124,800,000 AED

Cost of sales = (20,000 × 110) + (200,000 × 100) + (300,000 × 90) = 49,200,000 AED

Gross profit = 124,800,000 – 49,200,000 = 75,600,000 AED

Ending inventory = 80,000 × 110 = 8,800,000 AED

In this example, the ending inventory amount is lower under the perpetual system because only 20,000 kg of the oldest inventory with the highest cost is assumed to remain in inventory. The cost of sales is higher and the gross profit is lower under the perpetual system compared to the periodic system.

COMPARISON OF INVENTORY VALUATION METHODS

6

d calculate and explain how inflation and deflation of inventory costs affect the financial statements and ratios of companies that use different inventory valuation methods

As shown in Example 2, the allocation of the total cost of goods available for sale to cost of sales on the income statement and to ending inventory on the balance sheet varies under the different inventory valuation methods. In an environment of declining inventory unit costs and constant or increasing inventory quantities, FIFO (in comparison with weighted average cost or LIFO) will allocate a higher amount of the total cost of goods available for sale to cost of sales on the income statement and a lower amount to ending inventory on the balance sheet. Accordingly, because cost of sales will be higher under FIFO, a company's gross profit, operating profit, and income before taxes will be lower.

Conversely, in an environment of rising inventory unit costs and constant or increasing inventory quantities, FIFO (in comparison with weighted average cost or LIFO) will allocate a lower amount of the total cost of goods available for sale to cost of sales on the income statement and a higher amount to ending inventory on the balance sheet. Accordingly, because cost of sales will be lower under FIFO, a company's gross profit, operating profit, and income before taxes will be higher.

The carrying amount of inventories under FIFO will more closely reflect current replacement values because inventories are assumed to consist of the most recently purchased items. The cost of sales under LIFO will more closely reflect current replacement value. LIFO ending inventory amounts are typically not reflective of current replacement value because the ending inventory is assumed to be the oldest inventory and costs are allocated accordingly. Example 4 illustrates the different results obtained by using either the FIFO or LIFO methods to account for inventory.

EXAMPLE 4

Impact of Inflation Using LIFO Compared to FIFO

Company L and Company F are identical in all respects except that Company L uses the LIFO method and Company F uses the FIFO method. Each company has been in business for five years and maintains a base inventory of 2,000 units each year. Each year, except the first year, the number of units purchased equaled the number of units sold. Over the five year period, unit sales increased 10 percent each year and the unit purchase and selling prices increased at the beginning of each year to reflect inflation of 4 percent per year. In the first year, 20,000 units were sold at a price of $15.00 per unit and the unit purchase price was $8.00.

1. What was the end of year inventory, sales, cost of sales, and gross profit for each company for each of the five years?
2. Compare the inventory turnover ratios (based on ending inventory carrying amounts) and gross profit margins over the five year period and between companies.

Solution to 1:

Company L using LIFO	Year 1	Year 2	Year 3	Year 4	Year 5
Ending inventory[a]	\$16,000	\$16,000	\$16,000	\$16,000	\$16,000
Sales[b]	\$300,000	\$343,200	\$392,621	\$449,158	\$513,837
Cost of sales[c]	160,000	183,040	209,398	239,551	274,046
Gross profit	\$140,000	\$160,160	\$183,223	\$209,607	\$239,791

[a] Inventory is unchanged at \$16,000 each year (2,000 units × \$8). 2,000 of the units acquired in the first year are assumed to remain in inventory.

[b] Sales Year X = $(20{,}000 \times \$15)(1.10)^{X-1}(1.04)^{X-1}$. The quantity sold increases by 10 percent each year and the selling price increases by 4 percent each year.

[c] Cost of sales Year X = $(20{,}000 \times \$8)(1.10)^{X-1}(1.04)^{X-1}$. In Year 1, 20,000 units are sold with a cost of \$8. In subsequent years, the number of units purchased equals the number of units sold and the units sold are assumed to be those purchased in the year. The quantity purchased increases by 10 percent each year and the purchase price increases by 4 percent each year.

Note that if the company sold more units than it purchased in a year, inventory would decrease. This is referred to as LIFO liquidation. The cost of sales of the units sold in excess of those purchased would reflect the inventory carrying amount. In this example, each unit sold in excess of those purchased would have a cost of sales of \$8 and a higher gross profit.

Company F using FIFO	Year 1	Year 2	Year 3	Year 4	Year 5
Ending inventory[a]	\$16,000	\$16,640	\$17,306	\$17,998	\$18,718
Sales[b]	\$300,000	\$343,200	\$392,621	\$449,158	\$513,837
Cost of sales[c]	160,000	182,400	208,732	238,859	273,326
Gross profit	\$140,000	\$160,800	\$183,889	\$210,299	\$240,511

[a] Ending Inventory Year X = 2,000 units × Cost in Year X = 2,000 units $[\$8 \times (1.04)^{X-1}]$. 2,000 units of the units acquired in Year X are assumed to remain in inventory.

[b] Sales Year X = $(20{,}000 \times \$15)(1.10)^{X-1}(1.04)^{X-1}$

[c] Cost of sales Year 1 = \$160,000 (= 20,000 units × \$8). There was no beginning inventory.

Cost of sales Year X (where X ≠ 1) = Beginning inventory plus purchases less ending inventory

= (Inventory at Year X–1) + $[(20{,}000 \times \$8)(1.10)^{X-1}(1.04)^{X-1}]$ – (Inventory at Year X)

= $2{,}000(\$8)(1.04)^{X-2} + [(20{,}000 \times \$8)(1.10)^{X-1}(1.04)^{X-1}] - [2{,}000\ (\$8)(1.04)^{X-1}]$

For example, cost of sales Year 2 = 2,000(\$8) + [(20,000 × \$8)(1.10)(1.04)] – [2,000 (\$8)(1.04)] = \$16,000 + 183,040 – 16,640 = \$182,400

Solution to 2:

	Company L					Company F				
Year	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Inventory turnover	10.0	11.4	13.1	15.0	17.1	10.0	11.0	12.1	13.3	14.6
Gross profit margin (%)	46.7	46.7	46.7	46.7	46.7	46.7	46.9	46.8	46.8	46.8

Inventory turnover ratio = Cost of sales ÷ Ending inventory. The inventory turnover ratio increased each year for both companies because the units sold increased, whereas the units in ending inventory remained unchanged. The increase in the inventory turnover ratio is higher for Company L because Company L's cost of sales is increasing for inflation but the inventory carrying amount is unaffected by inflation. It might appear that a company using the LIFO method manages its inventory more effectively, but this is deceptive. Both companies have identical quantities and prices of purchases and sales and only differ in the inventory valuation method used.

Gross profit margin = Gross profit ÷ Sales. The gross profit margin is stable under LIFO because both sales and cost of sales increase at the same rate of inflation. The gross profit margin is slightly higher under the FIFO method after the first year because a proportion of the cost of sales reflects an older purchase price.

THE LIFO METHOD AND LIFO RESERVE

7

e explain LIFO reserve and LIFO liquidation and their effects on financial statements and ratios

f demonstrate the conversion of a company's reported financial statements from LIFO to FIFO for purposes of comparison

The potential income tax savings are a benefit of using the LIFO method when inventory costs are increasing. The higher cash flows due to lower income taxes may make the company more valuable because the value of a company is based on the present value of its future cash flows. Under the LIFO method, ending inventory is assumed to consist of those units that have been held the longest. This generally results in ending inventories with carrying amounts lower than current replacement costs because inventory costs typically increase over time. Cost of sales will more closely reflect current replacement costs.

If the purchase prices (purchase costs) or production costs of inventory are increasing, the income statement consequences of using the LIFO method compared to other methods will include higher cost of sales, and lower gross profit, operating profit, income tax expense, and net income. The balance sheet consequences include lower ending inventory, working capital, total assets, retained earnings, and shareholders' equity. The lower income tax paid will result in higher net cash flow from operating activities. Some of the financial ratio effects are a lower current ratio, higher debt-to-equity ratios, and lower profitability ratios.

If the purchase prices or production costs of inventory are decreasing, it is unlikely that a company will use the LIFO method for tax purposes (and therefore for financial reporting purposes due to the LIFO conformity rule) because this will result in lower cost of sales, and higher taxable income and income taxes. However, if the company had elected to use the LIFO method and cannot justify changing the inventory valuation method for tax and financial reporting purposes when inventory costs begin to decrease, the income statement, balance sheet, and ratio effects will be opposite to the effects during a period of increasing costs.

7.1 LIFO Reserve

For companies using the LIFO method, US GAAP requires disclosure, in the notes to the financial statements or on the balance sheet, of the amount of the LIFO reserve. The **LIFO reserve** is the difference between the reported LIFO inventory carrying amount and the inventory amount that would have been reported if the FIFO method had been used (in other words, the FIFO inventory value less the LIFO inventory value). The disclosure provides the information that analysts need to adjust a company's cost of sales (cost of goods sold) and ending inventory balance based on the LIFO method, to the FIFO method.

To compare companies using LIFO with companies not using LIFO, inventory is adjusted by adding the disclosed LIFO reserve to the inventory balance that is reported on the balance sheet. The reported inventory balance, using LIFO, plus the LIFO reserve equals the inventory that would have been reported under FIFO. Cost of sales is adjusted by subtracting the increase in the LIFO reserve during the period from the cost of sales amount that is reported on the income statement. If the LIFO reserve has declined during the period,[10] the decrease in the reserve is added to the cost of sales amount that is reported on the income statement. The LIFO reserve disclosure can be used to adjust the financial statements of a US company using the LIFO method to make them comparable with a similar company using the FIFO method.

8

LIFO LIQUIDATIONS

e explain LIFO reserve and LIFO liquidation and their effects on financial statements and ratios

In periods of rising inventory unit costs, the carrying amount of inventory under FIFO will always exceed the carrying amount of inventory under LIFO. The LIFO reserve may increase over time as the result of the increasing difference between the older costs used to value inventory under LIFO and the more recent costs used to value inventory under FIFO. Also, when the number of inventory units manufactured or purchased exceeds the number of units sold, the LIFO reserve may increase as the result of the addition of new LIFO layers (the quantity of inventory units is increasing and each increase in quantity creates a new LIFO layer).

When the number of units sold exceeds the number of units purchased or manufactured, the number of units in ending inventory is lower than the number of units in beginning inventory and a company using LIFO will experience a LIFO liquidation (some of the older units held in inventory are assumed to have been sold). If inventory unit costs have been rising from period to period and LIFO liquidation occurs, this will produce an inventory-related increase in gross profits. The increase in gross profits occurs because of the lower inventory carrying amounts of the liquidated units. The lower inventory carrying amounts are used for cost of sales and the sales are at the current prices. The gross profit on these units is higher than the gross profit that would be recognised using more current costs. These inventory profits caused by a LIFO liquidation, however, are one-time events and are not sustainable.

---

[10] This typically results from a reduction in inventory units and is referred to as LIFO liquidation. LIFO liquidation is discussed in the next section.

LIFO liquidations can occur for a variety of reasons. The reduction in inventory levels may be outside of management's control; for example, labour strikes at a supplier may force a company to reduce inventory levels to meet customer demands. In periods of economic recession or when customer demand is declining, a company may choose to reduce existing inventory levels rather than invest in new inventory. Analysts should be aware that management can potentially manipulate and inflate their company's reported gross profits and net income at critical times by intentionally reducing inventory quantities and liquidating older layers of LIFO inventory (selling some units of beginning inventory). During economic downturns, LIFO liquidation may result in higher gross profit than would otherwise be realised. If LIFO layers of inventory are temporarily depleted and not replaced by fiscal year-end, LIFO liquidation will occur resulting in unsustainable higher gross profits. Therefore, it is imperative to review the LIFO reserve footnote disclosures to determine if LIFO liquidation has occurred. A decline in the LIFO reserve from the prior period may be indicative of LIFO liquidation.

EXAMPLE 5

Inventory Conversion from LIFO to FIFO

Caterpillar Inc. (CAT), based in Peoria, Illinois, USA, is the largest maker of construction and mining equipment, diesel and natural gas engines, and industrial gas turbines in the world. Excerpts from CAT's consolidated financial statements are shown in Exhibits 1 and 2; notes pertaining to CAT's inventories are presented in Exhibit 3. CAT's Management Discussion and Analysis (MD&A) disclosure states that effective income tax rates were 28 percent for 2017 and 36 percent for 2016.

1. What inventory values would CAT report for 2017, 2016, and 2015 if it had used the FIFO method instead of the LIFO method?
2. What amount would CAT's cost of goods sold for 2017 and 2016 be if it had used the FIFO method instead of the LIFO method?
3. What net income (profit) would CAT report for 2017 and 2016 if it had used the FIFO method instead of the LIFO method?
4. By what amount would CAT's 2017 and 2016 net cash flow from operating activities decline if CAT used the FIFO method instead of the LIFO method?
5. What is the cumulative amount of income tax savings that CAT has generated through 2017 by using the LIFO method instead of the FIFO method?
6. What amount would be added to CAT's retained earnings (profit employed in the business) at 31 December 2017 if CAT had used the FIFO method instead of the LIFO method?
7. What would be the change in Cat's cash balance if CAT had used the FIFO method instead of the LIFO method?
8. Calculate and compare the following for 2017 under the LIFO method and the FIFO method: inventory turnover ratio, days of inventory on hand, gross profit margin, net profit margin, return on assets, current ratio, and total liabilities-to-equity ratio.

Exhibit 1 Caterpillar Inc. Consolidated Results of Operation (US$ millions)

For the years ended 31 December	2017	2016	2015
Sales and revenues:			
Sales of Machinery and Engines	42,676	35,773	44,147
Revenue of Financial Products	2,786	2,764	2,864
Total sales and revenues	45,462	38,537	47,011
Operating costs:			
Cost of goods sold	31,049	28,309	33,546
⋮	⋮	⋮	⋮
Interest expense of Financial Products	646	596	587
⋮	⋮	⋮	⋮
Total operating costs	41,056	38,039	43,226
Operating profit	4,406	498	3,785
Interest expense excluding Financial Products	531	505	507
Other income (expense)	207	146	161
Consolidated profit before taxes	4,082	139	4,439
Provision for income taxes	3,339	192	916
Profit (loss) of consolidated companies	743	(53)	2,523
Equity in profit (loss) of unconsolidated affiliated companies	16	(6)	—
Profit attributable to noncontrolling interests	5	8	11
Profit (loss)	754	(67)	2,512

Exhibit 2 Caterpillar Inc. Consolidated Financial Position (US$ millions)

31 December	2017	2016	2015
Assets			
Current assets:			
Cash and short-term investments	8,261	7,168	6,460
⋮	⋮	⋮	⋮
Inventories	10,018	8,614	9,700
Total current assets	36,244	31,967	33,508
⋮	⋮	⋮	⋮
Total assets	76,962	74,704	78,342
Liabilities			
Total current liabilities	26,931	26,132	26,242
⋮	⋮	⋮	⋮
Total liabilities	63,196	61,491	63,457
Stockholders’ equity			

Exhibit 2 (Continued)

31 December	2017	2016	2015
Common stock of $1.00 par value:			
Authorized shares: 2,000,000,000			
Issued shares (2017, 2016 and 2015 – 814,894,624) at paid-in amount	5,593	5,277	5,238
Treasury stock (2017 – 217,268,852 shares; 2016 – 228,408,600 shares and 2015 – 232,572,734 shares) at cost	(17,005)	(17,478)	(17,640)
Profit employed in the business	26,301	27,377	29,246
Accumulated other comprehensive income (loss)	(1,192)	(2,039)	(2,035)
Noncontrolling interests	69	76	76
Total stockholders' equity	13,766	13,213	14,885
Total liabilities and stockholders' equity	76,962	74,704	78,342

Exhibit 3 Caterpillar Inc. Selected Notes to Consolidated Financial Statements

Note 1. Operations and Summary of Significant Accounting Policies

D. Inventories

Inventories are stated at the lower of cost or net realizable value. Cost is principally determined using the last-in, first-out (LIFO) method. The value of inventories on the LIFO basis represented about 65% of total inventories at December 31, 2017 and about 60% of total inventories at December 31, 2016 and 2015.

If the FIFO (first-in, first-out) method had been in use, inventories would have been $1,924 million, $2,139 million and $2,498 million higher than reported at December 31, 2017, 2016 and 2015, respectively.

Note 7. Inventories

31 December (millions of dollars)	2017	2016	2015
Raw Materials	2,802	2,102	2,467
Work-in-process	2,254	1,719	1,857
Finished goods	4,761	4,576	5,122
Supplies	201	217	254
Total inventories	10,018	8,614	9,700

We had long-term material purchase obligations of approximately $813 million at December 31, 2017.

Solution to 1:

31 December (millions of dollars)	2017	2016	2015
Total inventories (LIFO method)	10,018	8,614	9,700
From Note 1.D (LIFO reserve)	1,924	2,139	2,498
Total inventories (FIFO method)	11,942	10,753	12,198

Note that the decrease in the LIFO reserve from 2015–2016 and again from 2016–2017 likely indicates a LIFO liquidation for both 2016 and 2017.

Solution to 2:

31 December (millions of dollars)	2017	2016
Cost of goods sold (LIFO method)	31,049	28,309
Plus: Decrease in LIFO reserve*	215	359
Cost of goods sold (FIFO method)	31,264	28,668

* From Note 1.D, the decrease in LIFO reserve for 2017 is 215 (1,924 – 2,139) and for 2016 is 359 (2,139 – 2,498).

Solution to 3:

31 December (millions of dollars)	2017	2016
Net income (loss) (LIFO method)	754	–67
Less: Increase in cost of goods sold (decrease in operating profit)	–215	–359
Tax reduction on decreased operating profit*	60	129
Net income (loss) (FIFO method)	599	–297

* The reduction in taxes on the decreased operating profit are 60 (215 × 28%) for 2017 and 129 (359 × 36%) for 2016.

Solution to 4:

The effect on a company's net cash flow from operating activities is limited to the impact of the change on income taxes paid; changes in allocating inventory costs to ending inventory and cost of goods sold does not change any cash flows except income taxes. Consequently, the effect of using FIFO on CAT's net operating cash flow from operating activities would be an increase of $60 million in 2017 and an increase of $129 million in 2016. These are the approximate incremental decreases in income taxes that CAT would have incurred if the FIFO method were used instead of the LIFO method (see solution to 3 above).

Solution to 5:

Using the previously mentioned effective tax rates of 28 percent for 2017 and 36 percent for 2016 (as well as for earlier years), the cumulative amount of income tax savings that CAT has generated by using the LIFO method instead of FIFO is approximately $710 million (–215 × 28% + 2,139 × 36%). Note 1.D indicates a LIFO reserve of $2,139 million at the end of 2016 and a decrease in the LIFO reserve of $215 million in 2017. Therefore, under the FIFO method, cumulative gross profits would have been $2,139 million higher as of the end of 2016 and $1,924 million higher as of the end of 2017. The estimated tax savings would be higher (lower) if income tax rates were assumed to be higher (lower).

Solution to 6:

The amount that would be added to CAT's retained earnings is $1,214 million (1,924 – 710) or (–215 × 72% + 2,139 × 64%). This represents the cumulative increase in operating profit due to the decrease in cost of goods sold (LIFO reserve of $1,924 million) less the assumed taxes on that profit ($710 million, see solution to 5 above). Some analysts advocate ignoring the tax consequences and suggest simply adjusting inventory and equity by the same amount. They argue that the reported equity of the firm is understated by the difference between the current value of its inventory (approximated by the value under FIFO) and its carrying value (value under LIFO).

Solution to 7:

Under the FIFO method, an additional $710 million is assumed to have been incurred for tax expenses. If CAT switched to FIFO, it would have an additional tax liability of $710 million as a consequence of the restatement of financial statements to the FIFO method. This illustrates the significant immediate income tax liabilities that may arise in the year of transition from the LIFO method to the FIFO method. If CAT switched to FIFO for tax purposes, there would be a cash outflow of $710 million for the additional taxes. However, because the company is not actually converting at this point for either tax or reporting purposes, it is appropriate to reflect a deferred tax liability rather than a reduction in cash. In this case for analysis purposes, under FIFO, inventory would increase by $1,924 million, equity by $1,214 million, and non-current liabilities by $710 million.

Solution to 8:

CAT's ratios for 2017 under the LIFO and FIFO methods are as follows:

	LIFO	FIFO
Inventory turnover	3.33	2.76
Days of inventory on hand	109.6 days	132.2 days
Gross profit margin	27.24%	26.74%
Net profit margin	1.66%	1.32%
Return on assets	0.99%	0.77%
Current ratio	1.35	1.42
Total liabilities-to-equity ratio	4.59	4.27

Inventory turnover ratio = Cost of goods sold ÷ Average inventory

LIFO = 3.33 = 31,049 ÷ [(10,018 + 8,614) ÷ 2]

FIFO = 2.76 = 31,264 ÷ [(11,942 + 10,753) ÷ 2]

The ratio is higher under LIFO because, given rising inventory costs, cost of goods sold will be higher and inventory carrying amounts will be lower under LIFO. If an analyst made no adjustment for the difference in inventory methods, it might appear that a company using the LIFO method manages its inventory more effectively.

Days of inventory on hand = Number of days in period ÷ Inventory turnover ratio

LIFO = 109.6 days = (365 days ÷ 3.33)

FIFO = 132.2 days = (365 days ÷ 2.76)

Without adjustment, a company using the LIFO method might appear to manage its inventory more effectively. This is primarily the result of the lower inventory carrying amounts under LIFO.

Gross profit margin = Gross profit ÷ Total revenue

LIFO = 27.24 percent = [(42,676 – 31,049) ÷ 42,676]

FIFO = 26.74 percent = [(42,676 – 31,264) ÷ 42,676]

Revenue of financial products is excluded from the calculation of gross profit. Gross profit is sales of machinery and engines less cost of goods sold. The gross profit margin is lower under FIFO because the cost of goods sold is higher from the LIFO reserve reduction.

Net profit margin = Net income ÷ Total revenue

LIFO = 1.66 percent = (754 ÷ 45,462)

FIFO = 1.32 percent = (599 ÷ 45,462]

The net profit margin is higher under LIFO because the cost of goods sold is lower due to the LIFO liquidation. The absolute percentage difference is less than that of the gross profit margin because of lower income taxes on the decreased income reported under FIFO and because net income is divided by total revenue including sales of machinery and engines and revenue of financial products. The company appears to be more profitable under LIFO.

Return on assets = Net income ÷ Average total assets

LIFO = 0.99 percent = 754 ÷ [(76,962 + 74,704) ÷ 2]

FIFO = 0.77 percent = 599 ÷ [(76,962 + 1,924) + (74,704 + 2,139) ÷ 2]

The total assets under FIFO are the LIFO total assets increased by the LIFO reserve. The return on assets is lower under FIFO because the of the lower net income due to the higher cost of goods sold as well as higher total assets due to the LIFO reserve adjustment. The company appears to be less profitable under FIFO.

Current ratio = Current assets ÷ Current liabilities

LIFO = 1.35 = (36,244 ÷ 26,931)

FIFO = 1.42 = [(36,244 + 1,924) ÷ 26,931]

The current ratio is lower under LIFO primarily because of lower inventory carrying amount. The company appears to be less liquid under LIFO.

Total liabilities-to-equity ratio = Total liabilities ÷ Total shareholders' equity

LIFO = 4.59 = (63,196 ÷ 13,766)

FIFO = 4.27 = [(63,196 + 710) ÷ (13,766 + 1,214)]

The ratio is higher under LIFO because the addition to retained earnings under FIFO reduces the ratio. The company appears to be more highly leveraged under LIFO.

In summary, the company appears to be more profitable, less liquid, and more highly leveraged under LIFO. Yet, because a company's value is based on the present value of future cash flows, LIFO will increase the company's value because the cash flows are higher in earlier years due to lower taxes. LIFO is primarily used for the tax benefits it provides.

EXAMPLE 6

LIFO Liquidation Illustration

Reliable Fans, Inc. (RF), a hypothetical company, sells high quality fans and has been in business since 2015. Exhibit 4 provides relevant data and financial statement information about RF's inventory purchases and sales of fan inventory for the years 2015 through 2018. RF uses the LIFO method and a periodic inventory system. What amount of RF's 2018 gross profit is due to LIFO liquidation?

Exhibit 4 RF Financial Statement Information under LIFO

	2015	2016	2017	2018
Fans units purchased	12,000	12,000	12,000	12,000
Purchase cost per fan	$100	$105	$110	$115
Fans units sold	10,000	12,000	12,000	13,000
Sales price per fan	$200	$205	$210	$215
LIFO Method				
Beginning inventory	$0	$200,000	$200,000	$200,000
Purchases	1,200,000	1,260,000	1,320,000	1,380,000
Goods available for sale	1,200,000	1,460,000	1,520,000	1,580,000
Ending inventory*	(200,000)	(200,000)	(200,000)	(100,000)
Cost of goods sold	$1,000,000	1,260,000	$1,320,000	$1,480,000
Income Statement				
Sales	$2,000,000	$2,460,000	$2,520,000	$2,795,000
Cost of goods sold	1,000,000	1,260,000	1,320,000	1,480,000
Gross profit	$1,000,000	$1,200,000	$1,200,000	$1,315,000
Balance Sheet				
Inventory	$200,000	$200,000	$200,000	$100,000

* Ending inventory 2015, 2016, and 2017 = (2,000 × $100); Ending inventory 2018 = (1,000 × $100).

Solution:

RF's reported gross profit for 2018 is $1,315,000. RF's 2018 gross profit due to LIFO liquidation is $15,000. If RF had purchased 13,000 fans in 2018 rather than 12,000 fans, the cost of goods sold under the LIFO method would have been $1,495,000 (13,000 fans sold at $115.00 purchase cost per fan), and the reported gross profit would have been $1,300,000 ($2,795,000 less $1,495,000). The gross profit due to LIFO liquidation is $15,000 ($1,315,000 reported gross profit less the $1,300,000 gross profit that would have been reported without the LIFO liquidation). The gross profit due to LIFO liquidation may also be determined by multiplying the number of units liquidated times the difference between the replacement cost of the units liquidated and their historical purchase cost. For RF, 1,000 units times $15 ($115 replacement cost per fan less the $100 historical cost per fan) equals the $15,000 gross profit due to LIFO liquidation.

9 INVENTORY METHOD CHANGES

b describe different inventory valuation methods (cost formulas)

f demonstrate the conversion of a company's reported financial statements from LIFO to FIFO for purposes of comparison

Companies on rare occasion change inventory valuation methods. Under IFRS, a change in method is acceptable only if the change "results in the financial statements providing reliable and more relevant information about the effects of transactions, other events, or conditions on the business entity's financial position, financial performance, or cash flows."[11] If the change is justifiable, then it is applied retrospectively.

This means that the change is applied to comparative information for prior periods as far back as is practicable. The cumulative amount of the adjustments relating to periods prior to those presented in the current financial statements is made to the opening balance of each affected component of equity (i.e., retained earnings or comprehensive income) of the earliest period presented. For example, if a company changes its inventory method in 2018 and it presents three years of comparative financial statements (2016, 2017, and 2018) in its annual report, it would retrospectively reflect this change as far back as possible. The change would be reflected in the three years of financial statements presented; the financial statements for 2016 and 2017 would be restated as if the new method had been used in these periods, and the cumulative effect of the change on periods prior to 2016 would be reflected in the 2016 opening balance of each affected component of equity. An exemption to the restatement applies when it is impracticable to determine either the period-specific effects or the cumulative effect of the change.

Under US GAAP, the conditions to make a change in accounting policy and the accounting for a change in inventory policy are similar to IFRS.[12] US GAAP, however, requires companies to thoroughly explain why the newly adopted inventory accounting method is superior and preferable to the old method. If a company decides to change from LIFO to another inventory method, US GAAP requires a retrospective restatement as described above. However, if a company decides to change to the LIFO method, it must do so on a prospective basis and retrospective adjustments are not made to the financial statements. The carrying amount of inventory under the old method becomes the initial LIFO layer in the year of LIFO adoption.

Analysts should carefully evaluate changes in inventory valuation methods. Although the stated reason for the inventory change may be to better match inventory costs with sales revenue (or some other plausible business explanation), the real underlying (and unstated) purpose may be to reduce income tax expense (if changing to LIFO from FIFO or average cost), or to increase reported profits (if changing from LIFO to FIFO or average cost). As always, the choice of inventory valuation method can have a significant impact on financial statements and the financial ratios that are derived from them. As a consequence, analysts must carefully consider the impact of the change in inventory valuation methods and the differences in inventory valuation methods when comparing a company's performance with that of its industry or its competitors.

---

[11] IAS 8 [Accounting Policies, Changes in Accounting Estimates and Errors].

[12] FASB ASC Topic 250 [Accounting Changes and Error Corrections].

INVENTORY ADJUSTMENTS

10

g describe the measurement of inventory at the lower of cost and net realisable value

h describe implications of valuing inventory at net realisable value for financial statements and ratios

Significant financial risk can result from the holding of inventory. The cost of inventory may not be recoverable due to spoilage, obsolescence, or declines in selling prices. IFRS state that inventories shall be measured (and carried on the balance sheet) at the lower of cost and net realisable value.[13] **Net realisable value** is the estimated selling price in the ordinary course of business less the estimated costs necessary to make the sale and estimated costs to get the inventory in condition for sale. The assessment of net realisable value is typically done item by item or by groups of similar or related items. In the event that the value of inventory declines below the carrying amount on the balance sheet, the inventory carrying amount must be written down to its net realisable value[14] and the loss (reduction in value) recognised as an expense on the income statement. This expense may be included as part of cost of sales or reported separately.

In each subsequent period, a new assessment of net realisable value is made. Reversal (limited to the amount of the original write-down) is required for a subsequent increase in value of inventory previously written down. The reversal of any write-down of inventories is recognised as a reduction in cost of sales (reduction in the amount of inventories recognised as an expense).

US GAAP used to specify the lower of cost or market to value inventories.[15] For fiscal years beginning after December 15, 2016, inventories measured using other than LIFO and retail inventory methods are measured at the lower of cost or net realisable value. This is broadly consistent with IFRS with one major difference: US GAAP prohibit the reversal of write-downs. For inventories measured using LIFO and retail inventory methods, market value is defined as current replacement cost subject to upper and lower limits. Market value cannot exceed net realisable value (selling price less reasonably estimated costs of completion and disposal). The lower limit of market value is net realisable value less a normal profit margin. Any write-down to market value or net realisable value reduces the value of the inventory, and the loss in value (expense) is generally reflected in the income statement in cost of goods sold.

An inventory write-down reduces both profit and the carrying amount of inventory on the balance sheet and thus has a negative effect on profitability, liquidity, and solvency ratios. However, activity ratios (for example, inventory turnover and total asset turnover) will be positively affected by a write-down because the asset base (denominator) is reduced. The negative impact on some key ratios, due to the decrease in profit, may result in the reluctance by some companies to record inventory write-downs unless there is strong evidence that the decline in the value of inventory is permanent. This is especially true under US GAAP where reversal of a write-down is prohibited.

---

[13] IAS 2 paragraphs 28–33 [Inventories – Net realisable value].

[14] Frequently, rather than writing inventory down directly, an inventory valuation allowance account is used. The allowance account is netted with the inventory accounts to arrive at the carrying amount that appears on the balance sheet.

[15] FASB ASC Section 330-10-35 [Inventory – Overall – Subsequent Measurement].

IAS 2 [Inventories] does not apply to the inventories of producers of agricultural and forest products and minerals and mineral products, nor to commodity broker–traders. These inventories may be measured at net realisable value (fair value less costs to sell and complete) according to well-established industry practices. If an active market exists for these products, the quoted market price in that market is the appropriate basis for determining the fair value of that asset. If an active market does not exist, a company may use market determined prices or values (such as the most recent market transaction price) when available for determining fair value. Changes in the value of inventory (increase or decrease) are recognised in profit or loss in the period of the change. US GAAP is similar to IFRS in its treatment of inventories of agricultural and forest products and mineral ores. Mark-to-market inventory accounting is allowed for bullion.

EXAMPLE 7

Accounting for Declines and Recoveries of Inventory Value

Hatsumei Enterprises, a hypothetical company, manufactures computers and prepares its financial statements in accordance with IFRS. In 2017, the cost of ending inventory was €5.2 million but its net realisable value was €4.9 million. The current replacement cost of the inventory is €4.7 million. This figure exceeds the net realisable value less a normal profit margin. In 2018, the net realisable value of Hatsumei's inventory was €0.5 million greater than the carrying amount.

1. What was the effect of the write-down on Hatsumei's 2017 financial statements? What was the effect of the recovery on Hatsumei's 2018 financial statements?
2. Under US GAAP, if Hatsumei used the LIFO method, what would be the effects of the write-down on Hatsumei's 2017 financial statements and of the recovery on Hatsumei's 2018 financial statements?
3. What would be the effect of the recovery on Hatsumei's 2018 financial statements if Hatsumei's inventory were agricultural products instead of computers?

Solution to 1:

For 2017, Hatsumei would write its inventory down to €4.9 million and record the change in value of €0.3 million as an expense on the income statement. For 2018, Hatsumei would increase the carrying amount of its inventory and reduce the cost of sales by €0.3 million (the recovery is limited to the amount of the original write-down).

Solution to 2:

Under US GAAP, for 2017, Hatsumei would write its inventory down to €4.7 million and typically include the change in value of €0.5 million in cost of goods sold on the income statement. For 2018, Hatsumei would not reverse the write-down.

Solution to 3:

If Hatsumei's inventory were agricultural products instead of computers, inventory would be measured at net realisable value and Hatsumei would, therefore, increase inventory by and record a gain of €0.5 million for 2018.

Analysts should consider the possibility of an inventory write-down because the impact on a company's financial ratios may be substantial. The potential for inventory write-downs can be high for companies in industries where technological obsolescence of inventories is a significant risk. Analysts should carefully evaluate prospective inventory impairments (as well as other potential asset impairments) and their potential effects on the financial ratios when debt covenants include financial ratio requirements. The breaching of debt covenants can have a significant impact on a company.

Companies that use specific identification, weighted average cost, or FIFO methods are more likely to incur inventory write-downs than companies that use the LIFO method. Under the LIFO method, the *oldest* costs are reflected in the inventory carrying amount on the balance sheet. Given increasing inventory costs, the inventory carrying amounts under the LIFO method are already conservatively presented at the oldest and lowest costs. Thus, it is far less likely that inventory write-downs will occur under LIFO—and if a write-down does occur, it is likely to be of a lesser magnitude.

EXAMPLE 8

Effect of Inventory Write-downs on Financial Ratios

The Volvo Group, based in Göteborg, Sweden, is a leading supplier of commercial transport products such as construction equipment, trucks, busses, and drive systems for marine and industrial applications as well as aircraft engine components.[16] Excerpts from Volvo's consolidated financial statements are shown in Exhibits 5 and 6. Notes pertaining to Volvo's inventories are presented in Exhibit 7.

1. What inventory values would Volvo have reported for 2017, 2016, and 2015 if it had no allowance for inventory obsolescence?
2. Assuming that any changes to the allowance for inventory obsolescence are reflected in the cost of sales, what amount would Volvo's cost of sales be for 2017 and 2016 if it had not recorded inventory write-downs in 2017 and 2016?
3. What amount would Volvo's profit (net income) be for 2017 and 2016 if it had not recorded inventory write-downs in 2017 and 2016? Volvo's effective income tax rate was reported as 25 percent for 2017 and 31 percent for 2016.
4. What would Volvo's 2017 profit (net income) have been if it had reversed all past inventory write-downs in 2017? This question is independent of 1, 2, and 3. The effective income tax rate was 25 percent for 2017.
5. Compare the following for 2017 based on the numbers as reported and those assuming no allowance for inventory obsolescence as in questions 1, 2, and 3: inventory turnover ratio, days of inventory on hand, gross profit margin, and net profit margin.

---

[16] The Volvo line of automobiles has not been under the control and management of the Volvo Group since 1999.

6 CAT (Example 5) has no disclosures indicative of either inventory write-downs or a cumulative allowance for inventory obsolescence in its 2017 financial statements. Provide a conceptual explanation as to why Volvo incurred inventory write-downs for 2017 but CAT did not.

Exhibit 5 Volvo Group Consolidated Income Statements (Swedish Krona in millions, except per share data)

For the years ended 31 December	2017	2016	2015
Net sales	334,748	301,914	312,515
Cost of sales	(254,581)	(231,602)	(240,653)
Gross income	80,167	70,312	71,862
⋮	⋮	⋮	⋮
Operating income	30,327	20,826	23,318
Interest income and similar credits	164	240	257
Income expenses and similar charges	(1,852)	(1,847)	(2,366)
Other financial income and expenses	(386)	11	(792)
Income after financial items	28,254	19,230	20,418
Income taxes	(6,971)	(6,008)	(5,320)
Income for the period	21,283	13,223	15,099
Attributable to:			
Equity holders of the parent company	20,981	13,147	15,058
Minority interests	302	76	41
Profit	21,283	13,223	15,099

Exhibit 6 Volvo Group Consolidated Balance Sheets (Swedish Krona in millions)

31 December	2017	2016	2015
Assets			
Total non-current assets	213,455	218,465	203,478
Current assets:			
Inventories	52,701	48,287	44,390
⋮	⋮	⋮	⋮
Cash and cash equivalents	36,092	23,949	21,048
Total current assets	199,039	180,301	170,687
Total assets	412,494	398,916	374,165
Shareholders' equity and liabilities			
Equity attributable to equity holders of the parent company	107,069	96,061	83,810
Minority interests	1,941	1,703	1,801
Total shareholders' equity	109,011	97,764	85,610
Total non-current provisions	29,147	29,744	26,704

Exhibit 6 (Continued)

31 December	2017	2016	2015
Total non-current liabilities	96,213	104,873	91,814
Total current provisions	10,806	11,333	14,176
Total current liabilities	167,317	155,202	155,860
Total shareholders' equity and liabilities	412,404	398,916	374,165

Exhibit 7 Volvo Group Selected Notes to Consolidated Financial Statements

Note 17. Inventories

Accounting Policy

Inventories are reported at the lower of cost and net realisable value. The cost is established using the first-in, first-out principle (FIFO) and is based on the standard cost method, including costs for all direct manufacturing expenses and the attributable share of capacity and other related manufacturing-related costs. The standard costs are tested regularly and adjustments are made based on current conditions. Costs for research and development, selling, administration and financial expenses are not included. Net realisable value is calculated as the selling price less costs attributable to the sale.

Sources of Estimation Uncertainty

Inventory obsolescence

If the net realisable value is lower than cost, a valuation allowance is established for inventory obsolescence. The total inventory value, net of inventory obsolescence allowance, was: SEK (in millions) 52,701 as of December 2017 and 48,287 as of 31 December 2016.

Inventory

31 December (millions of Krona)	2017	2016	2015
Finished products	32,304	31,012	27,496
Production materials, etc.	20,397	17,275	16,894
Total	**52,701**	**48,287**	**44,390**

Increase (decrease) in allowance for inventory obsolescence

31 December (millions of Krona)	2017	2016	2015
Opening balance	3,683	3,624	3,394
Change in allowance for inventory obsolescence charged to income	304	480	675
Scrapping	(391)	(576)	(435)
Translation differences	(116)	177	(29)

(continued)

31 December (millions of Krona)	2017	2016	2015
Reclassifications, etc.	8	(23)	20
Allowance for inventory obsolescence as of 31 December	3,489	3,683	3,624

Solution to 1:

31 December (Swedish krona in millions)	2017	2016	2015
Total inventories, net	52,701	48,287	44,390
From Note 17. (Allowance for obsolescence)	3,489	3,683	3,624
Total inventories (without allowance)	56,190	51,970	48,014

Solution to 2:

31 December (Swedish krona in millions)	2017	2016
Cost of sales	254,581	231,602
(Increase) decrease in allowance for obsolescence*	194	(59)
Cost of sales without allowance	254,775	231,543

* From Note 17, the decrease in allowance for obsolescence for 2017 is 194 (3,489 – 3,683) and the increase for 2016 is 59 (3,683 – 3,624).

Solution to 3:

31 December (Swedish krona in millions)	2017	2016
Profit (Net income)	21,283	13,223
Increase (reduction) in cost of sales	(194)	59
Taxes (tax reduction) on operating profit*	49	(18)
Profit (without allowance)	21,138	13,264

* Taxes (tax reductions) on the operating profit are assumed to be 49 (194 × 25%) for 2017 and −18 (−59 × 31%) for 2016.

Solution to 4:

31 December (Swedish krona in millions)	2017
Profit (Net income)	21,283
Reduction in cost of sales (increase in operating profit)	3,489

31 December (Swedish krona in millions)	2017
Taxes on increased operating profit*	–872
Profit (after recovery of previous write-downs)	23,900

* Taxes on the increased operating profit are assumed to be 872 (3,489 × 25%) for 2017.

Solution to 5:

The Volvo Group's financial ratios for 2017 with the allowance for inventory obsolescence and without the allowance for inventory obsolescence are as follows:

	With Allowance (As Reported)	Without Allowance (Adjusted)
Inventory turnover ratio	5.04	4.71
Days of inventory on hand	72.4	77.5
Gross profit margin	23.95%	23.89%
Net profit margin	6.36%	6.31%

Inventory turnover ratio = Cost of sales ÷ Average inventory

With allowance (as reported) = 5.04 = 254,581 ÷ [(52,701 + 48,287) ÷ 2]

Without allowance (adjusted) = 4.71 = 254,775 ÷ [(56,190 + 51,970) ÷ 2]

Inventory turnover is higher based on the numbers as reported because inventory carrying amounts will be lower with an allowance for inventory obsolescence. The company might appear to manage its inventory more efficiently when it has inventory write-downs.

Days of inventory on hand = Number of days in period ÷ Inventory turnover ratio

With allowance (as reported) = 72.4 days = (365 days ÷ 5.04)

Without allowance (adjusted) = 77.5 days = (365 days ÷ 4.71)

Days of inventory on hand are lower based on the numbers as reported because the inventory turnover is higher. A company with inventory write-downs might appear to manage its inventory more effectively. This is primarily the result of the lower inventory carrying amounts.

Gross profit margin = Gross income ÷ Net sales

With allowance (as reported) = 23.95 percent = (80,167 ÷ 334,748)

Without allowance (adjusted) = 23.89 percent = [(80,167 – 194) ÷ 334,748]

In this instance, the gross profit margin is slightly higher with inventory write-downs because the cost of sales is lower (due to the reduction in the allowance for inventory obsolescence). This assumes that inventory write-downs (and inventory write-down recoveries) are reported as part of cost of sales.

Net profit margin = Profit ÷ Net sales

With allowance (as reported) = 6.36 percent = (21,283 ÷ 334,748)

Without allowance (adjusted) = 6.31 percent = (21,138 ÷ 334,748)

In this instance, the net profit margin is higher with inventory write-downs because the cost of sales is lower (due to the reduction in the allowance for inventory obsolescence). The absolute percentage difference is less than that of the gross profit margin because of the income tax reduction on the decreased income without write-downs.

The profitability ratios (gross profit margin and net profit margin) for Volvo Group would have been slightly lower for 2017 if the company had not recorded inventory write-downs. The activity ratio (inventory turnover ratio) would appear less attractive without the write-downs. The inventory turnover ratio is slightly better (higher) with inventory write-downs because inventory write-downs decrease the average inventory (denominator), making inventory management appear more efficient with write-downs.

Solution to 6:

CAT uses the LIFO method whereas Volvo uses the FIFO method. Given increasing inventory costs, companies that use the FIFO inventory method are far more likely to incur inventory write-downs than those companies that use the LIFO method. This is because under the LIFO method, the inventory carrying amounts reflect the *oldest* costs and therefore the *lowest* costs given increasing inventory costs. Because inventory carrying amounts under the LIFO method are already conservatively presented, it is less likely that inventory write-downs will occur.

11 EVALUATION OF INVENTORY MANAGEMENT: DISCLOSURES & RATIOS

i. describe the financial statement presentation of and disclosures relating to inventories

j. explain issues that analysts should consider when examining a company's inventory disclosures and other sources of information

The choice of inventory valuation method impacts the financial statements. The financial statement items impacted include cost of sales, gross profit, net income, inventories, current assets, and total assets. Therefore, the choice of inventory valuation method also affects financial ratios that contain these items. Ratios such as current ratio, return on assets, gross profit margin, and inventory turnover are impacted. As a consequence, analysts must carefully consider inventory valuation method differences when evaluating a company's performance over time or when comparing its performance with the performance of the industry or industry competitors. Additionally, the financial statement items and ratios may be impacted by adjustments of inventory carrying amounts to net realisable value or current replacement cost.

11.1 Presentation and Disclosure

Disclosures are useful when analyzing a company. IFRS require the following financial statement disclosures concerning inventory:

a the accounting policies adopted in measuring inventories, including the cost formula (inventory valuation method) used;

b the total carrying amount of inventories and the carrying amount in classifications (for example, merchandise, raw materials, production supplies, work in progress, and finished goods) appropriate to the entity;

c the carrying amount of inventories carried at fair value less costs to sell;

d the amount of inventories recognised as an expense during the period (cost of sales);

e the amount of any write-down of inventories recognised as an expense in the period;

f the amount of any reversal of any write-down that is recognised as a reduction in cost of sales in the period;

g the circumstances or events that led to the reversal of a write-down of inventories; and

h the carrying amount of inventories pledged as security for liabilities.

Inventory-related disclosures under US GAAP are very similar to the disclosures above, except that requirements (f) and (g) are not relevant because US GAAP do not permit the reversal of prior-year inventory write-downs. US GAAP also require disclosure of significant estimates applicable to inventories and of any material amount of income resulting from the liquidation of LIFO inventory.

11.2 Inventory Ratios

Three ratios often used to evaluate the efficiency and effectiveness of inventory management are **inventory turnover**, **days of inventory on hand**, and **gross profit margin**.[17] These ratios are directly impacted by a company's choice of inventory valuation method. Analysts should be aware, however, that many other ratios are also affected by the choice of inventory valuation method, although less directly. These include the current ratio, because inventory is a component of current assets; the return-on-assets ratio, because cost of sales is a key component in deriving net income and inventory is a component of total assets; and even the debt-to-equity ratio, because the cumulative measured net income from the inception of a business is an aggregate component of retained earnings.

The inventory turnover ratio measures the number of times during the year a company sells (i.e., turns over) its inventory. The higher the turnover ratio, the more times that inventory is sold during the year and the lower the relative investment of resources in inventory. Days of inventory on hand can be calculated as days in the period divided by inventory turnover. Thus, inventory turnover and days of inventory on hand are inversely related. It may be that inventory turnover, however, is calculated using average inventory in the year whereas days of inventory on hand is based on the ending inventory amount. In general, inventory turnover and the number of days of inventory on hand should be benchmarked against industry norms and compared across years.

A high inventory turnover ratio and a low number of days of inventory on hand might indicate highly effective inventory management. Alternatively, a high inventory ratio and a low number of days of inventory on hand could indicate that the company does not carry an adequate amount of inventory or that the company has written down inventory values. Inventory shortages could potentially result in lost sales or production problems in the case of the raw materials inventory of a manufacturer. To assess which explanation is more likely, analysts can compare the company's inventory turnover and sales growth rate with those of the industry and review financial statement disclosures. Slower growth combined with higher inventory turnover could indicate inadequate inventory levels. Write-downs of inventory could reflect poor inventory management. Minimal write-downs and sales growth rates at or above the industry+'s growth rates would support the interpretation that the higher turnover reflects greater efficiency in managing inventory.

---

[17] *Days of inventory on hand* is also referred to as *days in inventory* and *average inventory days outstanding*.

A low inventory turnover ratio and a high number of days of inventory on hand relative to industry norms could be an indicator of slow-moving or obsolete inventory. Again, comparing the company's sales growth across years and with the industry and reviewing financial statement disclosures can provide additional insight.

The gross profit margin, the ratio of gross profit to sales, indicates the percentage of sales being contributed to net income as opposed to covering the cost of sales. Firms in highly competitive industries generally have lower gross profit margins than firms in industries with fewer competitors. A company's gross profit margin may be a function of its type of product. A company selling luxury products will generally have higher gross profit margins than a company selling staple products. The inventory turnover of the company selling luxury products, however, is likely to be much lower than the inventory turnover of the company selling staple products.

12

ILLUSTRATIONS OF INVENTORY ANALYSIS: ADJUSTING LIFO TO FIFO

k calculate and compare ratios of companies, including companies that use different inventory methods

l analyze and compare the financial statements of companies, including companies that use different inventory methods

IFRS and US GAAP require companies to disclose, either on the balance sheet or in the notes to the financial statements, the carrying amounts of inventories in classifications suitable to the company. For manufacturing companies, these classifications might include production supplies, raw materials, work in progress, and finished goods. For a retailer, these classifications might include significant categories of merchandise or the grouping of inventories with similar attributes. These disclosures may provide signals about a company's future sales and profits.

For example, a significant increase (attributable to increases in unit volume rather than increases in unit cost) in raw materials and/or work-in-progress inventories may signal that the company expects an increase in demand for its products. This suggests an anticipated increase in sales and profit. However, a substantial increase in finished goods inventories while raw materials and work-in-progress inventories are declining may signal a decrease in demand for the company's products and hence lower future sales and profit. This may also signal a potential future write down of finished goods inventory. Irrespective of the signal, an analyst should thoroughly investigate the underlying reasons for any significant changes in a company's raw materials, work-in-progress, and finished goods inventories.

Analysts also should compare the growth rate of a company's sales to the growth rate of its finished goods inventories, because this could also provide a signal about future sales and profits. For example, if the growth of inventories is greater than the growth of sales, this could indicate a decline in demand and a decrease in future earnings. The company may have to lower (mark down) the selling price of its products to reduce its inventory balances, or it may have to write down the value of its inventory because of obsolescence, both of which would negatively affect profits. Besides the potential for mark-downs or write-downs, having too much inventory on hand or the wrong type of inventory can have a negative financial effect on a company because it increases inventory related expenses such as insurance, storage costs, and taxes. In addition, it means that the company has less cash and working capital available to use for other purposes.

Inventory write-downs may have a substantial impact on a company's activity, profitability, liquidity, and solvency ratios. It is critical for the analyst to be aware of industry trends toward product obsolescence and to analyze the financial ratios for their sensitivity to potential inventory impairment. Companies can minimise the impact of inventory write-downs by better matching their inventory composition and growth with prospective customer demand. To obtain additional information about a company's inventory and its future sales, a variety of sources of information are available. Analysts should consider the Management Discussion and Analysis (MD&A) or similar sections of the company's financial reports, industry related news and publications, and industry economic data.

When conducting comparisons, differences in the choice of inventory valuation method can significantly affect the comparability of financial ratios between companies. A restatement from the LIFO method to the FIFO method is critical to make a valid comparison with companies using a method other than the LIFO method such as those companies reporting under IFRS. Analysts should seek out as much information as feasible when analyzing the performance of companies.

EXAMPLE 9

Comparative Illustration

1 Using CAT's LIFO numbers as reported and FIFO adjusted numbers (Example 5) and Volvo's numbers as reported (Example 8), compare the following for 2017: inventory turnover ratio, days of inventory on hand, gross profit margin, net profit margin, return on assets, current ratio, total liabilities-to-equity ratio, and return on equity. For the current ratio, include current provisions as part of current liabilities. For the total liabilities-to-equity ratio, include provisions in total liabilities.

2 How much do inventories represent as a component of total assets for CAT using LIFO numbers as reported and FIFO adjusted numbers, and for Volvo using reported numbers in 2017 and 2016? Discuss any changes that would concern an analyst.

3 Using the reported numbers, compare the 2016 and 2017 growth rates of CAT and Volvo for sales, finished goods inventory, and inventories other than finished goods.

Solution to 1:

The comparisons between Caterpillar and Volvo for 2017 are as follows:

	CAT (LIFO)	CAT (FIFO)	Volvo
Inventory turnover ratio	3.33	2.76	5.04
Days of inventory on hand	109.6 days	132.2 days	72.4 days
Gross profit margin	27.24%	26.74%	23.95%
Net profit margin	1.66%	1.32%	6.36%
Return on assets[a]	0.99%	0.77%	5.25%
Current ratio[b]	1.35	1.42	1.12

(continued)

	CAT (LIFO)	CAT (FIFO)	Volvo
Total liabilities-to-equity ratio[c]	4.59	4.27	2.78
Return on equity[d]	5.59%	4.05%	20.59%

Calculations for ratios previously calculated (see Examples 5 and 8) are not shown again.

[a] Return on assets = Net income ÷ Average total assets
Volvo = 5.25 percent = 21,283 ÷ [(412,494 + 398,916) ÷ 2]

[b] Current ratio = Current assets ÷ Current liabilities
Volvo = 1.12 = [199,039 ÷ (10,806 + 167,317)]
The question indicates to include current provisions in current liabilities.

[c] Total liabilities-to-equity ratio = Total liabilities ÷ Total shareholders' equity
Volvo = 2.78 = [(29,147 + 96,213 + 10,806 + 167,317) ÷ 109,011]
The question indicates to include provisions in total liabilities.

[d] Return on equity = Net income ÷ Average shareholders' equity
CAT (LIFO) = 5.59 percent = 754 ÷ [(13,766 + 13,213) ÷ 2]
CAT (FIFO) = 4.05 percent = 599 ÷ {[(13,766 + 1,924 – 710) + (13,213 + 2,139 – 770)] ÷ 2}
Volvo = 20.59 percent = 21,283 ÷ [(109,011 + 97,764) ÷ 2]

Comparing CAT (FIFO) and Volvo, it appears that Volvo manages its inventory more effectively. It has higher inventory turnover and fewer days of inventory on hand. Volvo appears to have superior profitability based on net profit margin. A primary reason for CAT's low profitability in 2017 was due to a substantial increase in the provision for income taxes. An analyst would likely further investigate CAT's increase in provision for income taxes, as well as other reported numbers, rather than reaching a conclusion based on ratios alone (in other words, try to identify the underlying causes of changes or differences in ratios).

Solution to 2:

The 2017 and 2016 inventory to total assets ratios for CAT using LIFO and adjusted to FIFO and for Volvo as reported, are as follows:

	CAT (LIFO)	CAT (FIFO)	Volvo
2017	13.02%	15.28%	12.78%
2016	11.53%	14.14%	12.10%

Inventory to total assets

CAT (LIFO) 2017 =13.02 percent = 10,018 ÷ 76,962
CAT (LIFO) 2016 = 11.53 percent = 8,614 ÷ 74,704
CAT (FIFO) 2017 = 15.28 percent = 11,942 ÷ (76,962 + 1,924 – 710)
CAT (FIFO) 2016 = 14.14 percent = 10,753 ÷ (74,704 + 2,139 – 770)
Volvo 2017 = 12.78 percent = 52,701 ÷ 412,494
Volvo 2016 = 12.10 percent = 48,287 ÷ 398,916

Based on the numbers as reported, CAT appears to have a similar percentage of assets tied up in inventory as Volvo. However, when CAT's inventory is adjusted to FIFO, it has a higher percentage of its assets tied up in inventory than Volvo.

The increase in inventory as a percentage of total assets is cause for some concern. Higher inventory typically results in higher maintenance costs (for example, storage and financing costs). A build-up of slow moving or obsolete inventories may result in future inventory write-downs. In Volvo's Note 17, the breakdown by inventory classification shows a significant increase in the inventory of production materials. Volvo may be planning on increasing production of more finished goods inventory (which has also increased). Looking at CAT's Note 7, all classifications of inventory seem to be increasing and because these are valued using the LIFO method, there is some cause for concern. The company must be increasing inventory quantities and adding new LIFO layers.

Solution to 3:

CAT's and Volvo's 2017 and 2016 growth rates for sales ("Sales of machinery and engines" for CAT and "Net sales" for Volvo), finished goods, and inventories other than finished goods" are as follows:

2017	CAT	Volvo
Sales	19.3%	10.9%
Finished goods	4.0%	4.2%
Inventories other than finished goods	30.2%	18.1%

2016	CAT	Volvo
Sales	–19.0%	–3.4%
Finished goods	–10.7%	12.8%
Inventories other than finished goods	–11.8%	2.3%

Growth rate = (Value for year – Value for previous year)/Value for previous year

2017 CAT

Sales = 19.3 percent = (42,676 – 35,773) ÷ 35,773

Finished goods = 4.0 percent = (4,761 – 4,576) ÷ 4,576

Inventories other than finished goods = 30.2 percent = [(2,802 + 2,254 + 201) – (2,102 + 1,719 + 217)] ÷ (2,102 + 1,719 + 217)

2017 Volvo

Sales = 10.9 percent = (334,748 – 301,914) ÷ 301,914

Finished products = 4.2 percent = (32,304 – 31,012) ÷ 31,012

Inventories other than finished products =18.1 percent = (20,397 – 17,275) ÷ 17,275

2016 CAT

Sales = –19.0 percent = (35,773 – 44,147) ÷ 44,147

Finished goods = –10.7 percent = (4,576 – 5,122) ÷ 5,122

Inventories other than finished goods = –11.8 percent = [(2,102+ 1,719 + 217) – (2,467 + 1,857 + 254)] ÷ (2,467 + 1,857 + 254)

2016 Volvo

Sales = – 3.4 percent = (301,914 – 312,515) ÷ 312,515

Finished products = 12.8 percent = (31,012 – 27,496) ÷ 27,496

Inventories other than finished products = 2.3 percent = (17,275 – 16,894) ÷ 16,894

For both companies, the growth rates in finished goods inventory exceeds the growth rate in sales; this could be indicative of accumulating excess inventory. Volvo's growth rate in finished goods compared to its growth rate in sales is significantly higher but the lower growth rates in finished goods inventory for

CAT is potentially a result of using the LIFO method versus the FIFO method. It appears Volvo is aware that an issue exists and is planning on cutting back production given the relatively small increase in inventories other than finished products. Regardless, an analyst should do further investigation before reaching any conclusion about a company's future prospects for sales and profit.

13 ILLUSTRATIONS OF INVENTORY ANALYSIS: IMPACTS OF WRITEDOWNS

k calculate and compare ratios of companies, including companies that use different inventory methods

l analyze and compare the financial statements of companies, including companies that use different inventory methods

EXAMPLE 10

Single Company Illustration

Selected excerpts from the consolidated financial statements and notes to consolidated financial statements for Jollof Inc., a hypothetical telecommunications company providing networking and communications solutions, are presented in Exhibits 8, 9, and 10. Exhibit 8 contains excerpts from the consolidated income statements, and Exhibit 9 contains excerpts from the consolidated balance sheets. Exhibit 10 contains excerpts from three of the notes to consolidated financial statements.

Note 1 (a) discloses that Jollof's finished goods inventories and work in progress are valued at the lower of cost or net realisable value. Note 2 (a) discloses that the impact of inventory and work in progress write-downs on Jollof's income before tax was a net reduction of €239 million in 2017, a net reduction of €156 million in 2016, and a net reduction of €65 million in 2015.[18] The inventory impairment loss amounts steadily increased from 2015 to 2017 and are included as a component, (additions)/reversals, of Jollof's change in valuation allowance as disclosed in Note 3 (b) from Exhibit 10. Observe also that Jollof discloses its valuation allowance at 31 December 2017, 2016, and 2015 in Note 3 (b) and details on the allocation of the allowance are included in Note 3 (a). The €549 million valuation allowance is the total of a €528 million allowance for inventories and a €21 million allowance for work in progress on construction contracts. Finally, observe that the €1,845 million net value for inventories (excluding construction contracts) at 31 December 2017 in Note 3 (a) reconciles with the balance sheet amount for inventories and work in progress, net, on 31 December 2017, as presented in Exhibit 9.

The inventory valuation allowance represents the total amount of inventory write-downs taken for the inventory reported on the balance sheet (which is measured at the lower of cost or net realisable value). Therefore, an analyst can determine the historical cost of the company's inventory by adding the inventory valuation allowance to the reported inventory carrying amount on the balance sheet. The valuation allowance increased in magnitude and as a percentage of gross inventory values from 2015 to 2017.

---

[18] This reduction is often referred to as a *charge.* An accounting charge is the recognition of a loss or expense. In this case, the charge is attributable to the impairment of assets.

Exhibit 8 Alcatel-Lucent Consolidated Income Statements (€ millions)

For years ended 31 December	2017	2016	2015
Revenues	14,267	14,945	10,317
Cost of sales	(9,400)	(10,150)	(6,900)
Gross profit	4,867	4,795	3,417
Administrative and selling expenses	(2,598)	(2,908)	(1,605)
Research and development costs	(2,316	(2,481)	(1,235)
Income from operating activities before restructuring costs, impairment of assets, gain/(loss) on disposal of consolidated entities, and post-retirement benefit plan amendments	(47)	(594)	577
Restructuring costs	(472)	(719)	(594)
Impairment of assets	(3,969)	(2,473)	(118)
Gain/(loss) on disposal of consolidated entities	(6)	—	13
Post-retirement benefit plan amendments	39	217	—
Income (loss) from operating activities	(4,455)	(3,569)	(122)
⋮	⋮	⋮	⋮
Income (loss) from continuing operations	(4,373)	(3,433)	(184)
Income (loss) from discontinued operations	28	512	133
Net income (loss)	(4,345)	(2,921)	51

Exhibit 9 Alcatel-Lucent Consolidated Balance Sheets (€ millions)

31 December	2017	2016	2015
⋮	⋮	⋮	⋮
Total non-current assets	10,703	16,913	21,559
Inventories and work in progress, net	1,845	1,877	1,898
Amounts due from customers on construction contracts	416	591	517
Trade receivables and related accounts, net	3,637	3,497	3,257
Advances and progress payments	83	92	73
⋮	⋮	⋮	⋮
Total current assets	12,238	11,504	13,629
Total assets	22,941	28,417	35,188
⋮	⋮	⋮	⋮
Retained earnings, fair value, and other reserves	(7,409)	(3,210)	(2,890)
⋮	⋮	⋮	⋮
Total shareholders' equity	4,388	9,830	13,711
Pensions, retirement indemnities, and other post-retirement benefits	4,038	3,735	4,577
Bonds and notes issued, long-term	3,302	3,794	4,117

(continued)

Exhibit 9 (Continued)

31 December	2017	2016	2015
Other long-term debt	56	40	123
Deferred tax liabilities	968	1,593	2,170
Other non-current liabilities	372	307	232
Total non-current liabilities	8,736	9,471	11,219
Provisions	2,036	2,155	1,987
Current portion of long-term debt	921	406	975
Customers' deposits and advances	780	711	654
Amounts due to customers on construction contracts	158	342	229
Trade payables and related accounts	3,840	3,792	3,383
Liabilities related to disposal groups held for sale	—	—	1,349
Current income tax liabilities	155	59	55
Other current liabilities	1,926	1,651	1,625
Total current liabilities	9,817	9,117	10,257
Total liabilities and shareholders' equity	22,941	28,417	35,188

Exhibit 10 Jollof Inc. Selected Notes to Consolidated Financial Statements

Note 1. Summary of Significant Accounting Policies

(a) Inventories and work in progress

Inventories and work in progress are valued at the lower of cost (including indirect production costs where applicable) or net realizable value.[19] Net realizable value is the estimated sales revenue for a normal period of activity less expected completion and selling costs.

Note 2. Principal uncertainties regarding the use of estimates

(a) Valuation allowance for inventories and work in progress

Inventories and work in progress are measured at the lower of cost or net realizable value. Valuation allowances for inventories and work in progress are calculated based on an analysis of foreseeable changes in demand, technology, or the market, in order to determine obsolete or excess inventories and work in progress.

The valuation allowances are accounted for in cost of sales or in restructuring costs, depending on the nature of the amounts concerned.

---

[19] *Cost* approximates cost on a first-in, first-out basis.

Exhibit 10 (Continued)

(€ millions)	31 December 2017	2016	2015
Valuation allowance for inventories and work in progress on construction contracts	(549)	(432)	318
Impact of inventory and work in progress write-downs on income (loss) before income tax related reduction of goodwill and discounted operations	(239)	(156)	(65)

Note 3. Inventories and work in progress
(a) Analysis of net value

(€ millions)	2017	2016	2015
Raw materials and goods	545	474	455
Work in progress excluding construction contracts	816	805	632
Finished goods	1,011	995	1,109
Gross value (excluding construction contracts)	2,373	2,274	2,196
Valuation allowance	(528)	(396)	(298)
Net value (excluding construction contracts)	1,845	1,877	1,898
Work in progress on construction contracts, gross*	184	228	291
Valuation allowance	(21)	(35)	(19)
Work in progress on construction contracts, net	163	193	272
Total, net	2,008	2,071	2,170

* Included in the amounts due from/to construction contracts.

(b) Change in valuation allowance

(€ millions)	2017	2016	2015
At 1 January	(432)	(318)	(355)
(Additions)/reversals	(239)	(156)	(65)
Utilization	58	32	45
Changes in consolidation group	—	—	45

(continued)

(€ millions)	2017	2016	2015
Net effect of exchange rate changes and other changes	63	10	12
At 31 December	(549)	(432)	(318)

Rounding differences may result in totals that are slightly different from the sum and from corresponding numbers in the note.

1 Calculate Jollof's inventory turnover, number of days of inventory on hand, gross profit margin, current ratio, debt-to-equity ratio, and return on total assets for 2017 and 2016 based on the numbers reported. Use an average for inventory and total asset amounts and year-end numbers for other ratio items. For debt, include only bonds and notes issued, long-term; other long-term debt; and current portion of long-term debt.

2 Based on the answer to Question 1, comment on the changes from 2016 to 2017.

3 If Jollof had used the weighted average cost method instead of the FIFO method during 2017, 2016, and 2015, what would be the effect on Jollof's reported cost of sales and inventory carrying amounts? What would be the directional impact on the financial ratios that were calculated for Jollof in Question 1?

Solution to 1:

The financial ratios are as follows:

	2017	2016
Inventory turnover ratio	5.05	5.38
Number of days of inventory on hand	72.3 days	67.8 days
Gross profit margin	34.1%	32.1%
Current ratio	1.25	1.26
Debt-to-equity ratio	0.98	0.43
Return on total assets	–16.9%	–9.2%

Inventory turnover ratio = Cost of sales ÷ Average inventory

2017 inventory turnover ratio = 5.05 = 9,400 ÷ [(1,845 + 1,877) ÷ 2]

2016 inventory turnover ratio = 5.38 = 10,150 ÷ [(1,877 + 1,898) ÷ 2]

Number of days of inventory = 365 days ÷ Inventory turnover ratio

2017 number of days of inventory = 72.3 days = 365 days ÷ 5.05

2016 number of days of inventory = 67.8 days = 365 days ÷ 5.38

Gross profit margin = Gross profit ÷ Total revenue

2017 gross profit margin = 34.1% = 4,867 ÷ 14,267

2016 gross profit margin = 32.1% = 4,795 ÷ 14,945

Current ratio = Current assets ÷ Current liabilities

2017 current ratio = 1.25 = 12,238 ÷ 9,817

2016 current ratio = 1.26 = 11,504 ÷ 9,117

Debt-to-equity ratio = Total debt ÷ Total shareholders’ equity

2017 debt-to-equity ratio = 0.98 = (3,302 + 56 + 921) ÷ 4,388

2016 debt-to-equity ratio = 0.43 = (3,794 + 40 + 406) ÷ 9,830

Return on assets = Net income ÷ Average total assets

2017 return on assets = −16.9% = −4,345 ÷ [(22,941 + 28,417) ÷ 2]

2016 return on assets = −9.2% = −2,921 ÷ [(28,417 + 35,188) ÷ 2]

Solution to 2:

From 2016 to 2017, the inventory turnover ratio declined and the number of days of inventory increased by 4.5 days. Jollof appears to be managing inventory less efficiently. The gross profit margin improved by 2.0 percent, from 32.1 percent in 2016 to 34.1 percent in 2017. The current ratio is relatively unchanged from 2016 to 2017. The debt-to-equity ratio has risen significantly in 2017 compared to 2016. Although Jollofn’s total debt has been relatively stable during this time period, the company’s equity has been declining rapidly because of the cumulative effect of its net losses on retained earnings.

The return on assets is negative and deteriorated in 2017 compared to 2016. A larger net loss and lower total assets in 2017 resulted in a higher negative return on assets. The analyst should investigate the underlying reasons for the sharp decline in Jollof’s return on assets. From Exhibit 8, it is apparent that Jollof’s gross profit margins were insufficient to cover the administrative and selling expenses and research and development costs in 2016 and 2017. Large restructuring costs and asset impairment losses contributed to the loss from operating activities in both 2016 and 2017.

Solution to 3:

If inventory replacement costs were increasing during 2015, 2016, and 2017 (and inventory quantity levels were stable or increasing), Jollof’s cost of sales would have been higher and its gross profit margin would have been lower under the weighted average cost inventory method than what was reported under the FIFO method (assuming no inventory write-downs that would otherwise neutralize the differences between the inventory valuation methods). FIFO allocates the oldest inventory costs to cost of sales; the reported cost of sales would be lower under FIFO given increasing inventory costs. Inventory carrying amounts would be higher under the FIFO method than under the weighted average cost method because the more recently purchased inventory items would be included in inventory at their higher costs (again assuming no inventory write-downs that would otherwise neutralize the differences between the inventory valuation methods). Consequently, Jollof’s reported gross profit, net income, and retained earnings would also be higher for those years under the FIFO method.

The effects on ratios are as follows:

- The inventory turnover ratios would all be higher under the weighted average cost method because the numerator (cost of sales) would be higher and the denominator (inventory) would be lower than what was reported by Jollof under the FIFO method.

- The number of days of inventory would be lower under the weighted average cost method because the inventory turnover ratios would be higher.
- The gross profit margin ratios would all be lower under the weighted average cost method because cost of sales would be higher under the weighted average cost method than under the FIFO method.
- The current ratios would all be lower under the weighted average cost method because inventory carrying values would be lower than under the FIFO method (current liabilities would be the same under both methods).
- The return-on-assets ratios would all be lower under the weighted average cost method because the incremental profit added to the numerator (net income) has a greater impact than the incremental increase to the denominator (total assets). By way of example, assume that a company has €3 million in net income and €100 million in total assets using the weighted average cost method. If the company reports another €1 million in net income by using FIFO instead of weighted average cost, it would then also report an additional €1 million in total assets (after tax). Based on this example, the return on assets is 3.00 percent (€3/€100) under the weighted average cost method and 3.96 percent (€4/€101) under the FIFO method.
- The debt-to-equity ratios would all be higher under the weighted average cost method because retained earnings would be lower than under the FIFO method (again assuming no inventory write-downs that would otherwise neutralize the differences between the inventory valuation methods).

Conversely, if inventory replacement costs were decreasing during 2015, 2016, and 2017 (and inventory quantity levels were stable or increasing), Jollof's cost of sales would have been lower and its gross profit and inventory would have been higher under the weighted average cost method than were reported under the FIFO method (assuming no inventory write-downs that would otherwise neutralize the differences between the inventory valuation methods). As a result, the ratio assessment that was performed above would result in directly opposite conclusions.

SUMMARY

The choice of inventory valuation method (cost formula or cost flow assumption) can have a potentially significant impact on inventory carrying amounts and cost of sales. These in turn impact other financial statement items, such as current assets, total assets, gross profit, and net income. The financial statements and accompanying notes provide important information about a company's inventory accounting policies that the analyst needs to correctly assess financial performance and compare it with that of other companies. Key concepts in this reading are as follows:

- Inventories are a major factor in the analysis of merchandising and manufacturing companies. Such companies generate their sales and profits through inventory transactions on a regular basis. An important consideration in determining profits for these companies is measuring the cost of sales when inventories are sold.

- The total cost of inventories comprises all costs of purchase, costs of conversion, and other costs incurred in bringing the inventories to their present location and condition. Storage costs of finished inventory and abnormal costs due to waste are typically treated as expenses in the period in which they occurred.
- The allowable inventory valuation methods implicitly involve different assumptions about cost flows. The choice of inventory valuation method determines how the cost of goods available for sale during the period is allocated between inventory and cost of sales.
- IFRS allow three inventory valuation methods (cost formulas): first-in, first-out (FIFO); weighted average cost; and specific identification. The specific identification method is used for inventories of items that are not ordinarily interchangeable and for goods or services produced and segregated for specific projects. US GAAP allow the three methods above plus the last-in, first-out (LIFO) method. The LIFO method is widely used in the United States for both tax and financial reporting purposes because of potential income tax savings.
- The choice of inventory method affects the financial statements and any financial ratios that are based on them. As a consequence, the analyst must carefully consider inventory valuation method differences when evaluating a company's performance over time or in comparison to industry data or industry competitors.
- A company must use the same cost formula for all inventories having a similar nature and use to the entity.
- The inventory accounting system (perpetual or periodic) may result in different values for cost of sales and ending inventory when the weighted average cost or LIFO inventory valuation method is used.
- Under US GAAP, companies that use the LIFO method must disclose in their financial notes the amount of the LIFO reserve or the amount that would have been reported in inventory if the FIFO method had been used. This information can be used to adjust reported LIFO inventory and cost of goods sold balances to the FIFO method for comparison purposes.
- LIFO liquidation occurs when the number of units in ending inventory declines from the number of units that were present at the beginning of the year. If inventory unit costs have generally risen from year to year, this will produce an inventory-related increase in gross profits.
- Consistency of inventory costing is required under both IFRS and US GAAP. If a company changes an accounting policy, the change must be justifiable and applied retrospectively to the financial statements. An exception to the retrospective restatement is when a company reporting under US GAAP changes to the LIFO method.
- Under IFRS, inventories are measured at the lower of cost and net realisable value. Net realisable value is the estimated selling price in the ordinary course of business less the estimated costs necessary to make the sale. Under US GAAP, inventories are measured at the lower of cost, market value, or net realisable value depending upon the inventory method used. Market value is defined as current replacement cost subject to an upper limit of net realizable value and a lower limit of net realizable value less a normal profit margin. Reversals of previous write-downs are permissible under IFRS but not under US GAAP.
- Reversals of inventory write-downs may occur under IFRS but are not allowed under US GAAP.

- Changes in the carrying amounts within inventory classifications (such as raw materials, work-in-process, and finished goods) may provide signals about a company's future sales and profits. Relevant information with respect to inventory management and future sales may be found in the Management Discussion and Analysis or similar items within the annual or quarterly reports, industry news and publications, and industry economic data.
- The inventory turnover ratio, number of days of inventory ratio, and gross profit margin ratio are useful in evaluating the management of a company's inventory.
- Inventory management may have a substantial impact on a company's activity, profitability, liquidity, and solvency ratios. It is critical for the analyst to be aware of industry trends and management's intentions.
- Financial statement disclosures provide information regarding the accounting policies adopted in measuring inventories, the principal uncertainties regarding the use of estimates related to inventories, and details of the inventory carrying amounts and costs. This information can greatly assist analysts in their evaluation of a company's inventory management.

PRACTICE PROBLEMS

1 Inventory cost is *least likely* to include:

A production-related storage costs.

B costs incurred as a result of normal waste of materials.

C transportation costs of shipping inventory to customers.

2 Mustard Seed PLC adheres to IFRS. It recently purchased inventory for €100 million and spent €5 million for storage prior to selling the goods. The amount it charged to inventory expense (€ millions) was *closest* to:

A €95.

B €100.

C €105.

3 Carrying inventory at a value above its historical cost would *most likely* be permitted if:

A the inventory was held by a producer of agricultural products.

B financial statements were prepared using US GAAP.

C the change resulted from a reversal of a previous write-down.

The following information relates to Questions 4 and 5

A retail company is comparing different approaches to valuing inventory. The company has one product that it sells for $50.

Table 1 Units Purchased and Sold (first quarter)

Date	Units Purchased	Purchase Price	Units Sold	Selling Price	Inventory Units on Hand
2 Jan	1,000	$20.00			1,000
17 Jan			500	$50.00	500
16 Feb	1,000	$18.00			1,500
3 Mar			1,200	$50.00	300
13 Mar	1,000	$17.00			1,300
23 Mar			500	$50.00	800
End of quarter totals:	3,000	$55,000	2,200	$110,000	

© 2019 CFA Institute. All rights reserved.

Table 2 Comparison of Inventory Methods and Models

End of Quarter Valuations 31 March	Perpetual LIFO	Periodic LIFO	Perpetual FIFO
Sales	$110,000	$110,000	$110,000
Ending inventory		$16,000	$13,600
Cost of goods sold		$39,000	$41,400
Gross profit		$71,000	$68,600
Inventory turnover ratio	279%		

Note: LIFO is last in, first out and FIFO is first in, first out.

4 What is the value of ending inventory for the first quarter if the company uses a perpetual LIFO inventory valuation method?

A $14,500

B $15,000

C $16,000

5 Which inventory accounting method results in the lowest inventory turnover ratio for the first quarter?

A Periodic LIFO

B Perpetual LIFO

C Perpetual FIFO

---

6 During periods of rising inventory unit costs, a company using the FIFO method rather than the LIFO method will report a lower:

A current ratio.

B inventory turnover.

C gross profit margin.

7 LIFO reserve is *most likely* to increase when inventory unit:

A costs are increasing.

B costs are decreasing.

C levels are decreasing.

8 If inventory unit costs are increasing from period-to-period, a LIFO liquidation is *most likely* to result in an increase in:

A gross profit.

B LIFO reserve.

C inventory carrying amounts.

9 A company using the LIFO method reports the following in £:

	2018	2017
Cost of goods sold (COGS)	50,800	48,500
Ending inventories	10,550	10,000
LIFO reserve	4,320	2,600

Cost of goods sold for 2018 under the FIFO method is *closest* to:

A £48,530.

B £49,080.

C £52,520.

10 Eric's Used Book Store prepares its financial statements in accordance with IFRS. Inventory was purchased for £1 million and later marked down to £550,000. One of the books, however, was later discovered to be a rare collectible item, and the inventory is now worth an estimated £3 million. The inventory is *most likely* reported on the balance sheet at:

A £550,000.

B £1,000,000.

C £3,000,000.

11 Fernando's Pasta purchased inventory and later wrote it down. The current net realisable value is higher than the value when written down. Fernando's inventory balance will *most likely* be:

A higher if it complies with IFRS.

B higher if it complies with US GAAP.

C the same under US GAAP and IFRS.

12 A write down of the value of inventory to its net realizable value will have a positive effect on the:

A balance sheet.

B income statement.

C inventory turnover ratio.

For questions 13–24, assume the companies use a periodic inventory system

13 Cinnamon Corp. started business in 2017 and uses the weighted average cost method. During 2017, it purchased 45,000 units of inventory at €10 each and sold 40,000 units for €20 each. In 2018, it purchased another 50,000 units at €11 each and sold 45,000 units for €22 each. Its 2018 cost of sales (€ thousands) was *closest* to:

A €490.

B €491.

C €495.

14 Zimt AG started business in 2017 and uses the FIFO method. During 2017, it purchased 45,000 units of inventory at €10 each and sold 40,000 units for €20 each. In 2018, it purchased another 50,000 units at €11 each and sold 45,000 units for €22 each. Its 2018 ending inventory balance (€ thousands) was *closest* to:

A €105.

B €109.

C €110.

15 Zimt AG uses the FIFO method, and Nutmeg Inc. uses the LIFO method. Compared to the cost of replacing the inventory, during periods of rising prices, the cost of sales reported by:

A Zimt is too low.

B Nutmeg is too low.

C Nutmeg is too high.

16 Zimt AG uses the FIFO method, and Nutmeg Inc. uses the LIFO method. Compared to the cost of replacing the inventory, during periods of rising prices the ending inventory balance reported by:

A Zimt is too high.

B Nutmeg is too low.

C Nutmeg is too high.

17 Like many technology companies, TechnoTools operates in an environment of declining prices. Its reported profits will tend to be *highest* if it accounts for inventory using the:

A FIFO method.

B LIFO method.

C weighted average cost method.

18 Compared to using the weighted average cost method to account for inventory, during a period in which prices are generally rising, the current ratio of a company using the FIFO method would *most likely* be:

A lower.

B higher.

C dependent upon the interaction with accounts payable.

19 Zimt AG wrote down the value of its inventory in 2017 and reversed the write-down in 2018. Compared to the ratios that would have been calculated if the write-down had never occurred, Zimt's reported 2017:

A current ratio was too high.

B gross margin was too high.

C inventory turnover was too high.

20 Zimt AG wrote down the value of its inventory in 2017 and reversed the write-down in 2018. Compared to the results the company would have reported if the write-down had never occurred, Zimt's reported 2018:

A profit was overstated.

B cash flow from operations was overstated.

C year-end inventory balance was overstated.

21 Compared to a company that uses the FIFO method, during periods of rising prices a company that uses the LIFO method will *most likely* appear more:

A liquid.

B efficient.

C profitable.

22 Nutmeg, Inc. uses the LIFO method to account for inventory. During years in which inventory unit costs are generally rising and in which the company purchases more inventory than it sells to customers, its reported gross profit margin will *most likely* be:

A lower than it would be if the company used the FIFO method.

B higher than it would be if the company used the FIFO method.

C about the same as it would be if the company used the FIFO method.

23 Compared to using the FIFO method to account for inventory, during periods of rising prices, a company using the LIFO method is *most likely* to report higher:

A net income.

B cost of sales.

C income taxes.

24 Carey Company adheres to US GAAP, whereas Jonathan Company adheres to IFRS. It is *least likely* that:

A Carey has reversed an inventory write-down.

B Jonathan has reversed an inventory write-down.

C Jonathan and Carey both use the FIFO inventory accounting method.

---

25 Company A adheres to US GAAP and Company B adheres to IFRS. Which of the following is *most likely* to be disclosed on the financial statements of both companies?

A Any material income resulting from the liquidation of LIFO inventory

B The amount of inventories recognized as an expense during the period

C The circumstances that led to the reversal of a write down of inventories

26 Which of the following *most likely* signals that a manufacturing company expects demand for its product to increase?

A Finished goods inventory growth rate higher than the sales growth rate

B Higher unit volumes of work in progress and raw material inventories

C Substantially higher finished goods, with lower raw materials and work-in-process

27 Compared with a company that uses the FIFO method, during a period of rising unit inventory costs, a company using the LIFO method will *most likely* appear more:

A liquid.

B efficient.

C profitable.

28 In a period of declining inventory unit costs and constant or increasing inventory quantities, which inventory method is *most likely* to result in a higher debt-to-equity ratio?

A LIFO

B FIFO

C Weighted average cost

The following information relates to Questions 29–36

Hans Annan, CFA, a food and beverage analyst, is reviewing Century Chocolate's inventory policies as part of his evaluation of the company. Century Chocolate, based in Switzerland, manufactures chocolate products and purchases and resells other confectionery products to complement its chocolate line. Annan visited Century Chocolate's manufacturing facility last year. He learned that cacao beans, imported from Brazil, represent the most significant raw material and that the work-in-progress inventory consists primarily of three items: roasted cacao beans, a thick paste produced from the beans (called chocolate liquor), and a sweetened mixture that needs to be "conched" to produce chocolate. On the tour, Annan learned that the conching process ranges from a few hours for lower-quality products to six days for the highest-quality chocolates. While there, Annan saw the facility's climate-controlled area where manufactured finished products (cocoa and chocolate) and purchased finished goods are stored prior to shipment to customers. After touring the facility, Annan had a discussion with Century Chocolate's CFO regarding the types of costs that were included in each inventory category.

Annan has asked his assistant, Joanna Kern, to gather some preliminary information regarding Century Chocolate's financial statements and inventories. He also asked Kern to calculate the inventory turnover ratios for Century Chocolate and another chocolate manufacturer for the most recent five years. Annan does not know Century Chocolate's most direct competitor, so he asks Kern to do some research and select the most appropriate company for the ratio comparison.

Kern reports back that Century Chocolate prepares its financial statements in accordance with IFRS. She tells Annan that the policy footnote states that raw materials and purchased finished goods are valued at purchase cost whereas work in progress and manufactured finished goods are valued at production cost. Raw material inventories and purchased finished goods are accounted for using the FIFO (first-in, first-out) method, and the weighted average cost method is used for other inventories. An allowance is established when the net realisable value of any inventory item is lower than the value calculated above.

Kern provides Annan with the selected financial statements and inventory data for Century Chocolate shown in Exhibits 1 through 5. The ratio exhibit Kern prepared compares Century Chocolate's inventory turnover ratios to those of Gordon's Goodies, a US-based company. Annan returns the exhibit and tells Kern to select a different competitor that reports using IFRS rather than US GAAP. During this initial review, Annan asks Kern why she has not indicated whether Century Chocolate uses a perpetual or a periodic inventory system. Kern replies that she learned that Century Chocolate uses a perpetual system but did not include this information in her report because inventory values would be the same under either a perpetual or periodic inventory system. Annan tells Kern she is wrong and directs her to research the matter.

While Kern is revising her analysis, Annan reviews the most recent month's Cocoa Market Review from the International Cocoa Organization. He is drawn to the statement that "the ICCO daily price, averaging prices in both futures markets, reached a 29-year high in US$ terms and a 23-year high in SDRs terms (the SDR unit comprises a basket of major currencies used in international trade: US$, euro, pound sterling and yen)." Annan makes a note that he will need to factor the potential continuation of this trend into his analysis.

Exhibit 1 Century Chocolate Income Statements (CHF Millions)

For Years Ended 31 December	2018	2017
Sales	95,290	93,248
Cost of sales	–41,043	–39,047
Marketing, administration, and other expenses	–35,318	–42,481
Profit before taxes	**18,929**	**11,720**
Taxes	–3,283	–2,962
Profit for the period	**15,646**	**8,758**

Exhibit 2 Century Chocolate Balance Sheets (CHF Millions)

31 December	2018	2017
Cash, cash equivalents, and short-term investments	6,190	8,252
Trade receivables and related accounts, net	11,654	12,910
Inventories, net	8,100	7,039
Other current assets	2,709	2,812
Total current assets	**28,653**	**31,013**
Property, plant, and equipment, net	18,291	19,130
Other non-current assets	45,144	49,875
Total assets	**92,088**	**100,018**
Trade and other payables	10,931	12,299
Other current liabilities	17,873	25,265
Total current liabilities	**28,804**	**37,564**
Non-current liabilities	15,672	14,963
Total liabilities	**44,476**	**52,527**
Equity		
Share capital	332	341
Retained earnings and other reserves	47,280	47,150
Total equity	**47,612**	**47,491**
Total liabilities and shareholders' equity	**92,088**	**100,018**

Exhibit 3 Century Chocolate Supplementary Footnote Disclosures: Inventories (CHF Millions)

31 December	2018	2017
Raw Materials	2,154	1,585
Work in Progress	1,061	1,027
Finished Goods	5,116	4,665
Total inventories before allowance	8,331	7,277
Allowance for write-downs to net realisable value	–231	–238
Total inventories net of allowance	8,100	7,039

Exhibit 4 Century Chocolate Inventory Record for Purchased Lemon Drops

Date		Cartons	Per Unit Amount (CHF)
	Beginning inventory	100	22
4 Feb 09	Purchase	40	25

(continued)

Exhibit 4 (Continued)

Date		Cartons	Per Unit Amount (CHF)
3 Apr 09	Sale	50	32
23 Jul 09	Purchase	70	30
16 Aug 09	Sale	100	32
9 Sep 09	Sale	35	32
15 Nov 09	Purchase	100	28

Exhibit 5 Century Chocolate Net Realisable Value Information for Black Licorice Jelly Beans

	2018	2017
FIFO cost of inventory at 31 December (CHF)	314,890	374,870
Ending inventory at 31 December (Kilograms)	77,750	92,560
Cost per unit (CHF)	4.05	4.05
Net Realisable Value (CHF per Kilograms)	4.20	3.95

29 The costs *least likely* to be included by the CFO as inventory are:

A storage costs for the chocolate liquor.

B excise taxes paid to the government of Brazil for the cacao beans.

C storage costs for chocolate and purchased finished goods awaiting shipment to customers.

30 What is the *most likely* justification for Century Chocolate's choice of inventory valuation method for its purchased finished goods?

A It is the preferred method under IFRS.

B It allocates the same per unit cost to both cost of sales and inventory.

C Ending inventory reflects the cost of goods purchased most recently.

31 In Kern's comparative ratio analysis, the 2018 inventory turnover ratio for Century Chocolate is *closest* to:

A 5.07.

B 5.42.

C 5.55.

32 The *most accurate* statement regarding Annan's reasoning for requiring Kern to select a competitor that reports under IFRS for comparative purposes is that under US GAAP:

A fair values are used to value inventory.

B the LIFO method is permitted to value inventory.

C the specific identification method is permitted to value inventory.

33 Annan's statement regarding the perpetual and periodic inventory systems is most significant when which of the following costing systems is used?

A LIFO.

B FIFO.

C Specific identification.

34 Using the inventory record for purchased lemon drops shown in Exhibit 4, the cost of sales for 2018 will be *closest* to:

A CHF 3,550.

B CHF 4,550.

C CHF 4,850.

35 Ignoring any tax effect, the 2018 net realisable value reassessment for the black licorice jelly beans will *most likely* result in:

A an increase in gross profit of CHF 7,775.

B an increase in gross profit of CHF 11,670.

C no impact on cost of sales because under IFRS, write-downs cannot be reversed.

36 If the trend noted in the ICCO report continues and Century Chocolate plans to maintain constant or increasing inventory quantities, the *most likely* impact on Century Chocolate's financial statements related to its raw materials inventory will be:

A a cost of sales that more closely reflects current replacement values.

B a higher allocation of the total cost of goods available for sale to cost of sales.

C a higher allocation of the total cost of goods available for sale to ending inventory.

The following information relates to Questions 37–42

John Martinson, CFA, is an equity analyst with a large pension fund. His supervisor, Linda Packard, asks him to write a report on Karp Inc. Karp prepares its financial statements in accordance with US GAAP. Packard is particularly interested in the effects of the company's use of the LIFO method to account for its inventory. For this purpose, Martinson collects the financial data presented in Exhibits 1 and 2.

Exhibit 1 Balance Sheet Information (US$ Millions)

As of 31 December	2018	2017
Cash and cash equivalents	172	157
Accounts receivable	626	458
Inventories	620	539
Other current assets	125	65
Total current assets	1,543	1,219
Property and equipment, net	3,035	2,972
Total assets	4,578	4,191
Total current liabilities	1,495	1,395
Long-term debt	644	604

(continued)

Exhibit 1 (Continued)

As of 31 December	2018	2017
Total liabilities	2,139	1,999
Common stock and paid in capital	1,652	1,652
Retained earnings	787	540
Total shareholders' equity	2,439	2,192
Total liabilities and shareholders' equity	4,578	4,191

Exhibit 2 Income Statement Information (US$ Millions)

For the Year Ended 31 December	2018	2017
Sales	4,346	4,161
Cost of goods sold	2,211	2,147
Depreciation and amortisation expense	139	119
Selling, general, and administrative expense	1,656	1,637
Interest expense	31	18
Income tax expense	62	48
Net income	247	192

Martinson finds the following information in the notes to the financial statements:

- The LIFO reserves as of 31 December 2018 and 2017 are $155 million and $117 million respectively, and
- The effective income tax rate applicable to Karp for 2018 and earlier periods is 20 percent.

37 If Karp had used FIFO instead of LIFO, the amount of inventory reported as of 31 December 2018 would have been *closest* to:

A $465 million.

B $658 million.

C $775 million.

38 If Karp had used FIFO instead of LIFO, the amount of cost of goods sold reported by Karp for the year ended 31 December 2018 would have been *closest* to:

A $2,056 million.

B $2,173 million.

C $2,249 million.

39 If Karp had used FIFO instead of LIFO, its reported net income for the year ended 31 December 2018 would have been higher by an amount *closest to*:

A $30 million.

B $38 million.

C $155 million.

40 If Karp had used FIFO instead of LIFO, Karp's retained earnings as of 31 December 2018 would have been higher by an amount *closest to*:

A $117 million.

B $124 million.

C $155 million.

41 If Karp had used FIFO instead of LIFO, which of the following ratios computed as of 31 December 2018 would *most likely* have been lower?

A Cash ratio.

B Current ratio.

C Gross profit margin.

42 If Karp had used FIFO instead of LIFO, its debt to equity ratio computed as of 31 December 2018 would have:

A increased.

B decreased.

C remained unchanged.

The following information relates to Questions 43–48

Robert Groff, an equity analyst, is preparing a report on Crux Corp. As part of his report, Groff makes a comparative financial analysis between Crux and its two main competitors, Rolby Corp. and Mikko Inc. Crux and Mikko report under US GAAP and Rolby reports under IFRS.

Groff gathers information on Crux, Rolby, and Mikko. The relevant financial information he compiles is in Exhibit 1. Some information on the industry is in Exhibit 2.

Exhibit 1 Selected Financial Information (US$ Millions)

	Crux	Rolby	Mikko
Inventory valuation method	LIFO	FIFO	LIFO
From the Balance Sheets			
As of 31 December 2018			
Inventory, gross	480	620	510
Valuation allowance	20	25	14
Inventory, net	460	595	496
Total debt	1,122	850	732
Total shareholders' equity	2,543	2,403	2,091
As of 31 December 2017			
Inventory, gross	465	602	401
Valuation allowance	23	15	12
Inventory, net	442	587	389
From the Income Statements			
Year Ended 31 December 2018			

(continued)

Exhibit 1 (Continued)

	Crux	Rolby	Mikko
Revenues	4,609	5,442	3,503
Cost of goods sold[a]	3,120	3,782	2,550
Net income	229	327	205
[a]Charges included in cost of goods sold for inventory write-downs*	13	15	15

* This does not match the change in the inventory valuation allowance because the valuation allowance is reduced to reflect the valuation allowance attached to items sold and increased for additional necessary write-downs.

LIFO Reserve			
As of 31 December 2018	55	0	77
As of 31 December 2017	72	0	50
As of 31 December 2016	96	0	43
Tax Rate			
Effective tax rate	30%	30%	30%

Exhibit 2 Industry Information

	2018	2017	2016
Raw materials price index	112	105	100
Finished goods price index	114	106	100

To compare the financial performance of the three companies, Groff decides to convert LIFO figures into FIFO figures, and adjust figures to assume no valuation allowance is recognized by any company.

After reading Groff's draft report, his supervisor, Rachel Borghi, asks him the following questions:

Question 1	Which company's gross profit margin would best reflect current costs of the industry?
Question 2	Would Rolby's valuation method show a higher gross profit margin than Crux's under an inflationary, a deflationary, or a stable price scenario?
Question 3	Which group of ratios usually appears more favorable with an inventory write-down?

43 Crux's inventory turnover ratio computed as of 31 December 2018, after the adjustments suggested by Groff, is *closest* to:

A 5.67.

B 5.83.

C 6.13.

44 Rolby's net profit margin for the year ended 31 December 2018, after the adjustments suggested by Groff, is *closest* to:

A 6.01%.

B 6.20%.

C 6.28%.

45 Compared with its unadjusted debt-to-equity ratio, Mikko's debt-to-equity ratio as of 31 December 2018, after the adjustments suggested by Groff, is:

A lower.

B higher.

C the same.

46 The *best* answer to Borghi's Question 1 is:

A Crux's.

B Rolby's.

C Mikko's.

47 The *best* answer to Borghi's Question 2 is:

A Stable.

B Inflationary.

C Deflationary.

48 The *best* answer to Borghi's Question 3 is:

A Activity ratios.

B Solvency ratios.

C Profitability ratios.

The following information relates to Questions 49–55

ZP Corporation is a (hypothetical) multinational corporation headquartered in Japan that trades on numerous stock exchanges. ZP prepares its consolidated financial statements in accordance with US GAAP. Excerpts from ZP's 2018 annual report are shown in Exhibits 1–3.

Exhibit 1 Consolidated Balance Sheets (¥ Millions)

31 December	2017	2018
Current Assets		
Cash and cash equivalents	¥542,849	¥814,760
⋮	⋮	⋮
Inventories	608,572	486,465
⋮	⋮	⋮
Total current assets	4,028,742	3,766,309
⋮	⋮	⋮
Total assets	**¥10,819,440**	**¥9,687,346**

(continued)

Exhibit 1 (Continued)

31 December	2017	2018
⋮	⋮	⋮
Total current liabilities	¥3,980,247	¥3,529,765
⋮	⋮	⋮
Total long-term liabilities	2,663,795	2,624,002
Minority interest in consolidated subsidiaries	218,889	179,843
Total shareholders' equity	3,956,509	3,353,736
Total liabilities and shareholders' equity	**¥10,819,440**	**¥9,687,346**

Exhibit 2 Consolidated Statements of Income (¥ Millions)

For the years ended 31 December	2016	2017	2018
Net revenues			
Sales of products	¥7,556,699	¥8,273,503	¥6,391,240
Financing operations	425,998	489,577	451,950
	7,982,697	8,763,080	6,843,190
Cost and expenses			
Cost of products sold	6,118,742	6,817,446	5,822,805
Cost of financing operations	290,713	356,005	329,128
Selling, general and administrative	827,005	832,837	844,927
⋮	⋮	⋮	⋮
Operating income (loss)	746,237	756,792	−153,670
⋮	⋮	⋮	⋮
Net income	¥548,011	¥572,626	−¥145,646

Exhibit 3 Selected Disclosures in the 2018 Annual Report

Management's Discussion and Analysis of Financial Condition and Results of Operations

Cost reduction efforts were offset by increased prices of raw materials, other production materials and parts … Inventories decreased during fiscal 2018 by ¥122.1 billion, or 20.1%, to ¥486.5 billion. This reflects the impacts of decreased sales volumes and fluctuations in foreign currency translation rates.

Management & Corporate Information

Risk Factors

Industry and Business Risks

Exhibit 3 (Continued)

The worldwide market for our products is highly competitive. ZP faces intense competition from other manufacturers in the respective markets in which it operates. Competition has intensified due to the worldwide deterioration in economic conditions. In addition, competition is likely to further intensify because of continuing globalization, possibly resulting in industry reorganization. Factors affecting competition include product quality and features, the amount of time required for innovation and development, pricing, reliability, safety, economy in use, customer service and financing terms. Increased competition may lead to lower unit sales and excess production capacity and excess inventory. This may result in a further downward price pressure.

ZP's ability to adequately respond to the recent rapid changes in the industry and to maintain its competitiveness will be fundamental to its future success in maintaining and expanding its market share in existing and new markets.

Notes to Consolidated Financial Statements

2. Summary of significant accounting policies:

Inventories. Inventories are valued at cost, not in excess of market. Cost is determined on the "average-cost" basis, except for the cost of finished products carried by certain subsidiary companies which is determined "last-in, first-out" ("LIFO") basis. Inventories valued on the LIFO basis totaled ¥94,578 million and ¥50,037 million at December 31, 2017 and 2018, respectively. Had the "first-in, first-out" basis been used for those companies using the LIFO basis, inventories would have been ¥10,120 million and ¥19,660 million higher than reported at December 31, 2017 and 2018, respectively.

9. Inventories:

Inventories consist of the following:

31 December (¥ Millions)	2017	2018
Finished goods	¥ 403,856	¥ 291,977
Raw materials	99,869	85,966
Work in process	79,979	83,890
Supplies and other	24,868	24,632
	¥ 608,572	**¥ 486,465**

49 The MD&A indicated that the prices of raw material, other production materials, and parts increased. Based on the inventory valuation methods described in Note 2, which inventory classification would *least accurately* reflect current prices?

A Raw materials.

B Finished goods.

C Work in process.

50 The 2017 inventory value as reported on the 2018 Annual Report if the company had used the FIFO inventory valuation method instead of the LIFO inventory valuation method for a portion of its inventory would be *closest* to:

A ¥104,698 million.

B ¥506,125 million.

C ¥618,692 million.

51 If ZP had prepared its financial statement in accordance with IFRS, the inventory turnover ratio (using average inventory) for 2018 would be:

A lower.

B higher.

C the same.

52 Inventory levels decreased from 2017 to 2018 for all of the following reasons *except*:

A LIFO liquidation.

B decreased sales volume.

C fluctuations in foreign currency translation rates.

53 Which observation is *most likely* a result of looking only at the information reported in Note 9?

A Increased competition has led to lower unit sales.

B There have been significant price increases in supplies.

C Management expects a further downturn in sales during 2019.

54 Note 2 indicates that, "Inventories valued on the LIFO basis totaled ¥94,578 million and ¥50,037 million at December 31, 2017 and 2018, respectively." Based on this, the LIFO reserve should *most likely*:

A increase.

B decrease.

C remain the same.

55 The Industry and Business Risk excerpt states that, "Increased competition may lead to lower unit sales and excess production capacity and excess inventory. This may result in a further downward price pressure." The downward price pressure could lead to inventory that is valued above current market prices or net realisable value. Any write-downs of inventory are *least likely* to have a significant effect on the inventory valued using:

A weighted average cost.

B first-in, first-out (FIFO).

C last-in, first-out (LIFO).

SOLUTIONS

1 C is correct. Transportation costs incurred to ship inventory to customers are an expense and may not be capitalized in inventory. (Transportation costs incurred to bring inventory to the business location can be capitalized in inventory.) Storage costs required as part of production, as well as costs incurred as a result of normal waste of materials, can be capitalized in inventory. (Costs incurred as a result of abnormal waste must be expensed.)

2 B is correct. Inventory expense includes costs of purchase, costs of conversion, and other costs incurred in bringing the inventories to their present location and condition. It does not include storage costs not required as part of production.

3 A is correct. IFRS allow the inventories of producers and dealers of agricultural and forest products, agricultural produce after harvest, and minerals and mineral products to be carried at net realisable value even if above historical cost. (US GAAP treatment is similar.)

4 A is correct. A perpetual inventory system updates inventory values and quantities and cost of goods sold continuously to reflect purchases and sales. The ending inventory of 800 units consists of 300 units at $20 and 500 units at $17.

$$(300 \times \$20) + (500 \times \$17) = \$14{,}500$$

5 A is correct. In an environment with falling inventory costs and declining inventory levels, periodic LIFO will result in a higher ending inventory value and lower cost of goods sold versus perpetual LIFO and perpetual FIFO methods. This results in a lower inventory turnover ratio, which is calculated as follows:

$$\text{Inventory turnover ratio} = \text{Cost of goods sold}/\text{Ending inventory}$$

The inventory turnover ratio using periodic LIFO is $39,000/$16,000 = 244% or 2.44 times.

The inventory turnover ratio using perpetual LIFO is 279% or 2.79 times, which is provided in Table 2 (= 40,500/14,500 from previous question).

The inventory turnover for perpetual FIFO is $41,400/$13,600 = 304% or 3.04 times.

6 B is correct. During a period of rising inventory costs, a company using the FIFO method will allocate a lower amount to cost of goods sold and a higher amount to ending inventory as compared with the LIFO method. The inventory turnover ratio is the ratio of cost of sales to ending inventory. A company using the FIFO method will produce a lower inventory turnover ratio as compared with the LIFO method. The current ratio (current assets/current liabilities) and the gross profit margin [gross profit/sales = (sales less cost of goods sold)/sales] will be higher under the FIFO method than under the LIFO method in periods of rising inventory unit costs.

7 A is correct. LIFO reserve is the FIFO inventory value less the LIFO inventory value. In periods of rising inventory unit costs, the carrying amount of inventory under FIFO will always exceed the carrying amount of inventory under LIFO. The LIFO reserve may increase over time as a result of the increasing difference between the older costs used to value inventory under LIFO and the

more recent costs used to value inventory under FIFO. When inventory unit levels are decreasing, the company will experience a LIFO liquidation, reducing the LIFO reserve.

8 A is correct. When the number of units sold exceeds the number of units purchased, a company using LIFO will experience a LIFO liquidation. If inventory unit costs have been rising from period-to-period and a LIFO liquidation occurs, it will produce an increase in gross profit as a result of the lower inventory carrying amounts of the liquidated units (lower cost per unit of the liquidated units).

9 B is correct. The adjusted COGS under the FIFO method is equal to COGS under the LIFO method less the increase in LIFO reserve:

COGS (FIFO) = COGS (LIFO) – Increase in LIFO reserve
COGS (FIFO) = £50,800 – (£4,320 – £2,600)
COGS (FIFO) = £49,080

10 B is correct. Under IFRS, the reversal of write-downs is required if net realisable value increases. The inventory will be reported on the balance sheet at £1,000,000. The inventory is reported at the lower of cost or net realisable value. Under US GAAP, inventory is carried at the lower of cost or market value. After a write-down, a new cost basis is determined and additional revisions may only reduce the value further. The reversal of write-downs is not permitted.

11 A is correct. IFRS require the reversal of inventory write-downs if net realisable values increase; US GAAP do not permit the reversal of write-downs.

12 C is correct. Activity ratios (for example, inventory turnover and total asset turnover) will be positively affected by a write down to net realizable value because the asset base (denominator) is reduced. On the balance sheet, the inventory carrying amount is written down to its net realizable value and the loss in value (expense) is generally reflected on the income statement in cost of goods sold, thus reducing gross profit, operating profit, and net income.

13 B is correct. Cinnamon uses the weighted average cost method, so in 2018, 5,000 units of inventory were 2017 units at €10 each and 50,000 were 2008 purchases at €11. The weighted average cost of inventory during 2008 was thus (5,000 × 10) + (50,000 × 11) = 50,000 + 550,000 = €600,000, and the weighted average cost was approximately €10.91 = €600,000/55,000. Cost of sales was €10.91 × 45,000, which is approximately €490,950.

14 C is correct. Zimt uses the FIFO method, and thus the first 5,000 units sold in 2018 depleted the 2017 inventory. Of the inventory purchased in 2018, 40,000 units were sold and 10,000 remain, valued at €11 each, for a total of €110,000.

15 A is correct. Zimt uses the FIFO method, so its cost of sales represents units purchased at a (no longer available) lower price. Nutmeg uses the LIFO method, so its cost of sales is approximately equal to the current replacement cost of inventory.

16 B is correct. Nutmeg uses the LIFO method, and thus some of the inventory on the balance sheet was purchased at a (no longer available) lower price. Zimt uses the FIFO method, so the carrying value on the balance sheet represents the most recently purchased units and thus approximates the current replacement cost.

17 B is correct. In a declining price environment, the newest inventory is the lowest-cost inventory. In such circumstances, using the LIFO method (selling the newer, cheaper inventory first) will result in lower cost of sales and higher profit.

18 B is correct. In a rising price environment, inventory balances will be higher for the company using the FIFO method. Accounts payable are based on amounts due to suppliers, not the amounts accrued based on inventory accounting.

19 C is correct. The write-down reduced the value of inventory and increased cost of sales in 2017. The higher numerator and lower denominator mean that the inventory turnover ratio as reported was too high. Gross margin and the current ratio were both too low.

20 A is correct. The reversal of the write-down shifted cost of sales from 2018 to 2017. The 2017 cost of sales was higher because of the write-down, and the 2018 cost of sales was lower because of the reversal of the write-down. As a result, the reported 2018 profits were overstated. Inventory balance in 2018 is the same because the write-down and reversal cancel each other out. Cash flow from operations is not affected by the non-cash write-down, but the higher profits in 2018 likely resulted in higher taxes and thus lower cash flow from operations.

21 B is correct. LIFO will result in lower inventory and higher cost of sales. Gross margin (a profitability ratio) will be lower, the current ratio (a liquidity ratio) will be lower, and inventory turnover (an efficiency ratio) will be higher.

22 A is correct. LIFO will result in lower inventory and higher cost of sales in periods of rising costs compared to FIFO. Consequently, LIFO results in a lower gross profit margin than FIFO.

23 B is correct. The LIFO method increases cost of sales, thus reducing profits and the taxes thereon.

24 A is correct. US GAAP do not permit inventory write-downs to be reversed.

25 B is correct. Both US GAAP and IFRS require disclosure of the amount of inventories recognized as an expense during the period. Only US GAAP allows the LIFO method and requires disclosure of any material amount of income resulting from the liquidation of LIFO inventory. US GAAP does not permit the reversal of prior-year inventory write downs.

26 B is correct. A significant increase (attributable to increases in unit volume rather than increases in unit cost) in raw materials and/or work-in-progress inventories may signal that the company expects an increase in demand for its products. If the growth of finished goods inventories is greater than the growth of sales, it could indicate a decrease in demand and a decrease in future earnings. A substantial increase in finished goods inventories while raw materials and work-in-progress inventories are declining may signal a decrease in demand for the company's products.

27 B is correct. During a period of rising inventory prices, a company using the LIFO method will have higher cost of cost of goods sold and lower inventory compared with a company using the FIFO method. The inventory turnover ratio will be higher for the company using the LIFO method, thus making it appear more efficient. Current assets and gross profit margin will be lower for the company using the LIFO method, thus making it appear less liquid and less profitable.

28 B is correct. In an environment of declining inventory unit costs and constant or increasing inventory quantities, FIFO (in comparison with weighted average cost or LIFO) will have higher cost of goods sold and lower net income and inventory. Because both inventory and net income are lower, total equity is lower, resulting in a higher debt-to-equity ratio.

29 C is correct. The storage costs for inventory awaiting shipment to customers are not costs of purchase, costs of conversion, or other costs incurred in bringing the inventories to their present location and condition and are not included in inventory. The storage costs for the chocolate liquor occur during the production process and are thus part of the conversion costs. Excise taxes are part of the purchase cost.

30 C is correct. The carrying amount of inventories under FIFO will more closely reflect current replacement values because inventories are assumed to consist of the most recently purchased items. FIFO is an acceptable, but not preferred, method under IFRS. Weighted average cost, not FIFO, is the cost formula that allocates the same per unit cost to both cost of sales and inventory.

31 B is correct. Inventory turnover = Cost of sales/Average inventory = 41,043/7,569.5 = 5.42. Average inventory is (8,100 + 7,039)/2 = 7,569.5.

32 B is correct. For comparative purposes, the choice of a competitor that reports under IFRS is requested because LIFO is permitted under US GAAP.

33 A is correct. The carrying amount of the ending inventory may differ because the perpetual system will apply LIFO continuously throughout the year, liquidating layers as sales are made. Under the periodic system, the sales will start from the last layer in the year. Under FIFO, the sales will occur from the same layers regardless of whether a perpetual or periodic system is used. Specific identification identifies the actual products sold and remaining in inventory, and there will be no difference under a perpetual or periodic system.

34 B is correct. The cost of sales is closest to CHF 4,550. Under FIFO, the inventory acquired first is sold first. Using Exhibit 4, a total of 310 cartons were available for sale (100 + 40 + 70 + 100) and 185 cartons were sold (50 + 100 + 35), leaving 125 in ending inventory. The FIFO cost would be as follows:

100 (beginning inventory) × 22 = 2,200

40 (4 February 2009) × 25 = 1,000

45 (23 July 2009) × 30 = 1,350

Cost of sales = 2,200 + 1,000 + 1,350 = CHF 4,550

35 A is correct. Gross profit will most likely increase by CHF 7,775. The net realisable value has increased and now exceeds the cost. The write-down from 2017 can be reversed. The write-down in 2017 was 9,256 [92,560 × (4.05 – 3.95)]. IFRS require the reversal of any write-downs for a subsequent increase in value of inventory previously written down. The reversal is limited to the lower of the subsequent increase or the original write-down. Only 77,750 kilograms remain in inventory; the reversal is 77,750 × (4.05 – 3.95) = 7,775. The amount of any reversal of a write-down is recognised as a reduction in cost of sales. This reduction results in an increase in gross profit.

36 C is correct. Using the FIFO method to value inventories when prices are rising will allocate more of the cost of goods available for sale to ending inventories (the most recent purchases, which are at higher costs, are assumed to remain in inventory) and less to cost of sales (the oldest purchases, which are at lower costs, are assumed to be sold first).

37 C is correct. Karp's inventory under FIFO equals Karp's inventory under LIFO plus the LIFO reserve. Therefore, as of 31 December 2018, Karp's inventory under FIFO equals:

$$\begin{aligned}\text{Inventory (FIFO method)} &= \text{Inventory (LIFO method)} + \text{LIFO reserve}\\ &= \$620 \text{ million} + 155 \text{ million}\\ &= \$775 \text{ million}\end{aligned}$$

38 B is correct. Karp's cost of goods sold (COGS) under FIFO equals Karp's cost of goods sold under LIFO minus the increase in the LIFO reserve. Therefore, for the year ended 31 December 2018, Karp's cost of goods sold under FIFO equals:

$$\begin{aligned}\text{COGS (FIFO method)} &= \text{COGS (LIFO method)} - \text{Increase in LIFO reserve}\\ &= \$2{,}211 \text{ million} - (155 \text{ million} - 117 \text{ million})\\ &= \$2{,}173 \text{ million}\end{aligned}$$

39 A is correct. Karp's net income (NI) under FIFO equals Karp's net income under LIFO plus the after-tax increase in the LIFO reserve. For the year ended 31 December 2018, Karp's net income under FIFO equals:

$$\begin{aligned}\text{NI (FIFO method)} &= \text{NI (LIFO method)} + \text{Increase in LIFO reserve} \times (1 - \text{Tax rate})\\ &= \$247 \text{ million} + 38 \text{ million} \times (1 - 20\%)\\ &= \$277.4 \text{ million}\end{aligned}$$

Therefore, the increase in net income is:

$$\begin{aligned}\text{Increase in NI} &= \text{NI (FIFO method)} - \text{NI (LIFO method)}\\ &= \$277 \text{ million} - 247 \text{ million}\\ &= \$30.4 \text{ million}\end{aligned}$$

40 B is correct. Karp's retained earnings (RE) under FIFO equals Karp's retained earnings under LIFO plus the after-tax LIFO reserve. Therefore, for the year ended 31 December 2018, Karp's retained earnings under FIFO equals:

$$\begin{aligned}\text{RE (FIFO method)} &= \text{RE (LIFO method)} + \text{LIFO reserve} \times (1 - \text{Tax rate})\\ &= \$787 \text{ million} + 155 \text{ million} \times (1 - 20\%)\\ &= \$911 \text{ million}\end{aligned}$$

Therefore, the increase in retained earnings is:

$$\begin{aligned}\text{Increase in RE} &= \text{RE (FIFO method)} - \text{RE (LIFO method)}\\ &= \$911 \text{ million} - 787 \text{ million}\\ &= \$124 \text{ million}\end{aligned}$$

41 A is correct. The cash ratio (cash and cash equivalents ÷ current liabilities) would be lower because cash would have been less under FIFO. Karp's income before taxes would have been higher under FIFO, and consequently taxes paid by Karp would have also been higher and cash would have been lower. There is no impact on current liabilities. Both Karp's current ratio and gross profit margin would have been higher if FIFO had been used. The current ratio would have been higher because inventory under FIFO increases by a larger amount than the cash decreases for taxes paid. Because the cost of goods sold under FIFO is lower than under LIFO, the gross profit margin would have been higher.

42 B is correct. If Karp had used FIFO instead of LIFO, the debt-to-equity ratio would have decreased. No change in debt would have occurred, but shareholders' equity would have increased as a result of higher retained earnings.

43 B is correct. Crux's adjusted inventory turnover ratio must be computed using cost of goods sold (COGS) under FIFO and excluding charges for increases in valuation allowances.

COGS (adjusted) = COGS (LIFO method) – Charges included in cost of goods sold for inventory write-downs – Change in LIFO reserve
= \$3,120 million – 13 million – (55 million – 72 million)
= \$3,124 million

Note: Minus the change in LIFO reserve is equivalent to plus the decrease in LIFO reserve. The adjusted inventory turnover ratio is computed using average inventory under FIFO.

Ending inventory (FIFO) = Ending inventory (LIFO) + LIFO reserve

Ending inventory 2018 (FIFO) = \$480 + 55 = \$535

Ending inventory 2017 (FIFO) = \$465 + 72 = \$537

Average inventory = (\$535 + 537)/2 = \$536

Therefore, adjusted inventory turnover ratio equals:

Inventory turnover ratio = COGS/Average inventory = \$3,124/\$536 = 5.83

44 B is correct. Rolby's adjusted net profit margin must be computed using net income (NI) under FIFO and excluding charges for increases in valuation allowances.

NI (adjusted) = NI (FIFO method) + Charges, included in cost of goods sold for inventory write-downs, after tax
= \$327 million + 15 million × (1 – 30%)
= \$337.5 million

Therefore, adjusted net profit margin equals:

Net profit margin = NI/Revenues = \$337.5/\$5,442 = 6.20%

45 A is correct. Mikko's adjusted debt-to-equity ratio is lower because the debt (numerator) is unchanged and the adjusted shareholders' equity (denominator) is higher. The adjusted shareholders' equity corresponds to shareholders' equity under FIFO, excluding charges for increases in valuation allowances. Therefore, adjusted shareholders' equity is higher than reported (unadjusted) shareholders' equity.

46 C is correct. Mikko's and Crux's gross margin ratios would better reflect the current gross margin of the industry than Rolby because both use LIFO. LIFO recognizes as cost of goods sold the cost of the most recently purchased units, therefore, it better reflects replacement cost. However, Mikko's gross margin ratio best reflects the current gross margin of the industry because Crux's LIFO reserve is decreasing. This could reflect a LIFO liquidation by Crux which would distort gross profit margin.

47 B is correct. The FIFO method shows a higher gross profit margin than the LIFO method in an inflationary scenario, because FIFO allocates to cost of goods sold the cost of the oldest units available for sale. In an inflationary environment, these units are the ones with the lowest cost.

48 A is correct. An inventory write-down increases cost of sales and reduces profit and reduces the carrying value of inventory and assets. This has a negative effect on profitability and solvency ratios. However, activity ratios appear positively affected by a write-down because the asset base, whether total assets or inventory (denominator), is reduced. The numerator, sales, in total asset turnover is unchanged, and the numerator, cost of sales, in inventory turnover is increased. Thus, turnover ratios are higher and appear more favorable as the result of the write-down.

49 B is correct. Finished goods least accurately reflect current prices because some of the finished goods are valued under the "last-in, first-out" ("LIFO") basis. The costs of the newest units available for sale are allocated to cost of goods sold, leaving the oldest units (at lower costs) in inventory. ZP values raw materials and work in process using the weighted average cost method. While not fully reflecting current prices, some inflationary effect will be included in the inventory values.

50 C is correct. FIFO inventory = Reported inventory + LIFO reserve = ¥608,572 + 10,120 = ¥618,692. The LIFO reserve is disclosed in Note 2 of the notes to consolidated financial statements.

51 A is correct. The inventory turnover ratio would be lower. The average inventory would be higher under FIFO and cost of products sold would be lower by the increase in LIFO reserve. LIFO is not permitted under IFRS.

Inventory turnover ratio = Cost of products sold ÷ Average inventory

2018 inventory turnover ratio as reported = 10.63 = ¥5,822,805/[(608,572 + 486,465)/2].

2018 inventory turnover ratio adjusted to FIFO as necessary = 10.34 = [¥5,822,805 – (19,660 – 10,120)]/[(608,572 + 10,120 + 486,465 + 19,660)/2].

52 A is correct. No LIFO liquidation occurred during 2018; the LIFO reserve increased from ¥10,120 million in 2008 to ¥19,660 million in 2018. Management stated in the MD&A that the decrease in inventories reflected the impacts of decreased sales volumes and fluctuations in foreign currency translation rates.

53 C is correct. Finished goods and raw materials inventories are lower in 2018 when compared to 2017. Reduced levels of inventory typically indicate an anticipated business contraction.

54 B is correct. The decrease in LIFO inventory in 2018 would typically indicate that more inventory units were sold than produced or purchased. Accordingly, one would expect a liquidation of some of the older LIFO layers and the LIFO reserve to decrease. In actuality, the LIFO reserve *increased* from ¥10,120 million in 2017 to ¥19,660 million in 2018. This is not to be expected and is likely caused by the increase in prices of raw materials, other production materials, and parts of foreign currencies as noted in the MD&A. An analyst should seek to confirm this explanation.

55 B is correct. If prices have been decreasing, write-downs under FIFO are least likely to have a significant effect because the inventory is valued at closer to the new, lower prices. Typically, inventories valued using LIFO are less likely to incur inventory write-downs than inventories valued using weighted average cost or FIFO. Under LIFO, the *oldest* costs are reflected in the inventory carrying value on the balance sheet. Given increasing inventory costs, the inventory carrying values under the LIFO method are already conservatively presented at the oldest and lowest costs. Thus, it is far less likely that inventory write-downs will occur under LIFO; and if a write-down does occur, it is likely to be of a lesser magnitude.

READING
22

Long-Lived Assets

by Elaine Henry, PhD, CFA, and Elizabeth A. Gordon, PhD, MBA, CPA

Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). Elizabeth A. Gordon, PhD, MBA, CPA, is at Temple University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** identify and contrast costs that are capitalised and costs that are expensed in the period in which they are incurred;
☐	**b.** compare the financial reporting of the following types of intangible assets: purchased, internally developed, acquired in a business combination;
☐	**c.** explain and evaluate how capitalising versus expensing costs in the period in which they are incurred affects financial statements and ratios;
☐	**d.** describe the different depreciation methods for property, plant, and equipment and calculate depreciation expense;
☐	**e.** describe how the choice of depreciation method and assumptions concerning useful life and residual value affect depreciation expense, financial statements, and ratios;
☐	**f.** describe the different amortisation methods for intangible assets with finite lives and calculate amortisation expense;
☐	**g.** describe how the choice of amortisation method and assumptions concerning useful life and residual value affect amortisation expense, financial statements, and ratios;
☐	**h.** describe the revaluation model;
☐	**i.** explain the impairment of property, plant, and equipment and intangible assets;
☐	**j.** explain the derecognition of property, plant, and equipment and intangible assets;
☐	**k.** explain and evaluate how impairment, revaluation, and derecognition of property, plant, and equipment and intangible assets affect financial statements and ratios;

(continued)

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**l.** describe the financial statement presentation of and disclosures relating to property, plant, and equipment and intangible assets;
☐	**m.** analyze and interpret financial statement disclosures regarding property, plant, and equipment and intangible assets;
☐	**n.** compare the financial reporting of investment property with that of property, plant, and equipment.

1 INTRODUCTION & ACQUISITION OF PROPERTY, PLANT AND EQUIPMENT

a identify and contrast costs that are capitalised and costs that are expensed in the period in which they are incurred;

1.1 Introduction

Long-lived assets, also referred to as non-current assets or long-term assets, are assets that are expected to provide economic benefits over a future period of time, typically greater than one year.[1] Long-lived assets may be tangible, intangible, or financial assets. Examples of long-lived tangible assets, typically referred to as **property, plant, and equipment** and sometimes as fixed assets, include land, buildings, furniture and fixtures, machinery and equipment, and vehicles; examples of long-lived **intangible assets** (assets lacking physical substance) include patents and trademarks; and examples of long-lived financial assets include investments in equity or debt securities issued by other entities. The scope of this reading is limited to long-lived tangible and intangible assets (hereafter, referred to for simplicity as long-lived assets).

The first issue in accounting for a long-lived asset is determining its cost at acquisition. The second issue is how to allocate the cost to expense over time. The costs of most long-lived assets are capitalised and then allocated as expenses in the profit or loss (income) statement over the period of time during which they are expected to provide economic benefits. The two main types of long-lived assets with costs that are typically *not* allocated over time are land, which is not depreciated, and those intangible assets with indefinite useful lives. Additional issues that arise are the treatment of subsequent costs incurred related to the asset, the use of the cost model versus the revaluation model, unexpected declines in the value of the asset, classification of the asset with respect to intent (for example, held for use or held for sale), and the derecognition of the asset.

This reading is organised as follows. Sections 1–5 describe and illustrate accounting for the acquisition of long-lived assets, with particular attention to the impact of capitalizing versus expensing expenditures. Sections 6–7 describe the allocation of the costs of long-lived assets over their useful lives. Section 8 discusses the revaluation model that is based on changes in the fair value of an asset. Section 9 covers the concepts of impairment (unexpected decline in the value of an asset). Section 10 describes accounting for the derecognition of long-lived assets. Sections 11–12 describe financial statement presentation, disclosures, and analysis of long-lived assets. Section 13 discusses differences in financial reporting of investment property compared with property, plant, and equipment. A summary is followed by practice problems.

---

[1] In some industries, inventory is held longer than one year but is nonetheless reported as a current asset.

1.2 Acquisition of Long-Lived Assets

Upon acquisition, property, plant, and equipment (tangible assets with an economic life of longer than one year and intended to be held for the company's own use) are recorded on the balance sheet at cost, which is typically the same as their fair value.[2] Accounting for an intangible asset depends on how the asset is acquired. If several assets are acquired as part of a group, the purchase price is allocated to each asset on the basis of its fair value. An asset's cost potentially includes expenditures additional to the purchase price.

A key concept in accounting for expenditures related to long-lived assets is whether and when such expenditures are capitalised (i.e., included in the asset shown on the balance sheet) versus expensed (i.e., treated as an expense of the period on the income statement). After examining the specific treatment of certain expenditures, we will consider the general financial statement impact of capitalising versus expensing and two analytical issues related to the decision—namely, the effects on an individual company's trend analysis and on comparability across companies.

1.3 Property, Plant, and Equipment

This section primarily discusses the accounting treatment for the acquisition of long-lived tangible assets (property, plant, and equipment) through purchase. Assets can be acquired by methods other than purchase.[3] When an asset is exchanged for another asset, the asset acquired is recorded at fair value if reliable measures of fair value exist. Fair value is the fair value of the asset given up unless the fair value of the asset acquired is more clearly evident. If there is no reliable measure of fair value, the acquired asset is measured at the carrying amount of the asset given up. In this case, the carrying amount of the assets is unchanged, and no gain or loss is reported.

Typically, accounting for the exchange involves removing the carrying amount of the asset given up, adding a fair value for the asset acquired, and reporting any difference between the carrying amount and the fair value as a gain or loss. A gain would be reported when the fair value used for the newly acquired asset exceeds the carrying amount of the asset given up. A loss would be reported when the fair value used for the newly acquired asset is less than the carrying amount of the asset given up.

When property, plant, or equipment is purchased, the buyer records the asset at cost. In addition to the purchase price, the buyer also includes, as part of the cost of an asset, all the expenditures necessary to get the asset ready for its intended use. For example, freight costs borne by the purchaser to get the asset to the purchaser's place of business and special installation and testing costs required to make the asset usable are included in the total cost of the asset.

---

[2] Fair value is defined in International Financial Reporting Standards (IFRS) and under US generally accepted accounting principles (US GAAP) in the Financial Accounting Standards Board (FASB) Accounting Standards Codification (ASC) as "the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants at the measurement date." [IFRS 13 and FASB ASC Topic 820]

[3] IAS 16 *Property, Plant and Equipment*, paragraphs 24–26 [Measurement of Cost]; IAS 38 *Intangible Assets*, paragraphs 45–47 [Exchange of Assets]; and FASB ASC Section 845-10-30 [Nonmonetary Transactions – Overall – Initial Measurement].

Subsequent expenditures related to long-lived assets are included as part of the recorded value of the assets on the balance sheet (i.e., capitalised) if they are expected to provide benefits beyond one year in the future and are expensed if they are not expected to provide benefits in future periods. Expenditures that extend the original life of the asset are typically capitalised. Example 1 illustrates the difference between costs that are capitalised and costs that are expensed in a period.

EXAMPLE 1

Acquisition of PPE

Assume a (hypothetical) company, Trofferini S.A., incurred the following expenditures to purchase a towel and tissue roll machine: €10,900 purchase price including taxes, €200 for delivery of the machine, €300 for installation and testing of the machine, and €100 to train staff on maintaining the machine. In addition, the company paid a construction team €350 to reinforce the factory floor and ceiling joists to accommodate the machine's weight. The company also paid €1,500 to repair the factory roof (a repair expected to extend the useful life of the factory by five years) and €1,000 to have the exterior of the factory and adjoining offices repainted for maintenance reasons. The repainting neither extends the life of factory and offices nor improves their usability.

1. Which of these expenditures will be capitalised and which will be expensed?
2. How will the treatment of these expenditures affect the company's financial statements?

Solution to 1:

The company will capitalise as part of the cost of the machine all costs that are necessary to get the new machine ready for its intended use: €10,900 purchase price, €200 for delivery, €300 for installation and testing, and €350 to reinforce the factory floor and ceiling joists to accommodate the machine's weight (which was necessary to use the machine and does not increase the value of the factory). The €100 to train staff is not necessary to get the asset ready for its intended use and will be expensed.

The company will capitalise the expenditure of €1,500 to repair the factory roof because the repair is expected to extend the useful life of the factory. The company will expense the €1,000 to have the exterior of the factory and adjoining offices repainted because the painting does not extend the life or alter the productive capacity of the buildings.

Solution to 2:

The costs related to the machine that are capitalised—€10,900 purchase price, €200 for delivery, €300 for installation and testing, and €350 to prepare the factory—will increase the carrying amount of the machine asset as shown on the balance sheet and will be included as investing cash outflows. The item related to the factory that is capitalised—the €1,500 roof repair—will increase the carrying amount of the factory asset as shown on the balance sheet and is an investing cash outflow. The expenditures of €100 to train staff and €1,000 to paint are expensed in the period and will reduce the amount of income reported on the company's income statement (and thus reduce retained earnings on the balance sheet) and the operating cash flow.

Example 1 describes capitalising versus expensing in the context of purchasing property, plant, and equipment. When a company constructs an asset (or acquires an asset that requires a long period of time to get ready for its intended use), borrowing costs incurred directly related to the construction are generally capitalised. Constructing a building, whether for sale (in which case, the building is classified as inventory) or for the company's own use (in which case, the building is classified as a long-lived asset), typically requires a substantial amount of time. To finance construction, any borrowing costs incurred prior to the asset being ready for its intended use are capitalised as part of the cost of the asset. The company determines the interest rate to use on the basis of its existing borrowings or, if applicable, on a borrowing specifically incurred for constructing the asset. If a company takes out a loan specifically to construct a building, the interest cost on that loan during the time of construction would be capitalised as part of the building's cost. Under IFRS, but not under US GAAP, income earned on temporarily investing the borrowed monies decreases the amount of borrowing costs eligible for capitalisation.

Thus, a company's interest costs for a period are included either on the balance sheet (to the extent they are capitalised as part of an asset) or on the income statement (to the extent they are expensed). If the interest expenditure is incurred in connection with constructing an asset for the company's own use, the capitalised interest appears on the balance sheet as a part of the relevant long-lived asset (i.e., property, plant, and equipment). The capitalised interest is expensed over time as the property is depreciated and is thus part of subsequent years' depreciation expense rather than interest expense of the current period. If the interest expenditure is incurred in connection with constructing an asset to sell (for example, by a home builder), the capitalised interest appears on the company's balance sheet as part of inventory. The capitalised interest is expensed as part of the cost of goods sold when the asset is sold. Interest payments made prior to completion of construction that are capitalised are classified as an investing cash outflow. Expensed interest may be classified as an operating or financing cash outflow under IFRS and is classified as an operating cash outflow under US GAAP.

EXAMPLE 2

Capitalised Borrowing Costs

BILDA S.A., a hypothetical company, borrows €1,000,000 at an interest rate of 10 percent per year on 1 January 2010 to finance the construction of a factory that will have a useful life of 40 years. Construction is completed after two years, during which time the company earns €20,000 by temporarily investing the loan proceeds.

1 What is the amount of interest that will be capitalised under IFRS, and how would that amount differ from the amount that would be capitalised under US GAAP?

2 Where will the capitalised borrowing cost appear on the company's financial statements?

Solution to 1:

The total amount of interest paid on the loan during construction is €200,000 (= €1,000,000 × 10% × 2 years). Under IFRS, the amount of borrowing cost eligible for capitalisation is reduced by the €20,000 interest income from temporarily investing the loan proceeds, so the amount to be capitalised is €180,000. Under US GAAP, the amount to be capitalised is €200,000.

Solution to 2:

The capitalised borrowing costs will appear on the company's balance sheet as a component of property, plant, and equipment. In the years prior to completion of construction, the interest paid will appear on the statement of cash flows as an investment activity. Over time, as the property is depreciated, the capitalised interest component is part of subsequent years' depreciation expense on the company's income statement.

2

ACQUISITION OF INTANGIBLE ASSETS

b compare the financial reporting of the following types of intangible assets: purchased, internally developed, acquired in a business combination;

Intangible assets are assets lacking physical substance. Intangible assets include items that involve exclusive rights, such as patents, copyrights, trademarks, and franchises. Under IFRS, identifiable intangible assets must meet three definitional criteria. They must be (1) identifiable (either capable of being separated from the entity or arising from contractual or legal rights), (2) under the control of the company, and (3) expected to generate future economic benefits. In addition, two recognition criteria must be met: (1) It is probable that the expected future economic benefits of the asset will flow to the company, and (2) the cost of the asset can be reliably measured. Goodwill, which is not considered an identifiable intangible asset,[4] arises when one company purchases another and the acquisition price exceeds the fair value of the net identifiable assets (both the tangible assets and the identifiable intangible assets, minus liabilities) acquired.

Accounting for an intangible asset depends on how it is acquired. The following sections describe accounting for intangible assets obtained in three ways: purchased in situations other than business combinations, developed internally, and acquired in business combinations.

2.1 Intangible Assets Purchased in Situations Other Than Business Combinations

Intangible assets purchased in situations other than business combinations, such as buying a patent, are treated at acquisition the same as long-lived tangible assets; they are recorded at their fair value when acquired, which is assumed to be equivalent to the purchase price. If several intangible assets are acquired as part of a group, the purchase price is allocated to each asset on the basis of its fair value.

In deciding how to treat individual intangible assets for analytical purposes, analysts are particularly aware that companies must use a substantial amount of judgment and numerous assumptions to determine the fair value of individual intangible assets. For analysis, therefore, understanding the types of intangible assets acquired can often be more useful than focusing on the values assigned to the individual assets. In other words, an analyst would typically be more interested in understanding what assets a company acquired (for example, franchise rights) than in the precise portion of the purchase price a company allocated to each asset. Understanding the types of assets a company acquires can offer insights into the company's strategic direction and future operating potential.

---

[4] The IFRS definition of an intangible asset as an "identifiable non-monetary asset without physical substance" applies to intangible assets not specifically dealt with in standards other than IAS 38. The definition of intangible assets under US GAAP—"assets (other than financial assets) that lack physical substance"—includes goodwill in the definition of an intangible asset.

2.2 Intangible Assets Developed Internally

In contrast with the treatment of construction costs of tangible assets, the costs to internally develop intangible assets are generally expensed when incurred. There are some situations, however, in which the costs incurred to internally develop an intangible asset are capitalised. The general analytical issues related to the capitalising-versus-expensing decision apply here—namely, comparability across companies and the effect on an individual company's trend analysis.

The general requirement that costs to internally develop intangible assets be expensed should be compared with capitalising the cost of acquiring intangible assets in situations other than business combinations. Because costs associated with internally developing intangible assets are usually expensed, a company that has internally developed such intangible assets as patents, copyrights, or brands through expenditures on R&D or advertising will recognise a lower amount of assets than a company that has obtained intangible assets through external purchase. In addition, on the statement of cash flows, costs of internally developing intangible assets are classified as operating cash outflows whereas costs of acquiring intangible assets are classified as investing cash outflows. Differences in strategy (developing versus acquiring intangible assets) can thus impact financial ratios.

IFRS require that expenditures on research (or during the research phase of an internal project) be expensed rather than capitalised as an intangible asset.[5] Research is defined as "original and planned investigation undertaken with the prospect of gaining new scientific or technical knowledge and understanding."[6] The "research phase of an internal project" refers to the period during which a company cannot demonstrate that an intangible asset is being created—for example, the search for alternative materials or systems to use in a production process. In contrast with the treatment of research-phase expenditures, IFRS allow companies to recognise an intangible asset arising from development expenditures (or the development phase of an internal project) if certain criteria are met, including a demonstration of the technical feasibility of completing the intangible asset and the intent to use or sell the asset. Development is defined as "the application of research findings or other knowledge to a plan or design for the production of new or substantially improved materials, devices, products, processes, systems or services before the start of commercial production or use."[7]

Generally, US GAAP require that both research and development costs be expensed as incurred but require capitalisation of certain costs related to software development.[8] Costs incurred to develop a software product for sale are expensed until the product's technological feasibility is established and are capitalised thereafter. Similarly, companies expense costs related to the development of software for internal use until it is probable that the project will be completed and that the software will be used as intended. Thereafter, development costs are capitalised. The probability that the project will be completed is easier to demonstrate than is technological feasibility. The capitalised costs, related directly to developing software for sale or internal use, include the costs of employees who help build and test the software. The treatment of software development costs under US GAAP is similar to the treatment of all costs of internally developed intangible assets under IFRS.

---

[5] IAS 38 *Intangible Assets*.

[6] IAS 38 *Intangible Assets*, paragraph 8 [Definitions].

[7] IAS 38 *Intangible Assets*, paragraph 8 [Definitions].

[8] FASB ASC Section 350-40-25 [Intangibles—Goodwill and Other – Internal-Use Software – Recognition] and FASB ASC Section 985-20-25 [Software – Costs of Software to be Sold, Leased, or Marketed – Recognition] specify US GAAP accounting for software development costs for software for internal use and for software to be sold, respectively.

EXAMPLE 3

Software Development Costs

Assume REH AG, a hypothetical company, incurs expenditures of €1,000 per month during the fiscal year ended 31 December 2019 to develop software for internal use. Under IFRS, the company must treat the expenditures as an expense until the software meets the criteria for recognition as an intangible asset, after which time the expenditures can be capitalised as an intangible asset.

1. What is the accounting impact of the company being able to demonstrate that the software met the criteria for recognition as an intangible asset on 1 February versus 1 December?
2. How would the treatment of expenditures differ if the company reported under US GAAP and it had established in 2018 that the project was likely to be completed and the software used to perform the function intended?

Solution to 1:

If the company is able to demonstrate that the software met the criteria for recognition as an intangible asset on 1 February, the company would recognise the €1,000 expended in January as an expense on the income statement for the fiscal year ended 31 December 2019. The other €11,000 of expenditures would be recognised as an intangible asset (on the balance sheet). Alternatively, if the company is not able to demonstrate that the software met the criteria for recognition as an intangible asset until 1 December, the company would recognise the €11,000 expended in January through November as an expense on the income statement for the fiscal year ended 31 December 2019, with the other €1,000 of expenditures recognised as an intangible asset.

Solution to 2:

Under US GAAP, the company would capitalise the entire €12,000 spent to develop software for internal use.

2.3 Intangible Assets Acquired in a Business Combination

When one company acquires another company, the transaction is accounted for using the **acquisition method** of accounting.[9] Under the acquisition method, the company identified as the acquirer allocates the purchase price to each asset acquired (and each liability assumed) on the basis of its fair value. If the purchase price exceeds the sum of the amounts that can be allocated to individual identifiable assets and liabilities, the excess is recorded as goodwill. Goodwill cannot be identified separately from the business as a whole.

---

[9] Both IFRS and US GAAP require the use of the acquisition method in accounting for business combinations (IFRS 3 and FASB ASC Section 805).

Under IFRS, the acquired individual assets include identifiable intangible assets that meet the definitional and recognition criteria.[10] Otherwise, if the item is acquired in a business combination and cannot be recognised as a tangible or identifiable intangible asset, it is recognised as goodwill. Under US GAAP, there are two criteria to judge whether an intangible asset acquired in a business combination should be recognised separately from goodwill: The asset must be either an item arising from contractual or legal rights or an item that can be separated from the acquired company. Examples of intangible assets treated separately from goodwill include the intangible assets previously mentioned that involve exclusive rights (patents, copyrights, franchises, licenses), as well as such items as internet domain names and video and audiovisual materials.

Exhibit 1 describes how AB InBev allocated the $103 billion purchase consideration in its 2016 acquisition of SABMiller Group. The combined company was renamed Anheuser-Busch InBev SA/NV. The majority of the intangible asset valuation relates to brands with indefinite life ($19.9 billion of the $20.0 billion total). Of $63.0 billion total assets acquired, assets to be divested were valued at $24.8 billion and assets to be held for were valued at $38.2 billion. In total, intangible assets represent 52 percent of the total assets to be held for use. In addition, $74.1 billion of goodwill was recognized in the transaction.

Exhibit 1 Acquisition of Intangible Assets through a Business Combination

Excerpt from the 2016 annual report of AB InBev:

> "On 10 October 2016, AB InBev announced the … successful completion of the business combination with the former SABMiller Group ("SAB").
>
> "The transaction resulted in 74.1 billion US dollar of goodwill provisionally allocated primarily to the businesses in Colombia, Ecuador, Peru, Australia, South Africa and other African, Asia Pacific and Latin American countries. The factors that contributed to the recognition of goodwill include the acquisition of an assembled workforce and the premiums paid for cost synergies expected to be achieved in SABMiller. Management's assessment of the future economic benefits supporting recognition of this goodwill is in part based on expected savings through the implementation of AB InBev best practices such as, among others, a zero based budgeting program and initiatives that are expected to bring greater efficiency and standardization, generate cost savings and maximize purchasing power. Goodwill also arises due to the recognition of deferred tax liabilities in relation to the preliminary fair value adjustments on acquired intangible assets for which the amortization does not qualify as a tax deductible expense. None of the goodwill recognized is deductible for tax purposes.
>
> "The majority of the intangible asset valuation relates to brands with indefinite life, valued for a total amount of 19.9 billion US dollar. The valuation of the brands with indefinite life is based on a series of factors, including the brand history, the operating plan and the countries in which the brands are sold. The fair value of brands was estimated by applying a combination of known valuation methodologies, such as the royalty relief and excess earnings valuation approaches.

(continued)

---

[10] As previously described, the definitional criteria are identifiability, control by the company, and expected future benefits. The recognition criteria are probable flows of the expected economic benefits to the company and measurability.

Exhibit 1 (Continued)

> “The intangibles with an indefinite life mainly include the Castle and Carling brand families in Africa, the Aguila and Poker brand families in Colombia, the Cristal and Pilsner brand families in Ecuador, and the Carlton brand family in Australia.
>
> “Assets held for sale were recognized in relation to the divestiture of SABMiller’s interests in the MillerCoors LLC joint venture and certain of SABMiller’s portfolio of Miller brands outside of the U.S. to Molson Coors Brewing company; the divestiture of SABMiller’s European premium brands to Asahi Group Holdings, Ltd and the divestiture of SABMiller’s interest in China Resources Snow Breweries Ltd. to China Resources Beer (Holdings) Co. Ltd. ….” [Excerpt]

The following is a summary of the provisional allocation of AB InBev’s purchase price of SABMiller:

Assets	**$ million**
Property, plant and equipment	9,060
Intangible assets	20,040
Investment in associates	4,386
Inventories	977
Trade and other receivables	1,257
Cash and cash equivalents	1,410
Assets held for sale	24,805
All other assets	*1,087*
Total assets	*63,022*
Total liabilities	*–27,769*
Net identified assets and liabilities	**35,253**
Non-controlling interests	–6,200
Goodwill on acquisition	**74,083**
Purchase consideration	**103,136**

Table is excerpted from the company’s 2016 Annual Report. Portions of detail are omitted, and subtotals are shown in italics.

Source: AB InBev 2016 Annual Report, pp. 82-85.

3 CAPITALIZATION VERSUS EXPENSING: IMPACT ON FINANCIAL STATEMENTS AND RATIOS

c explain and evaluate how capitalising versus expensing costs in the period in which they are incurred affects financial statements and ratios;

This section discusses the implications for financial statements and ratios of capitalising versus expensing costs in the period in which they are incurred. We first summarize the general financial statement impact of capitalising versus expensing and two analytical issues related to the decision—namely the effect on an individual company’s trend analysis and on comparability across companies.

In the period of the expenditure, an expenditure that is capitalised increases the amount of assets on the balance sheet and appears as an investing cash outflow on the statement of cash flows. After initial recognition, a company allocates the capitalised amount over the asset's useful life as depreciation or amortisation expense (except assets that are not depreciated, i.e., land, or amortised, e.g., intangible assets with indefinite lives). This expense reduces net income on the income statement and reduces the value of the asset on the balance sheet. Depreciation and amortisation are non-cash expenses and therefore, apart from their effect on taxable income and taxes payable, have no impact on the cash flow statement. In the section of the statement of cash flows that reconciles net income to operating cash flow, depreciation and amortisation expenses are added back to net income.

Alternatively, an expenditure that is expensed reduces net income by the after-tax amount of the expenditure in the period it is made. No asset is recorded on the balance sheet and thus no depreciation or amortisation occurs in subsequent periods. The lower amount of net income is reflected in lower retained earnings on the balance sheet. An expenditure that is expensed appears as an operating cash outflow in the period it is made. There is no effect on the financial statements of subsequent periods.

Example 4 illustrates the impact on the financial statements of capitalising versus expensing an expenditure.

EXAMPLE 4

General Financial Statement Impact of Capitalising Versus Expensing

Assume two identical (hypothetical) companies, CAP Inc. (CAP) and NOW Inc. (NOW), start with €1,000 cash and €1,000 common stock. Each year the companies recognise total revenues of €1,500 cash and make cash expenditures, excluding an equipment purchase, of €500. At the beginning of operations, each company pays €900 to purchase equipment. CAP estimates the equipment will have a useful life of three years and an estimated salvage value of €0 at the end of the three years. NOW estimates a much shorter useful life and expenses the equipment immediately. The companies have no other assets and make no other asset purchases during the three-year period. Assume the companies pay no dividends, earn zero interest on cash balances, have a tax rate of 30 percent, and use the same accounting method for financial and tax purposes.

The left side of Exhibit 2 shows CAP's financial statements; i.e., with the expenditure capitalised and depreciated at €300 per year based on the straight-line method of depreciation (€900 cost minus €0 salvage value equals €900, divided by a three-year life equals €300 per year). The right side of the exhibit shows NOW's financial statements, with the entire €900 expenditure treated as an expense in the first year. All amounts are in euro.

Exhibit 2 Capitalising versus Expensing

CAP Inc.				NOW Inc.			
Capitalise €900 as asset and depreciate				**Expense €900 immediately**			
For Year	**1**	**2**	**3**	**For Year**	**1**	**2**	**3**
Revenue	1,500	1,500	1,500	Revenue	1,500	1,500	1,500
Cash expenses	500	500	500	Cash expenses	1,400	500	500
Depreciation	300	300	300	Depreciation	0	0	0

(continued)

Exhibit 2 (Continued)

CAP Inc.

Capitalise €900 as asset and depreciate

For Year	1	2	3
Income before tax	700	700	700
Tax at 30%	210	210	210
Net income	490	490	490
Cash from operations	790	790	790
Cash used in investing	(900)	0	0
Total change in cash	(110)	790	790

NOW Inc.

Expense €900 immediately

For Year	1	2	3
Income before tax	100	1,000	1,000
Tax at 30%	30	300	300
Net income	70	700	700
Cash from operations	70	700	700
Cash used in investing	0	0	0
Total change in cash	70	700	700

As of	Time 0	End of Year 1	End of Year2	End of Year 3
Cash	1,000	890	1,680	2,470
PP & E (net)	—	600	300	—
Total Assets	1,000	1,490	1,980	2,470
Retained earnings	0	490	980	1,470
Common stock	1,000	1,000	1,000	1,000
Total shareholders' equity	1,000	1,490	1,980	2,470

Time	Time 0	End of Year 1	End of Year 2	End of Year 3
Cash	1,000	1,070	1,770	2,470
PP & E (net)	—	—	—	—
Total Assets	1,000	1,070	1,770	2,470
Retained earnings	0	70	770	1,470
Common stock	1,000	1,000	1,000	1,000
Total shareholders' equity	1,000	1,070	1,770	2,470

1 Which company reports higher net income over the three years? Total cash flow? Cash from operations?

2 Based on ROE and net profit margin, how does the profitability of the two companies compare?

3 Why does NOW report change in cash of €70 in Year 1, while CAP reports total change in cash of (€110)?

Solution to 1:

Neither company reports higher total net income or cash flow over the three years. The sum of net income over the three years is identical (€1,470 total) whether the €900 is capitalised or expensed. Also, the sum of the change in cash (€1,470 total) is identical under either scenario. CAP reports higher cash from operations by an amount of €900 because, under the capitalisation scenario, the €900 purchase is treated as an investing cash flow.

Note: Because the companies use the same accounting method for both financial and taxable income, absent the assumption of zero interest on cash balances, expensing the €900 would have resulted in higher income and cash flow for NOW because the lower taxes paid in the first year (€30 versus €210) would have allowed NOW to earn interest income on the tax savings.

Solution to 2:

In general, Ending shareholders' equity = Beginning shareholders' equity + Net income + Other comprehensive income − Dividends + Net capital contributions from shareholders. Because the companies in this example do not have other comprehensive income, did not pay dividends, and reported no capital contributions from shareholders, Ending retained earnings = Beginning retained earnings + Net income, and Ending shareholders' equity = Beginning shareholders' equity + Net income.

ROE is calculated as Net income divided by Average shareholders' equity, and Net profit margin is calculated as Net income divided by Total revenue. For example, CAP had Year 1 ROE of 39 percent (€490/[(€1,000 + €1,490)/2]), and Year 1 net profit margin of 33 percent (€490/€1,500).

CAP Inc.				**NOW Inc.**			
Capitalise €900 as asset and depreciate				**Expense €900 immediately**			
For year	**1**	**2**	**3**	**For year**	**1**	**2**	**3**
ROE	39%	28%	22%	ROE	7%	49%	33%
Net profit margin	33%	33%	33%	Net profit margin	5%	47%	47%

As shown, compared to expensing, capitalising results in higher profitability ratios (ROE and net profit margin) in the first year, and lower profitability ratios in subsequent years. For example, CAP's Year 1 ROE of 39 percent was higher than NOW's Year 1 ROE of 7 percent, but in Years 2 and 3, NOW reports superior profitability.

Note also that NOW's superior growth in net income between Year 1 and Year 2 is not attributable to superior performance compared to CAP but rather to the accounting decision to recognise the expense sooner than CAP. In general, all else equal, accounting decisions that result in recognising expenses sooner will give the appearance of greater subsequent growth. Comparison of the growth of the two companies' net incomes without an awareness of the difference in accounting methods would be misleading. As a corollary, NOW's income and profitability exhibit greater volatility across the three years, not because of more volatile performance but rather because of the different accounting decision.

Solution to 3:

NOW reports an increase in cash of €70 in Year 1, while CAP reports a decrease in cash of €110 because NOW's taxes were €180 lower than CAP's taxes (€30 versus €210).

Note that this problem assumes the accounting method used by each company for its tax purposes is identical to the accounting method used by the company for its financial reporting. In many countries, companies are allowed to use different depreciation methods for financial reporting and taxes, which may give rise to deferred taxes.

As shown, discretion regarding whether to expense or capitalise expenditures can impede comparability across companies. Example 4 assumes the companies purchase a single asset in one year. Because the sum of net income over the three-year period is identical whether the asset is capitalised or expensed, it illustrates that although capitalising results in higher profitability compared to expensing in the first year, it results in lower profitability in the subsequent years. Conversely, expensing results in lower profitability in the first year but higher profitability in later years, indicating a favorable trend.

Similarly, shareholders' equity for a company that capitalises the expenditure will be higher in the early years because the initially higher profits result in initially higher retained earnings. Example 4 assumes the companies purchase a single asset in one year and report identical amounts of total net income over the three-year period, so shareholders' equity (and retained earnings) for the firm that expenses will be identical to shareholders' equity (and retained earnings) for the capitalising firm at the end of the three-year period.

Although Example 4 shows companies purchasing an asset only in the first year, if a company continues to purchase similar or increasing amounts of assets each year, the profitability-enhancing effect of capitalising continues if the amount of the expenditures in a period continues to be more than the depreciation expense. Example 5 illustrates this point.

EXAMPLE 5

Impact of Capitalising Versus Expensing for Ongoing Purchases

A company buys a £300 computer in Year 1 and capitalises the expenditure. The computer has a useful life of three years and an expected salvage value of £0, so the annual depreciation expense using the straight-line method is £100 per year. Compared to expensing the entire £300 immediately, the company's pre-tax profit in Year 1 is £200 greater.

1 Assume that the company continues to buy an identical computer each year at the same price. If the company uses the same accounting treatment for each of the computers, when does the profit-enhancing effect of capitalising versus expensing end?

2 If the company buys another identical computer in Year 4, using the same accounting treatment as the prior years, what is the effect on Year 4 profits of capitalising versus expensing these expenditures?

Solution to 1:

The profit-enhancing effect of capitalising versus expensing would end in Year 3. In Year 3, the depreciation expense on each of the three computers bought in Years 1, 2, and 3 would total £300 (£100 + £100 + £100). Therefore, the total depreciation expense for Year 3 will be exactly equal to the capital expenditure in Year 3. The expense in Year 3 would be £300, regardless of whether the company capitalised or expensed the annual computer purchases.

Solution to 2:

There is no impact on Year 4 profits. As in the previous year, the depreciation expense on each of the three computers bought in Years 2, 3, and 4 would total £300 (£100 + £100 + £100). Therefore, the total depreciation expense for Year 4 will be exactly equal to the capital expenditure in Year 4. Pre-tax profits would be reduced by £300, regardless of whether the company capitalised or expensed the annual computer purchases.

Compared to expensing an expenditure, capitalising the expenditure typically results in greater amounts reported as cash from operations. Capitalised expenditures are typically treated as an investment cash outflow whereas expenses reduce operating cash flows. Because cash flow from operating activities is an important consideration in some valuation models, companies may try to maximise reported cash flow from operations by capitalising expenditures that should be expensed. Valuation models that use free cash flow will consider not only operating cash flows but also investing cash flows. Analysts should be alert to evidence of companies manipulating reported cash flow from operations by capitalising expenditures that should be expensed.

In summary, holding all else constant, capitalising an expenditure enhances current profitability and increases reported cash flow from operations. The profitability-enhancing effect of capitalising continues so long as capital expenditures exceed the depreciation expense. Profitability-enhancing motivations for decisions to capitalise should be considered when analyzing performance. For example, a company may choose to capitalise more expenditures (within the allowable bounds of accounting standards) to achieve earnings targets for a given period. Expensing a cost in the period reduces current period profits but enhances future profitability and thus enhances the profit trend. Profit trend-enhancing motivations should also be considered when analyzing performance. If the company is in a reporting environment which requires identical accounting methods for financial reporting and taxes (unlike the United States, which permits companies to use depreciation methods for reporting purposes that differ from the depreciation method required by tax purposes), then expensing will have a more favorable cash flow impact because paying lower taxes in an earlier period creates an opportunity to earn interest income on the cash saved.

In contrast with the relatively simple examples above, it is generally neither possible nor desirable to identify individual instances involving discretion about whether to capitalise or expense expenditures. An analyst can, however, typically identify significant items of expenditure treated differently across companies. The items of expenditure giving rise to the most relevant differences across companies will vary by industry. This cross-industry variation is apparent in the following discussion of the capitalisation of expenditures.

CAPITALISATION OF INTEREST COSTS

4

c explain and evaluate how capitalising versus expensing costs in the period in which they are incurred affects financial statements and ratios;

As noted above, companies generally must capitalise interest costs associated with acquiring or constructing an asset that requires a long period of time to get ready for its intended use.[11]

As a consequence of this accounting treatment, a company's interest costs for a period can appear either on the balance sheet (to the extent they are capitalised) or on the income statement (to the extent they are expensed).

---

[11] IAS 23 [Borrowing Costs] and FASB ASC Subtopic 835-20 [Interest – Capitalization of Interest] specify respectively IFRS and US GAAP for capitalisation of interest costs. Although the standards are not completely converged, the standards are in general agreement.

If the interest expenditure is incurred in connection with constructing an asset for the company's own use, the capitalised interest appears on the balance sheet as a part of the relevant long-lived asset. The capitalised interest is expensed over time as the property is depreciated—and is thus part of depreciation expense rather than interest expense. If the interest expenditure is incurred in connection with constructing an asset to sell, for example by a real estate construction company, the capitalised interest appears on the company's balance sheet as part of inventory. The capitalised interest is then expensed as part of the cost of sales when the asset is sold.

The treatment of capitalised interest poses certain issues that analysts should consider. First, capitalised interest appears as part of investing cash outflows, whereas expensed interest typically reduces operating cash flow. US GAAP reporting companies are required to categorize interest in operating cash flow, and IFRS reporting companies can categorize interest in operating, investing, or financing cash flows. Although the treatment is consistent with accounting standards, an analyst may want to examine the impact on reported cash flows. Second, interest coverage ratios are solvency indicators measuring the extent to which a company's earnings (or cash flow) in a period covered its interest costs. To provide a true picture of a company's interest coverage, the entire amount of interest expenditure, both the capitalised portion and the expensed portion, should be used in calculating interest coverage ratios. Additionally, if a company is depreciating interest that it capitalised in a previous period, income should be adjusted to eliminate the effect of that depreciation. Example 6 illustrates the calculations.

EXAMPLE 6

Effect of Capitalised Interest Costs on Coverage Ratios and Cash Flow

Melco Resorts & Entertainment Limited (NASDAQ: MLCO), a Hong Kong SAR based casino company which is listed on the NASDAQ stock exchange and prepares financial reports under US GAAP, disclosed the following information in one of the footnotes to its 2017 financial statements: "Interest and amortization of deferred financing costs associated with major development and construction projects is capitalized and included in the cost of the project. …. Total interest expenses incurred amounted to $267,065, $252,600, and $253,168, of which $37,483, $29,033, and $134,838 were capitalized during the years ended December 31, 2017, 2016, and 2015, respectively. Amortization of deferred financing costs of $26,182, $48,345, and $38,511, net of amortization capitalized of nil, nil, and $5,458, were recorded during the years ended December 31, 2017, 2016, and 2015, respectively." (Form 20-F filed 12 April 2018). Cash payments for deferred financing costs were reported in cash flows from financing activities.

Exhibit 3 Melco Resorts & Entertainment Limited Selected Data, as Reported (Dollars in thousands)

	2017	2016	2015
EBIT (from income statement)	544,865	298,663	58,553
Interest expense (from income statement)	229,582	223,567	118,330
Capitalized interest (from footnote)	37,483	29,033	134,838
Amortization of deferred financing costs (from footnote)	26,182	48,345	38,511

Exhibit 3 (Continued)

	2017	2016	2015
Net cash provided by operating activities	1,162,500	1,158,128	522,026
Net cash from (used) in investing activities	(410,226)	280,604	(469,656)
Net cash from (used) in financing activities	(1,046,041)	(1,339,717)	(29,688)

Notes: EBIT represents "Income (Loss) Before Income Tax" plus "Interest expenses, net of capitalized interest" from the income statement.

1 Calculate and interpret Melco's interest coverage ratio with and without capitalised interest.

2 Calculate Melco's percentage change in operating cash flow from 2016 to 2017. Assuming the financial reporting does not affect reporting for income taxes, what were the effects of capitalised interest on operating and investing cash flows?

Solution to 1:

Interest coverage ratios with and without capitalised interest were as follows:

For 2017

2.37 ($544,865 ÷ $229,582) without adjusting for capitalised interest; and

2.14 [($544,865 + $26,182) ÷ ($229,582 + $37,483)] including an adjustment to EBIT for depreciation of previously capitalised interest and an adjustment to interest expense for the amount of interest capitalised in 2017.

For 2016

1.34 ($298,663÷ $223,567) without adjusting for capitalised interest; and

1.37 [($298,663 + $48,345) ÷ ($223,567 + $29,033)] including an adjustment to EBIT for depreciation of previously capitalised interest and an adjustment to interest expense for the amount of interest capitalised in 2016.

For 2015

0.49 ($58,533÷ $118,330) without adjusting for capitalised interest; and

0.38 [($58,533 + $38,511) ÷ ($118,330+ $134,838)] including an adjustment to EBIT for depreciation of previously capitalised interest and an adjustment to interest expense for the amount of interest capitalised in 2015.

The above calculations indicate that Melco's interest coverage improved in 2017 compared to the previous two years. In both 2017 and 2015, the coverage ratio is lower when adjusted for capitalised interest.

Solution to 2:

If the interest had been expensed rather than capitalised, operating cash flows would have been lower in all three years. On an adjusted basis, but not an unadjusted basis, the company's operating cash flow declined in 2017 compared to 2016. On an unadjusted basis, for 2017 compared with 2016, Melco's operating cash flow increased by 0.4 percent in 2017 [($1,162,500 ÷ $1,158,128) – 1]. Including adjustments to expense all interest costs, Melco's operating cash flow also decreased by 0.4 percent in 2017 {[$1,162,500 – $37,483) ÷ ($1,158,128 – $29,033)] – 1}.

If the interest had been expensed rather than capitalised, financing cash flows would have been higher in all three years.

The treatment of capitalised interest raises issues for consideration by an analyst. First, capitalised interest appears as part of investing cash outflows, whereas expensed interest reduces operating or financing cash flow under IFRS and operating cash flow under US GAAP. An analyst may want to examine the impact on reported cash flows of interest expenditures when comparing companies. Second, interest coverage ratios are solvency indicators measuring the extent to which a company's earnings (or cash flow) in a period covered its interest costs. To provide a true picture of a company's interest coverage, the entire amount of interest, both the capitalised portion and the expensed portion, should be used in calculating interest coverage ratios.

Generally, including capitalised interest in the calculation of interest coverage ratios provides a better assessment of a company's solvency. In assigning credit ratings, rating agencies include capitalised interest in coverage ratios. For example, Standard & Poor's calculates the EBIT interest coverage ratio as EBIT divided by gross interest (defined as interest prior to deductions for capitalised interest or interest income).

Maintaining a minimum interest coverage ratio is a financial covenant often included in lending agreements, e.g., bank loans and bond indentures. The definition of the coverage ratio can be found in the company's credit agreement. The definition is relevant because treatment of capitalised interest in calculating coverage ratios would affect an assessment of how close a company's actual ratios are to the levels specified by its financial covenants and thus the probability of breaching those covenants.

5

CAPITALISATION OF INTEREST AND INTERNAL DEVELOPMENT COSTS

c explain and evaluate how capitalising versus expensing costs in the period in which they are incurred affects financial statements and ratios;

As noted above, accounting standards require companies to capitalise software development costs after a product's feasibility is established. Despite this requirement, judgment in determining feasibility means that companies' capitalisation practices may differ. For example, as illustrated in Exhibit 4, Microsoft judges product feasibility to be established very shortly before manufacturing begins and, therefore, effectively expenses—rather than capitalises—research and development costs.

Exhibit 4 Disclosure on Software Development Costs

Excerpt from Management's Discussion and Analysis (MD&A) of Microsoft Corporation, Application of Critical Accounting Policies, Research and Development Costs:

> "Costs incurred internally in researching and developing a computer software product are charged to expense until technological feasibility has been established for the product. Once technological feasibility is established, all software costs are capitalized until the product is available for general release to customers. Judgment is required in determining when technological feasibility of a product is established. We have determined that technological feasibility for our software products is reached after all high-risk development issues have been

Exhibit 4 (Continued)

resolved through coding and testing. Generally, this occurs shortly before the products are released to production. The amortization of these costs is included in cost of revenue over the estimated life of the products."

Source: Microsoft Corporation Annual Report on Form 10-K 2017, p. 45.

Expensing rather than capitalising development costs results in lower net income in the current period. Expensing rather than capitalising will continue to result in lower net income so long as the amount of the current-period development expenses is higher than the amortisation expense that would have resulted from amortising prior periods' capitalised development costs—the typical situation when a company's development costs are increasing. On the statement of cash flows, expensing rather than capitalising development costs results in lower net operating cash flows and higher net investing cash flows. This is because the development costs are reflected as operating cash outflows rather than investing cash outflows.

In comparing the financial performance of a company that expenses most or all software development costs, such as Microsoft, with another company that capitalises software development costs, adjustments can be made to make the two comparable. For the company that capitalises software development costs, an analyst can adjust (a) the income statement to include software development costs as an expense and to exclude amortisation of prior years' software development costs; (b) the balance sheet to exclude capitalised software (decrease assets and equity); and (c) the statement of cash flows to decrease operating cash flows and decrease cash used in investing by the amount of the current period development costs. Any ratios that include income, long-lived assets, or cash flow from operations—such as return on equity—will also be affected.

EXAMPLE 7

Software Development Costs

You are working on a project involving the analysis of JHH Software, a (hypothetical) software development company that established technical feasibility for its first product in 2017. Part of your analysis involves computing certain market-based ratios, which you will use to compare JHH to another company that expenses all of its software development expenditures. Relevant data and excerpts from the company's annual report are included in Exhibit 5.

Exhibit 5 JHH SOFTWARE (Dollars in Thousands, Except Per-Share Amounts)

CONSOLIDATED STATEMENT OF EARNINGS—abbreviated

For year ended 31 December:	2018	2017	2016
Total revenue	$91,424	$91,134	$96,293
Total operating expenses	78,107	78,908	85,624
Operating income	13,317	12,226	10,669
Provision for income taxes	3,825	4,232	3,172

(continued)

Exhibit 5 (Continued)

CONSOLIDATED STATEMENT OF EARNINGS—abbreviated

For year ended 31 December:	2018	2017	2016
Net income	$9,492	$7,994	$7,479
Earnings per share (EPS)	$1.40	$0.82	$0.68

STATEMENT OF CASH FLOWS—abbreviated

For year ended 31 December:	2018	2017	2016
Net cash provided by operating activities	$15,007	$14,874	$15,266
Net cash used in investing activities*	(11,549)	(4,423)	(5,346)
Net cash used in financing activities	(8,003)	(7,936)	(7,157)
Net change in cash and cash equivalents	($4,545)	$2,515	$2,763
**Includes software development expenses of and includes capital expenditures of*	*($6,000)* *($2,000)*	*($4,000)* *($1,600)*	*($2,000)* *($1,200)*

Additional information:

For year ended 31 December:	2018	2017	2016
Market value of outstanding debt	0	0	0
Amortisation of capitalised software development expenses	($2,000)	($667)	0
Depreciation expense	($2,200)	($1,440)	($1,320)
Market price per share of common stock	$42	$26	$17
Shares of common stock outstanding (thousands)	6,780	9,765	10,999

Footnote disclosure of accounting policy for software development:
Expenses that are related to the conceptual formulation and design of software products are expensed to research and development as incurred. The company capitalises expenses that are incurred to produce the finished product after technological feasibility has been established.

1 Compute the following ratios for JHH based on the reported financial statements for fiscal year ended 31 December 2018, with no adjustments. Next, determine the approximate impact on these ratios if the company had expensed rather than capitalised its investments in software. (Assume the financial reporting does not affect reporting for income taxes. There would be no change in the effective tax rate.)

A P/E: Price/Earnings per share

B P/CFO: Price/Operating cash flow per share

C EV/EBITDA: Enterprise value/EBITDA, where enterprise value is defined as the total market value of all sources of a company's financing, including equity and debt, and EBITDA is earnings before interest, taxes, depreciation, and amortisation.

2 Interpret the changes in the ratios.

Solution to 1:

(Dollars are in thousands, except per-share amounts.) JHH's 2019 ratios are presented in the following table:

	Ratios	As reported	As adjusted
A	P/E ratio	30.0	42.9
B	P/CFO	19.0	31.6
C	EV/EBITDA	16.3	24.7

A Based on the information as reported, the P/E ratio was 30.0 ($42 ÷ $1.40). Based on EPS adjusted to expense software development costs, the P/E ratio was 42.9 ($42 ÷ $0.98).

- Price: Assuming that the market value of the company's equity is based on its fundamentals, the price per share is $42, regardless of a difference in accounting.
- EPS: As reported, EPS was $1.40. Adjusted EPS was $0.98. Expensing software development costs would have reduced JHH's 2018 operating income by $6,000, but the company would have reported no amortisation of prior years' software costs, which would have increased operating income by $2,000. The net change of $4,000 would have reduced operating income from the reported $13,317 to $9,317. The effective tax rate for 2018 ($3,825 ÷ $13,317) is 28.72%, and using this effective tax rate would give an adjusted net income of $6,641 [$9,317 × (1 − 0.2872)], compared to $9,492 before the adjustment. The EPS would therefore be reduced from the reported $1.40 to $0.98 (adjusted net income of $6,641 divided by 6,780 shares).

B Based on information as reported, the P/CFO was 19.0 ($42 ÷ $2.21). Based on CFO adjusted to expense software development costs, the P/CFO was 31.6 ($42 ÷ $1.33).

- Price: Assuming that the market value of the company's equity is based on its fundamentals, the price per share is $42, regardless of a difference in accounting.
- CFO per share, as reported, was $2.21 (total operating cash flows $15,007 ÷ 6,780 shares).
- CFO per share, as adjusted, was $1.33. The company's $6,000 expenditure on software development costs was reported as a cash outflow from investing activities, so expensing those costs would reduce cash from operating activities by $6,000, from the reported $15,007 to $9,007. Dividing adjusted total operating cash flow of $9,007 by 6,780 shares results in cash flow per share of $1.33.

C Based on information as reported, the EV/EBITDA was 16.3 ($284,760 ÷ $17,517). Based on EBITDA adjusted to expense software development costs, the EV/EBITDA was 24.7 ($284,760 ÷ $11,517).

- Enterprise Value: Enterprise value is the sum of the market value of the company's equity and debt. JHH has no debt, and therefore the enterprise value is equal to the market value of its equity. The market value of its equity is $284,760 ($42 per share × 6,780 shares).
- EBITDA, as reported, was $17,517 (earnings before interest and taxes of $13,317 plus $2,200 depreciation plus $2,000 amortisation).
- EBITDA, adjusted for expensing software development costs by the inclusion of $6,000 development expense and the exclusion of $2,000 amortisation of prior expense, would be $11,517 (earnings before interest and taxes of $9,317 plus $2,200 depreciation plus $0 amortisation).

Solution to 2:

Expensing software development costs would decrease historical profits, operating cash flow, and EBITDA, and would thus increase all market multiples. So JHH's stock would appear more expensive if it expensed rather than capitalised the software development costs.

If the unadjusted market-based ratios were used in the comparison of JHH to its competitor that expenses all software development expenditures, then JHH might appear to be under-priced when the difference is solely related to accounting factors. JHH's adjusted market-based ratios provide a better basis for comparison.

For the company in Example 7, current period software development expenditures exceed the amortisation of prior periods' capitalised software development expenditures. As a result, expensing rather than capitalising software development costs would have the effect of lowering income. If, however, software development expenditures slowed such that current expenditures were lower than the amortisation of prior periods' capitalised software development expenditures, then expensing software development costs would have the effect of increasing income relative to capitalising it.

This section illustrated how decisions about capitalising versus expensing impact financial statements and ratios. Earlier expensing lowers current profits but enhances trends, whereas capitalising now and expensing later enhances current profits. Having described the accounting for acquisition of long-lived assets, we now turn to the topic of measuring long-lived assets in subsequent periods.

6

DEPRECIATION OF LONG-LIVED ASSETS: METHODS AND CALCULATION

d describe the different depreciation methods for property, plant, and equipment and calculate depreciation expense;

e describe how the choice of depreciation method and assumptions concerning useful life and residual value affect depreciation expense, financial statements, and ratios;

k explain and evaluate how impairment, revaluation, and derecognition of property, plant, and equipment and intangible assets affect financial statements and ratios;

Under the cost model of reporting long-lived assets, which is permitted under IFRS and required under US GAAP, the capitalised costs of long-lived tangible assets (other than land, which is not depreciated) and intangible assets with finite useful lives are allocated to subsequent periods as depreciation and amortisation expenses. Depreciation and amortisation are effectively the same concept, with the term depreciation referring to the process of allocating tangible assets' costs and the term amortisation referring to the process of allocating intangible assets' costs.[12] The alternative model of reporting long-lived assets is the **revaluation model**, which is permitted under IFRS but not under US GAAP. Under the revaluation model, a company reports the long-lived asset at fair value rather than at acquisition cost (historical cost) less accumulated depreciation or amortisation, as in the cost model.

An asset's carrying amount is the amount at which the asset is reported on the balance sheet. Under the cost model, at any point in time, the carrying amount (also called carrying value or net book value) of a long-lived asset is equal to its historical cost minus the amount of depreciation or amortisation that has been accumulated since the asset's purchase (assuming that the asset has not been impaired, a topic which will be addressed in Section 9). Companies may present on the balance sheet the total net amount of property, plant, and equipment and the total net amount of intangible assets. However, more detail is disclosed in the notes to financial statements. The details disclosed typically include the acquisition costs, the depreciation and amortisation expenses, the accumulated depreciation and amortisation amounts, the depreciation and amortisation methods used, and information on the assumptions used to depreciate and amortise long-lived assets.

6.1 Depreciation Methods and Calculation of Depreciation Expense

Depreciation methods include the **straight-line method**, in which the cost of an asset is allocated to expense evenly over its useful life; **accelerated methods**, in which the allocation of cost is greater in earlier years; and the **units-of-production method**, in which the allocation of cost corresponds to the actual use of an asset in a particular period. The choice of depreciation method affects the amounts reported on the financial statements, including the amounts for reported assets and operating and net income. This, in turn, affects a variety of financial ratios, including fixed asset turnover, total asset turnover, operating profit margin, operating return on assets, and return on assets.

Using the straight-line method, depreciation expense is calculated as depreciable cost divided by estimated useful life and is the same for each period. Depreciable cost is the historical cost of the tangible asset minus the estimated residual (salvage) value.[13] A commonly used accelerated method is the declining balance method, in which the amount of depreciation expense for a period is calculated as some percentage of the carrying amount (i.e., cost net of accumulated depreciation at the beginning of the period). When an accelerated method is used, depreciable cost is not used to calculate the depreciation expense but the carrying amount should not be reduced below the estimated residual value. In the units-of-production method, the amount of depreciation expense for a period is based on the proportion of the asset's production during the period compared with the total estimated productive capacity of the asset over its useful life. The depreciation expense is calculated as depreciable cost times production in the period divided by estimated productive capacity over the life of the asset. Equivalently, the company may estimate a depreciation cost per unit (depreciable cost divided by estimated productive capacity) and calculate depreciation expense as depreciation cost per unit times production in the period. Regardless of the depreciation method used, the carrying amount of the asset is not reduced below the estimated residual value. Example 8 provides an example of these depreciation methods.

---

[12] Depletion is the term applied to a similar concept for natural resources; costs associated with those resources are allocated to a period on the basis of the usage or extraction of those resources.

[13] The residual value is the estimated amount that an entity will obtain from disposal of the asset at the end of its useful life.

EXAMPLE 8

Alternative Depreciation Methods

You are analyzing three hypothetical companies: EVEN-LI Co., SOONER Inc., and AZUSED Co. At the beginning of Year 1, each company buys an identical piece of box manufacturing equipment for $2,300 and has the same assumptions about useful life, estimated residual value, and productive capacity. The annual production of each company is the same, but each company uses a different method of depreciation. As disclosed in each company's notes to the financial statements, each company's depreciation method, assumptions, and production are as follows:

Depreciation method

- EVEN-LI Co.: straight-line method
- SOONER Inc.: double-declining balance method (the rate applied to the carrying amount is double the depreciation rate for the straight-line method)
- AZUSED Co.: units-of-production method

Assumptions and production

- Estimated residual value: $100
- Estimated useful life: 4 years
- Total estimated productive capacity: 800 boxes
- Production in each of the four years: 200 boxes in the first year, 300 in the second year, 200 in the third year, and 100 in the fourth year

1 Using the following template for each company, record its beginning and ending net book value (carrying amount), end-of-year accumulated depreciation, and annual depreciation expense for the box manufacturing equipment.

Template:

	Beginning Net Book Value	Depreciation Expense	Accumulated Depreciation	Ending Net Book Value
Year 1				
Year 2				
Year 3				
Year 4				

2 Explain the significant differences in the timing of the recognition of the depreciation expense.

3 For each company, assume that sales, earnings before interest, taxes, depreciation, and amortization, and assets other than the box manufacturing equipment are as shown in the following table. Calculate the total asset turnover ratio, the operating profit margin, and the operating return on assets for each company for each of the four years. Discuss the ratios, comparing results within and across companies.

	Sales	Earnings before Interest, Taxes, Depreciation, and Amortization	Carrying Amount of Total Assets, Excluding the Box Manufacturing Equipment, at Year End*
Year 1	$300,000	$36,000	$30,000
Year 2	320,000	38,400	32,000
Year 3	340,000	40,800	34,000
Year 4	360,000	43,200	36,000

* Assume that total assets at the beginning of Year 1, *including* the box manufacturing equipment, had a value of $30,300. Assume that depreciation expense on assets other than the box manufacturing equipment totaled $1,000 per year.

Solution to 1:

For *each* company, the following information applies: Beginning net book value in Year 1 equals the purchase price of $2,300; accumulated year-end depreciation equals the balance from the previous year plus the current year's depreciation expense; ending net book value (carrying amount) equals original cost minus accumulated year-end depreciation (which is the same as beginning net book value minus depreciation expense); and beginning net book value in Years 2, 3, and 4 equals the ending net book value of the prior year. The following text and filled-in templates describe how depreciation *expense* is calculated for each company.

EVEN-LI Co. uses the straight-line method, so depreciation expense in each year equals $550, which is calculated as ($2,300 original cost – $100 residual value)/4 years. The net book value at the end of Year 4 is the estimated residual value of $100.

EVEN-LI Co.	Beginning Net Book Value	Depreciation Expense	Accumulated Year-End Depreciation	Ending Net Book Value
Year 1	$2,300	$550	$550	$1,750
Year 2	1,750	550	1,100	1,200
Year 3	1,200	550	1,650	650
Year 4	650	550	2,200	100

SOONER Inc. uses the double-declining balance method. The depreciation rate for the double-declining balance method is double the depreciation rate for the straight-line method. The depreciation rate under the straight-line method is 25 percent (100 percent divided by 4 years). Thus, the depreciation rate for the double-declining balance method is 50 percent (2 times 25 percent). The depreciation expense for the first year is $1,150 (50 percent of $2,300). Note that under this method, the depreciation rate of 50 percent is applied to the carrying amount (net book value) of the asset, without adjustment for expected residual value. Because the carrying amount of the asset is not depreciated below its estimated residual value, however, the depreciation expense in the final year of depreciation decreases the ending net book value (carrying amount) to the estimated residual value.

SOONER Inc.	Beginning Net Book Value	Depreciation Expense	Accumulated Year-End Depreciation	Ending Net Book Value
Year 1	$2,300	$1,150	$1,150	$1,150
Year 2	1,150	575	1,725	575
Year 3	575	288	2,013	287
Year 4	287	187	2,200	100

Another common approach (not required in this question) is to use an accelerated method, such as the double-declining method, for some period (a year or more) and then to change to the straight-line method for the remaining life of the asset. If SOONER had used the double-declining method for the first year and then switched to the straight-line method for Years 2, 3, and 4, the depreciation expense would be $350 [($1,150 – $100 estimated residual value)/3 years] a year for Years 2, 3, and 4. The results for SOONER under this alternative approach are shown below.

SOONER Inc.	Beginning Net Book Value	Depreciation Expense	Accumulated Year-End Depreciation	Ending Net Book Value
Year 1	$2,300	$1,150	$1,150	$1,150
Year 2	1,150	350	1,500	800
Year 3	800	350	1,850	450
Year 4	450	350	2,200	100

AZUSED Co. uses the units-of-production method. Dividing the equipment's total depreciable cost by its total productive capacity gives a cost per unit of $2.75, calculated as ($2,300 original cost – $100 residual value)/800. The depreciation expense recognised each year is the number of units produced times $2.75. For Year 1, the amount of depreciation expense is $550 (200 units times $2.75). For Year 2, the amount is $825 (300 units times $2.75). For Year 3, the amount is $550. For Year 4, the amount is $275.

AZUSED Co.	Beginning Net Book Value	Depreciation Expense	Accumulated Year-End Depreciation	Ending Net Book Value
Year 1	$2,300	$550	$550	$1,750
Year 2	1,750	825	1,375	925
Year 3	925	550	1,925	375
Year 4	375	275	2,200	100

Solution to 2:

All three methods result in the same total amount of accumulated depreciation over the life of the equipment. The significant differences are simply in the timing of the recognition of the depreciation expense. The straight-line method recognises the expense evenly, the accelerated method recognises most of the expense in the first year, and the units-of-production method recognises the expense on the basis of production (or use of the asset). Under all three methods, the ending net book value is $100.

Solution to 3:

Total asset turnover ratio = Total revenue ÷ Average total assets

Operating profit margin = Earnings before interest and taxes ÷ Total revenue

Operating return on assets = Earnings before interest and taxes ÷ Average total assets

Ratios are shown in the table below, and details of the calculations for Years 1 and 2 are described after discussion of the ratios.

	EVEN-LI Co.			SOONER Inc.			AZUSED Co.		
Ratio*	**AT**	**PM (%)**	**ROA (%)**	**AT**	**PM (%)**	**ROA (%)**	**AT**	**PM (%)**	**ROA (%)**
Year 1	9.67	11.48	111.04	9.76	11.28	110.17	9.67	11.48	111.04
Year 2	9.85	11.52	113.47	10.04	11.51	115.57	9.90	11.43	113.10
Year 3	10.02	11.54	115.70	10.17	11.62	118.21	10.10	11.54	116.64
Year 4	10.18	11.57	117.74	10.23	11.67	119.42	10.22	11.65	118.98

* AT = Total asset turnover ratio. PM = Operating profit margin. ROA = Operating return on assets.

For all companies, the asset turnover ratio increased over time because sales grew at a faster rate than that of the assets. SOONER had consistently higher asset turnover ratios than the other two companies, however, because higher depreciation expense in the earlier periods decreased its average total assets. In addition, the higher depreciation in earlier periods resulted in SOONER having lower operating profit margin and operating ROA in the first year and higher operating profit margin and operating ROA in the later periods. SOONER appears to be more efficiently run, on the basis of its higher asset turnover and greater increases in profit margin and ROA over time; however, these comparisons reflect differences in the companies' choice of depreciation method. In addition, an analyst might question the sustainability of the extremely high ROAs for all three companies because such high profitability levels would probably attract new competitors, which would likely put downward pressure on the ratios.

EVEN-LI Co.

Year 1:

Total asset turnover ratio = 300,000/[(30,300 + 30,000 + 1,750)/2]
= 300,000/31,025 = 9.67
Operating profit margin = (36,000 – 1,000 – 550)/300,000
= 34,450/300,000 = 11.48%
Operating ROA = 34,450/31,025 = 111.04%

Year 2:

Total asset turnover ratio = 320,000/[(30,000 + 1,750 + 32,000 + 1,200)/2]
= 320,000/32,475 = 9.85
Operating profit margin = (38,400 – 1,000 – 550)/320,000
= 36,850/320,000 = 11.52%
Operating ROA = 36,850/32,475 = 113.47%

SOONER Inc.

Year 1:

$$\begin{aligned}
\text{Total asset turnover ratio} &= 300{,}000/[(30{,}300 + 30{,}000 + 1{,}150)/2] \\
&= 300{,}000/30{,}725 = 9.76 \\
\text{Operating profit margin} &= (36{,}000 - 1{,}000 - 1{,}150)/300{,}000 \\
&= 33{,}850/300{,}000 = 11.28\% \\
\text{Operating ROA} &= 33{,}850/30{,}725 = 110.17\%
\end{aligned}$$

Year 2:

$$\begin{aligned}
\text{Total asset turnover ratio} &= 320{,}000/[(30{,}000 + 1{,}150 + 32{,}000 + 575)/2] \\
&= 320{,}000/31{,}862.50 = 10.04 \\
\text{Operating profit margin} &= (38{,}400 - 1{,}000 - 575)/320{,}000 \\
&= 36{,}825/320{,}000 = 11.51\% \\
\text{Operating ROA} &= 36{,}825/31{,}862.50 = 115.57\%
\end{aligned}$$

AZUSED Co.

Year 1:

$$\begin{aligned}
\text{Total asset turnover ratio} &= 300{,}000/[(30{,}300 + 30{,}000 + 1{,}750)/2] \\
&= 300{,}000/31{,}025 = 9.67 \\
\text{Operating profit margin} &= (36{,}000 - 1{,}000 - 550)/300{,}000 \\
&= 34{,}450/300{,}000 = 11.48\% \\
\text{Operating ROA} &= 34{,}450/31{,}025 = 111.04\%
\end{aligned}$$

Year 2:

$$\begin{aligned}
\text{Total asset turnover ratio} &= 320{,}000/[(30{,}000 + 1{,}750 + 32{,}000 + 925)/2] \\
&= 320{,}000/32{,}337.50 = 9.90 \\
\text{Operating profit margin} &= (38{,}400 - 1{,}000 - 825)/320{,}000 \\
&= 36{,}575/320{,}000 = 11.43\% \\
\text{Operating ROA} &= 36{,}575/32{,}337.50 = 113.10\%
\end{aligned}$$

In many countries, a company must use the same depreciation methods for both financial and tax reporting. In other countries, including the United States, a company need not use the same depreciation method for financial reporting and taxes. As a result of using different depreciation methods for financial and tax reporting, pre-tax income on the income statement and taxable income on the tax return may differ. Thus, the amount of tax expense computed on the basis of pre-tax income and the amount of taxes actually owed on the basis of taxable income may differ. Although these differences eventually reverse because the total depreciation is the same regardless of the timing of its recognition in financial statements versus on tax returns, during the period of the difference, the balance sheet will show what is known as deferred taxes. For instance, if a company uses straight-line depreciation for financial reporting and an accelerated depreciation method for tax purposes, the company's financial statements will report lower depreciation expense and higher pre-tax income in the first year, compared with the amount of depreciation expense and taxable income in its tax reporting. (Compare the depreciation expense in Year 1 for EVEN-LI Co. and SOONER Inc. in the previous example.) Tax expense calculated on the basis of the financial statements' pre-tax income will be higher than taxes payable on the basis of taxable income; the difference between the two amounts represents a deferred tax liability. The deferred tax liability will be reduced as the difference reverses (i.e., when depreciation for financial reporting is higher than the depreciation for tax purposes) and the income tax is paid.

Significant estimates required for calculating depreciation include the useful life of the asset (or its total lifetime productive capacity) and its expected residual value at the end of that useful life. A longer useful life and higher expected residual value decrease the amount of annual depreciation expense relative to a shorter useful life and lower expected residual value. Companies should review their estimates periodically to ensure they remain reasonable. IFRS require companies to review estimates annually.

Although no significant differences exist between IFRS and US GAAP with respect to the definition of depreciation and the acceptable depreciation methods, IFRS require companies to use a component method of depreciation.[14] Companies are required to separately depreciate the significant components of an asset (parts of an item with a cost that is significant in relation to the total cost and/or with different useful lives) and thus require additional estimates for the various components. For instance, it may be appropriate to depreciate separately the engine, frame, and interior furnishings of an aircraft. Under US GAAP, the component method of depreciation is allowed but is seldom used in practice.[15] The following example illustrates depreciating components of an asset.

EXAMPLE 9

Illustration of Depreciating Components of an Asset

CUTITUP Co., a hypothetical company, purchases a milling machine, a type of machine used for shaping metal, at a total cost of $10,000. $2,000 was estimated to represent the cost of the rotating cutter, a significant component of the machine. The company expects the machine to have a useful life of eight years and a residual value of $3,000 and that the rotating cutter will need to be replaced every two years. Assume the entire residual value is attributable to the milling machine itself, and assume the company uses straight-line depreciation for all assets.

1. How much depreciation expense would the company report in Year 1 if it uses the component method of depreciation, and how much depreciation expense would the company report in Year 1 if it does not use the component method?
2. Assuming a new cutter with an estimated two-year useful life is purchased at the end of Year 2 for $2,000, what depreciation expenses would the company report in Year 3 if it uses the component method and if it does not use the component method?
3. Assuming replacement of the cutter every two years at a price of $2,000, what is the total depreciation expense over the eight years if the company uses the component method compared with the total depreciation expense if the company does not use the component method?
4. How many different items must the company estimate in the first year to compute depreciation expense for the milling machine if it uses the component method, and how does this compare with what would be required if it does not use the component method?

---

[14] IAS 16 *Property, Plant and Equipment*, paragraphs 43–47 [Depreciation].

[15] According to KPMG's *IFRS Compared to US GAAP*, December 2017, kpmg.com.

Solution to 1:

Depreciation expense in Year 1 under the component method would be $1,625. For the portion of the machine excluding the cutter, the depreciable base is total cost minus the cost attributable to the cutter minus the estimated residual value = $10,000 – $2,000 – $3,000 = $5,000. Depreciation expense for the machine excluding the cutter in the first year equals $625 (depreciable cost divided by the useful life of the machine = $5,000/8 years). For the cutter, the depreciation expense equals $1,000 (depreciable cost divided by the useful life of the cutter = $2,000/2 years). Thus, the total depreciation expense for Year 1 under the component method is $1,625 (the sum of the depreciation expenses of the two components = $625 + $1,000). Depreciation expense in Year 2 would also be $1,625.

If the company does not use the component method, depreciation expense in Year 1 is $875 (the depreciable cost of the total milling machine divided by its useful life = [$10,000 – $3,000]/8 years). Depreciation expense in Year 2 would also be $875.

Solution to 2:

Assuming that at the end of Year 2, the company purchases a new cutter for $2,000 with an estimated two-year life, under the component method, the depreciation expense in Year 3 will remain at $1,625. If the company does not use the component method and purchases a new cutter with an estimated two-year life for $2,000 at the end of Year 2, the depreciation expense in Year 3 will be $1,875 [$875 + ($2,000/2) = $875 + $1,000].

Solution to 3:

Over the eight years, assuming replacement of the cutters every two years at a price of $2,000, the total depreciation expense will be $13,000 [$1,625 × 8 years] when the component method is used. When the component method is not used, the total depreciation expense will also be $13,000 [$875 × 2 years + $1,875 × 6 years]. This amount equals the total expenditures of $16,000 [$10,000 + 3 cutters × $2,000] less the residual value of $3,000.

Solution to 4:

The following table summarizes the estimates required in the first year to compute depreciation expense if the company does or does not use the component method:

Estimate	Required using component method?	Required if not using component method?
Useful life of milling machine	Yes	Yes
Residual value of milling machine	Yes	Yes
Portion of machine cost attributable to cutter	Yes	No
Portion of residual value attributable to cutter	Yes	No
Useful life of cutter	Yes	No

Total depreciation expense may be allocated between the cost of sales and other expenses. Within the income statement, depreciation expense of assets used in production is usually allocated to the cost of sales, and the depreciation expense of assets not used in production may be allocated to some other expense category. For instance, depreciation expense may be allocated to selling, general, and administrative expenses if depreciable assets are used in those functional areas. Notes to the financial statements sometimes disclose information regarding which income statement line items include depreciation expense, although the exact amount of detail disclosed by individual companies varies.

AMORTISATION OF LONG-LIVED ASSETS: METHODS AND CALCULATION

7

f describe the different amortisation methods for intangible assets with finite lives and calculate amortisation expense;

g describe how the choice of amortisation method and assumptions concerning useful life and residual value affect amortisation expense, financial statements, and ratios;

k explain and evaluate how impairment, revaluation, and derecognition of property, plant, and equipment and intangible assets affect financial statements and ratios;

Amortisation is similar in concept to depreciation. The term amortisation applies to intangible assets, and the term depreciation applies to tangible assets. Both terms refer to the process of allocating the cost of an asset over the asset's useful life. Only those intangible assets assumed to have finite useful lives are amortised over their useful lives, following the pattern in which the benefits are used up. Acceptable amortisation methods are the same as the methods acceptable for depreciation. Assets assumed to have an indefinite useful life (in other words, without a finite useful life) are not amortised. An intangible asset is considered to have an indefinite useful life when there is "no foreseeable limit to the period over which the asset is expected to generate net cash inflows" for the company.[16]

Intangible assets with finite useful lives include an acquired customer list expected to provide benefits to a direct-mail marketing company for two to three years, an acquired patent or copyright with a specific expiration date, an acquired license with a specific expiration date and no right to renew the license, and an acquired trademark for a product that a company plans to phase out over a specific number of years. Examples of intangible assets with indefinite useful lives include an acquired license that, although it has a specific expiration date, can be renewed at little or no cost and an acquired trademark that, although it has a specific expiration, can be renewed at a minimal cost and relates to a product that a company plans to continue selling for the foreseeable future.

As with depreciation for a tangible asset, the calculation of amortisation for an intangible asset requires the original amount at which the intangible asset is recognised and estimates of the length of its useful life and its residual value at the end of its useful life. Useful lives are estimated on the basis of the expected use of the asset, considering any factors that may limit the life of the asset, such as legal, regulatory, contractual, competitive, or economic factors.

---

[16] IAS 38 *Intangible Assets*, paragraph 88.

EXAMPLE 10

Amortisation Expense

IAS 38 *Intangible Assets* provides illustrative examples regarding the accounting for intangible assets, including the following:

> A direct-mail marketing company acquires a customer list and expects that it will be able to derive benefit from the information on the list for at least one year, but no more than three years. The customer list would be amortised over management's best estimate of its useful life, say 18 months. Although the direct-mail marketing company may intend to add customer names and other information to the list in the future, the expected benefits of the acquired customer list relate only to the customers on that list at the date it was acquired.

In this example, in what ways would management's decisions and estimates affect the company's financial statements?

Solution:

Because the acquired customer list is expected to generate future economic benefits for a period greater than one year, the cost of the list should be capitalised and not expensed. The acquired customer list is determined to not have an indefinite life and must be amortised. Management must estimate the useful life of the customer list and must select an amortisation method. In this example, the list appears to have no residual value. Both the amortisation method and the estimated useful life affect the amount of the amortisation expense in each period. A shorter estimated useful life, compared with a longer estimated useful life, results in a higher amortisation expense each year over a shorter period, but the *total* accumulated amortisation expense over the life of the intangible asset is unaffected by the estimate of the useful life. Similarly, the total accumulated amortisation expense over the life of the intangible asset is unaffected by the choice of amortisation method. The amortisation expense per period depends on the amortisation method. If the straight-line method is used, the amortisation expense is the same for each year of useful life. If an accelerated method is used, the amortisation expense will be higher in earlier years.

8

THE REVALUATION MODEL

h describe the revaluation model;

k explain and evaluate how impairment, revaluation, and derecognition of property, plant, and equipment and intangible assets affect financial statements and ratios;

The revaluation model is an alternative to the cost model for the periodic valuation and reporting of long-lived assets. IFRS permit the use of either the revaluation model or the cost model, but the revaluation model is not allowed under US GAAP. Revaluation changes the carrying amounts of classes of long-lived assets to fair value (the fair value must be measured reliably). Under the cost model, carrying amounts are historical costs less accumulated depreciation or amortisation. Under the revaluation model, carrying amounts are the fair values at the date of revaluation less any subsequent accumulated depreciation or amortisation.

IFRS allow companies to value long-lived assets either under a cost model at historical cost minus accumulated depreciation or amortisation or under a revaluation model at fair value. In contrast, US accounting standards require that the cost model be used. A key difference between the two models is that the cost model allows only decreases in the values of long-lived assets compared with historical costs but the revaluation model may result in increases in the values of long-lived assets to amounts greater than historical costs.

IFRS allow a company to use the cost model for some classes of assets and the revaluation model for others, but the company must apply the same model to all assets within a particular class of assets and must revalue all items within a class to avoid selective revaluation. Examples of different classes of assets include land, land and buildings, machinery, motor vehicles, furniture and fixtures, and office equipment. The revaluation model may be used for classes of intangible assets but only if an active market for the assets exists, because the revaluation model may only be used if the fair values of the assets can be measured reliably. For practical purposes, the revaluation model is rarely used for either tangible or intangible assets, but its use is especially rare for intangible assets.

Under the revaluation model, whether an asset revaluation affects earnings depends on whether the revaluation initially increases or decreases an asset class' carrying amount. If a revaluation initially decreases the carrying amount of the asset class, the decrease is recognised in profit or loss. Later, if the carrying amount of the asset class increases, the increase is recognised in profit or loss to the extent that it reverses a revaluation decrease of the same asset class previously recognised in profit or loss. Any increase in excess of the reversal amount will not be recognised in the income statement but will be recorded directly to equity in a revaluation surplus account. An upward revaluation is treated the same as the amount in excess of the reversal amount. In other words, if a revaluation initially increases the carrying amount of the asset class, the increase in the carrying amount of the asset class bypasses the income statement and goes directly to equity under the heading of revaluation surplus. Any subsequent decrease in the asset's value first decreases the revaluation surplus and then goes to income. When an asset is retired or disposed of, any related amount of revaluation surplus included in equity is transferred directly to retained earnings.

Asset revaluations offer several considerations for financial statement analyses. First, an increase in the carrying amount of depreciable long-lived assets increases total assets and shareholders' equity, so asset revaluations that increase the carrying amount of an asset can be used to reduce reported leverage. Defining leverage as average total assets divided by average shareholders' equity, increasing both the numerator (assets) and denominator (equity) by the same amount leads to a decline in the ratio. (Mathematically, when a ratio is greater than one, as in this case, an increase in both the numerator and the denominator by the same amount leads to a decline in the ratio.) Therefore, the leverage motivation for the revaluation should be considered in analysis. For example, a company may revalue assets up if it is seeking new capital or approaching leverage limitations set by financial covenants.

Second, assets revaluations that decrease the carrying amount of the assets reduce net income. In the year of the revaluation, profitability measures such as return on assets and return on equity decline. However, because total assets and shareholders' equity are also lower, the company may appear more profitable in future years. Additionally, reversals of downward revaluations also go through income, thus increasing earnings. Managers can then opportunistically time the reversals to manage earnings and increase income. Third, asset revaluations that increase the carrying amount of an asset initially increase depreciation expense, total assets, and shareholders' equity. Therefore, profitability measures, such as return on assets and return on equity, would decline. Although upward asset revaluations also generally decrease income (through higher depreciation expense), the increase in the value of the long-lived asset is presumably based on increases in the operating capacity of the asset, which will likely be evidenced in increased future revenues.

Finally, an analyst should consider who did the appraisal—i.e. an independent external appraiser or management—and how often revaluations are made. Appraisals of the fair value of long-lived assets involve considerable judgment and discretion. Presumably, appraisals of assets from independent external sources are more reliable. How often assets are revalued can provide an indicator of whether their reported value continues to be representative of their fair values.

The next two examples illustrate revaluation of long-lived assets under IFRS.

EXAMPLE 11

Revaluation Resulting in an Increase in Carrying Amount Followed by Subsequent Revaluation Resulting in a Decrease in Carrying Amount

UPFIRST, a hypothetical manufacturing company, has elected to use the revaluation model for its machinery. Assume for simplicity that the company owns a single machine, which it purchased for €10,000 on the first day of its fiscal period, and that the measurement date occurs simultaneously with the company's fiscal period end.

1 At the end of the first fiscal period after acquisition, assume the fair value of the machine is determined to be €11,000. How will the company's financial statements reflect the asset?

2 At the end of the second fiscal period after acquisition, assume the fair value of the machine is determined to be €7,500. How will the company's financial statements reflect the asset?

Solution to 1:

At the end of the first fiscal period, the company's balance sheet will show the asset at a value of €11,000. The €1,000 increase in the value of the asset will appear in other comprehensive income and be accumulated in equity under the heading of revaluation surplus.

Solution to 2:

At the end of the second fiscal period, the company's balance sheet will show the asset at a value of €7,500. The total decrease in the carrying amount of the asset is €3,500 (€11,000 – €7,500). Of the €3,500 decrease, the first €1,000 will reduce the amount previously accumulated in equity under the heading of revaluation surplus. The other €2,500 will be shown as a loss on the income statement.

EXAMPLE 12

Revaluation Resulting in a Decrease in Asset's Carrying Amount Followed by Subsequent Revaluation Resulting in an Increase in Asset's Carrying Amount

DOWNFIRST, a hypothetical manufacturing company, has elected to use the revaluation model for its machinery. Assume for simplicity that the company owns a single machine, which it purchased for €10,000 on the first day of its fiscal period, and that the measurement date occurs simultaneously with the company's fiscal period end.

1. At the end of the first fiscal period after acquisition, assume the fair value of the machine is determined to be €7,500. How will the company's financial statements reflect the asset?
2. At the end of the second fiscal period after acquisition, assume the fair value of the machine is determined to be €11,000. How will the company's financial statements reflect the asset?

Solution to 1:

At the end of the first fiscal period, the company's balance sheet will show the asset at a value of €7,500. The €2,500 decrease in the value of the asset will appear as a loss on the company's income statement.

Solution to 2:

At the end of the second fiscal period, the company's balance sheet will show the asset at a value of €11,000. The total increase in the carrying amount of the asset is an increase of €3,500 (€11,000 – €7,500). Of the €3,500 increase, the first €2,500 reverses a previously reported loss and will be reported as a gain on the income statement. The other €1,000 will bypass profit or loss and be reported as other comprehensive income and be accumulated in equity under the heading of revaluation surplus.

Exhibit 6 provides two examples of disclosures concerning the revaluation model. The first disclosure is an excerpt from the 2006 annual report of KPN, a Dutch telecommunications and multimedia company. The report was produced at a time during which any IFRS-reporting company with a US stock exchange listing was required to explain differences between its reporting under IFRS and its reporting if it had used US GAAP.[17] One of these differences, as previously noted, is that US GAAP do not allow revaluation of fixed assets held for use. KPN's disclosure states that the company elected to report a class of fixed assets (cables) at fair value and explained that under US GAAP, using the cost model, the value of the asset class would have been €350 million lower. The second disclosure is an excerpt from the 2017 annual report of Avianca Holdings S.A., a Latin American airline that reports under IFRS and uses the revaluation model for one component of its fixed assets.

---

[17] On 15 November 2007, the SEC approved rule amendments under which financial statements from foreign private issuers in the United States will be accepted without reconciliation to US GAAP if the financial statements are prepared in accordance with IFRS as issued by the International Accounting Standards Board. The rule took effect for the 2007 fiscal year. As a result, companies such as KPN no longer need to provide reconciliations to US GAAP.

Exhibit 6 Impact of Revaluation

1 Excerpt from the annual report of Koninklijke KPN N.V. explaining certain differences between IFRS and US GAAP regarding "Deemed cost fixed assets":

> KPN elected the exemption to revalue certain of its fixed assets upon the transition to IFRS to fair value and to use this fair value as their deemed cost. KPN applied the depreciated replacement cost method to determine this fair value. The revalued assets pertain to certain cables, which form part of property, plant & equipment. Under US GAAP, this revaluation is not allowed and therefore results in a reconciling item. As a result, the value of these assets as of December 31, 2006 under US GAAP is EUR 350 million lower (2005: EUR 415 million; 2004: EUR 487 million) than under IFRS.
>
> *Source*: KPN's Form 20-F, p. 168, filed 1 March 2007.

2 The 2017 annual report of Avianca Holdings S.A. and Subsidiaries shows $58.4 million of "Revaluation and Other Reserves" as a component of Equity on its balance sheet and $31.0 million in Other Comprehensive Income for the current year's "Revaluation of Administrative Property". The relevant footnote disclosure explains:

> "Administrative property in Bogota, Medellín, El Salvador, and San Jose is recorded at fair value less accumulated depreciation on buildings and impairment losses recognized at the date of revaluation. Valuations are performed with sufficient frequency to ensure that the fair value of a revalued asset does not differ materially from its carrying amount. A revaluation reserve is recorded in other comprehensive income and credited to the asset revaluation reserve in equity. However, to the extent that it reverses a revaluation deficit of the same asset previously recognized in profit or loss, the increase is recognized in profit and loss. A revaluation deficit is recognized in the income statement, except to the extent that it offsets an existing surplus on the same asset recognized in the asset revaluation reserve. Upon disposal, any revaluation reserve relating to the particular asset being sold is transferred to retained earnings.
>
> *Source*: AVIANCA HOLDINGS S.A. Form 20-F filed 01 May 2018.

Clearly, the use of the revaluation model as opposed to the cost model can have a significant impact on the financial statements of companies. This has potential consequences for comparing financial performance using financial ratios of companies that use different models.

9 IMPAIRMENT OF ASSETS

i. explain the impairment of property, plant, and equipment and intangible assets;

k. explain and evaluate how impairment, revaluation, and derecognition of property, plant, and equipment and intangible assets affect financial statements and ratios;

In contrast with depreciation and amortisation charges, which serve to allocate the depreciable cost of a long-lived asset over its useful life, impairment charges reflect an unanticipated decline in the value of an asset. Both IFRS and US GAAP require companies to write down the carrying amount of impaired assets. Impairment reversals for identifiable, long-lived assets are permitted under IFRS but typically not under US GAAP.

An asset is considered to be impaired when its carrying amount exceeds its recoverable amount. Although IFRS and US GAAP define recoverability differently (as described below), in general, impairment losses are recognised when the asset's carrying amount is not recoverable. The following paragraphs describe accounting for impairment for different categories of assets.

9.1 Impairment of Property, Plant, and Equipment

Accounting standards do not require that property, plant, and equipment be tested annually for impairment. Rather, at the end of each reporting period (generally, a fiscal year), a company assesses whether there are indications of asset impairment. If there is no indication of impairment, the asset is not tested for impairment. If there is an indication of impairment, such as evidence of obsolescence, decline in demand for products, or technological advancements, the recoverable amount of the asset should be measured in order to test for impairment. For property, plant, and equipment, impairment losses are recognised when the asset's carrying amount is not recoverable; the carrying amount is more than the recoverable amount. The amount of the impairment loss will reduce the carrying amount of the asset on the balance sheet and will reduce net income on the income statement. The impairment loss is a non-cash item and will not affect cash from operations.

IFRS and US GAAP differ somewhat both in the guidelines for determining that impairment has occurred and in the measurement of an impairment loss. Under IAS 36, an impairment loss is measured as the excess of carrying amount over the recoverable amount of the asset. The recoverable amount of an asset is defined as "the higher of its fair value less costs to sell and its value in use." Value in use is based on the present value of expected future cash flows. Under US GAAP, assessing recoverability is separate from measuring the impairment loss. The carrying amount of an asset "group" is considered not recoverable when it exceeds the undiscounted expected future cash flows of the group. If the asset's carrying amount is considered not recoverable, the impairment loss is measured as the difference between the asset's fair value and carrying amount.

EXAMPLE 13

Impairment of Property, Plant, and Equipment

Sussex, a hypothetical manufacturing company in the United Kingdom, has a machine it uses to produce a single product. The demand for the product has declined substantially since the introduction of a competing product. The company has assembled the following information with respect to the machine:

Carrying amount	£18,000
Undiscounted expected future cash flows	£19,000
Present value of expected future cash flows	£16,000
Fair value if sold	£17,000
Costs to sell	£2,000

1 Under IFRS, what would the company report for the machine?

2 Under US GAAP, what would the company report for the machine?

Solution to 1:

Under IFRS, the company would compare the carrying amount (£18,000) with the higher of its fair value less costs to sell (£15,000) and its value in use (£16,000). The carrying amount exceeds the value in use, the higher of the two amounts, by £2,000. The machine would be written down to the recoverable amount of £16,000, and a loss of £2,000 would be reported in the income statement. The carrying amount of the machine is now £16,000. A new depreciation schedule based on the carrying amount of £16,000 would be developed.

Solution to 2:

Under US GAAP, the carrying amount (£18,000) is compared with the undiscounted expected future cash flows (£19,000). The carrying amount is less than the undiscounted expected future cash flows, so the carrying amount is considered recoverable. The machine would continue to be carried at £18,000, and no loss would be reported.

In Example 13, a write down in the value of a piece of property, plant, and equipment occurred under IFRS but not under US GAAP. In Example 14, a write down occurs under both IFRS and US GAAP.

EXAMPLE 14

Impairment of Property, Plant, and Equipment

Essex, a hypothetical manufacturing company, has a machine it uses to produce a single product. The demand for the product has declined substantially since the introduction of a competing product. The company has assembled the following information with respect to the machine:

Carrying amount	£18,000
Undiscounted expected future cash flows	£16,000
Present value of expected future cash flows	£14,000
Fair value if sold	£10,000
Costs to sell	£2,000

1 Under IFRS, what would the company report for the machine?

2 Under US GAAP, what would the company report for the machine?

Solution to 1:

Under IFRS, the company would compare the carrying amount (£18,000) with the higher of its fair value less costs to sell (£8,000) and its value in use (£14,000). The carrying amount exceeds the value in use, the higher of the two amounts, by £4,000. The machine would be written down to the recoverable amount of £14,000, and a loss of £4,000 would be reported in the income statement. The carrying amount of the machine is now £14,000. A new depreciation schedule based on the carrying amount of £14,000 would be developed.

Solution to 2:

Under US GAAP, the carrying amount (£18,000) is compared with the undiscounted expected future cash flows (£16,000). The carrying amount exceeds the undiscounted expected future cash flows, so the carrying amount is considered not recoverable. The machine would be written down to fair value of £10,000, and a loss of £8,000 would be reported in the income statement. The carrying amount of the machine is now £10,000. A new depreciation schedule based on the carrying amount of £10,000 would be developed.

Example 14 shows that the write down to value in use under IFRS can be less than the write down to fair value under US GAAP. The difference in recognition of impairment losses is ultimately reflected in difference in book value of equity.

9.2 Impairment of Intangible Assets with a Finite Life

Intangible assets with a finite life are amortised (carrying amount decreases over time) and may become impaired. As is the case with property, plant, and equipment, the assets are not tested annually for impairment. Instead, they are tested only when significant events suggest the need to test. The company assesses at the end of each reporting period whether a significant event suggesting the need to test for impairment has occurred. Examples of such events include a significant decrease in the market price or a significant adverse change in legal or economic factors. Impairment accounting for intangible assets with a finite life is essentially the same as for tangible assets; the amount of the impairment loss will reduce the carrying amount of the asset on the balance sheet and will reduce net income on the income statement.

9.3 Impairment of Intangibles with Indefinite Lives

Intangible assets with indefinite lives are not amortised. Instead, they are carried on the balance sheet at historical cost but are tested at least annually for impairment. Impairment exists when the carrying amount exceeds its fair value.

9.4 Impairment of Long-Lived Assets Held for Sale

A long-lived (non-current) asset is reclassified as held for sale rather than held for use when management's intent is to sell it and its sale is highly probable. (Additionally, accounting standards require that the asset must be available for immediate sale in its present condition.)[18] For instance, assume a building is no longer needed by a company and management's intent is to sell it, if the transaction meets the accounting criteria, the building is reclassified from property, plant, and equipment to non-current assets held for sale. At the time of reclassification, assets previously held for use are tested for impairment. If the carrying amount at the time of reclassification exceeds the fair value less costs to sell, an impairment loss is recognised and the asset is written down to fair value less costs to sell. Long-lived assets held for sale cease to be depreciated or amortised.

---

[18] IFRS 5 *Non-current Assets Held for Sale and Discontinued Operations.*

9.5 Reversals of Impairments of Long-Lived Assets

After an asset has been deemed impaired and an impairment loss has been reported, the asset's recoverable amount could potentially increase. For instance, a lawsuit appeal may successfully challenge a patent infringement by another company, with the result that a patent previously written down has a higher recoverable amount. IFRS permit impairment losses to be reversed if the recoverable amount of an asset increases regardless of whether the asset is classified as held for use or held for sale. Note that IFRS permit the reversal of impairment losses only. IFRS do not permit the revaluation to the recoverable amount if the recoverable amount exceeds the previous carrying amount. Under US GAAP, the accounting for reversals of impairments depends on whether the asset is classified as held for use or held for sale.[19] Under US GAAP, once an impairment loss has been recognised for assets held for use, it cannot be reversed. In other words, once the value of an asset held for use has been decreased by an impairment charge, it cannot be increased. For assets held for sale, if the fair value increases after an impairment loss, the loss can be reversed.

10 DERECOGNITION

j explain the derecognition of property, plant, and equipment and intangible assets;

k explain and evaluate how impairment, revaluation, and derecognition of property, plant, and equipment and intangible assets affect financial statements and ratios;

A company derecognises an asset (i.e., removes it from the financial statements) when the asset is disposed of or is expected to provide no future benefits from either use or disposal. A company may dispose of a long-lived operating asset by selling it, exchanging it, abandoning it, or distributing it to existing shareholders. As previously described, non-current assets that management intends to sell or to distribute to existing shareholders and which meet the accounting criteria (immediately available for sale in current condition and the sale is highly probable) are reclassified as non-current assets held for sale.

10.1 Sale of Long-Lived Assets

The gain or loss on the sale of long-lived assets is computed as the sales proceeds minus the carrying amount of the asset at the time of sale. An asset's carrying amount is typically the net book value (i.e., the historical cost minus accumulated depreciation), unless the asset's carrying amount has been changed to reflect impairment and/or revaluation, as previously discussed.

---

[19] FASB ASC Section 360-10-35 [Property, Plant, and Equipment – Overall – Subsequent Measurement].

EXAMPLE 15

Calculation of Gain or Loss on the Sale of Long-Lived Assets

Moussilauke Diners Inc., a hypothetical company, as a result of revamping its menus to focus on healthier food items, sells 450 used pizza ovens and reports a gain on the sale of $1.2 million. The ovens had a carrying amount of $1.9 million (original cost of $5.1 million less $3.2 million of accumulated depreciation). At what price did Moussilauke sell the ovens?

A $0.7 million

B $3.1 million

C $6.3 million

Solution:

B is correct. The ovens had a carrying amount of $1.9 million, and Moussilauke recognised a gain of $1.2 million. Therefore, Moussilauke sold the ovens at a price of $3.1 million. The gain on the sale of $1.2 million is the selling price of $3.1 million minus the carrying amount of $1.9 million. Ignoring taxes, the cash flow from the sale is $3.1 million, which would appear as a cash inflow from investing.

A gain or loss on the sale of an asset is disclosed on the income statement, either as a component of other gains and losses or in a separate line item when the amount is material. A company typically discloses further detail about the sale in the management discussion and analysis and/or financial statement footnotes. In addition, a statement of cash flows prepared using the indirect method adjusts net income to remove any gain or loss on the sale from operating cash flow and to include the amount of proceeds from the sale in cash from investing activities. Recall that the indirect method of the statement of cash flows begins with net income and makes all adjustments to arrive at cash from operations, including removal of gains or losses from non-operating activities.

10.2 Long-Lived Assets Disposed of Other Than by a Sale

Long-lived assets to be disposed of other than by a sale (e.g., abandoned, exchanged for another asset, or distributed to owners in a spin-off) are classified as held for use until disposal or until they meet the criteria to be classified as held for sale or held for distribution.[20] Thus, the long-lived assets continue to be depreciated and tested for impairment, unless their carrying amount is zero, as required for other long-lived assets owned by the company.

When an asset is retired or abandoned, the accounting is similar to a sale, except that the company does not record cash proceeds. Assets are reduced by the carrying amount of the asset at the time of retirement or abandonment, and a loss equal to the asset's carrying amount is recorded.

When an asset is exchanged, accounting for the exchange typically involves removing the carrying amount of the asset given up, adding a fair value for the asset acquired, and reporting any difference between the carrying amount and the fair value as a gain or loss. The fair value used is the fair value of the asset given up unless the fair value of the asset acquired is more clearly evident. If no reliable measure of fair value exists, the acquired asset is measured at the carrying amount of the asset given up. A gain is reported when the fair value used for the newly acquired asset exceeds the carrying amount of the asset given up. A loss is reported when the fair value used for the newly acquired asset is less than the carrying amount of the asset given up. If the acquired asset is valued at the carrying amount of the asset given up because no reliable measure of fair value exists, no gain or loss is reported.

---

[20] In a spin-off, shareholders of the parent company receive a proportional number of shares in a new, separate entity.

When a spin-off occurs, typically, an entire cash generating unit of a company with all its assets is spun off. As an illustration of a spin-off, Fiat Chrysler Automobiles (FCA) spun off its ownership of Ferrari in 2016. Prior to the spinoff, FCA had sold 10 percent of its ownership of Ferrari in an IPO and recognized an increase in Shareholders' equity of € 873 million (the difference between the consideration it received in the IPO of € 866 million and the carrying amount of the equity interest sold of € 7 million.) In contrast, the spin-off, in which FCA distributed its ownership in Ferrari to the existing FCA shareholders, did not result in any gain or loss.

FCA's spinoff was completed on 3 January 2016, with each FCA shareholder receiving one common share of Ferrari N.V. for every ten common shares of FCA. In its financial statements for the prior fiscal year, FCA shows the assets and liabilities of Ferrari as held for distribution. Specifically, its balance sheet includes € 3,650 million Assets Held for Distribution as a component of current assets and € 3,584 million Liabilities Held for Distribution. Exhibit 7 includes excerpts from the company's 31 December 2015 annual report.

Exhibit 7 Fiat Chrysler Automobiles (FCA) Excerpts from Notes to the Consolidated Financial Statements - 2015 Annual Report

Ferrari Spin-off and Discontinued Operations

"As the spin-off of Ferrari N.V. became highly probable with the aforementioned shareholders' approval and since it was available for immediate distribution at that date, the Ferrari segment met the criteria to be classified as a disposal group held for distribution to owners and a discontinued operation pursuant to IFRS 5 - *Non-current Assets Held for Sale and Discontinued Operations*."

The following assets and liabilities of the Ferrari segment were classified as held for distribution at December 31, 2015:

	At December 31, 2015
Assets classified as held for distribution	(€ million)
Goodwill	786
Other intangible assets	297
Property, plant and equipment	627
Other non-current assets	134
Receivables from financing activities	1,176
Cash and cash equivalents	182
Other current assets	448
Total Assets held for distribution	**3,650**
Liabilities classified as held for distribution	
Provisions	224

Exhibit 7 (Continued)

	At December 31, 2015
Debt	2,256
Other current liabilities	624
Trade payables	480
Total Liabilities held for distribution	**3,584**

Source: Fiat Chrysler Automobiles (FCA)'s Form 20-F for the year ending 31 December 2015.

PRESENTATION AND DISCLOSURE REQUIREMENTS

11

l. describe the financial statement presentation of and disclosures relating to property, plant, and equipment and intangible assets;

Under IFRS, for each class of property, plant, and equipment, a company must disclose the measurement bases, the depreciation method, the useful lives (or, equivalently, the depreciation rate) used, the gross carrying amount and the accumulated depreciation at the beginning and end of the period, and a reconciliation of the carrying amount at the beginning and end of the period.[21] In addition, disclosures of restrictions on title and pledges as security of property, plant, and equipment and contractual agreements to acquire property, plant, and equipment are required. If the revaluation model is used, the date of revaluation, details of how the fair value was obtained, the carrying amount under the cost model, and the revaluation surplus must be disclosed.

The disclosure requirements under US GAAP are less exhaustive.[22] A company must disclose the depreciation expense for the period, the balances of major classes of depreciable assets, accumulated depreciation by major classes or in total, and a general description of the depreciation method(s) used in computing depreciation expense with respect to the major classes of depreciable assets.

Under IFRS, for each class of intangible assets, a company must disclose whether the useful lives are indefinite or finite. If finite, for each class of intangible asset, a company must disclose the useful lives (or, equivalently, the amortisation rate) used, the amortisation methods used, the gross carrying amount and the accumulated amortisation at the beginning and end of the period, where amortisation is included on the income statement, and a reconciliation of the carrying amount at the beginning and end of the period.[23] If an asset has an indefinite life, the company must disclose the carrying amount of the asset and why it is considered to have an indefinite life. Similar to property, plant, and equipment, disclosures of restrictions on title and pledges as security of intangible assets and contractual agreements to acquire intangible assets are required. If the revaluation model is used, the date of revaluation, details of how the fair value was obtained, the carrying amount under the cost model, and the revaluation surplus must be disclosed.

---

[21] IAS 16 *Property, Plant and Equipment*, paragraphs 73–79 [Disclosure].

[22] FASB ASC Section 360-10-50 [Property, Plant, and Equipment – Overall – Disclosure].

[23] IAS 38 *Intangible Assets*, paragraphs 118–128 [Disclosure].

Under US GAAP, companies are required to disclose the gross carrying amounts and accumulated amortisation in total and by major class of intangible assets, the aggregate amortisation expense for the period, and the estimated amortisation expense for the next five fiscal years.[24]

The disclosures related to impairment losses also differ under IFRS and US GAAP. Under IFRS, a company must disclose for each class of assets the amounts of impairment losses and reversals of impairment losses recognised in the period and where those are recognised on the financial statements.[25] The company must also disclose in aggregate the main classes of assets affected by impairment losses and reversals of impairment losses and the main events and circumstances leading to recognition of these impairment losses and reversals of impairment losses. Under US GAAP, there is no reversal of impairment losses for assets held for use. The company must disclose a description of the impaired asset, what led to the impairment, the method of determining fair value, the amount of the impairment loss, and where the loss is recognised on the financial statements.[26]

Disclosures about long-lived assets appear throughout the financial statements: in the balance sheet, the income statement, the statement of cash flows, and the notes. The balance sheet reports the carrying value of the asset. For the income statement, depreciation expense may or may not appear as a separate line item. Under IFRS, whether the income statement discloses depreciation expense separately depends on whether the company is using a 'nature of expense' method or a 'function of expense' method. Under the nature of expense method, a company aggregates expenses "according to their nature (for example, depreciation, purchases of materials, transport costs, employee benefits and advertising costs), and does not reallocate them among functions within the entity."[27] Under the function of expense method, a company classifies expenses according to the function, for example as part of cost of sales or of SG&A (selling, general, and administrative expenses). At a minimum, a company using the function of expense method must disclose cost of sales, but the other line items vary.

The statement of cash flows reflects acquisitions and disposals of fixed assets in the investing section. In addition, when prepared using the indirect method, the statement of cash flows typically shows depreciation expense (or depreciation plus amortisation) as a line item in the adjustments of net income to cash flow from operations. The notes to the financial statements describe the company's accounting method(s), the range of estimated useful lives, historical cost by main category of fixed asset, accumulated depreciation, and annual depreciation expense.

To illustrate financial statement presentation and disclosures, the following example provides excerpts relating to intangible assets and property, plant, and equipment from the annual report of Orange SA for the year ended 31 December 2017.

EXAMPLE 16

Financial Statement Presentation and Disclosures for Long-Lived Assets

The following exhibits include excerpts from the annual report for the year ended 31 December 2017 of Orange SA, an international telecommunications company based in France.

---

[24] FASB ASC Section 350-30-50 [Intangibles – General – Disclosure].

[25] IAS 36 *Impairment of Assets*, paragraphs 126–137 [Disclosure].

[26] FASB ASC Section 360-10-50 [Property, Plant, and Equipment – Overall – Disclosure] and FASB ASC Section 350-30-50 [Intangibles – General – Disclosure].

[27] IAS 1 paragraph 102.

Exhibit 8 Orange SA
Excerpts from the 2017 Consolidated Financial Statements

(Note that only selected line items/data are shown for illustrative purposes)

Excerpt from Consolidated income statement
EUR (€) € in Millions

	12 Months Ended		
	Dec. 31, 2017	**Dec. 31, 2016**	**Dec. 31, 2015**
Revenues	€41,096	€40,918	€40,236
…	…	…	…
Depreciation and amortization	(6,846)	(6,728)	(6,465)
…	…	…	…
Impairment of goodwill	(20)	(772)	
Impairment of fixed assets	(190)	(207)	(38)
…	…	…	…
Operating income	4,917	4,077	4,742
…	…	…	…
Consolidated net income of continuing operations	2,114	1,010	2,510
Consolidated net income of discontinued operations (EE)	29	2,253	448
Consolidated net income	**2,143**	**3,263**	**2,958**
Net income attributable to owners of the parent company	1,906	2,935	2,652
Non-controlling interests	€237	€328	€306

Excerpt from the Consolidated statement of financial position
EUR (€) € in Millions

Assets	**Dec. 31, 2017**	**Dec. 31, 2016**	**Dec. 31, 2015**
Goodwill	€27,095	€27,156	€27,071
Other intangible assets	14,339	14,602	14,327
Property, plant and equipment	26,665	25,912	25,123
…	…	…	…
Total non-current assets	74,035	74,819	71,330
…	…	…	…
Total current assets	20,679	19,849	14,312
Assets held for sale			5,788
Total assets	**94,714**	**94,668**	**91,430**
Equity and liabilities			
…	…	…	…
Total equity	32,942	33,174	33,267
…	…	…	…
Total non-current liabilities	32,736	35,590	36,537

(continued)

Exhibit 8 (Continued)

Excerpt from the Consolidated statement of financial position
EUR (€) € in Millions

Assets	Dec. 31, 2017	Dec. 31, 2016	Dec. 31, 2015
…	…	…	…
Total current liabilities	29,036	25,904	21,626
Total equity and liabilities	**94,714**	**94,668**	**91,430**

Exhibit 9 Orange
Excerpts from the 2017 Notes to the Consolidated Financial Statements

Excerpt from Note 7.2 Goodwill

[Excerpt] Reconciliation of Changes in Goodwill (€ in Millions)

	12 Months Ended		
	Dec. 31, 2017	Dec. 31, 2016	Dec. 31, 2015
Gross Value in the opening balance	€32,689	€32,606	€30,271
Acquisitions	38	904	2,333
Disposals	0	(6)	(69)
Translation adjustment	(40)	(815)	73
Reclassifications and other items	0	0	(2)
Reclassification to assets held for sale	0	0	0
Gross Value in Closing Balance	32,687	32,689	32,606
Accumulated Impairment losses in the opening balance	(5,533)	(5,535)	(5,487)
Impairment	(20)	(772)	0
Disposals	0	0	0
Translation adjustment	(39)	774	(48)
Reclassifications and other items	0	0	0
Reclassification to assets held for sale	0	0	0
Accumulated Impairment losses in the closing balance	€(5,592)	€(5,533)	€(5,535)
Net book value of goodwill	**27,095**	**27,156**	**27,071**

Excerpt* from Note 7.3 Key assumptions used to determine recoverable amounts as of 31 December 2017

The parameters used for the determination of recoverable amount of the main consolidated operations are set forth below:

	France	Spain	Poland	Belgium	Romania
Perpetuity growth rate	0.8%	1.5%	1.0%	0.5%	2.3%
Post-tax discount rate	5.5%	8.6%	8.3%	6.8%	8.8%

Excerpt* from Note 7.4 Sensitivity of recoverable amounts as of 31 December 2017

The level of sensitivity presented allows readers of the financial statements to estimate the impact in their own assessment.

(in billions of euros)	France	Spain	Poland	Belgium	Romania
Decrease by 1% in perpetuity growth rate	10.4	1.6	0.6	0.3	0.3
An increase by 1% in post-tax discount rate	11.4	2.0	0.6	0.3	0.3

* Table extracted presents only selected assumptions and selected countries.

The company's annual report provides more detail.

Goodwill is not amortized. It is tested for impairment at least annually and more frequently when there is an indication that it may be impaired .... These tests are performed at the level of each Cash Generating Unit (CGU) (or group of CGUs)... To determine whether an impairment loss should be recognized, the carrying value of the assets and liabilities of the CGUs or groups of CGUs is compared to recoverable amount, for which Orange uses mostly the value in use.... Value in use is the present value of the future expected cash flows. Cash flow projections are based on economic and regulatory assumptions, license renewal assumptions and forecast trading and investment activity drawn up by the Group's management...

Excerpt from Note 8.3 Other intangible assets – Net book value

	December 31		
(in millions of euros)	2017	2016	2015
Telecommunications licenses	6,233	6,440	5,842
Orange brand	3,133	3,133	3,133
Other brands	88	102	137
Customer bases	555	703	729
Software	3,946	3,781	3,815
Other intangible assets	384	443	671
Total	€14,339	€14,602	€14,327

Excerpt from Note 8.4 Property, plant and equipment – Net book value

	December 31		
(in millions of euros)	**2017**	**2016**	**2015**
Land and buildings	2,535	2,661	2,733
Network and terminals	22,880	21,984	21,194
IT equipment	802	784	787
Other property, plant and equipment	448	483	409
Total	€26,665	€25,912	€25,123

Exhibit 10 **Orange**
Excerpt from the 2017 Analysis of the Group's financial position and earnings

"Orange group operating income stood at 4,077 million euros in 2016, compared with 4,742 million euros in 2015 on a historical basis, a drop of 14.0% or 665 million euros. This drop on a historical basis was largely attributable to:

- the recognition, in 2016, of 772 million euros in impairment loss of goodwill … and 207 million euros in impairment loss of fixed assets … primarily relating to:
  - Poland for 507 million euros. This impairment loss mainly reflects a decline in competitiveness in the ADSL market, a deterioration in revenue assumptions in the mobile market and an increase in the post-tax discount rate due to the downgrading of the country's sovereign rating by the rating agencies,
  - Egypt for 232 million euros. This impairment loss reflects the financial terms of the 4G license awarded in 2016, the sharp depreciation of the Egyptian pound and increased political and economic uncertainty,
  - in the Congo (DRC), for 109 million euros. This impairment loss reflects political and economic uncertainty, a decline in purchasing power with a knock-on effect on the consumption of telecommunications products and services and an increased regulatory burden (particularly connected with the implementation of customer identification),
  - Cameroon for 90 million euros. This impairment loss reflects a decline in voice revenues following the surge in messaging services and in VoIP of Over-The-Top (OTT) providers and heightened competition in the mobile market,

Exhibit 10 (Continued)

  - and Niger for 26 million euros;
- and the 263 million euro increase in depreciation and amortization

  …

1 What proportion of Orange's total assets as of December 31, 2017, is represented by goodwill and other intangible assets?

2 What is the largest component of the company's impairment losses during the year ending December 2016?

3 The company discloses that it determines whether an impairment loss should be recognized by comparing the carrying value of a unit's assets and liabilities to the "recoverable amount" for which the company uses mostly the value in use. How does the company determine value in use?

4 By what amount would the estimated recoverable value of the company's operations in France, Spain, Poland, Belgium and Romania change if the company decreased its estimate of the perpetuity growth rate by 1%? By what amount would the estimated recoverable value of these operations change if the company increased its estimate of the post-tax discount rate by 1%?

5 What are the largest components of other intangible assets as of December 31, 2017? What is the largest component of property, plant and equipment as of December 31, 2017?

Solution to 1:

As of 31 December 2017, goodwill represents 28% (= 27,095 ÷ 97,714) of Orange's total assets. Other intangible assets represent 15% (= 14,339 ÷ 97,714). Data are from the company's balance sheet in Exhibit 8.

Solution to 2:

The largest component of the € 772 impairment loss on goodwill and the € 207 million impairment loss of fixed assets related to a € 507 million loss in Poland. The company attributed the loss to a decline in the competitiveness of the market for its ADSL technology, a reduction in revenue assumptions, and an increase in the discount rate resulting from the downgrading of the country's debt rating. From Exhibit 10.

[The company's financial statements define ADSL (Asymmetrical Digital Subscriber Line) as a "broadband data transmission technology on the traditional telephone network. It enables broadband data transmission (first and foremost Internet access) via twisted paired copper cable (the most common type of telephone line found in buildings)."]

Solution to 3:

The company determines value in use – which it uses as a unit's assets and liabilities "recoverable amount" in impairment testing – as the present value of the future expected cash flows. The cash flow projections are based on management's assumptions. From Note 7.4 in Exhibit 9.

Solution to 4:

If the company decreased its estimate of the perpetuity growth rate by 1%, the estimated recoverable value of the company's operations in France, Spain, Poland, Belgium and Romania would change by €13.2 billion (=10.4 + 1.6 +

0.6 + 0.3 + 0.3). A decrease in estimated growth decreases the present value of the cash flows. If the company increased its estimate of the post-tax discount rate by 1%, the estimated recoverable value of these operations would change by €14.6 billion (=11.4 + 2.0 + 0.6 + 0.3 + 0.3). An increase in the discount rate decreases the present value of cash flows. Data are from Note 7.4 in Exhibit 9.

Solution to 5:

The largest components of other intangible assets as of December 31, 2017, are telecommunications licenses, software, and the Orange brand, reported at €6,233 million, €3,946 million, and €3,133 million, respectively. The largest component of property, plant and equipment as of December 31, 2017, is network and terminals (€22,880 million.) Data are from Note 8.3 and 8.4 in Exhibit 9.

Note that the exhibits in the previous example contain relatively brief excerpts from the company's disclosures. The complete text of the disclosures concerning the company's non-current assets spans numerous different footnotes, some of which are several pages long. Overall, an analyst can use the disclosures to understand a company's investments in tangible and intangible assets, how those investments changed during a reporting period, how those changes affected current performance, and what those changes might indicate about future performance.

12

USING DISCLOSURES IN ANALYSIS

m analyze and interpret financial statement disclosures regarding property, plant, and equipment and intangible assets;

Ratios used in analyzing fixed assets include the fixed asset turnover ratio and several asset age ratios. The fixed asset turnover ratio (total revenue divided by average net fixed assets) reflects the relationship between total revenues and investment in PPE. The higher this ratio, the higher the amount of sales a company is able to generate with a given amount of investment in fixed assets. A higher asset turnover ratio is often interpreted as an indicator of greater efficiency.

Asset age ratios generally rely on the relationship between historical cost and depreciation. Under the revaluation model (permitted under IFRS but not US GAAP), the relationship between carrying amount, accumulated depreciation, and depreciation expense will differ when the carrying amount differs significantly from the depreciated historical cost. Therefore, the following discussion of asset age ratios applies primarily to PPE reported under the cost model.

Asset age and remaining useful life, two asset age ratios, are important indicators of a company's need to reinvest in productive capacity. The older the assets and the shorter the remaining life, the more a company may need to reinvest to maintain productive capacity. The average age of a company's asset base can be estimated as accumulated depreciation divided by depreciation expense. The average remaining life of a company's asset base can be estimated as net PPE divided by depreciation expense. These estimates simply reflect the following relationships for assets accounted for on a historical cost basis: total historical cost minus accumulated depreciation equals net PPE; and, under straight-line depreciation, total historical cost less salvage value divided by estimated useful life equals annual depreciation expense. Equivalently, total historical cost less salvage value divided by annual depreciation expense equals estimated useful life. Assuming straight-line depreciation and no salvage value (for simplicity), we have the following:

Estimated total useful life = Time elapsed since purchase (Age) + Estimated remaining life

Historical cost ÷ annual depreciation expense = Estimated total useful life

Historical cost = Accumulated depreciation + Net PPE

Equivalently,

Estimated total useful life = Estimated age of equipment + Estimated remaining life

Historical cost ÷ annual depreciation expense = Accumulated depreciation ÷ annual depreciation expense + Net PPE ÷ annual depreciation expense

The application of these estimates can be illustrated by a hypothetical example of a company with a single depreciable asset. Assume the asset initially cost $100, had an estimated useful life of 10 years, and an estimated salvage value of $0. Each year, the company records a depreciation expense of $10, so accumulated depreciation will equal $10 times the number of years since the asset was acquired (when the asset is 7 years old, accumulated depreciation will be $70). Equivalently, the age of the asset will equal accumulated depreciation divided by the annual depreciation expense.

In practice, such estimates are difficult to make with great precision. Companies use depreciation methods other than the straight-line method and have numerous assets with varying useful lives and salvage values, including some assets that are fully depreciated, so this approach produces an estimate only. Moreover, fixed asset disclosures are often quite general. Consequently, these estimates may be primarily useful to identify areas for further investigation.

One further measure compares a company's current reinvestment in productive capacity. Comparing annual capital expenditures to annual depreciation expense provides an indication of whether productive capacity is being maintained. It is a very general indicator of the rate at which a company is replacing its PPE relative to the rate at which PPE is being depreciated.

EXAMPLE 17

Using Fixed Asset Disclosure to Compare Companies' Fixed Asset Turnover and Average Age of Depreciable Assets

You are analyzing the property, plant, and equipment of three international telecommunications companies:

- Orange SA, which we discussed previously, has been listed on Euronext Paris (symbol ORA) and on the New York Stock Exchange (symbol ORAN) since 1997. At December 31, 2017, the French government retained 22.95% of the share capital.

- BCE Inc., Canada's largest communications company, provides wireless, wireline, Internet, TV and business communications across Canada. BCE's shares are publicly traded on the Toronto Stock Exchange and on the New York Stock Exchange (TSX, NYSE: BCE).
- Verizon Communications Inc. is a US-based global provider of communications, information and entertainment products and services to consumers, businesses and governmental agencies. Verizon's shares are listed on the New York Stock Exchange and the NASDAQ Global Select Market (symbol VZ).

Exhibit 11 presents selected information from the companies' financial statements.

Exhibit 11

	Orange	BCE Inc	Verizon
Currency, Millions of:	**Euro €**	**Canadian $**	**US $**
Historical cost total PPE, end of year	€97,092	$69,230	$246,498
Accumulated depreciation, end of year	70,427	45,197	157,930
Net PPE, end of year	26,665	24,033	88,568
Net PPE, beginning of year	25,912	22,346	84,751
Revenues	41,096	22,719	126,034
Annual depreciation expense	4,708	3,037	14,741
Capital expenditure	5,677	4,149	17,247
Land included in PPE	Not separated	Not separated	806
Accounting standards	IFRS	IFRS	US GAAP
PPE measurement	Historical cost	Historical cost	Historical cost
Depreciation method	Straight-line	Straight-line	Straight-line

Sources: Companies' 2017 Annual Financial Reports.

1 Based on the above data for each company, estimate the total useful life, age, and remaining useful life of PPE.

2 Interpret the estimates. What items might affect comparisons across these companies?

3 How does each company's 2017 depreciation expense compare to its capital expenditures for the year?

4 Calculate and compare fixed asset turnover for each company.

Solution to 1:

The following table presents the estimated total useful life, estimated age, and estimated remaining useful life of PPE for each of the companies.

Estimates	Orange	BCE Inc	Verizon
Estimated total useful life (years)	20.6	22.8	16.7
Estimated age (years)	15.0	14.9	10.7
Estimated remaining life (years)	5.7	7.9	6.0

The computations are demonstrated using Verizon's data ($ millions). The estimated total useful life of PPE is total historical cost of PPE of $246,498 divided by annual depreciation expense of $14,741, giving 16.7 years. Estimated age and estimated remaining life are obtained by dividing accumulated depreciation of $157,930 and net PPE of $88,568 by the annual depreciation expense of $14,741, giving 10.7 years and 6.0 years, respectively.

Ideally, the estimates of asset lives illustrated in this example should exclude land, which is not depreciable, when the information is available; however, both Orange and BCE report Land and Buildings as a combined amount. We will use Verizon, for which land appeared to be disclosed separately in the above table, to illustrate the estimates with adjusting for land. As an illustration of the calculations to exclude land, excluding Verizon's land would give an estimated total useful life for the non-land PPE of 16.7 years [(total cost €246,498 minus land cost of $806) divided by annual depreciation expense of €14,741 million]. The estimate is essentially unchanged from the estimate including land because land represents such a small component of Verizon's PPE.

Solution to 2:

The estimated total useful life suggests that Orange and BCE depreciate PPE over a much longer period than Verizon: 20.6 and 22.8 years for Orange and BCE, respectively, versus 16.7 years for Verizon.

The estimated age of the equipment suggests that Verizon has the newest PPE with an estimated age of 10.7 years. Additionally, the estimates suggest that around 73 percent of Orange's assets' useful lives have passed (15.0 years ÷ 20.6 years, or equivalently, €70,427 million ÷ €97,092 million). In comparison, around 65 and 64 percent of the useful lives of the PPE of BCE and Verizon, respectively, have passed.

Items that can affect comparisons across the companies include business differences, such as differences in composition of the companies' operations and differences in acquisition and divestiture activity. This result can be compared, to an extent, to the useful lives and asset mix disclosed in the companies' footnotes; however, differences in disclosures, e.g. in the categories of assets disclosed, can affect comparisons.

Solution to 3:

All three companies' capital expenditure exceeds its depreciation expense. Rounding to the nearest 10%, capital expenditure as a percentage of depreciation is 120 percent for Orange, 140 percent for BCE, and 120 percent for Verizon. All three companies are replacing PPE at a faster rate than the PPE is being depreciated, consistent with the companies' somewhat older asset base.

Solution to 4:

Fixed asset turnover is calculated as total revenues divided by average net PPE. Orange's fixed asset turnover is 1.6 (= 41,096/((26,665 + 25,912)/2). BCE's fixed asset turnover is 1.0, and Verizon's fixed asset turnover is 1.5.

Orange's and Verizon's higher levels of fixed asset turnover indicate these companies, compared to BCE, are able to generate more sales per unit of investment in fixed assets.

13 INVESTMENT PROPERTY

- compare the financial reporting of investment property with that of property, plant, and equipment.

Investment property is defined under IFRS as property that is owned (or, in some cases, leased under a **finance lease**) for the purpose of earning rentals or capital appreciation or both.[28] An example of investment property is a building owned by a company and leased out to tenants. In contrast, other long-lived tangible assets (i.e., property considered to be property, plant, and equipment) are owner-occupied properties used for producing the company's goods and services or for housing the company's administrative activities. Investment properties do not include long-lived tangible assets held for sale in the ordinary course of business. For example, the houses and property owned by a housing construction company are considered to be its inventory.

Under IFRS, companies are allowed to value investment properties using either a cost model or a fair value model. The cost model is identical to the cost model used for property, plant, and equipment. If the cost model is used, the fair value of investment property must be disclosed.[29] The fair value model, however, differs from the revaluation model used for property, plant, and equipment. Under the revaluation model, whether an asset revaluation affects net income depends on whether the revaluation initially increases or decreases the carrying amount of the asset. In contrast, under the fair value model, all changes in the fair value of the asset affect net income. To use the fair value model, a company must be able to reliably determine the property's fair value on a continuing basis.

Example 18 presents an excerpt from the annual report of a property company reporting under IFRS.

EXAMPLE 18

Financial Statement Presentation and Disclosures for Long-Lived Assets

The following exhibit presents information and excerpts from the annual report for the year ended 31 December 2017 of intu properties plc, a property company headquartered in London that owns, develops and manages shopping centres in the United Kingdom and Spain. Its common stock is listed in London and Johannesburg.

Exhibit 12 Information and excerpts from the Annual Report of intu properties plc (Currency in £ millions)

Financial Information

Financial Statement	Item Label	Amount 2017	Amount 2016
Balance Sheet	Investment and development property	9,179.4	9,212.1
Balance Sheet	Plant and equipment	12.2	7.6

---

[28] IAS 40 *Investment Property* prescribes the accounting treatment for investment property.

[29] Ibid., paragraph 32.

Exhibit 12 (Continued)

Financial Statement	Item Label	Amount 2017	Amount 2016
Balance Sheet	Total assets	10,794.5	10,369.2
Income Statement	Net rental income	423.4	406.1
Income Statement	Revaluation of investment and development property	30.8	(78.0)

Excerpt from Note 2 Accounting policies

Investment and development property

Investment and development property is owned or leased by the Group and held for long-term rental income and capital appreciation.

The Group has elected to use the fair value model. Properties are initially recognised at cost and subsequently revalued at the balance sheet date to fair value as determined by professionally qualified external valuers on the basis of market value with the exception of certain development land where an assessment of fair value may be made internally. External valuations are received for significant development land once required planning permissions are obtained. The cost of investment and development property includes capitalised interest and other directly attributable outgoings incurred during development. Interest is capitalised on the basis of the average interest rate on the relevant debt outstanding. Interest ceases to be capitalised on the date of practical completion.

Gains or losses arising from changes in the fair value of investment and development property are recognised in the income statement. Depreciation is not provided in respect of investment and development property. Gains or losses arising on the sale of investment and development property are recognised when the significant risks and rewards of ownership have been transferred to the buyer. The gain or loss recognised is the proceeds received less the carrying value of the property and costs directly associated with the sale.

Plant and equipment

Plant and equipment consists of vehicles, fixtures, fittings and other equipment. Plant and equipment is stated at cost less accumulated depreciation and any accumulated impairment losses. Depreciation is charged to the income statement on a straight- line basis over an asset's estimated useful life up to a maximum of five years.

Excerpt from Note 14 Investment and development property

The market value of investment and development property at 31 December 2017 includes £8,831.9 million (31 December 2016: £9,088.6 million) in respect of investment property and £376.5 million (31 December 2016: £153.2 million) in respect of development property. …All the Group's significant investment and development property relates to prime shopping centres which are of a similar nature and share characteristics and risks.…

Valuation methodology

The fair value of the Group's investment and development property at 31 December 2017 was determined by independent external valuers … Fair values for investment properties are calculated using the present value income approach. …The key driver of the property valuations is the terms

> of the leases in place at the valuation date. These determine the majority of the cash flow profile of the property for a number of years and therefore form the base of the valuation…

1. How do the assets included in the balance sheet line item "Investment and development property" differ from the assets included in the balance sheet line item "Plant and equipment"?
2. How does the valuation model used by the company for its investment and development property differ from the valuation model used for its plant and equipment?
3. How does accounting for depreciation differ for investment and development property versus plant and equipment?
4. Do the revaluation gains and losses on investment and development properties indicate that the properties have been sold?

Solution to 1:

The assets included in the balance sheet line item "Investment and development property" are shopping centres which the company holds for long-term rental income and capital appreciation. In 2017, the company reported net rental income of £423.4 million. The balance sheet line item "Plant and equipment" includes vehicles, fixtures, fittings, and other equipment used by the company in its operations.

Solution to 2:

The valuation model used by the company for its investment and development property is the fair value model, in which properties are initially recognised at cost and subsequently revalued and shown on the balance sheet at fair value. All changes in the fair value of the asset affect net income. The company employs external valuation experts to determine the fair value, which is based on expected future cash flow from rental income.

The valuation model used for its plant and equipment is the historical cost model in which properties are shown on the balance sheet at cost minus accumulated depreciation and any impairment losses.

Solution to 3:

Depreciation in accounting refers to the allocation of the cost of a long-lived asset over its useful life. No depreciation is recorded for investment and development property. Depreciation expense for plant and equipment is calculated on a straight-line basis over the asset's estimated useful life.

Solution to 4:

No. The revaluation gains and losses on investment properties arise from changes in the fair value of properties that are owned by the company. The company reported a revaluation gain of £30.8 million in 2017 and a revaluation loss of £78.0 million in 2016.

Sales of property would have resulted in a gain or loss on disposal, calculated as the proceeds minus the carrying value of the property and related selling costs.

In general, a company must apply its chosen model (cost or fair value) to all of its investment property. If a company chooses the fair value model for its investment property, it must continue to use the fair value model until it disposes of the property or changes its use such that it is no longer considered investment property (e.g., it becomes owner-occupied property or part of inventory). The company must continue to use the fair value model for that property even if transactions on comparable properties, used to estimate fair value, become less frequent.

Certain valuation issues arise when a company changes the use of property such that it moves from being an investment property to owner-occupied property or part of inventory. If a company's chosen model for investment property is the cost model, such transfers do not change the carrying amount of the property transferred. If a company's chosen model is the fair value model, transfers from investment property to owner-occupied property or to inventory are made at fair value. In other words, the property's fair value at the time of transfer is considered to be its cost for ongoing accounting for the property. If a company's chosen model for investment property is the fair value model and it transfers a property from owner-occupied to investment property, the change in measurement of the property from depreciated cost to fair value is treated like a revaluation. If a company's chosen model is the fair value model and it transfers a property from inventory to investment property, any difference between the inventory carrying amount and the property's fair value at the time of transfer is recognised as profit or loss.

Investment property appears as a separate line item on the balance sheet. Companies are required to disclose whether they use the fair value model or the cost model for their investment property. If the company uses the fair value model, it must make additional disclosures about how it determines fair value and must provide reconciliation between the beginning and ending carrying amounts of investment property. If the company uses the cost model, it must make additional disclosures similar to those for property, plant, and equipment—for example, the depreciation method and useful lives must be disclosed. In addition, if the company uses the cost model, it must also disclose the fair value of investment property.

Under US GAAP, there is no specific definition of investment property. Most operating companies and real estate companies in the United States that hold investment-type property use the historical cost model.

SUMMARY

Understanding the reporting of long-lived assets at inception requires distinguishing between expenditures that are capitalised (i.e., reported as long-lived assets) and those that are expensed. Once a long-lived asset is recognised, it is reported under the cost model at its historical cost less accumulated depreciation (amortisation) and less any impairment or under the revaluation model at its fair value. IFRS permit the use of either the cost model or the revaluation model, whereas US GAAP require the use of the cost model. Most companies reporting under IFRS use the cost model. The choice of different methods to depreciate (amortise) long-lived assets can create challenges for analysts comparing companies.

Key points include the following:

- Expenditures related to long-lived assets are capitalised as part of the cost of assets if they are expected to provide future benefits, typically beyond one year. Otherwise, expenditures related to long-lived assets are expensed as incurred.
- Although capitalising expenditures, rather than expensing them, results in higher reported profitability in the initial year, it results in lower profitability in subsequent years; however, if a company continues to purchase similar or increasing amounts of assets each year, the profitability-enhancing effect of capitalisation continues.

- Capitalising an expenditure rather than expensing it results in a greater amount reported as cash from operations because capitalised expenditures are classified as an investing cash outflow rather than an operating cash outflow.
- Companies must capitalise interest costs associated with acquiring or constructing an asset that requires a long period of time to prepare for its intended use.
- Including capitalised interest in the calculation of interest coverage ratios provides a better assessment of a company's solvency.
- IFRS require research costs be expensed but allow all development costs (not only software development costs) to be capitalised under certain conditions. Generally, US accounting standards require that research and development costs be expensed; however, certain costs related to software development are required to be capitalised.
- When one company acquires another company, the transaction is accounted for using the acquisition method of accounting in which the company identified as the acquirer allocates the purchase price to each asset acquired (and each liability assumed) on the basis of its fair value. Under acquisition accounting, if the purchase price of an acquisition exceeds the sum of the amounts that can be allocated to individual identifiable assets and liabilities, the excess is recorded as goodwill.
- The capitalised costs of long-lived tangible assets and of intangible assets with finite useful lives are allocated to expense in subsequent periods over their useful lives. For tangible assets, this process is referred to as depreciation, and for intangible assets, it is referred to as amortisation.
- Long-lived tangible assets and intangible assets with finite useful lives are reviewed for impairment whenever changes in events or circumstances indicate that the carrying amount of an asset may not be recoverable.
- Intangible assets with an indefinite useful life are not amortised but are reviewed for impairment annually.
- Impairment disclosures can provide useful information about a company's expected cash flows.
- Methods of calculating depreciation or amortisation expense include the straight-line method, in which the cost of an asset is allocated to expense in equal amounts each year over its useful life; accelerated methods, in which the allocation of cost is greater in earlier years; and the units-of-production method, in which the allocation of cost corresponds to the actual use of an asset in a particular period.
- Estimates required for depreciation and amortisation calculations include the useful life of the equipment (or its total lifetime productive capacity) and its expected residual value at the end of that useful life. A longer useful life and higher expected residual value result in a smaller amount of annual depreciation relative to a shorter useful life and lower expected residual value.
- IFRS permit the use of either the cost model or the revaluation model for the valuation and reporting of long-lived assets, but the revaluation model is not allowed under US GAAP.
- Under the revaluation model, carrying amounts are the fair values at the date of revaluation less any subsequent accumulated depreciation or amortisation.
- In contrast with depreciation and amortisation charges, which serve to allocate the cost of a long-lived asset over its useful life, impairment charges reflect an unexpected decline in the fair value of an asset to an amount lower than its carrying amount.

- IFRS permit impairment losses to be reversed, with the reversal reported in profit. US GAAP do not permit the reversal of impairment losses.
- The gain or loss on the sale of long-lived assets is computed as the sales proceeds minus the carrying amount of the asset at the time of sale.
- Estimates of average age and remaining useful life of a company's assets reflect the relationship between assets accounted for on a historical cost basis and depreciation amounts.
- The average remaining useful life of a company's assets can be estimated as net PPE divided by depreciation expense, although the accounting useful life may not necessarily correspond to the economic useful life.
- Long-lived assets reclassified as held for sale cease to be depreciated or amortised. Long-lived assets to be disposed of other than by a sale (e.g., by abandonment, exchange for another asset, or distribution to owners in a spin-off) are classified as held for use until disposal. Thus, they continue to be depreciated and tested for impairment.
- Investment property is defined as property that is owned (or, in some cases, leased under a finance lease) for the purpose of earning rentals, capital appreciation, or both.
- Under IFRS, companies are allowed to value investment properties using either a cost model or a fair value model. The cost model is identical to the cost model used for property, plant, and equipment, but the fair value model differs from the revaluation model used for property, plant, and equipment. Unlike the revaluation model, under the fair value model, all changes in the fair value of investment property affect net income.
- Under US GAAP, investment properties are generally measured using the cost model.

PRACTICE PROBLEMS

1 JOOVI Inc. has recently purchased and installed a new machine for its manufacturing plant. The company incurred the following costs:

Purchase price	$12,980
Freight and insurance	$1,200
Installation	$700
Testing	$100
Maintenance staff training costs	$500

The total cost of the machine to be shown on JOOVI's balance sheet is *closest* to:

A $14,180.

B $14,980.

C $15,480.

2 Which costs incurred with the purchase of property and equipment are expensed?

A Delivery charges

B Installation and testing

C Training required to use the property and equipment

3 When constructing an asset for sale, directly related borrowing costs are *most likely*:

A expensed as incurred.

B capitalized as part of inventory.

C capitalized as part of property, plant, and equipment.

4 BAURU, S.A., a Brazilian corporation, borrows capital from a local bank to finance the construction of its manufacturing plant. The loan has the following conditions:

Borrowing date	1 January 2009
Amount borrowed	500 million Brazilian real (BRL)
Annual interest rate	14 percent
Term of the loan	3 years
Payment method	Annual payment of interest only. Principal amortization is due at the end of the loan term.

The construction of the plant takes two years, during which time BAURU earned BRL 10 million by temporarily investing the loan proceeds. Which of the following is the amount of interest related to the plant construction (in BRL million) that can be capitalized in BAURU's balance sheet?

A 130.

B 140.

C 210.

5 After reading the financial statements and footnotes of a company that follows IFRS, an analyst identified the following intangible assets:

- product patent expiring in 40 years;

© 2019 CFA Institute. All rights reserved.

- copyright with no expiration date; and
- goodwill acquired 2 years ago in a business combination.

Which of these assets is an intangible asset with a finite useful life?

	Product Patent	Copyright	Goodwill
A	Yes	Yes	No
B	Yes	No	No
C	No	Yes	Yes

6 Intangible assets with finite useful lives *mostly* differ from intangible assets with infinite useful lives with respect to accounting treatment of:

A revaluation.

B impairment.

C amortization.

7 Costs incurred for intangible assets are generally expensed when they are:

A internally developed.

B individually acquired.

C acquired in a business combination.

8 Under US GAAP, when assets are acquired in a business combination, goodwill *most likely* arises from:

A contractual or legal rights.

B assets that can be separated from the acquired company.

C assets that are neither tangible nor identifiable intangible assets.

9 All else equal, in the fiscal year when long-lived equipment is purchased:

A depreciation expense increases.

B cash from operations decreases.

C net income is reduced by the amount of the purchase.

10 Companies X and Z have the same beginning-of-the-year book value of equity and the same tax rate. The companies have identical transactions throughout the year and report all transactions similarly except for one. Both companies acquire a £300,000 printer with a three-year useful life and a salvage value of £0 on 1 January of the new year. Company X capitalizes the printer and depreciates it on a straight-line basis, and Company Z expenses the printer. The following year-end information is gathered for Company X.

	Company X **As of 31 December**
Ending shareholders' equity	£10,000,000
Tax rate	25%
Dividends	£0.00
Net income	£750,000

Based on the information given, Company Z's return on equity using year-end equity will be *closest* to:

A 5.4%.

B 6.1%.

C 7.5%.

11 A financial analyst is studying the income statement effect of two alternative depreciation methods for a recently acquired piece of equipment. She gathers the following information about the equipment's expected production life and use:

	Year 1	Year 2	Year 3	Year 4	Year 5	Total
Units of production	2,000	2,000	2,000	2,000	2,500	10,500

Compared with the units-of-production method of depreciation, if the company uses the straight-line method to depreciate the equipment, its net income in Year 1 will *most likely* be:

A lower.

B higher.

C the same.

12 A company purchases a piece of equipment for €1,500. The equipment is expected to have a useful life of five years and no residual value. In the first year of use, the units of production are expected to be 15% of the equipment's lifetime production capacity and the equipment is expected to generate €1,500 of revenue and incur €500 of cash expenses.

The depreciation method yielding the lowest operating profit on the equipment in the first year of use is:

A straight line.

B units of production.

C double-declining balance.

13 Juan Martinez, CFO of VIRMIN, S.A., is selecting the depreciation method to use for a new machine. The machine has an expected useful life of six years. Production is expected to be relatively low initially but to increase over time. The method chosen for tax reporting must be the same as the method used for financial reporting. If Martinez wants to minimize tax payments in the first year of the machine's life, which of the following depreciation methods is Martinez *most likely* to use?

A Straight-line method.

B Units-of-production method.

C Double-declining balance method.

The following information relates to Questions 14–15

Miguel Rodriguez of MARIO S.A., an Uruguayan corporation, is computing the depreciation expense of a piece of manufacturing equipment for the fiscal year ended 31 December 2009. The equipment was acquired on 1 January 2009. Rodriguez gathers the following information (currency in Uruguayan pesos, UYP):

Cost of the equipment	UYP 1,200,000
Estimated residual value	UYP 200,000
Expected useful life	8 years
Total productive capacity	800,000 units

Production in FY 2009	135,000 units
Expected production for the next 7 years	95,000 units each year

14 If MARIO uses the straight-line method, the amount of depreciation expense on MARIO's income statement related to the manufacturing equipment is *closest* to:

A 125,000.

B 150,000.

C 168,750.

15 If MARIO uses the units-of-production method, the amount of depreciation expense (in UYP) on MARIO's income statement related to the manufacturing equipment is *closest* to:

A 118,750.

B 168,750.

C 202,500.

---

16 Which of the following amortization methods is *most likely* to evenly distribute the cost of an intangible asset over its useful life?

A Straight-line method.

B Units-of-production method.

C Double-declining balance method.

17 Which of the following will cause a company to show a lower amount of amortization of intangible assets in the first year after acquisition?

A A higher residual value.

B A higher amortization rate.

C A shorter useful life.

18 A company purchases equipment for $200,000 with a five-year useful life and salvage value of zero. It uses the double-declining balance method of depreciation for two years, then shifts to straight-line depreciation at the beginning of Year 3. Compared with annual depreciation expense under the double-declining balance method, the resulting annual depreciation expense in Year 4 is:

A smaller.

B the same.

C greater.

19 An analyst in the finance department of BOOLDO S.A., a French corporation, is computing the amortization of a customer list, an intangible asset, for the fiscal year ended 31 December 2009. She gathers the following information about the asset:

Acquisition cost	€2,300,000
Acquisition date	1 January 2008
Expected residual value at time of acquisition	€500,000

The customer list is expected to result in extra sales for three years after acquisition. The present value of these expected extra sales exceeds the cost of the list.

If the analyst uses the straight-line method, the amount of accumulated amortization related to the customer list as of 31 December 2009 is *closest* to:

A €600,000.

B €1,200,000.

C €1,533,333.

20 A financial analyst is analyzing the amortization of a product patent acquired by MAKETTI S.p.A., an Italian corporation. He gathers the following information about the patent:

Acquisition cost	€5,800,000
Acquisition date	1 January 2009
Patent expiration date	31 December 2015
Total plant capacity of patented product	40,000 units per year
Production of patented product in fiscal year ended 31 December 2009	20,000 units
Expected production of patented product during life of the patent	175,000 units

If the analyst uses the units-of-production method, the amortization expense on the patent for fiscal year 2009 is *closest* to:

A €414,286.

B €662,857.

C €828,571.

21 A company acquires a patent with an expiration date in six years for ¥100 million. The company assumes that the patent will generate economic benefits that will decline over time and decides to amortize the patent using the double-declining balance method. The annual amortization expense in Year 4 is closest to:

A ¥6.6 million.

B ¥9.9 million.

C ¥19.8 million.

22 A company is comparing straight-line and double-declining balance amortization methods for a non-renewable six-year license, acquired for €600,000. The difference between the Year 4 ending net book values using the two methods is *closest to*:

A €81,400.

B €118,600.

C €200,000.

23 MARU S.A. de C.V., a Mexican corporation that follows IFRS, has elected to use the revaluation model for its property, plant, and equipment. One of MARU's machines was purchased for 2,500,000 Mexican pesos (MXN) at the beginning of the fiscal year ended 31 March 2010. As of 31 March 2010, the machine has a fair value of MXN 3,000,000. Should MARU show a profit for the revaluation of the machine?

A Yes.

B No, because this revaluation is recorded directly in equity.

C No, because value increases resulting from revaluation can never be recognized as a profit.

24 An analyst is studying the impairment of the manufacturing equipment of WLP Corp., a UK-based corporation that follows IFRS. He gathers the following information about the equipment:

Fair value	£16,800,000
Costs to sell	£800,000
Value in use	£14,500,000
Net carrying amount	£19,100,000

The amount of the impairment loss on WLP Corp.'s income statement related to its manufacturing equipment is *closest* to:

A £2,300,000.

B £3,100,000.

C £4,600,000.

25 Under IFRS, an impairment loss on a property, plant, and equipment asset is measured as the excess of the carrying amount over the asset's:

A fair value.

B recoverable amount.

C undiscounted expected future cash flows.

26 A financial analyst at BETTO S.A. is analyzing the result of the sale of a vehicle for 85,000 Argentine pesos (ARP) on 31 December 2009. The analyst compiles the following information about the vehicle:

Acquisition cost of the vehicle	ARP 100,000
Acquisition date	1 January 2007
Estimated residual value at acquisition date	ARP 10,000
Expected useful life	9 years
Depreciation method	Straight-line

The result of the sale of the vehicle is *most likely*:

A a loss of ARP 15,000.

B a gain of ARP 15,000.

C a gain of ARP 18,333.

27 CROCO S.p.A sells an intangible asset with a historical acquisition cost of €12 million and an accumulated amortization of €2 million and reports a loss on the sale of €3.2 million. Which of the following amounts is *most likely* the sale price of the asset?

A €6.8 million

B €8.8 million

C €13.2 million

28 The impairment of intangible assets with finite lives affects:

A the balance sheet but not the income statement.

B the income statement but not the balance sheet.

C both the balance sheet and the income statement.

29 The gain or loss on a sale of a long-lived asset to which the revaluation model has been applied is *most likely* calculated using sales proceeds less:

A carrying amount.

B carrying amount adjusted for impairment.

C historical cost net of accumulated depreciation.

30 According to IFRS, all of the following pieces of information about property, plant, and equipment must be disclosed in a company's financial statements and footnotes *except for*:

A useful lives.

B acquisition dates.

C amount of disposals.

31 According to IFRS, all of the following pieces of information about intangible assets must be disclosed in a company's financial statements and footnotes *except for*:

A fair value.

B impairment loss.

C amortization rate.

32 Which of the following is a required financial statement disclosure for long-lived intangible assets under US GAAP?

A The useful lives of assets

B The reversal of impairment losses

C Estimated amortization expense for the next five fiscal years

33 Which of the following characteristics is *most likely* to differentiate investment property from property, plant, and equipment?

A It is tangible.

B It earns rent.

C It is long-lived.

34 If a company uses the fair value model to value investment property, changes in the fair value of the asset are *least likely* to affect:

A net income.

B net operating income.

C other comprehensive income.

35 Investment property is *most likely* to:

A earn rent.

B be held for resale.

C be used in the production of goods and services.

36 A company is *most likely* to:

A use a fair value model for some investment property and a cost model for other investment property.

B change from the fair value model when transactions on comparable properties become less frequent.

C change from the fair value model when the company transfers investment property to property, plant, and equipment.

37 Under the revaluation model for property, plant, and equipment and the fair model for investment property:

A fair value of the asset must be able to be measured reliably.

B net income is affected by all changes in the fair value of the asset.

C net income is never affected if the asset increases in value from its carrying amount.

38 Under IFRS, what must be disclosed under the cost model of valuation for investment properties?

A Useful lives

B The method for determining fair value

C Reconciliation between beginning and ending carrying amounts of investment property

The following information relates to Questions 39–42

Melanie Hart, CFA, is a transportation analyst. Hart has been asked to write a research report on Altai Mountain Rail Company (AMRC). Like other companies in the railroad industry, AMRC's operations are capital intensive, with significant investments in such long-lived tangible assets as property, plant, and equipment. In November of 2008, AMRC's board of directors hired a new team to manage the company. In reviewing the company's 2009 annual report, Hart is concerned about some of the accounting choices that the new management has made. These choices differ from those of the previous management and from common industry practice. Hart has highlighted the following statements from the company's annual report:

Statement 1	"In 2009, AMRC spent significant amounts on track replacement and similar improvements. AMRC expensed rather than capitalised a significant proportion of these expenditures."
Statement 2	"AMRC uses the straight-line method of depreciation for both financial and tax reporting purposes to account for plant and equipment."
Statement 3	"In 2009, AMRC recognized an impairment loss of €50 million on a fleet of locomotives. The impairment loss was reported as 'other income' in the income statement and reduced the carrying amount of the assets on the balance sheet."

Exhibits 1 and 2 contain AMRC's 2009 consolidated income statement and balance sheet. AMRC prepares its financial statements in accordance with International Financial Reporting Standards.

Exhibit 1 Consolidated Statement of Income

	2009		2008	
For the Years Ended 31 December	**€ Millions**	**% Revenues**	**€ Millions**	**% Revenues**
Operating revenues	2,600	100.0	2,300	100.0
Operating expenses				
Depreciation	(200)	(7.7)	(190)	(8.3)
Other operating expense	(1,590)	(61.1)	(1,515)	(65.9)
Total operating expenses	(1,790)	(68.8)	(1,705)	(74.2)
Operating income	810	31.2	595	25.8
Other income	(50)	(1.9)	—	0.0
Interest expense	(73)	(2.8)	(69)	(3.0)
Income before taxes	687	26.5	526	22.8

(continued)

Exhibit 1 (Continued)

	2009		2008	
For the Years Ended 31 December	**€ Millions**	**% Revenues**	**€ Millions**	**% Revenues**
Income taxes	(272)	(10.5)	(198)	(8.6)
Net income	415	16	328	14.2

Exhibit 2 Consolidated Balance Sheet

As of 31 December	2009		2008	
Assets	**€ Millions**	**% Assets**	**€ Millions**	**% Assets**
Current assets	500	9.4	450	8.5
Property & equipment:				
Land	700	13.1	700	13.2
Plant & equipment	6,000	112.1	5,800	109.4
Total property & equipment	6,700	125.2	6,500	122.6
Accumulated depreciation	(1,850)	(34.6)	(1,650)	(31.1)
Net property & equipment	4,850	90.6	4,850	91.5
Total assets	5,350	100.0	5,300	100.0
Liabilities and Shareholders' Equity				
Current liabilities	480	9.0	430	8.1
Long-term debt	1,030	19.3	1,080	20.4
Other long-term provisions and liabilities	1,240	23.1	1,440	27.2
Total liabilities	2,750	51.4	2,950	55.7
Shareholders' equity				
Common stock and paid-in-surplus	760	14.2	760	14.3
Retained earnings	1,888	35.5	1,600	30.2
Other comprehensive losses	(48)	(0.9)	(10)	(0.2)
Total shareholders' equity	2,600	48.6	2,350	44.3
Total liabilities & shareholders' equity	5,350	100.0	5,300	100.0

39 With respect to Statement 1, which of the following is the *most likely* effect of management's decision to expense rather than capitalise these expenditures?

A 2009 net profit margin is higher than if the expenditures had been capitalised.

B 2009 total asset turnover is lower than if the expenditures had been capitalised.

C Future profit growth will be higher than if the expenditures had been capitalised.

40 With respect to Statement 2, what would be the *most likely* effect in 2010 if AMRC were to switch to an accelerated depreciation method for both financial and tax reporting?

A Net profit margin would increase.

B Total asset turnover would decrease.

C Cash flow from operating activities would increase.

41 With respect to Statement 3, what is the *most likely* effect of the impairment loss?

A Net income in years prior to 2009 was likely understated.

B Net profit margins in years after 2009 will likely exceed the 2009 net profit margin.

C Cash flow from operating activities in 2009 was likely lower due to the impairment loss.

42 Based on Exhibits 1 and 2, the *best estimate* of the average remaining useful life of the company's plant and equipment at the end of 2009 is:

A 20.75 years.

B 24.25 years.

C 30.00 years.

The following information relates to Questions 43–48

Brian Jordan is interviewing for a junior equity analyst position at Orion Investment Advisors. As part of the interview process, Mary Benn, Orion's Director of Research, provides Jordan with information about two hypothetical companies, Alpha and Beta, and asks him to comment on the information on their financial statements and ratios. Both companies prepare their financial statements in accordance with International Financial Reporting Standards (IFRS) and are identical in all respects except for their accounting choices.

Jordan is told that at the beginning of the current fiscal year, both companies purchased a major new computer system and began building new manufacturing plants for their own use. Alpha capitalised and Beta expensed the cost of the computer system; Alpha capitalised and Beta expensed the interest costs associated with the construction of the manufacturing plants.

Benn asks Jordan, "What was the impact of these decisions on each company's current fiscal year financial statements and ratios?"

Jordan responds, "Alpha's decision to capitalise the cost of its new computer system instead of expensing it results in lower net income, lower total assets, and higher cash flow from operating activities in the current fiscal year. Alpha's decision to capitalise its interest costs instead of expensing them results in a lower fixed asset turnover ratio and a higher interest coverage ratio."

Jordan is told that Alpha uses the straight-line depreciation method and Beta uses an accelerated depreciation method; both companies estimate the same useful lives for long-lived assets. Many companies in their industry use the units-of-production method.

Benn asks Jordan, "What are the financial statement implications of each depreciation method, and how do you determine a company's need to reinvest in its productive capacity?"

Jordan replies, “All other things being equal, the straight-line depreciation method results in the least variability of net profit margin over time, while an accelerated depreciation method results in a declining trend in net profit margin over time. The units-of-production can result in a net profit margin trend that is quite variable. I use a three-step approach to estimate a company’s need to reinvest in its productive capacity. First, I estimate the average age of the assets by dividing net property, plant, and equipment by annual depreciation expense. Second, I estimate the average remaining useful life of the assets by dividing accumulated depreciation by depreciation expense. Third, I add the estimates of the average remaining useful life and the average age of the assets in order to determine the total useful life.”

Jordan is told that at the end of the current fiscal year, Alpha revalued a manufacturing plant; this increased its reported carrying amount by 15 percent. There was no previous downward revaluation of the plant. Beta recorded an impairment loss on a manufacturing plant; this reduced its carrying by 10 percent.

Benn asks Jordan “What was the impact of these decisions on each company’s current fiscal year financial ratios?”

Jordan responds, “Beta’s impairment loss increases its debt to total assets and fixed asset turnover ratios, and lowers its cash flow from operating activities. Alpha’s revaluation increases its debt to capital and return on assets ratios, and reduces its return on equity.”

At the end of the interview, Benn thanks Jordan for his time and states that a hiring decision will be made shortly.

43 Jordan’s response about the financial statement impact of Alpha’s decision to capitalise the cost of its new computer system is most likely *correct* with respect to:

A lower net income.

B lower total assets.

C higher cash flow from operating activities.

44 Jordan’s response about the ratio impact of Alpha’s decision to capitalise interest costs is most likely *correct* with respect to the:

A interest coverage ratio.

B fixed asset turnover ratio.

C interest coverage and fixed asset turnover ratios.

45 Jordan’s response about the impact of the different depreciation methods on net profit margin is most likely *incorrect* with respect to:

A accelerated depreciation.

B straight-line depreciation.

C units-of-production depreciation.

46 Jordan’s response about his approach to estimating a company’s need to reinvest in its productive capacity is most likely *correct* regarding:

A estimating the average age of the asset base.

B estimating the total useful life of the asset base.

C estimating the average remaining useful life of the asset base.

47 Jordan’s response about the effect of Beta’s impairment loss is most likely *incorrect* with respect to the impact on its:

A debt to total assets.

B fixed asset turnover.

C cash flow from operating activities.

48 Jordan's response about the effect of Alpha's revaluation is most likely *correct* with respect to the impact on its:

A return on equity.

B return on assets.

C debt to capital ratio.

SOLUTIONS

1 B is correct. Only costs necessary for the machine to be ready to use can be capitalized. Therefore, Total capitalized costs = 12,980 + 1,200 + 700 + 100 = \$14,980.

2 C is correct. When property and equipment are purchased, the assets are recorded on the balance sheet at cost. Costs for the assets include all expenditures required to prepare the assets for their intended use. Any other costs are expensed. Costs to train staff for using the machine are not required to prepare the property and equipment for their intended use, and these costs are expensed.

3 B is correct. When a company constructs an asset, borrowing costs incurred directly related to the construction are generally capitalized. If the asset is constructed for sale, the borrowing costs are classified as inventory.

4 A is correct. Borrowing costs can be capitalized under IFRS until the tangible asset is ready for use. Also, under IFRS, income earned on temporarily investing the borrowed monies decreases the amount of borrowing costs eligible for capitalization. Therefore, Total capitalized interest = (500 million × 14% × 2 years) – 10 million = 130 million.

5 B is correct. A product patent with a defined expiration date is an intangible asset with a finite useful life. A copyright with no expiration date is an intangible asset with an indefinite useful life. Goodwill is no longer considered an intangible asset under IFRS and is considered to have an indefinite useful life.

6 C is correct. An intangible asset with a finite useful life is amortized, whereas an intangible asset with an indefinite useful life is not.

7 A is correct. The costs to internally develop intangible assets are generally expensed when incurred.

8 C is correct. Under both International Financial Reporting Standards (IFRS) and US GAAP, if an item is acquired in a business combination and cannot be recognized as a tangible asset or identifiable intangible asset, it is recognized as goodwill. Under US GAAP, assets arising from contractual or legal rights and assets that can be separated from the acquired company are recognized separately from goodwill.

9 A is correct. In the fiscal year when long-lived equipment is purchased, the assets on the balance sheet increase and depreciation expense on the income statement increases because of the new long-lived asset.

10 B is correct. Company Z's return on equity based on year-end equity value will be 6.1%. Company Z will have an additional £200,000 of expenses compared with Company X. Company Z expensed the printer for £300,000 rather than capitalizing the printer and having a depreciation expense of £100,000 like Company X. Company Z's net income and shareholders' equity will be £150,000 lower (= £200,000 × 0.75) than that of Company X.

$$\begin{aligned}\text{ROE} &= \left(\frac{\text{Net income}}{\text{Shareholders' Equity}}\right)\\ &= £600{,}000/£9{,}850{,}000\\ &= 0.61 = 6.1\%\end{aligned}$$

11 A is correct. If the company uses the straight-line method, the depreciation expense will be one-fifth (20 percent) of the depreciable cost in Year 1. If it uses the units-of-production method, the depreciation expense will be 19 percent (2,000/10,500) of the depreciable cost in Year 1. Therefore, if the company uses the straight-line method, its depreciation expense will be higher and its net income will be lower.

12 C is correct. The operating income or earnings before interest and taxes will be lowest for the method that results in the highest depreciation expense. The double-declining balance method results in the highest depreciation expense in the first year of use.

Depreciation expense:

Straight line = €1,500/5 = €300.

Double-declining balance = €1,500 × 0.40 = €600.

Units of production = €1,500 × 0.15 = €225.

13 C is correct. If Martinez wants to minimize tax payments in the first year of the machine's life, he should use an accelerated method, such as the double-declining balance method.

14 A is correct. Using the straight-line method, depreciation expense amounts to

Depreciation expense = (1,200,000 – 200,000)/8 years = 125,000.

15 B is correct. Using the units-of-production method, depreciation expense amounts to

Depreciation expense = (1,200,000 – 200,000) × (135,000/800,000) = 168,750.

16 A is correct. The straight-line method is the method that evenly distributes the cost of an asset over its useful life because amortization is the same amount every year.

17 A is correct. A higher residual value results in a lower total depreciable cost and, therefore, a lower amount of amortization in the first year after acquisition (and every year after that).

18 C is correct. Shifting at the end of Year 2 from double-declining balance to straight-line depreciation methodology results in depreciation expense being the same in each of Years 3, 4, and 5. Shifting to the straight-line methodology at the beginning of Year 3 results in a greater depreciation expense in Year 4 than would have been calculated using the double-declining balance method.

Depreciation expense Year 4 (Using double-declining balance method all five years)

= 2 × Annual depreciation % using straight-line method × Carrying amount at end of Year 3
= 40% × $43,200

Depreciation expense Year 4 with switch to straight-line method in Year 3

= ⅓ × Remaining depreciable cost at start of Year 3
= ⅓ × $72,000
= $24,000

19 B is correct. Using the straight-line method, accumulated amortization amounts to

$$\begin{aligned}\text{Accumulated amortization} &= [(2{,}300{,}000 - 500{,}000)/3\text{ years}] \times 2\text{ years}\\ &= 1{,}200{,}000\end{aligned}$$

20 B is correct. Using the units-of-production method, depreciation expense amounts to

$$\text{Depreciation expense} = 5{,}800{,}000 \times (20{,}000/175{,}000) = 662{,}857$$

21 B is correct. As shown in the following calculations, under the double-declining balance method, the annual amortization expense in Year 4 is closest to ¥9.9 million.

Annual amortization expense = 2 × Straight-line amortization rate × Net book value.

Amortization expense Year 4 = 33.3% × ¥29.6 million = ¥9.9 million.

22 A is correct. As shown in the following calculations, at the end of Year 4, the difference between the net book values calculated using straight-line versus double-declining balance is closest to €81,400.

Net book value end of Year 4 using straight-line method = €600,000 – [4 × (€600,000/6)] = €200,000.

Net book value end of Year 4 using double-declining balance method = €600,000 $(1 - 33.33\%)^4 \approx$ €118,600.

23 B is correct. In this case, the value increase brought about by the revaluation should be recorded directly in equity. The reason is that under IFRS, an increase in value brought about by a revaluation can only be recognized as a profit to the extent that it reverses a revaluation decrease of the same asset previously recognized in the income statement.

24 B is correct. The impairment loss equals £3,100,000.

$$\begin{aligned}\text{Impairment} &= \max(\text{Fair value less costs to sell; Value in use}) - \text{Net carrying amount}\\ &= \max(16{,}800{,}000 - 800{,}000;\ 14{,}500{,}000) - 19{,}100{,}000\\ &= -3{,}100{,}000.\end{aligned}$$

25 B is correct. Under IFRS, an impairment loss is measured as the excess of the carrying amount over the asset's recoverable amount. The recoverable amount is the higher of the asset's fair value less costs to sell and its value in use. Value in use is a discounted measure of expected future cash flows. Under US GAAP, assessing recoverability is separate from measuring the impairment loss. If the asset's carrying amount exceeds its undiscounted expected future cash flows, the asset's carrying amount is considered unrecoverable and the impairment loss is measured as the excess of the carrying amount over the asset's fair value.

26 B is correct. The result on the sale of the vehicle equals

$$\begin{aligned}\text{Gain or loss on the sale} &= \text{Sale proceeds} - \text{Carrying amount}\\ &= \text{Sale proceeds} - (\text{Acquisition cost} - \text{Accumulated depreciation})\\ &= 85{,}000 - \{100{,}000 - [((100{,}000 - 10{,}000)/9\text{ years}) \times 3\text{ years}]\}\\ &= 15{,}000.\end{aligned}$$

27 A is correct. Gain or loss on the sale = Sale proceeds – Carrying amount. Rearranging this equation, Sale proceeds = Carrying amount + Gain or loss on sale. Thus, Sale price = (12 million – 2 million) + (–3.2 million) = 6.8 million.

28 C is correct. The carrying amount of the asset on the balance sheet is reduced by the amount of the impairment loss, and the impairment loss is reported on the income statement.

29 A is correct. The gain or loss on the sale of long-lived assets is computed as the sales proceeds minus the carrying amount of the asset at the time of sale. This is true under the cost and revaluation models of reporting long-lived assets. In the absence of impairment losses, under the cost model, the carrying amount will equal historical cost net of accumulated depreciation.

30 B is correct. IFRS do not require acquisition dates to be disclosed.

31 A is correct. IFRS do not require fair value of intangible assets to be disclosed.

32 C is correct. Under US GAAP, companies are required to disclose the estimated amortization expense for the next five fiscal years. Under US GAAP, there is no reversal of impairment losses. Disclosure of the useful lives—finite or indefinite and additional related details—is required under IFRS.

33 B is correct. Investment property earns rent. Investment property and property, plant, and equipment are tangible and long-lived.

34 C is correct. When a company uses the fair value model to value investment property, changes in the fair value of the property are reported in the income statement—not in other comprehensive income.

35 A is correct. Investment property earns rent. Inventory is held for resale, and property, plant, and equipment are used in the production of goods and services.

36 C is correct. A company will change from the fair value model to either the cost model or revaluation model when the company transfers investment property to property, plant, and equipment.

37 A is correct. Under both the revaluation model for property, plant, and equipment and the fair model for investment property, the asset's fair value must be able to be measured reliably. Under the fair value model, net income is affected by all changes in the asset's fair value. Under the revaluation model, any increase in an asset's value to the extent that it reverses a previous revaluation decrease will be recognized on the income statement and increase net income.

38 A is correct. Under IFRS, when using the cost model for its investment properties, a company must disclose useful lives. The method for determining fair value, as well as reconciliation between beginning and ending carrying amounts of investment property, is a required disclosure when the fair value model is used.

39 C is correct. Expensing rather than capitalising an investment in long-term assets will result in higher expenses and lower net income and net profit margin in the current year. Future years' incomes will not include depreciation expense related to these expenditures. Consequently, year-to-year growth in profitability will be higher. If the expenses had been capitalised, the carrying amount of the assets would have been higher and the 2009 total asset turnover would have been lower.

40 C is correct. In 2010, switching to an accelerated depreciation method would increase depreciation expense and decrease income before taxes, taxes payable, and net income. Cash flow from operating activities would increase because of the resulting tax savings.

41 B is correct. 2009 net income and net profit margin are lower because of the impairment loss. Consequently, net profit margins in subsequent years are likely to be higher. An impairment loss suggests that insufficient depreciation expense was recognized in prior years, and net income was overstated in prior years. The impairment loss is a non-cash item and will not affect operating cash flows.

42 A is correct. The estimated average remaining useful life is 20.75 years.

$$\begin{aligned}\text{Estimate of remaining useful life} &= \text{Net plant and equipment} \div \text{Annual depreciation expense}\\ \text{Net plant and equipment} &= \text{Gross P \& E} - \text{Accumulated depreciation}\\ &= €6000 - €1850 = €4150\\ \text{Estimate of remaining useful life} &= \text{Net P \& E} \div \text{Depreciation expense}\\ &= €4150 \div €200 = 20.75\end{aligned}$$

43 C is correct. The decision to capitalise the costs of the new computer system results in higher cash flow from operating activities; the expenditure is reported as an outflow of investing activities. The company allocates the capitalised amount over the asset's useful life as depreciation or amortisation expense rather than expensing it in the year of expenditure. Net income and total assets are higher in the current fiscal year.

44 B is correct. Alpha's fixed asset turnover will be lower because the capitalised interest will appear on the balance sheet as part of the asset being constructed. Therefore, fixed assets will be higher and the fixed asset turnover ratio (total revenue/average net fixed assets) will be lower than if it had expensed these costs. Capitalised interest appears on the balance sheet as part of the asset being constructed instead of being reported as interest expense in the period incurred. However, the interest coverage ratio should be based on interest payments, not interest expense (earnings before interest and taxes/interest payments), and should be unchanged. To provide a true picture of a company's interest coverage, the entire amount of interest expenditure, both the capitalised portion and the expensed portion, should be used in calculating interest coverage ratios.

45 A is correct. Accelerated depreciation will result in an improving, not declining, net profit margin over time, because the amount of depreciation expense declines each year. Under straight-line depreciation, the amount of depreciation expense will remain the same each year. Under the units-of-production method, the amount of depreciation expense reported each year varies with the number of units produced.

46 B is correct. The estimated average total useful life of a company's assets is calculated by adding the estimates of the average remaining useful life and the average age of the assets. The average age of the assets is estimated by dividing accumulated depreciation by depreciation expense. The average remaining useful life of the asset base is estimated by dividing net property, plant, and equipment by annual depreciation expense.

47 C is correct. The impairment loss is a non-cash charge and will not affect cash flow from operating activities. The debt to total assets and fixed asset turnover ratios will increase, because the impairment loss will reduce the carrying amount of fixed assets and therefore total assets.

48 A is correct. In an asset revaluation, the carrying amount of the assets increases. The increase in the asset's carrying amount bypasses the income statement and is reported as other comprehensive income and appears in equity under the heading of revaluation surplus. Therefore, shareholders' equity will increase but net income will not be affected, so return on equity will decline. Return on assets and debt to capital ratios will also decrease.

READING

23

Income Taxes

by Elbie Louw, PhD, CFA, CIPM, and Michael A. Broihahn, CPA, CIA, CFA

Elbie Louw, PhD, CFA, CIPM (South Africa). Michael A. Broihahn, CPA, CIA, CFA, is at Barry University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the differences between accounting profit and taxable income and define key terms, including deferred tax assets, deferred tax liabilities, valuation allowance, taxes payable, and income tax expense;
☐	**b.** explain how deferred tax liabilities and assets are created and the factors that determine how a company's deferred tax liabilities and assets should be treated for the purposes of financial analysis;
☐	**c.** calculate the tax base of a company's assets and liabilities;
☐	**d.** calculate income tax expense, income taxes payable, deferred tax assets, and deferred tax liabilities, and calculate and interpret the adjustment to the financial statements related to a change in the income tax rate;
☐	**e.** evaluate the effect of tax rate changes on a company's financial statements and ratios;
☐	**f.** identify and contrast temporary versus permanent differences in pre-tax accounting income and taxable income;
☐	**g.** describe the valuation allowance for deferred tax assets—when it is required and what effect it has on financial statements;
☐	**h.** explain recognition and measurement of current and deferred tax items;
☐	**i.** analyze disclosures relating to deferred tax items and the effective tax rate reconciliation and explain how information included in these disclosures affects a company's financial statements and financial ratios;
☐	**j.** identify the key provisions of and differences between income tax accounting under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (GAAP).

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION

For those companies reporting under International Financial Reporting Standards (IFRS), IAS 12 [Income Taxes] covers accounting for a company's income taxes and the reporting of deferred taxes. For those companies reporting under United States generally accepted accounting principles (US GAAP), FASB ASC Topic 740 [Income Taxes] is the primary source for information on accounting for income taxes. Although IFRS and US GAAP follow similar conventions on many income tax issues, there are some key differences that will be discussed in the reading.

Differences between how and when transactions are recognized for financial reporting purposes relative to tax reporting can give rise to differences in tax expense and related tax assets and liabilities. To reconcile these differences, companies that report under either IFRS or US GAAP create a provision on the balance sheet called deferred tax assets or deferred tax liabilities, depending on the nature of the situation.

Deferred tax assets or liabilities usually arise when accounting standards and tax authorities recognize the timing of revenues and expenses at different times. Because timing differences such as these will eventually reverse over time, they are called "temporary differences." Deferred tax assets represent taxes that have been recognized for tax reporting purposes (or often the carrying forward of losses from previous periods) but have not yet been recognized on the income statement prepared for financial reporting purposes. Deferred tax liabilities represent tax expense that has appeared on the income statement for financial reporting purposes, but has not yet become payable under tax regulations.

This reading provides a primer on the basics of income tax accounting and reporting. The reading is organized as follows. Sections 2 and 3 describe the differences between taxable income and accounting profit. Sections 4 and 5 explain the determination of tax base, which relates to the valuation of assets and liabilities for tax purposes. Sections 6 and 7 discuss several types of timing differences between the recognition of taxable and accounting profit. Section 8 examines unused tax losses and tax credits. Section 9 describes the recognition and measurement of current and deferred tax. Section 10 discusses the disclosure and presentation of income tax information on companies' financial statements and illustrates its practical implications for financial analysis. Section 11 provides an overview of the similarities and differences for income-tax reporting between IFRS and US GAAP. A summary of the key points and practice problems in the CFA Institute multiple-choice format conclude the reading.

2 DIFFERENCES BETWEEN ACCOUNTING PROFIT AND TAXABLE INCOME

a describe the differences between accounting profit and taxable income and define key terms, including deferred tax assets, deferred tax liabilities, valuation allowance, taxes payable, and income tax expense;

A company's **accounting profit** is reported on its income statement in accordance with prevailing accounting standards. Accounting profit (also referred to as income before taxes or pretax income) does not include a provision for income tax expense.[1] A company's **taxable income** is the portion of its income that is subject to income taxes under the tax laws of its jurisdiction. Because of different guidelines for how income is reported on a company's financial statements and how it is measured for income tax purposes, accounting profit and taxable income may differ.

---

[1] As defined under IAS 12, paragraph 5.

A company's taxable income is the basis for its **income tax payable** (a liability) or recoverable (an asset), which is calculated on the basis of the company's tax rate and appears on its balance sheet. A company's **tax expense**, or tax benefit in the case of a recovery, appears on its income statement and is an aggregate of its income tax payable (or recoverable in the case of a tax benefit) and any changes in deferred tax assets and liabilities.

When a company's taxable income is greater than its accounting profit, then its income taxes payable will be higher than what would have otherwise been the case had the income taxes been determined based on accounting profit. **Deferred tax assets**, which appear on the balance sheet, arise when an excess amount is paid for income taxes (taxable income higher than accounting profit) and the company expects to recover the difference during the course of future operations. Actual income taxes payable will thus exceed the financial accounting income tax expense (which is reported on the income statement and is determined based on accounting profit). Related to deferred tax assets is a **valuation allowance**, which is a reserve created against deferred tax assets. The valuation allowance is based on the likelihood of realizing the deferred tax assets in future accounting periods. **Deferred tax liabilities**, which also appear on the balance sheet, arise when a deficit amount is paid for income taxes and the company expects to eliminate the deficit over the course of future operations. In this case, financial accounting income tax expense exceeds income taxes payable.

Income tax paid in a period is the actual amount paid for income taxes (not a provision, but the actual cash outflow). The income tax paid may be less than the income tax expense because of payments in prior periods or refunds received in the current period. Income tax paid reduces the income tax payable, which is carried on the balance sheet as a liability.

The **tax base** of an asset or liability is the amount at which the asset or liability is valued for tax purposes, whereas the **carrying amount** is the amount at which the asset or liability is valued according to accounting principles.[2] Differences between the tax base and the carrying amount also result in differences between accounting profit and taxable income. These differences can carry through to future periods. For example, a **tax loss carry forward** occurs when a company experiences a loss in the current period that may be used to reduce future taxable income. The company's tax expense on its income statement must not only reflect the taxes payable based on taxable income, but also the effect of these differences.

CURRENT AND DEFERRED TAX ASSETS AND LIABILITIES

3

b explain how deferred tax liabilities and assets are created and the factors that determine how a company's deferred tax liabilities and assets should be treated for the purposes of financial analysis;

d calculate income tax expense, income taxes payable, deferred tax assets, and deferred tax liabilities, and calculate and interpret the adjustment to the financial statements related to a change in the income tax rate;

---

[2] The terms "tax base" and "tax basis" are interchangeable. "Tax basis" is more commonly used in the United States. Similarly, "carrying amount" and "book value" refer to the same concept.

A company's current tax liability is the amount payable in taxes and is based on current taxable income. If the company expects to receive a refund for some portion previously paid in taxes, the amount recoverable is referred to as a current tax asset. The current tax liability or asset may, however, differ from what the liability would have been if it was based on accounting profit rather than taxable income for the period. Differences in accounting profit and taxable income are the result of the application of different rules. Such differences between accounting profit and taxable income can occur in several ways, including:

- Revenues and expenses may be recognized in one period for accounting purposes and a different period for tax purposes;
- Specific revenues and expenses may be either recognized for accounting purposes and not for tax purposes; or not recognized for accounting purposes but recognized for tax purposes;
- The carrying amount and tax base of assets and/or liabilities may differ;
- The deductibility of gains and losses of assets and liabilities may vary for accounting and income tax purposes;
- Subject to tax rules, tax losses of prior years might be used to reduce taxable income in later years, resulting in differences in accounting and taxable income (tax loss carryforward); and
- Adjustments of reported financial data from prior years might not be recognized equally for accounting and tax purposes or might be recognized in different periods.

3.1 Deferred Tax Assets and Liabilities

Deferred tax assets represent taxes that have been paid (or often the carrying forward of losses from previous periods) but have not yet been recognized on the income statement. Deferred tax liabilities occur when financial accounting income tax expense is greater than regulatory income tax expense. Deferred tax assets and liabilities usually arise when accounting standards and tax authorities recognize the timing of taxes due at different times; for example, when a company uses accelerated depreciation when reporting to the tax authority (to increase expense and lower tax payments in the early years) but uses the straight-line method on the financial statements. Although not similar in treatment on a year-to-year basis (e.g., depreciation of 5 percent on a straight-line basis may be permitted for accounting purposes whereas 10 percent is allowed for tax purposes) over the life of the asset, both approaches allow for the total cost of the asset to be depreciated (or amortized). Because these timing differences will eventually reverse or self-correct over the course of the asset's depreciable life, they are called "temporary differences."

Any deferred tax asset or liability is based on temporary differences that result in an excess or a deficit amount paid for taxes, which the company expects to recover from future operations. Because taxes will be recoverable or payable at a future date, it is only a temporary difference and a deferred tax asset or liability is created. Changes in the deferred tax asset or liability on the balance sheet reflect the difference between the amounts recognized in the previous period and the current period. The changes in deferred tax assets and liabilities are added to income tax payable to determine the company's income tax expense (or credit) as it is reported on the income statement.

At the end of each fiscal year, deferred tax assets and liabilities are recalculated by comparing the tax bases and carrying amounts of the balance sheet items. Identified temporary differences should be assessed on whether the difference will result in future economic benefits. For example, Pinto Construction (a hypothetical company) depreciates equipment on a straight-line basis of 10 percent per year. The tax authorities allow depreciation of 15 percent per year. At the end of the fiscal year, the carrying amount of the equipment for accounting purposes would be greater than the tax base of the equipment thus resulting in a temporary difference. A deferred tax item may only be created if it is not doubtful that the company will realize economic benefits in the future. In our example, the equipment is used in the core business of Pinto Construction. If the company is a going concern and stable, there should be no doubt that future economic benefits will result from the equipment and it would be appropriate to create the deferred tax item.

Should it be doubtful that future economic benefits will be realized from a temporary difference (such as Pinto Construction being under liquidation), the temporary difference will not lead to the creation of a deferred tax asset or liability. If a deferred tax asset or liability resulted in the past, but the criteria of economic benefits is not met on the current balance sheet date, then, under IFRS, an existing deferred tax asset or liability related to the item will be reversed. Under US GAAP, a valuation allowance is established. In assessing future economic benefits, much is left to the discretion of the auditor in assessing the temporary differences and the issue of future economic benefits.

EXAMPLE 1

The following information pertains to a hypothetical company, Reston Partners:

Reston Partners Consolidated Income Statement

Period Ending 31 March (£ Millions)	Year 3	Year 2	Year 1
Revenue	£40,000	£30,000	£25,000
Other net gains	2,000	0	0
Changes in inventories of finished goods and work in progress	400	180	200
Raw materials and consumables used	(5,700)	(4,000)	(8,000)
Depreciation expense	(2,000)	(2,000)	(2,000)
Other expenses	(6,000)	(5,900)	(4,500)
Interest expense	(2,000)	(3,000)	(6,000)
Profit before tax	£26,700	£15,280	£4,700

The financial performance and accounting profit of Reston Partners on this income statement is based on accounting principles appropriate for the jurisdiction in which Reston Partners operates. The principles used to calculate accounting profit (profit before tax in the example above) may differ from the principles applied for tax purposes (the calculation of taxable income). For illustrative purposes, however, assume that all income and expenses on the income statement are treated identically for tax and accounting purposes *except* depreciation.

The depreciation is related to equipment owned by Reston Partners. For simplicity, assume that the equipment was purchased at the beginning of the Year 1. Depreciation should thus be calculated and expensed for the full year. Assume that accounting standards permit equipment to be depreciated on a straight-line basis over a 10-year period, whereas the tax standards in the jurisdiction specify that equipment should be depreciated on a straight-line basis over a 7-year period. For simplicity, assume a salvage value of £0 at the end of the equipment's useful life. Both methods will result in the full depreciation of the asset over the respective tax or accounting life.

The equipment was originally purchased for £20,000. In accordance with accounting standards, over the next 10 years the company will recognize annual depreciation of £2,000 (£20,000 ÷ 10) as an expense on its income statement and for the determination of accounting profit. For tax purposes, however, the company will recognize £2,857 (£20,000 ÷ 7) in depreciation each year. Each fiscal year the depreciation expense related to the use of the equipment will, therefore, differ for tax and accounting purposes (tax base vs. carrying amount), resulting in a difference between accounting profit and taxable income.

The previous income statement reflects accounting profit (depreciation at £2,000 per year). The following table shows the taxable income for each fiscal year.

Taxable Income (£ Millions)	**Year 3**	**Year 2**	**Year 1**
Revenue	£40,000	£30,000	£25,000
Other net gains	2,000	0	0
Changes in inventories of finished goods and work in progress	400	180	200
Raw materials and consumables used	(5,700)	(4,000)	(8,000)
Depreciation expense	(2,857)	(2,857)	(2,857)
Other expenses	(6,000)	(5,900)	(4,500)
Interest expense	(2,000)	(3,000)	(6,000)
Taxable income	£25,843	£14,423	£3,843

The carrying amount and tax base for the equipment is as follows:

(£ Millions)	**Year 3**	**Year 2**	**Year 1**
Equipment value for accounting purposes (*carrying amount*) (depreciation of £2,000/year)	£14,000	£16,000	£18,000
Equipment value for tax purposes (*tax base*) (depreciation of £2,857/year)	£11,429	£14,286	£17,143
Difference	£2,571	£1,714	£857

At each balance sheet date, the tax base and carrying amount of all assets and liabilities must be determined. The income tax payable by Reston Partners will be based on the taxable income of each fiscal year. If a tax rate of 30 percent is assumed, then the income taxes payable for Year 1, Year 2, and Year 3 are £1,153 (30% × 3,843), £4,327 (30% × 14,423), and £7,753 (30% × 25,843).

Remember, though, that if the tax obligation is calculated based on accounting profits, it will differ because of the differences between the tax base and the carrying amount of equipment. The difference in each fiscal year is reflected in the table above. In each fiscal year the carrying amount of the equipment exceeds its tax base. For tax purposes, therefore, the asset tax base is less than its carrying value under financial accounting principles. The difference results in a deferred tax liability.

(£ Millions)	**Year 3**	**Year 2**	**Year 1**
Deferred tax liability	£771	£514	£257

(Difference between tax base and carrying amount) × tax rate

Year 1: £(18,000 – 17,143) × 30% = 257

Year 2: £(16,000 – 14,286) × 30% = 514

Year 3: £(14,000 – 11,429) × 30% = 771

The comparison of the tax base and carrying amount of equipment shows what the deferred tax liability should be on a particular balance sheet date. In each fiscal year, only the change in the deferred tax liability should be included in the calculation of the income tax expense reported on the income statement prepared for accounting purposes.

On the income statement, the company's income tax expense will be the sum of change in the deferred tax liability and the income tax payable.

(£ Millions)	Year 3	Year 2	Year 1
Income tax payable (based on tax accounting)	£7,753	£4,327	£1,153
Change in deferred tax liability	257	257	257
Income tax (based on financial accounting)	£8,010	£4,584	£1,410

Note that because the different treatment of depreciation is a temporary difference, the income tax on the income statement is 30 percent of the accounting profit, although only a part is income tax payable and the rest is a deferred tax liability.

The consolidated income statement of Reston Partners including income tax is presented as follows:

Reston Partners Consolidated Income Statement

Period Ending 31 March (£ Millions)	Year 3	Year 2	Year 1
Revenue	£40,000	£30,000	£25,000
Other net gains	2,000	0	0
Changes in inventories of finished goods and work in progress	400	180	200
Raw materials and consumables used	(5,700)	(4,000)	(8,000)
Depreciation expense	(2,000)	(2,000)	(2,000)
Other expenses	(6,000)	(5,900)	(4,500)
Interest expense	(2,000)	(3,000)	(6,000)
Profit before tax	£26,700	£15,280	£4,700
Income tax	(8,010)	(4,584)	(1,410)
Profit after tax	£18,690	£10,696	£3,290

Any amount paid to the tax authorities will reduce the liability for income tax payable and be reflected on the statement of cash flows of the company.

DETERMINING THE TAX BASE OF ASSETS AND LIABILITIES

4

c calculate the tax base of a company's assets and liabilities;

As mentioned in Sections 2 and 3, temporary differences arise from a difference in the tax base and carrying amount of assets and liabilities. The tax base of an asset or liability is the amount attributed to the asset or liability for tax purposes, whereas the carrying amount is based on accounting principles. Such a difference is considered temporary if it is expected that the taxes will be recovered or payable at a future date.

4.1 Determining the Tax Base of an Asset

The tax base of an asset is the amount that will be deductible for tax purposes in future periods as the economic benefits become realized and the company recovers the carrying amount of the asset.

For example, our previously mentioned Reston Partners (from Example 1) depreciates equipment on a straight-line basis at a rate of 10 percent per year. The tax authorities allow depreciation of approximately 15 percent per year. At the end of the fiscal year, the carrying amount of equipment for accounting purposes is greater than the asset tax base thus resulting in a temporary difference.

EXAMPLE 2

Determining the Tax Base of an Asset

The following information pertains to Entiguan Sports, a hypothetical developer of products used to treat sports-related injuries. (The treatment of items for accounting and tax purposes is based on hypothetical accounting and tax standards and is not specific to a particular jurisdiction.) Calculate the tax base and carrying amount for each item.

1 *Dividends receivable*: On its balance sheet, Entiguan Sports reports dividends of €1 million receivable from a subsidiary. Assume that dividends are not taxable.

2 *Development costs*: Entiguan Sports capitalized development costs of €3 million during the year. Entiguan amortized €500,000 of this amount during the year. For tax purposes amortization of 25 percent per year is allowed.

3 *Research costs*: Entiguan incurred €500,000 in research costs, which were all expensed in the current fiscal year for financial reporting purposes. Assume that applicable tax legislation requires research costs to be expensed over a four-year period rather than all in one year.

4 *Accounts receivable*: Included on the income statement of Entiguan Sports is a provision for doubtful debt of €125,000. The accounts receivable amount reflected on the balance sheet, after taking the provision into account, amounts to €1,500,000. The tax authorities allow a deduction of 25 percent of the gross amount for doubtful debt.

Solutions:

	Carrying Amount (€)	Tax Base (€)	Temporary Difference (€)
1. Dividends receivable	1,000,000	1,000,000	0
2. Development costs	2,500,000	2,250,000	250,000
3. Research costs	0	375,000	(375,000)
4. Accounts receivable	1,500,000	1,218,750	281,250

Comments:

1 *Dividends receivable*: Although the dividends received are economic benefits from the subsidiary, we are assuming that dividends are not taxable. Therefore, the carrying amount equals the tax base for dividends receivable.

2 *Development costs*: First, we assume that development costs will generate economic benefits for Entiguan Sports. Therefore, it may be included as an asset on the balance sheet for the purposes of this example. Second, the amortization allowed by the tax authorities exceeds the amortization accounted for based on accounting rules. Therefore, the carrying amount of the asset exceeds its tax base. The carrying amount is (€3,000,000 – €500,000) = €2,500,000 whereas the tax base is [€3,000,000 – (25% × €3,000,000)] = €2,250,000.

3 *Research costs*: We assume that research costs will result in future economic benefits for the company. If this were not the case, creation of a deferred tax asset or liability would not be allowed. The tax base of research costs exceeds their carrying amount. The carrying amount is €0 because the full amount has been expensed for financial reporting purposes in the year in which it was incurred. Therefore, there would not have been a balance sheet item "Research costs" for tax purposes, and only a proportion may be deducted in the current fiscal year. The tax base of the asset is (€500,000 – €500,000/4) = €375,000.

4 *Accounts receivable*: The economic benefits that should have been received from accounts receivable have already been included in revenues included in the calculation of the taxable income when the sales occurred. Because the receipt of a portion of the accounts receivable is doubtful, the provision is allowed. The provision, based on tax legislation, results in a greater amount allowed in the current fiscal year than would be the case under accounting principles. This results in the tax base of accounts receivable being lower than its carrying amount. Note that the example specifically states that the balance sheet amount for accounts receivable after the provision for accounting purposes amounts to €1,500,000. Therefore, accounts receivable before any provision was €1,500,000 + €125,000 = €1,625,000. The tax base is calculated as (€1,500,000 + €125,000) – [25% × (€1,500,000 + €125,000)] = €1,218,750.

4.2 Determining the Tax Base of a Liability

The tax base of a liability is the carrying amount of the liability less any amounts that will be deductible for tax purposes in the future. With respect to payments from customers received in advance of providing the goods and services, the tax base of such a liability is the carrying amount less any amount of the revenue that will not be taxable in future. Keep in mind the following fundamental principle: In general, a company will recognize a deferred tax asset or liability when recovery/settlement of the carrying amount will affect future tax payments by either increasing or reducing the taxable profit. Remember, an analyst is not only evaluating the difference between the carrying amount and the tax base, but the relevance of that difference on future profits and losses and thus by implication future taxes.

IFRS offers specific guidelines with regard to revenue received in advance: IAS 12 states that the tax base is the carrying amount less any amount of the revenue that will not be taxed at a future date. Under US GAAP, an analysis of the tax base would result in a similar outcome. The tax legislation within the jurisdiction will determine the amount recognized on the income statement and whether the liability (revenue received in advance) will have a tax base greater than zero. This will depend on how tax legislation recognizes revenue received in advance.

EXAMPLE 3

Determining the Tax Base of a Liability

The following information pertains to the hypothetical company Entiguan Sports for the fiscal year -end. The treatment of items for accounting and tax purposes is based on fictitious accounting and tax standards and is not specific to a particular jurisdiction. Calculate the tax base and carrying amount for each item.

1. *Donations*: Entiguan Sports made donations of €100,000 in the current fiscal year. The donations were expensed for financial reporting purposes, but are not tax deductible based on applicable tax legislation.
2. *Interest received in advance*: Entiguan Sports received in advance interest of €300,000. The interest is taxed because tax authorities recognize the interest to accrue to the company (part of taxable income) on the date of receipt.
3. *Rent received in advance*: Entiguan recognized €10 million for rent received in advance from a lessee for an unused warehouse building. Rent received in advance is deferred for accounting purposes but taxed on a cash basis.
4. *Loan*: Entiguan Sports secured a long-term loan for €550,000 in the current fiscal year. Interest is charged at 13.5 percent per annum and is payable at the end of each fiscal year.

Solutions:

	Carrying Amount (€)	Tax Base (€)	Temporary Difference (€)
1. Donations	0	0	0
2. Interest received in advance	300,000	0	(300,000)
3. Rent received in advance	10,000,000	0	(10,000,000)
4. Loan (capital)	0	0	0
Interest paid	0	0	0

Comments:

1. *Donations*: The amount of €100,000 was immediately expensed on Entiguan's income statement; therefore, the carrying amount is €0. Tax legislation does not allow donations to be deducted for tax purposes, so the tax base of the donations equals the carrying amount. Note that while the carrying amount and tax base are the same, the difference in the treatment of donations for accounting and tax purposes (expensed for accounting purposes, but not deductible for tax purposes) represents a permanent difference (a difference that will not be reversed in future). Permanent and temporary differences are elaborated on in Sections 6 and 7 and it will refer to this particular case with an expanded explanation.
2. *Interest received in advance*: Based on the information provided, for tax purposes, interest is deemed to accrue to the company on the date of receipt. For tax purposes, it is thus irrelevant whether it is for the current or a future accounting period; it must be included in taxable income in the financial year received. Interest received in advance is, for accounting purposes though, included in the financial period in which it is deemed to have been earned. For this reason, the interest income received in advance

is a balance sheet liability. It was not included on the income statement because the income relates to a future financial year. Because the full €300,000 is included in taxable income in the current fiscal year, the tax base is €300,000 − 300,000 = €0. Note that although interest received in advance and rent received in advance are both taxed, the timing depends on how the particular item is treated in tax legislation.

3 *Rent received in advance*: The result is similar to interest received in advance. The carrying amount of rent received in advance would be €10,000,000 while the tax base is €0.

4 *Loan*: Repayment of the loan has no tax implications. The repayment of the capital amount does not constitute an income or expense. The interest paid is included as an expense in the calculation of taxable income as well as accounting income. Therefore, the tax base and carrying amount is €0. For clarity, the interest paid that would be included on the income statement for the year amounts to 13.5% × €550,000 = €74,250 if the loan was acquired at the beginning of the current fiscal year.

CHANGES IN INCOME TAX RATES

5

d calculate income tax expense, income taxes payable, deferred tax assets, and deferred tax liabilities, and calculate and interpret the adjustment to the financial statements related to a change in the income tax rate;

e evaluate the effect of tax rate changes on a company's financial statements and ratios;

The measurement of deferred tax assets and liabilities is based on current tax law. But if there are subsequent changes in tax laws or new income tax rates, existing deferred tax assets and liabilities must be adjusted for the effects of these changes. The resulting effects of the changes are also included in determining accounting profit in the period of change.

When income tax rates change, the deferred tax assets and liabilities are adjusted to the new tax rate. If income tax rates increase, deferred taxes (that is, the deferred tax assets and liabilities) will also increase. Likewise, if income tax rates decrease, deferred taxes will decrease. A decrease in tax rates decreases deferred tax liabilities, which reduces future tax payments to the taxing authorities. A decrease in tax rates will also decrease deferred tax assets, which reduces their value toward the offset of future tax payments to the taxing authorities.

To illustrate the effect of a change in tax rate, consider Example 1 again. In that illustration, the timing difference that led to the recognition of a deferred tax liability for Reston Partners was attributable to differences in the method of depreciation and the related effects on the accounting carrying value and the asset tax base. The relevant information is restated below.

The carrying amount and tax base for the equipment is:

(£ Millions)	Year 3	Year 2	Year 1
Equipment value for accounting purposes (*carrying amount*) (depreciation of £2,000/year)	£14,000	£16,000	£18,000
Equipment value for tax purposes (*tax base*) (depreciation of £2,857/year)	£11,429	£14,286	£17,143
Difference	£2,571	£1,714	£857

At a 30 percent income tax rate, the deferred tax liability was then determined as follows:

(£ Millions)	Year 3	Year 2	Year 1
Deferred tax liability	£771	£514	£257
(Difference between tax base and carrying amount)			
Year 1: £(18,000 – 17,143) × 30% = £257			
Year 2: £(16,000 – 14,286) × 30% = £514			
Year 3: £(14,000 – 11,429) × 30% = £771			

For this illustration, assume that the taxing authority has changed the income tax rate to 25 percent for Year 3. Although the difference between the carrying amount and the tax base of the depreciable asset are the same, the deferred tax liability for 2017 will be £643 (instead of £771 or a reduction of £128 in the liability). 2017: £(14,000 – 11,429) × 25% = £643.

Reston Partners' provision for income tax expense is also affected by the change in tax rates. Taxable income for Year 3 will now be taxed at a rate of 25 percent. The benefit of the Year 3 accelerated depreciation tax shield is now only £214 (£857 × 25%) instead of the previous £257 (a reduction of £43). In addition, the reduction in the beginning carrying value of the deferred tax liability for Year 3 (the year of change) further reduces the income tax expense for Year 3. The reduction in income tax expense attributable to the change in tax rate is £86. Year 3: (30% – 25%) × £1,714 = £86. Note that these two components together account for the reduction in the deferred tax liability (£43 + £86 = £129).

As may be seen from this discussion, changes in the income tax rate have an effect on a company's deferred tax asset and liability carrying values as well as an effect on the measurement of income tax expense in the year of change. The analyst must thus note that proposed changes in tax law can have a quantifiable effect on these accounts (and any related financial ratios that are derived from them) if the proposed changes are subsequently enacted into law.

6

TEMPORARY AND PERMANENT DIFFERENCES BETWEEN TAXABLE AND ACCOUNTING PROFIT

f identify and contrast temporary versus permanent differences in pre-tax accounting income and taxable income;

Temporary differences arise from a difference between the tax base and the carrying amount of assets and liabilities. The creation of a deferred tax asset or liability from a temporary difference is only possible if the difference reverses itself at some future date and to such an extent that the balance sheet item is expected to create future economic benefits for the company. IFRS and US GAAP both prescribe the balance sheet liability method for recognition of deferred tax. This balance sheet method focuses on the recognition of a deferred tax asset or liability should there be a temporary difference between the carrying amount and tax base of balance sheet items.[3]

Permanent differences are differences between tax and financial reporting of revenue (expenses) that *will not* be reversed at some future date. Because they will not be reversed at a future date, these differences do not give rise to deferred tax. These items typically include

- Income or expense items not allowed by tax legislation, and
- Tax credits for some expenditures that directly reduce taxes.

Because no deferred tax item is created for permanent differences, all permanent differences result in a difference between the company's effective tax rate and statutory tax rate. The effective tax rate is also influenced by different statutory taxes should an entity conduct business in more than one tax jurisdiction. The formula for the reported effective tax rate is thus equal to:

$$\text{Reported effective tax rate} = \text{Income tax expense} \div \text{Pretax income (accounting profit)}$$

The net change in deferred tax during a reporting period is the difference between the balance of the deferred tax asset or liability for the current period and the balance of the previous period.

6.1 Taxable Temporary Differences

Temporary differences are further divided into two categories, namely taxable temporary differences and deductible temporary differences. **Taxable temporary differences** are temporary differences that result in a taxable amount in a future period when determining the taxable profit as the balance sheet item is recovered or settled. Taxable temporary differences result in a deferred tax liability when the carrying amount of an asset exceeds its tax base and, in the case of a liability, when the tax base of the liability exceeds its carrying amount.

Under US GAAP, a deferred tax asset or liability is not recognized for unamortizable goodwill. Under IFRS, a deferred tax account is not recognized for goodwill arising in a business combination. Since goodwill is a residual, the recognition of a deferred tax liability would increase the carrying amount of goodwill. Discounting deferred tax assets or liabilities is generally not allowed for temporary differences related to business combinations as it is for other temporary differences.

IFRS provides an exemption (that is, deferred tax is not provided on the temporary difference) for the initial recognition of an asset or liability in a transaction that: a) is not a business combination (e.g., joint ventures, branches and unconsolidated investments); and b) affects neither accounting profit nor taxable profit at the time of the transaction. US GAAP does not provide an exemption for these circumstances.

---

[3] Previously, IAS 12 required recognition of deferred tax based on the deferred method (also known as the income statement method), which focused on timing differences. Timing differences are differences in the recognition of income and expenses for accounting and tax purposes that originate in one period and will reverse in a future period. Given the definition of timing differences, all timing differences are temporary differences, such as the different treatment of depreciation for tax and accounting purposes (although the timing is different with regard to the allowed depreciation for tax and accounting purposes, the asset will eventually be fully depreciated).

As a simple example of a temporary difference with no recognition of deferred tax liability, assume that a holding company of various leisure related businesses and holiday resorts buys an interest in a hotel in the current financial year. The goodwill related to the transaction will be recognized on the financial statements, but the related tax liability will not, as it relates to the initial recognition of goodwill.

6.2 Deductible Temporary Differences

Deductible temporary differences are temporary differences that result in a reduction or deduction of taxable income in a future period when the balance sheet item is recovered or settled. Deductible temporary differences result in a deferred tax asset when the tax base of an asset exceeds its carrying amount and, in the case of a liability, when the carrying amount of the liability exceeds its tax base. The recognition of a deferred tax asset is only allowed to the extent there is a reasonable expectation of future profits against which the asset or liability (that gave rise to the deferred tax asset) can be recovered or settled.

To determine the probability of sufficient future profits for utilization, one must consider the following: 1) Sufficient taxable temporary differences must exist that are related to the same tax authority and the same taxable entity; and 2) The taxable temporary differences that are expected to reverse in the same periods as expected for the reversal of the deductible temporary differences.

As with deferred tax liabilities, IFRS states that deferred tax assets should not be recognized in cases that would arise from the initial recognition of an asset or liability in transactions that are not a business combination and when, at the time of the transaction, there is no impact on either accounting or taxable profit. Subsequent to initial recognition under IFRS and US GAAP, any deferred tax assets that arise from investments in subsidiaries, branches, associates, and interests in joint ventures are recognized as a deferred tax asset.

IFRS and US GAAP allow the creation of a deferred tax asset in the case of tax losses and tax credits. These two unique situations will be further elaborated on in Section 9. IAS 12 *does not* allow the creation of a deferred tax asset arising from negative goodwill. Negative goodwill arises when the amount that an entity pays for an interest in a business is less than the net fair market value of the portion of assets and liabilities of the acquired company, based on the interest of the entity.

6.3 Examples of Taxable and Deductible Temporary Differences

Exhibit 1 summarizes how differences between the tax bases and carrying amounts of assets and liabilities give rise to deferred tax assets or deferred tax liabilities.

Exhibit 1 Treatment of Temporary Differences

Balance Sheet Item	Carrying Amount vs. Tax Base	Results in Deferred Tax Asset/Liability
Asset	Carrying amount > tax base	Deferred tax liability
Asset	Carrying amount < tax base	Deferred tax asset
Liability	Carrying amount > tax base	Deferred tax asset
Liability	Carrying amount < tax base	Deferred tax liability

EXAMPLE 4

Taxable and Deductible Temporary Differences

Examples 2 and 3 illustrated how to calculate the tax base of assets and liabilities, respectively. Based on the information provided in Examples 2 and 3, indicate whether the difference in the tax base and carrying amount of the assets and liabilities are temporary or permanent differences and whether a deferred tax asset or liability will be recognized based on the difference identified.

Solution to Example 2:

	Carrying Amount (€)	Tax Base (€)	Temporary Difference (€)	*Will Result in Deferred Tax Asset/Liability*
1. Dividends receivable	1,000,000	1,000,000	0	*N/A*
2. Development costs	2,500,000	2,250,000	250,000	*Deferred tax liability*
3. Research costs	0	375,000	(375,000)	*Deferred tax asset*
4. Accounts receivable	1,500,000	1,218,750	281,250	*Deferred tax liability*

Example 2 included comments on the calculation of the carrying amount and tax base of the assets.

1 *Dividends receivable*: As a result of non-taxability, the carrying amount equals the tax base of dividends receivable. This constitutes a permanent difference and will not result in the recognition of any deferred tax asset or liability. A temporary difference constitutes a difference that will, at some future date, be reversed. Although the timing of recognition is different for tax and accounting purposes, in the end the full carrying amount will be expensed/recognized as income. A permanent difference will never be reversed. Based on tax legislation, dividends from a subsidiary are not recognized as income. Therefore, no amount will be reflected as dividend income when calculating the taxable income, and the tax base of dividends receivable must be the total amount received, namely €1,000,000. The taxable income and accounting profit will permanently differ with the amount of dividends receivable, even on future financial statements as an effect on the retained earnings reflected on the balance sheet.

2 *Development costs*: The difference between the carrying amount and tax base is a temporary difference that, in the future, will reverse. In this fiscal year, it will result in a deferred tax liability.

3 *Research costs*: The difference between the carrying amount and tax base is a temporary difference that results in a deferred tax asset. Remember the explanation in Section 3 for deferred tax assets—a deferred tax asset arises because of an excess amount paid for taxes (when taxable income is greater than accounting profit), which is expected to be recovered from future operations. Based on accounting principles, the full amount was deducted resulting in a lower accounting profit, while the taxable income by implication, should be greater because of the lower amount expensed.

4 *Accounts receivable*: The difference between the carrying amount and tax base of the asset is a temporary difference that will result in a deferred tax liability.

Solution to Example 3:

	Carrying Amount (€)	Tax Base (€)	Temporary Difference (€)	*Will Result in Deferred Tax Asset/Liability*
1. Donations	0	0	0	*N/A*
2. Interest received in advance	300,000	0	(300,000)	*Deferred tax asset*
3. Rent received in advance	10,000,000	0	(10,000,000)	*Deferred tax asset*
4. Loan (capital)	550,000	550,000	0	*N/A*
Interest paid	0	0	0	*N/A*

Example 3 included extensive comments on the calculation of the carrying amount and tax base of the liabilities.

1. *Donations*: It was assumed that tax legislation does not allow donations to be deducted for tax purposes. No temporary difference results from donations, and thus a deferred tax asset or liability will not be recognized. This constitutes a permanent difference.
2. *Interest received in advance*: Interest received in advance results in a temporary difference that gives rise to a deferred tax asset. A deferred tax asset arises because of an excess amount paid for taxes (when taxable income is greater than accounting profit), which is expected to be recovered from future operations.
3. *Rent received in advance*: The difference between the carrying amount and tax base is a temporary difference that leads to the recognition of a deferred tax asset.
4. *Loan*: There are no temporary differences as a result of the loan or interest paid, and thus no deferred tax item is recognized.

7

EXCEPTIONS TO THE USUAL RULES FOR TEMPORARY DIFFERENCES

f identify and contrast temporary versus permanent differences in pre-tax accounting income and taxable income;

In some situations the carrying amount and tax base of a balance sheet item may vary at initial recognition. For example, a company may deduct a government grant from the initial carrying amount of an asset or liability that appears on the balance sheet. For tax purposes, such grants may not be deducted when determining the tax base of the balance sheet item. In such circumstances, the carrying amount of the asset or liability will be lower than its tax base. Differences in the tax base of an asset or liability as a result of the circumstances described above may not be recognized as deferred tax assets or liabilities.

For example, a government may offer grants to Small, Medium, and Micro Enterprises (SMME) in an attempt to assist these entrepreneurs in their endeavors that contribute to the country's GDP and job creation. Assume that a particular grant is offered for infrastructure needs (office furniture, property, plant, and equipment, etc.). In these circumstances, although the carrying amount will be lower than the tax base of the asset, the related deferred tax may not be recognized. As mentioned earlier, deferred tax assets and liabilities should not be recognized in cases that would arise from the initial recognition of an asset or liability in transactions that are not a business combination and when, at the time of the transaction, there is no impact on either accounting or taxable profit.

A deferred tax liability will also not be recognized at the initial recognition of goodwill. Although goodwill may be treated differently across tax jurisdictions, which may lead to differences in the carrying amount and tax base of goodwill, IAS 12 does not allow the recognition of such a deferred tax liability. Any impairment that an entity should, for accounting purposes, impose on goodwill will again result in a temporary difference between its carrying amount and tax base. Any impairment that an entity should, for accounting purposes, impose on goodwill and if part of the goodwill is related to the initial recognition, that part of the difference in tax base and carrying amount should not result in any deferred taxation because the initial deferred tax liability was not recognized. Any future differences between the carrying amount and tax base as a result of amortization and the deductibility of a portion of goodwill constitute a temporary difference for which provision should be made.

7.1 Business Combinations and Deferred Taxes

The fair value of assets and liabilities acquired in a business combination is determined on the acquisition date and may differ from the previous carrying amount. It is highly probable that the values of acquired intangible assets, including goodwill, would differ from their carrying amounts. This temporary difference will affect deferred taxes as well as the amount of goodwill recognized as a result of the acquisition.

7.2 Investments in Subsidiaries, Branches, Associates and Interests in Joint Ventures

Investments in subsidiaries, branches, associates and interests in joint ventures may lead to temporary differences on the consolidated versus the parent's financial statements. The related deferred tax liabilities as a result of temporary differences will be recognized unless both of the following criterion are satisfied:

- The parent is in a position to control the timing of the future reversal of the temporary difference, and
- It is probable that the temporary difference will not reverse in the future.

With respect to deferred tax assets related to subsidiaries, branches, and associates and interests, deferred tax assets will only be recognized if the following criteria are satisfied:

- The temporary difference will reverse in the future, and
- Sufficient taxable profits exist against which the temporary difference can be used.

UNUSED TAX LOSSES AND TAX CREDITS

8

h explain recognition and measurement of current and deferred tax items;

IAS 12 allows the recognition of unused tax losses and tax credits only to the extent that it is probable that in the future there will be taxable income against which the unused tax losses and credits can be applied. Under US GAAP, a deferred tax asset is recognized in full but is then reduced by a valuation allowance if it is more likely than not that some or all of the deferred tax asset will not be realized. The same requirements for creation of a deferred tax asset as a result of deductible temporary differences also apply to unused tax losses and tax credits. The existence of tax losses may indicate that the entity cannot reasonably be expected to generate sufficient future taxable income. All other things held constant, the greater the history of tax losses, the greater the concern regarding the company's ability to generate future taxable profits.

Should there be concerns about the company's future profitability, then the deferred tax asset may not be recognized until it is realized. When assessing the probability that sufficient taxable profit will be generated in the future, the following criteria can serve as a guide:

- If there is uncertainty as to the probability of future taxable profits, a deferred tax asset as a result of unused tax losses or tax credits is only recognized to the extent of the available taxable temporary differences;
- Assess the probability that the entity will in fact generate future taxable profits before the unused tax losses and/or credits expire pursuant to tax rules regarding the carry forward of the unused tax losses;
- Verify that the above is with the same tax authority and based on the same taxable entity;
- Determine whether the past tax losses were a result of specific circumstances that are unlikely to be repeated; and
- Discover if tax planning opportunities are available to the entity that will result in future profits. These may include changes in tax legislation that is phased in over more than one financial period to the benefit of the entity.

It is imperative that the timing of taxable and deductible temporary differences also be considered before creating a deferred tax asset based on unused tax credits.

9

RECOGNITION AND MEASUREMENT OF CURRENT AND DEFERRED TAX

g describe the valuation allowance for deferred tax assets—when it is required and what effect it has on financial statements;

h explain recognition and measurement of current and deferred tax items;

Current taxes payable or recoverable from tax authorities are based on the applicable tax rates at the balance sheet date. Deferred taxes should be measured at the tax rate that is expected to apply when the asset is realized or the liability settled. With respect to the income tax for a current or prior period not yet paid, it is recognized as a tax liability until paid. Any amount paid in excess of any tax obligation is recognized as an asset. The income tax paid in excess or owed to tax authorities is separate from deferred taxes on the company's balance sheet.

When measuring deferred taxes in a jurisdiction, there are different forms of taxation such as income tax, capital gains tax (any capital gains made), or secondary tax on companies (tax payable on the dividends that a company declares) and possibly different tax bases for a balance sheet item (as in the case of government grants influencing the tax base of an asset such as property). In assessing which tax laws should apply, it is dependent on how the related asset or liability will be settled. It would be prudent to use the tax rate and tax base that is consistent with how it is expected the tax base will be recovered or settled.

Although deferred tax assets and liabilities are related to temporary differences expected to be recovered or settled at some future date, neither are discounted to present value in determining the amounts to be booked. Both must be adjusted for changes in tax rates.

Deferred taxes as well as income taxes should always be recognized on the income statement of an entity unless it pertains to:

- Taxes or deferred taxes charged directly to equity, or
- A possible provision for deferred taxes relates to a business combination.

The carrying amount of the deferred tax assets and liabilities should also be assessed. The carrying amounts may change even though there may have been no change in temporary differences during the period evaluated. This can result from:

- Changes in tax rates;
- Reassessments of the recoverability of deferred tax assets; or
- Changes in the expectations for how an asset will be recovered and what influences the deferred tax asset or liability.

All unrecognized deferred tax assets and liabilities must be reassessed at the balance sheet date and measured against the criteria of probable future economic benefits. If such a deferred asset is likely to be recovered, it may be appropriate to recognize the related deferred tax asset.

Different jurisdictions have different requirements for determining tax obligations that can range from different forms of taxation to different tax rates based on taxable income. When comparing financial statements of entities that conduct business in different jurisdictions subject to different tax legislation, the analyst should be cautious in reaching conclusions because of the potentially complex tax rules that may apply.

9.1 Recognition of a Valuation Allowance

Deferred tax assets must be assessed at each balance sheet date. If there is any doubt whether the deferral will be recovered, then the carrying amount should be reduced to the expected recoverable amount. Should circumstances subsequently change and suggest the future will lead to recovery of the deferral, the reduction may be reversed.

Under US GAAP, deferred tax assets are reduced by creating a valuation allowance. Establishing a valuation allowance reduces the deferred tax asset and income in the period in which the allowance is established. Should circumstances change to such an extent that a deferred tax asset valuation allowance may be reduced, the reversal will increase the deferred tax asset and operating income. Because of the subjective judgment involved, an analyst should carefully scrutinize any such changes.

9.2 Recognition of Current and Deferred Tax Charged Directly to Equity

In general, IFRS and US GAAP require that the recognition of deferred tax liabilities and current income tax should be treated similarly to the asset or liability that gave rise to the deferred tax liability or income tax based on accounting treatment. Should an item that gives rise to a deferred tax liability be taken directly to equity, the same should hold true for the resulting deferred tax.

The following are examples of such items:

- Revaluation of property, plant, and equipment (revaluations are not permissible under US GAAP);
- Long-term investments at fair value;

- Changes in accounting policies;
- Errors corrected against the opening balance of retained earnings;
- Initial recognition of an equity component related to complex financial instruments; and
- Exchange rate differences arising from the currency translation procedures for foreign operations.

Whenever it is determined that a deferred tax liability will not be reversed, an adjustment should be made to the liability. The deferred tax liability will be reduced and the amount by which it is reduced should be taken directly to equity. Any deferred taxes related to a business combination must also be recognized in equity.

Depending on the items that gave rise to the deferred tax liabilities, an analyst should exercise judgment regarding whether the taxes should be included with deferred tax liabilities or whether it should be taken directly to equity. It may be more appropriate simply to ignore deferred taxes.

EXAMPLE 5

Taxes Charged Directly to Equity

The following information pertains to Khaleej Company (a hypothetical company). A building owned by Khaleej Company was originally purchased for €1,000,000 on 1 January 20x1. For accounting purposes, buildings are depreciated at 5 percent a year on a straight-line basis, and depreciation for tax purposes is 10 percent a year on a straight-line basis. On the first day of 20x3, the building is revalued at €1,200,000. It is estimated that the remaining useful life of the building from the date of revaluation is 20 years. *Important*: For tax purposes the revaluation of the building is not recognized.

Based on the information provided, the following illustrates the difference in treatment of the building for accounting and tax purposes.

	Carrying Amount of Building	Tax Base of Building
Balance on 1 January 20x1	€1,000,000	€1,000,000
Depreciation 20x1	(50,000)	(100,000)
Balance on 31 December 20x1	€950,000	€900,000
Depreciation 20x2	(50,000)	(100,000)
Balance on 31 December 20x2	€900,000	€800,000
Revaluation on 1 January 20x3	300,000	n/a
Balance on 1 January 20x3	€1,200,000	€800,000
Depreciation 20x3	(60,000)	(100,000)
Balance on 31 December 20x3	€1,140,000	€700,000
Accumulated depreciation		
Balance on 1 January 20x1	€0	€0
Depreciation 20x1	50,000	100,000
Balance on 31 December 20x1	€50,000	€100,000
Depreciation 20x2	50,000	100,000
Balance on 31 December 20x2	€100,000	€200,000
Revaluation at 1 January 20x3	(100,000)	n/a
Balance on 1 January 20x3	€0	€200,000
Depreciation 20x3	60,000	100,000
Balance on 30 November 20x3	€60,000	€300,000

	Carrying Amount	Tax Base
On 31 December 20x1	€950,000	€900,000
On 31 December 20x2	€900,000	€800,000
On 31 December 20x3	€1,140,000	€700,000

31 December 20x1: On 31 December 20x1, different treatments for depreciation expense result in a temporary difference that gives rise to a deferred tax liability. The difference in the tax base and carrying amount of the building was a result of different depreciation amounts for tax and accounting purposes. Depreciation appears on the income statement. For this reason the deferred tax liability will also be reflected on the income statement. If we assume that the applicable tax rate in 20x1 was 40 percent, then the resulting deferred tax liability will be 40% × (€950,000 – €900,000) = €20,000.

31 December 20x2: As of 31 December 20x2, the carrying amount of the building remains greater than the tax base. The temporary difference again gives rise to a deferred tax liability. Again, assuming the applicable tax rate to be 40 percent, the deferred tax liability from the building is 40% × (€900,000 – €800,000) = €40,000.

31 December 20x3: On 31 December 20x3, the carrying amount of the building again exceeds the tax base. This is not the result of disposals or additions, but is a result of the revaluation at the beginning of the 20x3 fiscal year and the different rates of depreciation. The deferred tax liability would seem to be 40% × (€1,140,000 – €700,000) = €176,000, *but* the treatment is different than it was for the 20x1 and 20x2. In 20x3, revaluation of the building gave rise to a balance sheet equity account, namely "Revaluation Surplus" in the amount of €300,000, which is not recognized for tax purposes.

The deferred tax liability would usually have been calculated as follows:

	20x3	20x2	20x1
Deferred tax liability (closing balance at end of fiscal year)	€176,000	€40,000	€20,000
(Difference between tax base and carrying amount)			
20x1: €(950,000 – 900,000) × 40% = 20,000			
20x2: €(900,000 – 800,000) × 40% = 40,000			
20x3: €(1,140,000 – 700,000) × 40% = 176,000			

The change in the deferred tax liability in 20x1 is €20,000, in 20x2: €20,000 (€40,000 – €20,000) and, it would seem, in 20x3: €136,000 (€176,000 – €40,000). In 20x3, although it would seem that the balance for deferred tax liability should be €176,000, the revaluation is not recognized for tax purposes. Only the portion of the difference between the tax base and carrying amount that is not a result of the revaluation is recognized as giving rise to a deferred tax liability.

The effect of the revaluation surplus and the associated tax effects are accounted for in a direct adjustment to equity. The revaluation surplus is reduced by the tax provision associated with the excess of the fair value over the carry value and it affects retained earnings (€300,000 × 40% = €120,000).

The deferred tax liability that should be reflected on the balance sheet is thus not €176,000 but only €56,000 (€176,000 – €120,000). Given the balance of deferred tax liability at the beginning of the 20x3 fiscal year in the amount of €40,000, the change in the deferred tax liability is only €56,000 – €40,000 = €16,000.

In the future, at the end of each year, an amount equal to the depreciation as a result of the revaluation minus the deferred tax effect will be transferred from the revaluation reserve to retained earnings. In 20x3 this will amount to a portion of depreciation resulting from the revaluation, €15,000 (€300,000 ÷ 20), minus the deferred tax effect of €6,000 (€15,000 × 40%), thus €9,000.

10

PRESENTATION AND DISCLOSURE

i. analyze disclosures relating to deferred tax items and the effective tax rate reconciliation and explain how information included in these disclosures affects a company's financial statements and financial ratios;

We will discuss the presentation and disclosure of income tax related information by way of example. The Consolidated Statements of Operations (Income Statements) and Consolidated Balance Sheets for Micron Technology (MU), a global technology company based in the US, are provided in Exhibits 2 and 3, respectively. Exhibit 4 provides the income tax note disclosures for MU for the 2015, 2016, and 2017 fiscal years.

MU's income tax provision (i.e., income tax expense) for fiscal year 2017 is $114 million (see Exhibit 2). The income tax note disclosure in Exhibit 4 reconciles how the income tax provision was determined beginning with MU's reported income before taxes (shown in Exhibit 4 as $5,196 million for fiscal year 2017). The note disclosure then denotes the income tax provision for 2017 that is current ($153 million), which is then offset by the deferred tax benefit for foreign taxes ($39 million), for a net income tax provision of $114 million. Exhibit 4 further shows a reconciliation of how the income tax provision was derived from the US federal statutory rate. Many public companies comply with this required disclosure by displaying the information in percentage terms, but MU has elected to provide the disclosure in absolute dollar amounts. From this knowledge, for 2017 we can see that the dollar amount shown for US federal income tax provision at the statutory rate ($1,819 million) was determined by multiplying MU's income before taxes by the 35 percent US federal statutory rate ($5,196 × 0.35 = $1,819).

In addition, the note disclosure in Exhibit 4 provides detailed information about the derivation of the deferred tax assets ($766 million for 2017) and deferred tax liabilities ($17 million for 2017). These deferred tax assets are shown separately on MU's consolidated balance sheet for fiscal year 2017 with noncurrent assets (see Exhibit 3), while the deferred tax liabilities are included in other noncurrent liabilities (also see Exhibit 3).

Exhibit 2 Micron Technology, Inc. Consolidated Statements of Operations (Amounts in US$ Millions except Per Share)

For the Year Ended	31 Aug. 2017	1 Sept. 2016	3 Sept. 2015
Net sales	20,322	$12,399	$16,192
Cost of goods sold	11,886	9,894	10,977
Gross margin	8,436	2,505	5,215
Selling, general and administrative	743	659	719
Research and development	1,824	1,617	1,540
Restructure and asset impairments	18	67	3

Exhibit 2 (Continued)

For the Year Ended	31 Aug. 2017	1 Sept. 2016	3 Sept. 2015
Other operating (income) expense, net	(17)	(6)	(45)
Operating income	5,868	168	2,998
Interest income (expense), net	(560)	(395)	(336)
Other non-operating income (expense), net	(112)	(54)	(53)
Income tax (provision) benefit	(114)	(19)	(157)
Equity in net income (loss) of equity method investees	8	25	447
Net income (loss) attributable to non-controlling interests	(1)	(1)	—
Net income (loss) attributable to Micron	$5,089	$(276)	$2,899
Earnings (loss) per share:			
Basic	$4.67	$(0.27)	$2.71
Diluted	$4.41	$(0.27)	$2.47
Number of shares used in per share calculations:			
Basic	1,089	1,036	1,070
Diluted	1,154	1,036	1,170

Exhibit 3 Micron Technology, Inc. Consolidated Balance Sheets (Amounts in US$ Millions)

As of	31 Aug. 2017	1 Sept. 2016
Assets		
Cash and equivalents	$5,109	$4,140
Short-term investments	319	258
Receivables	3,759	2,068
Inventories	3,123	2,889
Other current assets	147	140
Total current assets	12,457	9,495
Long-term marketable investments	617	414
Property, plant and equipment, net	19,431	14,686
Equity method investments	16	1,364
Intangible assets, net	387	464
Deferred tax assets	766	657
Other noncurrent assets	1,662	460
Total assets	$35,336	$27,540
Liabilities and shareholders' equity		
Accounts payable and accrued expenses	$3,664	$3,879
Deferred income	408	200
Current debt	1,262	756

(continued)

Exhibit 3 (Continued)

As of	31 Aug. 2017	1 Sept. 2016
Total current liabilities	5,334	4,835
Long-term debt	9,872	9,154
Other noncurrent liabilities	639	623
Total liabilities	15,845	14,612
Redeemable convertible notes	21	—
Micron shareholder’s equity		
Common stock of $0.10 par value, 3,000 shares authorized, 1,116 shares issued and 1,112 shares outstanding (1,094 issued and 1,040 outstanding as of September 1, 2016)	112	109
Additional capital	8,287	7,736
Retained earnings	10,260	5,299
Treasury stock, 4 shares held (54 as of September 1, 2016)	(67)	(1,029)
Accumulated other comprehensive income (loss)	29	(35)
Total Micron shareholders’ equity	18,621	12,080
Noncontrolling interests in subsidiaries	849	848
Total equity	19,470	12,928
Total liabilities and shareholders’ equity	35,336	$27,540

Exhibit 4 Micron Technology, Inc. Income Taxes Note to the Consolidated Financial Statements

Income (loss) before taxes and the income tax (provision) benefit consisted of the following:

(in US$ Millions)	2017	2016	2015
Income (loss) before income taxes, net income (loss) attributable to noncontrolling interests, and equity in net income (loss) of equity method investees			
Foreign	$5,252	$(353)	$2,431
US	(56)	72	178
	$5,196	$(281)	$2,609
Income tax (provision) benefit:			
Current:			
Foreign	$(152)	$(27)	$(93)
State	(1)	(1)	(1)
US federal	—	—	6
	(153)	(28)	(88)
Deferred:			

Exhibit 4 (Continued)

(in US$ Millions)	2017	2016	2015
US federal	—	39	15
State	—	2	1
Foreign	39	(32)	(85)
	39	9	(69)
Income tax (provision)	$(114)	$(19)	$(157)

The company's income tax (provision) computed using the US federal statutory rate and the company's income tax (provision) benefit is reconciled as follows:

(US$ Millions)	2017	2016	2015
US federal income tax (provision) benefit at statutory rate	$(1,819)	$98	$(913)
Foreign tax rate differential	1,571	(300)	515
Change in valuation allowance	64	63	260
Change in unrecognized tax benefits	12	52	(118)
Tax credits	66	48	53
Noncontrolling investment transactions	—	—	57
Other	(8)	20	(11)
Income tax (provision) benefit	(114)	$(19)	$(157)

State taxes reflect investment tax credits of $233 million as at 31 August 2017. Deferred income taxes reflect the net tax effects of temporary differences between the bases of assets and liabilities for financial reporting and income tax purposes. The company's deferred tax assets and liabilities consist of the following as of the end of the periods shown below:

(US$ Millions)	2017	2016
Deferred tax assets:		
Net operating loss and tax credit carryforwards	$3,426	$3,014
Accrued salaries, wages, and benefits	211	142
Other accrued liabilities	59	76
Other	86	65
Gross deferred assets	3,782	3,297
Less valuation allowance	(2,321)	(2,107)
Deferred tax assets, net of valuation allowance	1,461	1,190
Deferred tax liabilities:		
Debt discount	(145)	(170)
Property, plant, and equipment	(300)	(135)
Unremitted earnings on certain subsidiaries	(123)	(121)
Product and process technology	(85)	(81)
Other	(59)	(28)
Deferred tax liabilities	(712)	(535)
Net deferred tax assets	$749	$655
Reported as:		

(continued)

Exhibit 4 (Continued)

(US$ Millions)	2017	2016
Current deferred tax assets (included in other current assets)	$—	$—
Deferred tax assets	766	657
Current deferred tax liabilities (included in accounts payable and accrued expenses)	—	—
Deferred tax liabilities (included in other noncurrent liabilities)	(17)	(2)
Net deferred tax assets	$749	$655

The company has a valuation allowance against substantially all of its US net deferred tax assets. As of 31 August 2017, the company had aggregate US tax net operating loss carryforwards of $3.88 billion and unused US tax credit carryforwards of $416 million. The company also has unused state tax net operating loss carryforwards of $1.95 billion and unused state tax credits of $233 million. The net operating loss carryforwards and the tax credit carryforwards expire between 2018 to 2037.

The changes in valuation allowance of $64 million and $63 million in 2017 and 2016, respectively, are primarily a result of uncertainties of realizing certain US and foreign net operating losses and certain tax credit carryforwards.

Provision has been made for deferred taxes on undistributed earnings of non-US subsidiaries to the extent that dividend payments from such companies are expected to result in additional tax liability. Remaining undistributed earnings of $12.91 billion as of 31 August 2017 have been indefinitely reinvested. Determination of the amount of unrecognized deferred tax liability on these unremitted earnings is not practicable.

EXAMPLE 6

Financial Analysis Example

Use the financial statement information and disclosures provided by MU in Exhibits 2, 3, and 4 to answer the following questions:

1. MU discloses a valuation allowance of $2,321 million (see Exhibit 4) against gross deferred assets of $3,782 million in 2017. Does the existence of this valuation allowance have any implications concerning MU's future earnings prospects?
2. How would MU's deferred tax assets and deferred tax liabilities be affected if the federal statutory tax rate was changed to 21 percent?
3. How would reported earnings have been affected if MU were not using a valuation allowance?
4. How would MU's $3.88 billion in net operating loss carryforwards in 2017 (see Exhibit 4) affect the valuation that an acquiring company would be willing to offer?
5. Under what circumstances should the analyst consider MU's deferred tax liability as debt or as equity? Under what circumstances should the analyst exclude MU's deferred tax liability from both debt and equity when calculating the debt-to-equity ratio?

Solution to 1:

According to Exhibit 4, MU's deferred tax assets expire gradually until 2037 (2018 to 2037 for the net operating loss carryforwards and the tax credit carryforwards).

Because the company is still relatively young, it is likely that most of these expirations occur toward the end of that period. Because cumulative US tax net operating loss carryforwards total $3.88 billion, the valuation allowance could imply that MU is not reasonably expected to earn $3.88 billion over the next 20 years. However, as we can see in Exhibit 2, MU earned a profit for 2017 and 2015, thereby showing that the allowance could be adjusted downward if the company continues to generate profits in the future, and making it more likely than not that the deferred tax asset would be recognized.

Solution to 2:

MU's total deferred tax assets exceed total deferred tax liabilities by $749 million. A change in the federal statutory tax rate to 21 percent from the current rate of 35 percent would make these net deferred assets less valuable. Also, because it is possible that the deferred tax asset valuation allowance could be adjusted downward in the future (see discussion to solution 1), the impact could be far greater in magnitude.

Solution to 3:

The disclosure in Exhibit 4 shows that the increase in the valuation allowance increased the income tax provision as reported on the income statement by $64 million in 2017. Additional potential reductions in the valuation allowance could similarly reduce reported income taxes (actual income taxes would not be affected by a valuation allowance established for financial reporting) in future years (see discussion to solution 1).

Solution to 4:

If an acquiring company is profitable, it may be able to use MU's tax loss carryforwards to offset its own tax liabilities. The value to an acquirer would be the present value of the carryforwards, based on the acquirer's tax rate and expected timing of realization. The higher the acquiring company's tax rate, and the more profitable the acquirer, the sooner it would be able to benefit. Therefore, an acquirer with a high current tax rate would theoretically be willing to pay more than an acquirer with a lower tax rate.

Solution to 5:

The analyst should classify the deferred tax liability as debt if the liability is expected to reverse with subsequent tax payment. If the liability is not expected to reverse, there is no expectation of a cash outflow and the liability should be treated as equity. By way of example, future company losses may preclude the payment of any income taxes, or changes in tax laws could result in taxes that are never paid. The deferred tax liability should be excluded from both debt and equity when both the amounts and timing of tax payments resulting from the reversals of temporary differences are uncertain.

11

COMPARISON OF IFRS AND US GAAP

j identify the key provisions of and differences between income tax accounting under International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (GAAP).

As mentioned earlier, though IFRS and US GAAP follow similar conventions on many tax issues, there are some notable differences. Exhibit 5 summarizes many of the key similarities and differences between IFRS and US GAAP. Though both frameworks require a provision for deferred taxes, there are differences in the methodologies.

Exhibit 5 Deferred Income Tax Issues IFRS and US GAAP Methodology Similarities and Differences

IFRS	US GAAP
Introduction	
The objective in accounting for income taxes is to recognize the amount of taxes +currently payable or refundable and deferred taxes. Income tax expense is the current tax expense (or recovery) plus the period change in deferred taxes (net of tax arising from a business combination or recorded outside profit or loss).	Similar to IFRS.
Unpaid taxes for current and prior periods are recognized as a liability, and an asset is recognised if the amount already paid exceeds the amount due. A prior tax loss benefit used to recover previous period current tax is also an asset.	The approach to calculating current taxes is similar to IFRS with some exceptions, such as the treatment of taxes on the elimination of intercompany profits.
In general, deferred taxes are recognized using an asset and liability approach which focuses on temporary differences arising between the tax base of an asset or liability and its carrying amount in the statement of financial position. Deferred taxes are recognized for the future tax consequences of events that have been recognized in an entity's financial statements or tax returns.	US GAAP also follows an asset and liability approach to calculating deferred taxes, although there are some differences in the application relative to IFRS.
Deferred taxes are not recognized for:	Deferred taxes are not recognized for:
■ The initial recognition of goodwill	■ Goodwill for which amortization is not deductible for tax purposes
■ The initial recognition of an asset or liability in a non-business combination transaction and where accounting profit or taxable profit (tax loss) is not affected	■ Unlike IFRS, US GAAP does not have a similar exception
■ Taxable temporary differences from investments in subsidiaries/branches/ associates, and interests in joint ventures in which the parent etc. is able to control the timing of the reversal of the temporary difference, and it is probable that the temporary difference will not reverse in the foreseeable future	■ An excess of the amount for the financial reporting over the tax basis of an investment in a foreign subsidiary or a foreign corporate joint venture that is essentially permanent in duration, unless it becomes apparent that those temporary differences will reverse in the foreseeable future. Unlike IFRS, this exception does not apply to domestic subsidiaries and corporate joint venture and investments in equity investees.
	Unlike IFRS, recognition of deferred taxes is prohibited for differences that are remeasured from the local currency into the functional currency using historical exchange rates and that result from changes in exchange rates or indexing for tax purposes.

Exhibit 5 (Continued)

IFRS	US GAAP
Deferred taxes should be recognized for the difference between the carrying amount determined by using the historical exchange rate and the relevant tax base, which may have been affected by exchange rate changes or tax indexing.	
Recognition and measurement	
Current tax liabilities and assets for the current and prior periods are measured at amounts expected to be paid to (recovered from) the taxation authorities based on tax rates (and tax laws) that have been enacted or substantially enacted by the end of the reporting period. Deferred tax assets are measured at the tax rates that are expected to apply when the asset is realized or the liability is settled, based on tax rates (and tax laws) that have been enacted or substantively enacted by the end of the reporting period.	Similar to IFRS. Measurement of current and deferred tax assets and liabilities is based on enacted tax law. Deferred tax assets and liabilities are measured using enacted tax rate(s) expected to apply to taxable income in periods in which deferred tax is expected to be settled or realized. Unlike IFRS, use of substantially enacted tax rates is not permitted.
Deferred tax assets are recognized to the extent that it is probable (more likely than not) that taxable profit will be available to utilize the deductible temporary difference or carryforward of unused tax losses or tax credits. End of reporting period reviews may reduce the carrying amount if sufficient taxable profit is no longer probable such as to allow the utilization benefit of all or part of that deferred tax asset, and any such reduction is reversed if it subsequently becomes probable again.	Unlike IFRS, deferred tax assets are recognized in full and reduced by a valuation allowance if it is more likely than not that some portion or all of the deferred tax assets will not be realized.
Current and deferred taxes are recognized outside profit or loss if the tax relates to items that are recognized, in the same or a different period, in other comprehensive income (OCI), or directly to equity.	Similar to IFRS, the tax effects of certain items occurring during the year are charged or credited directly to OCI or to related components of shareholders' equity.
Deferred tax assets and liabilities are not discounted.	Similar to IFRS.
Presentation and disclosure	
Deferred tax assets and liabilities are offset if the entity has a legally enforceable right to offset current tax assets against current tax liabilities and the deferred tax assets and deferred tax liabilities relate to income taxes levied by the same taxing authority on either the same taxable entity, or different taxable entities that intend either to settle current tax assets and liabilities on a net basis or to simultaneously realize/settle the asset/liability.	All deferred taxes are offset and presented as a single amount.

(continued)

Exhibit 5 (Continued)

IFRS	US GAAP
Deferred tax assets and liabilities are presented as separate line items in the statement of financial position. If a classified statement of financial position is used, deferred taxes are classified as noncurrent.	Deferred tax assets and deferred tax liabilities are presented as noncurrent in a classified statement of financial position, which aligns with IFRS.
All entities must disclose an explanation of the relationship between tax expense and accounting profit using either or both of the following formats: ■ A numerical reconciliation between tax expense (income) and the product of accounting profit multiplied by the applicable tax rate(s) including disclosure of the basis on which the applicable rate is computed. ■ A numerical reconciliation between the average effective tax rate and the applicable tax rate, including disclosure of the basis on which the applicable tax rate is computed.	Public companies must disclose a reconciliation using percentages or dollar amounts of the reported amount of income tax expense attributable to continuing operations for the year to that amount of income tax expense that would result from applying domestic federal statutory tax rates to pretax income from continuing operations. Nonpublic enterprises must disclose the nature of significant reconciling items but may omit a numerical reconciliation.

Sources: IFRS: IAS 12 and 32. US GAAP: ASC 740. "Comparison between US GAAP and IFRS Standards," Section 5.3 Taxation, Grant Thornton, April 2017. "IFRS and US GAAP: similarities and differences", PricewaterhouseCoopers LLC, 2018.

SUMMARY

Income taxes are a significant category of expense for profitable companies. Analyzing income tax expenses is often difficult for the analyst because there are many permanent and temporary timing differences between the accounting that is used for income tax reporting and the accounting that is used for financial reporting on company financial statements. The financial statements and notes to the financial statements of a company provide important information that the analyst needs to assess financial performance and to compare a company's financial performance with other companies. Key concepts in this reading are as follows:

- Differences between the recognition of revenue and expenses for tax and accounting purposes may result in taxable income differing from accounting profit. The discrepancy is a result of different treatments of certain income and expenditure items.
- The tax base of an asset is the amount that will be deductible for tax purposes as an expense in the calculation of taxable income as the company expenses the tax basis of the asset. If the economic benefit will not be taxable, the tax base of the asset will be equal to the carrying amount of the asset.
- The tax base of a liability is the carrying amount of the liability less any amounts that will be deductible for tax purposes in the future. With respect to revenue received in advance, the tax base of such a liability is the carrying amount less any amount of the revenue that will not be taxable in the future.
- Temporary differences arise from recognition of differences in the tax base and carrying amount of assets and liabilities. The creation of a deferred tax asset or liability as a result of a temporary difference will only be allowed if the difference reverses itself at some future date and to the extent that it is expected that the balance sheet item will create future economic benefits for the company.

- Permanent differences result in a difference in tax and financial reporting of revenue (expenses) that will not be reversed at some future date. Because it will not be reversed at a future date, these differences do not constitute temporary differences and do not give rise to a deferred tax asset or liability.
- Current taxes payable or recoverable are based on the applicable tax rates on the balance sheet date of an entity; in contrast, deferred taxes should be measured at the tax rate that is expected to apply when the asset is realized or the liability settled.
- All unrecognized deferred tax assets and liabilities must be reassessed on the appropriate balance sheet date and measured against their probable future economic benefit.
- Deferred tax assets must be assessed for their prospective recoverability. If it is probable that they will not be recovered at all or partly, the carrying amount should be reduced. Under US GAAP, this is done through the use of a valuation allowance.

PRACTICE PROBLEMS

1. Using the straight-line method of depreciation for reporting purposes and accelerated depreciation for tax purposes would *most likely* result in a:
   - **A** valuation allowance.
   - **B** deferred tax asset.
   - **C** temporary difference.
2. In early 2018 Sanborn Company must pay the tax authority €37,000 on the income it earned in 2017. This amount was recorded on the company's 31 December 2017 financial statements as:
   - **A** taxes payable.
   - **B** income tax expense.
   - **C** a deferred tax liability.
3. Income tax expense reported on a company's income statement equals taxes payable, plus the net increase in:
   - **A** deferred tax assets and deferred tax liabilities.
   - **B** deferred tax assets, less the net increase in deferred tax liabilities.
   - **C** deferred tax liabilities, less the net increase in deferred tax assets.
4. Analysts should treat deferred tax liabilities that are expected to reverse as:
   - **A** equity.
   - **B** liabilities.
   - **C** neither liabilities nor equity.
5. Deferred tax liabilities should be treated as equity when:
   - **A** they are not expected to reverse.
   - **B** the timing of tax payments is uncertain.
   - **C** the amount of tax payments is uncertain.
6. When both the timing and amount of tax payments are uncertain, analysts should treat deferred tax liabilities as:
   - **A** equity.
   - **B** liabilities.
   - **C** neither liabilities nor equity.
7. When accounting standards require recognition of an expense that is not permitted under tax laws, the result is a:
   - **A** deferred tax liability.
   - **B** temporary difference.
   - **C** permanent difference.
8. When certain expenditures result in tax credits that directly reduce taxes, the company will *most likely* record:
   - **A** a deferred tax asset.
   - **B** a deferred tax liability.
   - **C** no deferred tax asset or liability.

© 2019 CFA Institute. All rights reserved.

9 When accounting standards require an asset to be expensed immediately but tax rules require the item to be capitalized and amortized, the company will *most likely* record:

A a deferred tax asset.

B a deferred tax liability.

C no deferred tax asset or liability.

10 A company incurs a capital expenditure that may be amortized over five years for accounting purposes, but over four years for tax purposes. The company will *most likely* record:

A a deferred tax asset.

B a deferred tax liability.

C no deferred tax asset or liability.

11 A company receives advance payments from customers that are immediately taxable but will not be recognized for accounting purposes until the company fulfills its obligation. The company will *most likely* record:

A a deferred tax asset.

B a deferred tax liability.

C no deferred tax asset or liability.

The following information relates to Questions 12–14

Note I
Income Taxes

The components of earnings before income taxes are as follows ($ thousands):

	Year 3	Year 2	Year 1
Earnings before income taxes:			
United States	$88,157	$75,658	$59,973
Foreign	116,704	113,509	94,760
Total	$204,861	$189,167	$154,733

The components of the provision for income taxes are as follows ($ thousands):

	Year 3	Year 2	Year 1
Income taxes			
Current:			
Federal	$30,632	$22,031	$18,959
Foreign	28,140	27,961	22,263
	$58,772	$49,992	$41,222
Deferred:			
Federal	($4,752)	$5,138	$2,336
Foreign	124	1,730	621

(continued)

	Year 3	Year 2	Year 1
	(4,628)	6,868	2,957
Total	$54,144	$56,860	$44,179

12 In Year 3, the company's US GAAP income statement recorded a provision for income taxes *closest* to:

A $30,632.

B $54,144.

C $58,772.

13 The company's effective tax rate was *highest* in:

A Year 1.

B Year 2.

C Year 3.

14 Compared to the company's effective tax rate on US income, its effective tax rate on foreign income was:

A lower in each year presented.

B higher in each year presented.

C higher in some periods and lower in others.

---

15 Zimt AG presents its financial statements in accordance with US GAAP. In Year 3, Zimt discloses a valuation allowance of $1,101 against total deferred tax assets of $19,201. In Year 2, Zimt disclosed a valuation allowance of $1,325 against total deferred tax assets of $17,325. The change in the valuation allowance *most likely* indicates that Zimt's:

A deferred tax liabilities were reduced in Year 3.

B expectations of future earning power has increased.

C expectations of future earning power has decreased.

16 Cinnamon, Inc. recorded a total deferred tax asset in Year 3 of $12,301, offset by a $12,301 valuation allowance. Cinnamon *most likely*:

A fully utilized the deferred tax asset in Year 3.

B has an equal amount of deferred tax assets and deferred tax liabilities.

C expects not to earn any taxable income before the deferred tax asset expires.

The following information relates to Questions 17–19

The tax effects of temporary differences that give rise to deferred tax assets and liabilities are as follows ($ thousands):

	Year 3	Year 2
Deferred tax assets:		
Accrued expenses	$8,613	$7,927
Tax credit and net operating loss carryforwards	2,288	2,554

	Year 3	Year 2
LIFO and inventory reserves	5,286	4,327
Other	2,664	2,109
Deferred tax assets	18,851	16,917
Valuation allowance	(1,245)	(1,360)
Net deferred tax assets	$17,606	$15,557
Deferred tax liabilities:		
Depreciation and amortization	$(27,338)	$(29,313)
Compensation and retirement plans	(3,831)	(8,963)
Other	(1,470)	(764)
Deferred tax liabilities	(32,639)	(39,040)
Net deferred tax liability	($15,033)	($23,483)

17 A reduction in the statutory tax rate would *most likely* benefit the company's:

A income statement and balance sheet.

B income statement but not the balance sheet.

C balance sheet but not the income statement.

18 If the valuation allowance had been the same in Year 3 as it was in Year 2, the company would have reported $115 *higher*:

A net income.

B deferred tax assets.

C income tax expense.

19 Compared to the provision for income taxes in Year 3, the company's cash tax payments were:

A lower.

B higher.

C the same.

The following information relates to Questions 20–22

A company's provision for income taxes resulted in effective tax rates attributable to loss from continuing operations before cumulative effect of change in accounting principles that varied from the statutory federal income tax rate of 34 percent, as summarized in the table below.

Year Ended 30 June	Year 3	Year 2	Year 1
Expected federal income tax expense (benefit) from continuing operations at 34 percent	($112,000)	$768,000	$685,000
Expenses not deductible for income tax purposes	357,000	32,000	51,000
State income taxes, net of federal benefit	132,000	22,000	100,000

(continued)

Year Ended 30 June	Year 3	Year 2	Year 1
Change in valuation allowance for deferred tax assets	(150,000)	(766,000)	(754,000)
Income tax expense	$227,000	$56,000	$82,000

20 In Year 3, the company's net income (loss) was *closest* to:

A ($217,000).

B ($329,000).

C ($556,000).

21 The $357,000 adjustment in Year 3 *most likely* resulted in:

A an increase in deferred tax assets.

B an increase in deferred tax liabilities.

C no change to deferred tax assets and liabilities.

22 Over the three years presented, changes in the valuation allowance for deferred tax assets were *most likely* indicative of:

A decreased prospect for future profitability.

B increased prospects for future profitability.

C assets being carried at a higher value than their tax base.

SOLUTIONS

1. C is correct. Because the differences between tax and financial accounting will correct over time, the resulting deferred tax liability, for which the expense was charged to the income statement but the tax authority has not yet been paid, will be a temporary difference. A valuation allowance would only arise if there was doubt over the company's ability to earn sufficient income in the future to require paying the tax.
2. A is correct. The taxes a company must pay in the immediate future are taxes payable.
3. C is correct. Higher reported tax expense relative to taxes paid will increase the deferred tax liability, whereas lower reported tax expense relative to taxes paid increases the deferred tax asset.
4. B is correct. If the liability is expected to reverse (and thus require a cash tax payment) the deferred tax represents a future liability.
5. A is correct. If the liability will not reverse, there will be no required tax payment in the future and the "liability" should be treated as equity.
6. C is correct. The deferred tax liability should be excluded from both debt and equity when both the amounts and timing of tax payments resulting from the reversals of temporary differences are uncertain.
7. C is correct. Accounting items that are not deductible for tax purposes will not be reversed and thus result in permanent differences.
8. C is correct. Tax credits that directly reduce taxes are a permanent difference, and permanent differences do not give rise to deferred tax.
9. A is correct. The capitalization will result in an asset with a positive tax base and zero carrying value. The amortization means the difference is temporary. Because there is a temporary difference on an asset resulting in a higher tax base than carrying value, a deferred tax asset is created.
10. B is correct. The difference is temporary, and the tax base will be lower (because of more rapid amortization) than the carrying value of the asset. The result will be a deferred tax liability.
11. A is correct. The advances represent a liability for the company. The carrying value of the liability exceeds the tax base (which is now zero). A deferred tax asset arises when the carrying value of a liability exceeds its tax base.
12. B is correct. The income tax provision in Year 3 was $54,144, consisting of $58,772 in current income taxes, of which $4,628 were deferred.
13. B is correct. The effective tax rate of 30.1 percent ($56,860/$189,167) was higher than the effective rates in Year 2 and Year 3.
14. A is correct. In Year 3 the effective tax rate on foreign operations was 24.2 percent [($28,140 + $124)/$116,704] and the effective US tax rate was [($30,632 – $4,752)/$88,157] = 29.4 percent. In Year 2 the effective tax rate on foreign operations was 26.2 percent and the US rate was 35.9 percent. In Year 1 the foreign rate was 24.1 percent and the US rate was 35.5 percent.
15. B is correct. The valuation allowance is taken against deferred tax assets to represent uncertainty that future taxable income will be sufficient to fully utilize the assets. By decreasing the allowance, Zimt is signaling greater likelihood that future earnings will be offset by the deferred tax asset.

16 C is correct. The valuation allowance is taken when the company will "more likely than not" fail to earn sufficient income to offset the deferred tax asset. Because the valuation allowance equals the asset, by extension the company expects *no* taxable income prior to the expiration of the deferred tax assets.

17 A is correct. A lower tax rate would increase net income on the income statement, and because the company has a net deferred tax liability, the net liability position on the balance sheet would also improve (be smaller).

18 C is correct. The reduction in the valuation allowance resulted in a corresponding reduction in the income tax provision.

19 B is correct. The net deferred tax liability was smaller in Year 3 than it was in Year 2, indicating that in addition to meeting the tax payments provided for in Year 3 the company also paid taxes that had been deferred in prior periods.

20 C is correct. The income tax provision at the statutory rate of 34 percent is a benefit of $112,000, suggesting that the pre-tax income was a loss of $112,000/0.34 = ($329,412). The income tax provision was $227,000. ($329,412) – $227,000 = ($556,412).

21 C is correct. Accounting expenses that are not deductible for tax purposes result in a permanent difference, and thus do not give rise to deferred taxes.

22 B is correct. Over the three-year period, changes in the valuation allowance reduced cumulative income taxes by $1,670,000. The reductions to the valuation allowance were a result of the company being "more likely than not" to earn sufficient taxable income to offset the deferred tax assets.

READING

24

Non-Current (Long-Term) Liabilities

by Elizabeth A. Gordon, PhD, MBA, CPA, and Elaine Henry, PhD, CFA

Elizabeth A. Gordon, PhD, MBA, CPA, is at Temple University (USA). Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** determine the initial recognition, initial measurement and subsequent measurement of bonds;
☐	**b.** describe the effective interest method and calculate interest expense, amortisation of bond discounts/premiums, and interest payments;
☐	**c.** explain the derecognition of debt;
☐	**d.** describe the role of debt covenants in protecting creditors;
☐	**e.** describe the financial statement presentation of and disclosures relating to debt;
☐	**f.** explain motivations for leasing assets instead of purchasing them;
☐	**g.** explain the financial reporting of leases from a lessee's perspective;
☐	**h.** explain the financial reporting of leases from a lessor's perspective;
☐	**i.** compare the presentation and disclosure of defined contribution and defined benefit pension plans;
☐	**j.** calculate and interpret leverage and coverage ratios.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2021 CFA Institute. All rights reserved.

1 INTRODUCTION

A non-current liability (long-term liability) broadly represents a probable sacrifice of economic benefits in periods generally greater than one year in the future. Common types of **non-current liabilities** reported in a company's financial statements include long-term debt (e.g., bonds payable, long-term notes payable), leases, pension liabilities, and deferred tax liabilities. This reading focuses on bonds payable, leases, and pension liabilities.

This reading is organised as follows. Section 2 introduces bonds and the accounting for their issuance. Section 3 discusses the recording of interest expense and interest payments as well as the amortisation of discount or premium. Section 4 describes fair value accounting for bonds, an alternative to the amortised cost approach. Section 5 discusses the repayment of principal when bonds are redeemed or reach maturity, which requires derecognition from the financial statements. Section 6 covers debt covenants. Section 7 describes the financial statement presentation and disclosures about debt financings. Section 8 discusses leases, including the benefits of leasing and accounting for leases by both lessees and lessors. Section 9 introduces pension accounting and the resulting non-current liabilities. Section 10 discusses the use of leverage and coverage ratios in evaluating solvency. Section 11 concludes and summarises the reading.

2 BONDS PAYABLE & ACCOUNTING FOR BOND ISSUANCE

a determine the initial recognition, initial measurement and subsequent measurement of bonds

This section discusses accounting for bonds payable—a common form of long-term debt. In some contexts (e.g., some government debt obligations), the word "bond" is used only for a debt security with a maturity of 10 years or longer; "note" refers to a debt security with a maturity between 2 and 10 years; and "bill" refers to a debt security with a maturity of less than 2 years. In this reading, we use the terms bond and note interchangeably because the accounting treatments of bonds payable and long-term notes payable are similar. In the following sections, we discuss bond issuance (initial recognition and measurement); bond amortisation, interest expense, and interest payments; market rates and fair value (subsequent measurement); repayment of bonds, including retirements and redemptions (derecognition); and other issues concerning disclosures related to debt. We also discuss debt covenants.

2.1 Accounting for Bond Issuance

Bonds are contractual promises made by a company (or other borrowing entity) to pay cash in the future to its lenders (i.e., bondholders) in exchange for receiving cash in the present. The terms of a bond contract are contained in a document called an indenture. The cash or sales proceeds received by a company when it issues bonds is based on the value (price) of the bonds at the time of issue; the price at the time of issue is determined as the present value of the future cash payments promised by the company in the bond agreement.

Ordinarily, bonds contain promises of two types of future cash payments: 1) the face value of the bonds, and 2) periodic interest payments. The **face value** of the bonds is the amount of cash payable by the company to the bondholders when the bonds mature. The face value is also referred to as the principal, par value, stated value, or maturity value. The date of maturity of the bonds (the date on which the face value is paid to bondholders) is stated in the bond contract and typically is a number of years in the future. Periodic interest payments are made based on the interest rate promised in the bond contract applied to the bonds' face value. The interest rate promised in the contract, which is the rate used to calculate the periodic interest payments, is referred to as the **coupon rate**, nominal rate, or stated rate. Similarly, the periodic interest payment is referred to as the coupon payment or simply the coupon. For fixed rate bonds (the primary focus of our discussion here), the coupon rate remains unchanged throughout the life of the bonds. The frequency with which interest payments are made is also stated in the bond contract. For example, bonds paying interest semi-annually will make two interest payments per year.[1]

The future cash payments are discounted to the present to arrive at the market value of the bonds. The **market rate of interest** is the rate demanded by purchasers of the bonds given the risks associated with future cash payment obligations of the particular bond issue. The market rate of interest at the time of issue often differs from the coupon rate because of interest rate fluctuations that occur between the time the issuer establishes the coupon rate and the day the bonds are actually available to investors. If the market rate of interest when the bonds are issued equals the coupon rate, the market value (price) of the bonds will equal the face value of the bonds. Thus, ignoring

issuance costs, the issuing company will receive sales proceeds (cash) equal to the face value of the bonds. When a bond is issued at a price equal to its face value, the bond is said to have been issued at par.

If the coupon rate when the bonds are issued is higher than the market rate, the market value of the bonds—and thus the amount of cash the company receives—will be higher than the face value of the bonds. In other words, the bonds will sell at a premium to face value because they are offering an attractive coupon rate compared to current market rates. If the coupon rate is lower than the market rate, the market value and thus the sale proceeds from the bonds will be less than the face value of the bonds; the bond will sell at a discount to face value. The market rate at the time of issuance is the **effective interest rate** or borrowing rate that the company incurs on the debt. The effective interest rate is the discount rate that equates the present value of the two types of promised future cash payments to their selling price. For the issuing company, interest expense reported for the bonds in the financial statements is based on the effective interest rate.

On the issuing company's statement of cash flows, the cash received (sales proceeds) from issuing bonds is reported as a financing cash inflow. On the issuing company's balance sheet at the time of issue, bonds payable normally are measured and reported at the sales proceeds. In other words, the bonds payable are initially reported at the face value of the bonds minus any discount, or plus any premium.

Using a three-step approach, the following two examples illustrate accounting for bonds issued at face value and then accounting for bonds issued at a discount to face value. Accounting for bonds issued at a premium involves steps similar to the steps followed in the examples below. For simplicity, these examples assume a flat interest rate yield curve (i.e., that the market rate of interest is the same for each period). More-precise bond valuations use the interest rate applicable to each time period in which a payment of interest or principal occurs.

---

[1] Interest rates are stated on an annual basis regardless of the frequency of payment.

EXAMPLE 1

Bonds Issued at Face Value

Debond Corp. (a hypothetical company) issues £1,000,000 worth of five-year bonds, dated 1 January 2018, when the market interest rate on bonds of comparable risk and terms is 5 percent per annum. The bonds pay 5 percent interest annually on 31 December. What are the sales proceeds of the bonds when issued, and how is the issuance reflected in the financial statements?

Solution:

Calculating the value of the bonds at issuance and thus the sales proceeds involves three steps: 1) identifying key features of the bonds and the market interest rate, 2) determining future cash outflows, and 3) discounting the future cash flows to the present.

First, identify key features of the bonds and the market interest rate necessary to determine sales proceeds:

Face value (principal):	£1,000,000	
Time to maturity:	5 years	
Coupon rate:	5%	
Market rate at issuance:	5%	
Frequency of interest payments:	annual	
Interest payment:	£50,000	Each annual interest payment is the face value times the coupon rate (£1,000,000 × 5%). If interest is paid other than annually, adjust the interest rate to match the interest payment period (e.g., divide the annual coupon rate by two for semi-annual interest payments).

Second, determine future cash outflows. Debond will pay bondholders £1,000,000 when the bonds mature in five years. On 31 December of each year until the bonds mature, Debond will make an interest payment of £50,000.

Third, sum the present value[2] of the future payments of interest and principal to obtain the value of the bonds and thus the sales proceeds from issuing the bonds. In this example, the sum is £1,000,000 = (£216,474 + £783,526).

Date	Interest Payment	Present Value at Market Rate (5%)	Face Value Payment	Present Value at Market Rate (5%)	Total Present Value
31 December 2018	£50,000	£47,619			
31 December 2019	50,000	45,352			
31 December 2020	50,000	43,192			
31 December 2021	50,000	41,135			
31 December 2022	50,000	39,176	£1,000,000	£783,526	
Total		£216,474		£783,526	£1,000,000
					Sales Proceeds

---

[2] Alternative ways to calculate the present value include 1) to treat the five annual interest payments as an annuity and use the formula for finding the present value of an annuity and then add the present value of the principal payment, or 2) to use a financial calculator to calculate the total present value.

The sales proceeds of the bonds when issued are £1,000,000. There is no discount or premium because these bonds are issued at face value. The issuance is reflected on the balance sheet as an increase of cash and an increase in a long-term liability, bonds payable, of £1,000,000. The issuance is reflected in the statement of cash flows as a financing cash inflow of £1,000,000.

The price of bonds is often expressed as a percentage of face value. For example, the price of bonds issued at par, as in Example 1, is 100 (i.e., 100 percent of face value). In Example 2, in which bonds are issued at a discount, the price is 95.79 (i.e., 95.79 percent of face value).

EXAMPLE 2

Bonds Issued at a Discount

Debond Corp. issues £1,000,000 worth of five-year bonds, dated 1 January 2018, when the market interest rate on bonds of comparable risk and terms is 6 percent. The bonds pay 5 percent interest annually on 31 December. What are the sales proceeds of the bonds when issued, and how is the issuance reflected in the financial statements?

Solution:

The key features of the bonds and the market interest rate are:

Face value (principal):	£1,000,000	
Time to maturity:	5 years	
Coupon rate:	5%	
Market rate at issuance:	6%	
Frequency of interest payments:	annual	
Interest payment:	£50,000	Each annual interest payment is the face value times the coupon rate (£1,000,000 × 5%).

The future cash outflows (interest payments and face value payment), the present value of the future cash outflows, and the total present value are:

Date	Interest Payment	Present Value at Market Rate (6%)	Face Value Payment	Present Value at Market Rate (6%)	Total Present Value
31 December 2018	£50,000	£47,170			
31 December 2019	50,000	44,500			
31 December 2020	50,000	41,981			
31 December 2021	50,000	39,605			
31 December 2022	50,000	37,363	£1,000,000	£747,258	
Total		£210,618		£747,258	£957,876
					Sales Proceeds

The sales proceeds of the bonds when issued are £957,876. The bonds sell at a discount of £42,124 = (£1,000,000 − £957,876) because the market rate when the bonds are issued (6 percent) is greater than the bonds' coupon rate (5 percent). The issuance is reflected on the balance sheet as an increase of cash and an increase in a long-term liability, bonds payable, of £957,876. The bonds payable is composed of the face value of £1,000,000 minus a discount of £42,124. The issuance is reflected in the statement of cash flows as a financing cash inflow of £957,876.

In Example 2, the bonds were issued at a discount to face value because the bonds' coupon rate of 5 percent was less than the market rate. Bonds are issued at a premium to face value when the bonds' coupon rate exceeds the market rate.

Bonds issued with a coupon rate of zero (zero-coupon bonds) are issued at a discount to face value if the market rate is greater than zero. If the market rate is zero or negative, zero-coupon bonds will be issued at par or at premium, respectively. The value of zero-coupon bonds is based on the present value of the principal payment only because there are no periodic interest payments.

Such issuance costs as printing, legal fees, commissions, and other types of charges are costs incurred when bonds are issued. Under International Financial Reporting Standards (IFRS), all debt issuance costs are included in the measurement of the liability, bonds payable. Under US generally accepted accounting principles (US GAAP), companies generally used to show these debt issuance costs as an asset (a deferred charge), which was amortised on a straight-line basis to the relevant expense (e.g., legal fees) over the life of the bonds. Under US GAAP, debt issuance costs are deducted from the related debt liability. Companies reporting under US GAAP may still report debt issuance costs for lines of credit as an asset because the SEC indicated that it would not object to this treatment. Under IFRS and US GAAP, cash outflows related to bond issuance costs are included in the financing section of the statement of cash flows, usually netted against bond proceeds.

3

ACCOUNTING FOR BOND AMORTISATION, INTEREST EXPENSE, AND INTEREST PAYMENTS

a determine the initial recognition, initial measurement and subsequent measurement of bonds

b describe the effective interest method and calculate interest expense, amortisation of bond discounts/premiums, and interest payments

In this section, we discuss accounting and reporting for bonds after they are issued. Most companies maintain the historical cost (sales proceeds) of the bonds after issuance, and they amortise any discount or premium over the life of the bond. The amount reported on the balance sheet for bonds is thus the historical cost plus or minus the cumulative amortisation, which is referred to as amortised cost. Companies also have the option to report the bonds at their current fair values.

The rationale for reporting the bonds at amortised historical cost is the company's intention to retain the debt until it matures. Therefore, changes in the underlying economic value of the debt are not relevant from the issuing company's perspective. From an investor's perspective, however, analysis of a company's underlying economic liabilities and solvency is more difficult when debt is reported at amortised historical cost. The rest of this section illustrates accounting and reporting of bonds at amortised historical cost. Section 2.3 discusses the alternative of reporting bonds at fair value.

Companies initially report bonds as a liability on their balance sheet at the amount of the sales proceeds net of issuance costs under both IFRS and US GAAP. The amount at which bonds are reported on the company's balance sheet is referred to as the carrying amount, carrying value, book value, or net book value. If the bonds are issued at par, the initial carrying amount will be identical to the face value, and usually the carrying amount will not change over the life of the bonds.[3] For bonds issued at face value, the amount of periodic interest *expense* will be the same as the amount of periodic interest *payment* to bondholders.

If, however, the market rate differs from the bonds' coupon rate at issuance such that the bonds are issued at a premium or discount, the premium or discount is amortised systematically over the life of the bonds as a component of interest expense. For bonds issued at a premium to face value, the carrying amount of the bonds is initially greater than the face value. As the premium is amortised, the carrying amount (amortised cost) of the bonds will decrease to the face value. The reported interest expense will be less than the coupon payment. For bonds issued at a discount to face value, the carrying amount of the bonds is initially less than the face value. As the discount is amortised, the carrying amount (amortised cost) of the bonds will increase to the face value. The reported interest expense will be higher than the coupon payment.

The accounting treatment for bonds issued at a discount reflects the fact that the company essentially paid some of its borrowing costs at issuance by selling its bonds at a discount. Rather than there being an actual cash transfer in the future, this "payment" was made in the form of accepting less than the face value for the bonds at the date of issuance. The remaining borrowing cost occurs as a cash interest payment to investors each period. The total interest expense reflects both components of the borrowing cost: the periodic interest payments plus the amortisation of the discount. The accounting treatment for bonds issued at a premium reflects the fact that the company essentially received a reduction on its borrowing costs at issuance by selling its bonds at a premium. Rather than there being an actual reduced cash transfer in the future, this "reduction" was made in the form of receiving more than face value for the bonds at the date of issuance. The total interest expense reflects both components of the borrowing cost: the periodic interest payments less the amortisation of the premium. When the bonds mature, the carrying amount will be equal to the face value regardless of whether the bonds were issued at face value, a discount, or a premium.

Two methods for amortising the premium or discount of bonds that were issued at a price other than par are the effective interest rate method and the straight-line method. The effective interest rate method is required under IFRS and preferred under US GAAP because it better reflects the economic substance of the transaction. The effective interest rate method applies the market rate in effect when the bonds were issued (historical market rate or effective interest rate) to the current amortised cost (carrying amount) of the bonds to obtain interest expense for the period. The difference between the interest expense (based on the effective interest rate and amortised cost) and the interest payment (based on the coupon rate and face value) is the **amortisation** of the discount or premium. The straight-line method of amortisation evenly amortises the premium or discount over the life of the bond, similar to straight-line depreciation on long-lived assets. Under either method, as the bond approaches maturity, the amortised cost approaches face value.

Example 3 illustrates both methods of amortisation for bonds issued at a discount. Example 4 shows amortisation for bonds issued at a premium.

---

[3] If a company reports debt at fair value, rather than amortised cost, the carrying value may change.

EXAMPLE 3

Amortising a Bond Discount

Debond Corp. issues £1,000,000 face value of five-year bonds, dated 1 January 2017, when the market interest rate is 6 percent. The sales proceeds are £957,876. The bonds pay 5 percent interest annually on 31 December.

1 What is the interest *payment* on the bonds each year?

2 What amount of interest *expense* on the bonds would be reported in 2017 and 2018 using the effective interest rate method?

3 Determine the reported value of the bonds (i.e., the carrying amount) at 31 December 2017 and 2018, assuming the effective interest rate method is used to amortise the discount.

4 What amount of interest expense on the bonds would be reported under the straight-line method of amortising the discount?

Solution to 1:

The interest payment equals £50,000 annually (£1,000,000 × 5%).

Solution to 2:

The sales proceeds of £957,876 are less than the face value of £1,000,000; the bonds were issued at a discount of £42,124. The bonds are initially reported as a long-term liability, bonds payable, of £957,876, which comprises the face value of £1,000,000 minus a discount of £42,124. The discount is amortised over time, ultimately, increasing the carrying amount (amortised cost) to face value.

Under the effective interest rate method, interest expense on the bonds is calculated as the bonds' carrying amount times the market rate in effect when the bonds are issued (effective interest rate). For 2017, interest expense is £57,473 = (£957,876 × 6%). The amount of the discount amortised in 2017 is the difference between the interest expense of £57,473 and the interest payment of £50,000 (i.e., £7,473). The bonds' carrying amount increases by the discount amortisation; at 31 December 2017, the bonds' carrying amount is £965,349 (beginning balance of £957,876 plus £7,473 discount amortisation). At this point, the carrying amount reflects a remaining unamortised discount of £34,651 (£42,124 discount at issuance minus £7,473 amortised).

For 2018, interest expense is £57,921 = (£965,349 × 6%), the carrying amount of the bonds on 1 January 2018 times the effective interest rate. The amount of the discount amortised in 2018 is the difference between the interest expense of £57,921 and the interest payment of £50,000 (i.e., £7,921). At 31 December 2018, the bonds' carrying amount is £973,270 (beginning balance of £965,349 plus £7,921 discount amortisation).

The following table illustrates interest expense, discount amortisation, and carrying amount (amortised cost) over the life of the bonds.

Year	Carrying Amount (beginning of year)	Interest Expense (at effective interest rate of 6%)	Interest Payment (at coupon rate of 5%)	Amortisation of Discount	Carrying Amount (end of year)
	(a)	(b)	(c)	(d)	(e)
2017	£957,876	£57,473	£50,000	£7,473	£965,349
2018	965,349	57,921	50,000	7,921	973,270
2019	973,270	58,396	50,000	8,396	981,666

Year	Carrying Amount (beginning of year)	Interest Expense (at effective interest rate of 6%)	Interest Payment (at coupon rate of 5%)	Amortisation of Discount	Carrying Amount (end of year)
2020	981,666	58,900	50,000	8,900	990,566
2021	990,566	59,434	50,000	9,434	1,000,000
Total		£292,124	£250,000	£42,124	

Solution to 3:

The carrying amounts of the bonds at 31 December 2017 and 2018 are £965,349 and £973,270, respectively. Observe that the carrying amount of the bonds issued at a discount increases over the life of the bonds. At maturity, 31 December 2021, the carrying amount of the bonds equals the face value of the bonds. The carrying amount of the bonds will be reduced to zero when the principal payment is made.

Solution to 4:

Under the straight-line method, the discount (or premium) is evenly amortised over the life of the bonds. In this example, the £42,124 discount would be amortised by £8,424.80 (£42,124 divided by 5 years) each year under the straight-line method. So, the annual interest expense under the straight-line method would be £58,424.80 (£50,000 plus £8,424.80).

The accounting and reporting for zero-coupon bonds is similar to the example above except that no interest payments are made; thus, the amount of interest expense each year is the same as the amount of the discount amortisation for the year.

EXAMPLE 4

Amortising a Bond Premium

Prembond Corp. issues £1,000,000 face value of five-year bonds, dated 1 January 2017, when the market interest rate is 4 percent. The sales proceeds are £1,044,518. The bonds pay 5 percent interest annually on 31 December.

1. What is the interest *payment* on the bonds each year?
2. What amount of interest *expense* on the bonds would be reported in 2017 and 2018 using the effective interest rate method?
3. Determine the reported value of the bonds (i.e., the carrying amount) at 31 December 2017 and 2018, assuming the effective interest rate method is used to amortise the premium.
4. What amount of interest expense on the bonds would be reported under the straight-line method of amortising the premium?

Solution to 1:

The interest payment equals £50,000 annually (£1,000,000 × 5%).

Solution to 2:

The sales proceeds of £1,044,518 are more than the face value of £1,000,000; the bonds were issued at a premium of £44,518. The bonds are initially reported as a long-term liability, bonds payable, of £1,044,518, which comprises the face value of £1,000,000 plus a premium of £44,518. The premium is amortised over time, ultimately decreasing the carrying amount (amortised cost) to face value.

Under the effective interest rate method, interest expense on the bonds is calculated as the bonds' carrying amount times the market rate in effect when the bonds are issued (effective interest rate). For 2017, interest expense is £41,781 = (£1,044,518 × 4%). The amount of the premium amortised in 2017 is the difference between the interest expense of £41,781 and the interest payment of £50,000 (i.e., £8,219). The bonds' carrying amount decreases by the premium amortisation; at 31 December 2017, the bonds' carrying amount is £1,036,299 (beginning balance of £1,044,518 less £8,219 premium amortisation). At this point, the carrying amount reflects a remaining unamortised premium of £36,299 (£44,518 premium at issuance minus £8,219 amortised).

For 2018, interest expense is £41,452 = (£1,036,299 × 4%). The amount of the premium amortised in 2011 is the difference between the interest expense of £41,452 and the interest payment of £50,000 (i.e., £8,548). At 31 December 2018, the bonds' carrying amount is £1,027,751 (beginning balance of £1,036,299 less £8,548 premium amortisation).

The following table illustrates interest expense, premium amortisation, and carrying amount (amortised cost) over the life of the bonds.

Year	Carrying Amount (beginning of year)	Interest Expense (at effective interest rate of 4%)	Interest Payment (at coupon rate of 5%)	Amortisation of Premium	Carrying Amount (end of year)
	(a)	(b)	(c)	(d)	(e)
2017	£1,044,518	£41,781	£50,000	£8,219	£1,036,299
2018	1,036,299	41,452	50,000	8,548	1,027,751
2019	1,027,751	41,110	50,000	8,890	1,018,861
2020	1,018,861	40,754	50,000	9,246	1,009,615
2021	1,009,615	40,385	50,000	9,615	1,000,000
Total				£44,518	

Solution to 3:

The carrying amounts of the bonds at 31 December 2017 and 2018 are £1,036,299 and £1,027,751, respectively. Observe that the carrying amount of the bonds issued at a premium decreases over the life of the bonds. At maturity, 31 December 2021, the carrying amount of the bonds equals the face value of the bonds. The carrying amount of the bonds will be reduced to zero when the principal payment is made.

Solution to 4:

Under the straight-line method, the premium is evenly amortised over the life of the bonds. In this example, the £44,518 premium would be amortised by £8,903.64 (£44,518 divided by 5 years) each year under the straight-line method. So, the annual interest expense under the straight-line method would be £41,096.36 (£50,000 less £8,903.64).

The reporting of interest payments on the statement of cash flows can differ under IFRS and US GAAP. Under IFRS, interest payments on bonds can be included as an outflow in either the operating section or the financing section of the statement of cash flows. US GAAP requires interest payments on bonds to be included as an operating cash outflow. (Some financial statement users consider the placement of interest payments in the operating section to be inconsistent with the placement of bond issue proceeds in the financing section of the statement of cash flows.) Typically, cash interest paid is not shown directly on the statement of cash flows, but companies are required to disclose interest paid separately.

Amortisation of a discount (premium) is a non-cash item and thus, apart from its effect on pretax income, has no effect on cash flow. In the section of the statement of cash flows that reconciles net income to operating cash flow, amortisation of a discount (premium) is added back to (subtracted from) net income.

ACCOUNTING FOR BONDS AT FAIR VALUE

4

a determine the initial recognition, initial measurement and subsequent measurement of bonds

Reporting bonds at amortised historical costs (historical cost plus or minus the cumulative amortisation) reflects the market rate at the time the bonds were *issued* (i.e., historical market rate or effective interest rate). As market interest rates change, the bonds' carrying amount diverges from the bonds' fair market value. When market interest rates decline, the fair value of a bond with a fixed coupon rate increases. As a result, a company's economic liabilities may be higher than its reported debt based on amortised historical cost. Conversely, when market interest rates increase, the fair value of a bond with a fixed coupon rate decreases and the company's economic liability may be lower than its reported debt. Using financial statement amounts based on amortised cost may underestimate (or overestimate) a company's debt-to-total-capital ratio and similar leverage ratios.

Companies have the option to report financial liabilities at fair value. Financial liabilities reported at fair value are designated as financial liabilities at fair value through profit or loss under IFRS, or, equivalently under US GAAP, as liabilities under the fair value option. Even if a company does not opt to report financial liabilities at fair value, the availability of fair value information in the financial statements has increased. IFRS and US GAAP require fair value disclosures in the financial statements unless the carrying amount approximates fair value or the fair value cannot be reliably measured.[4]

A company electing to measure a liability at fair value will report decreases in the liability's fair value as income and increases in the liability's fair value as losses. Because changes in a liability's fair value can result from changes in market rates and/or changes in the credit quality of the issuing company, accounting standards require companies to present separately the portion of the change resulting from changes in their own credit risk. Specifically, the company will report the portion of the change in value attributable to changes in their credit risk in other comprehensive income. Only the portion of the change in value not attributable to changes in their credit risk will be recognised in profit or loss.[5]

As of the end of 2018, few companies have selected the option to report financial liabilities at fair value. Those that have are primarily companies in the financial sector. Reporting standards for financial investments and derivatives already required these companies to report a significant portion of their assets at fair values. Measuring financial liabilities at other than fair value, when financial assets are measured at fair value, results in earnings volatility. This volatility is the result of using different bases of measurement for financial assets and financial liabilities. Furthermore, when a liability is related to a specific asset, using different measurement bases creates an accounting mismatch. Goldman Sachs elected to account for some financial liabilities at fair value under the fair value option. In its fiscal year 2017 10-K filing (page 136), Goldman explains this choice:

---

[4] IFRS (IAS 32, IFRS 7, IFRS 9) and US GAAP (FASB ASC 820 and 825).

[5] IFRS 9, US GAAP (FASB ASC 825 and ASU 2016-01).

> "The primary reasons for electing the fair value option are to:
>
> - Reflect economic events in earnings on a timely basis;
> - Mitigate volatility in earnings from using different measurement attributes (e.g., transfers of financial instruments owned accounted for as financings are recorded at fair value, whereas the related secured financing would be recorded on an accrual basis absent electing the fair value option); and
> - Address simplification and cost-benefit considerations (e.g., accounting for hybrid financial instruments at fair value in their entirety versus bifurcation of embedded derivatives and hedge accounting for debt hosts)."

Most companies, as required under IFRS and US GAAP, disclose the fair values of financial liabilities. The primary exception to the disclosure occurs when fair value cannot be reliably measured. Example 5 illustrates ING Group's fair value disclosures, including the fair values of long-term debt.

EXAMPLE 5

Fair Value Disclosures of Debt and Financial Instruments

ING Group 2017 Form 20-F

ING Group [Condensed] Balance Sheet as of 31 December 2017 and 2016 [Liabilities Only]

Amounts in billions of euros	2017	2016
Deposits from banks	36.8	32.0
Customer deposits	539.8	522.9
Financial liabilities at fair value through profit or loss	87.2	99.0
Other liabilities	18.9	20.1
Debt securities in issue/subordinated loans	112.1	120.4
Total liabilities	**794.8**	**794.4**

The following are excerpts from the footnotes to ING Group's financial statements.

Excerpt from Note 1 Accounting Policies

> **Financial assets and liabilities at fair value through profit or loss**
>
> … Financial liabilities at fair value through profit or loss comprise the following sub-categories: trading liabilities, non-trading derivatives, and other financial liabilities designated at fair value through profit or loss by management. Trading liabilities include equity securities, debt securities, funds on deposit, and derivatives.
>
> A financial asset or financial liability is classified at fair value through profit or loss if acquired principally for the purpose of selling in the short term or if designated by management as such.

Management will designate a financial asset or a financial liability as such only if this eliminates a measurement inconsistency or if the related assets and liabilities are managed on a fair value basis....

Financial liabilities at amortised cost

Financial liabilities at amortised cost include the following subcategories: preference shares classified as debt, debt securities in issue, subordinated loans, and deposits from banks and customer deposits.

Financial liabilities at amortised cost are recognised initially at their issue proceeds (fair value of consideration received) net of transaction costs incurred. Liabilities in this category are subsequently stated at amortised cost; any difference between proceeds, net of transaction costs, and the redemption value is recognised in the statement of profit or loss over the period of the liability using the effective interest method....

Excerpt from Note 16 Debt securities in issue

Debt securities in issue relate to debentures and other issued debt securities with either fixed interest rates or interest rates based on floating interest rate levels, such as certificates of deposit and accepted bills issued by ING Group, except for subordinated items. Debt securities in issue do not include debt securities presented as Financial liabilities at fair value through profit or loss.

Excerpt from Note 37 Fair value of assets and liabilities

Fair Value of Financial Liabilities as of 31 December 2017 and 2016

	Estimated Fair Value		Statement of Financial Position Value	
Amounts in millions of euros	**2017**	**2016**	**2017**	**2016**
Financial liabilities				
Deposits from banks	36,868	32,352	36,821	31,964
Customer deposits	540,547	523,850	539,828	522,908
Financial liabilities at fair value through profit or loss				
• trading liabilities	73,596	83,167	73,596	83,167
• non-trading derivatives	2,331	3,541	2,331	3,541
• designated as at fair value through profit or loss	11,215	12,266	11,215	12,266
Other liabilities	14,488	15,247	14,488	15,247
Debt securities in issue	96,736	103,559	96,086	103,234
Subordinated loans	16,457	17,253	15,968	17,223
	792,238	**791,235**	**790,333**	**789,550**

Use the condensed balance sheet and excerpts from the notes to ING Group's financial statements shown above to address the following questions:

1 As of 31 December 2017, what proportion of the amount of liabilities on ING Group's balance sheet is reported at fair value through profit or loss?

2 As of 31 December 2017 and 2016, what is the percent difference between the carrying amount and fair value of the debt securities that are shown on ING Group's balance sheet at amortised cost?

Solution to 1:

Of ING Group's total €794.8 billion liabilities, 11 percent (=€87.2 billion/€794.8 billion) are reported at fair value through profit or loss.

Solution to 2:

ING's debt securities that are shown on the balance sheet at amortised cost appear in the line labeled "Debt securities in issue". Note 1 states that "Debt securities in issue" are reported at amortised cost. Note 16 indicates that this line item relates to debentures and other issued debt securities, and thus we exclude subordinated loans and deposits from banks and customer deposits in this case (there are no preference shares classified as debt listed in the Note 37 excerpt).

According to the above excerpt from Note 37, in each year the fair value of ING's debt securities is slightly higher than its carrying amount. The difference is 0.7% [= (96,736/96,086) – 1] on 31 December 2017 and 0.3% [= (103,559/103,234) – 1] on 31 December 2016.

5 DERECOGNITION OF DEBT

c explain the derecognition of debt

Once bonds are issued, a company may leave the bonds outstanding until maturity or redeem the bonds before maturity either by calling the bonds (if the bond issue includes a call provision) or by purchasing the bonds in the open market. If the bonds remain outstanding until the maturity date, the company pays bondholders the face value of the bonds at maturity. The discount or premium on the bonds would be fully amortised at maturity; the carrying amount would equal face value. Upon repayment, the bonds payable account is reduced by the carrying amount at maturity (face value) of the bonds, and cash is reduced by an equal amount. Repayment of the bonds appears in the statement of cash flows as a financing cash outflow.

If a company decides to redeem bonds before maturity and thus extinguish the liability early, the bonds payable account is reduced by the carrying amount of the redeemed bonds. The difference between the cash required to redeem the bonds and the carrying amount of the bonds is a gain or loss on the extinguishment of debt. Under IFRS, debt issuance costs are included in the measurement of the liability and are thus part of its carrying amount. Under US GAAP, debt issuance costs are accounted for separately from bonds payable and are amortised over the life of the bonds. Any unamortised debt issuance costs must be written off at the time of redemption and included in the gain or loss on debt extinguishment.

For example, a company reporting under IFRS has a £10 million bond issuance with a carrying amount equal to its face value and five years remaining until maturity. The company redeems the bonds at a call price of 103. The redemption cost is

£10.3 million (= £10 million × 103%). The company's loss on redemption would be £300 thousand (£10 million carrying amount minus £10.3 million cash paid to redeem the callable bonds).

A gain or loss on the extinguishment of debt is reported on the income statement, in a separate line item, when the amount is material. A company typically discloses further detail about the extinguishment in the management discussion and analysis (MD&A) and/or notes to the financial statements.[6] In addition, in a statement of cash flows prepared using the indirect method, net income is adjusted to remove any gain or loss on the extinguishment of debt from operating cash flows and the cash paid to redeem the bonds is classified as cash used for financing activities. (Recall that the indirect method of the statement of cash flows begins with net income and makes necessary adjustments to arrive at cash from operations, including removal of gains or losses from non-operating activities.)

To illustrate the financial statement impact of the extinguishment of debt, consider the notes payable repurchase in Example 6 below.

EXAMPLE 6

Debt Extinguishment Disclosure

The following excerpts are from the 2018 annual report of Monte Rock Inc. (a hypothetical company). In its statement of cash flows, the company uses the indirect method to reconcile net income with net cash (used in) provided by operations.

***Excerpt from* Consolidated Statements of Income**
For the years ended 31 December 2018, 2017, and 2016

	2018	2017	2016
Revenues:			
⋮	⋮	⋮	⋮
Total revenues	104,908,900	112,416,800	96,879,000
⋮	⋮	⋮	⋮
Total operating expenses	100,279,900	96,140,600	71,018,900
Income from operations	4,629,000	16,276,200	25,860,100
Other income (expense):			
⋮	⋮	⋮	⋮
Gain on debt extinguishment	2,345,000	—	—
⋮	⋮	⋮	⋮
Total other income (expense), net	11,236,100	(14,257,000)	(7,085,800)
Net income	$15,865,100	$2,019,200	$18,774,300

---

[6] We use the term MD&A generally to refer to any management commentary provided on a company's financial condition, changes in financial condition, and results of operations. In the United States, the Securities and Exchange Commission (SEC) requires a management discussion and analysis for companies listed on US public markets. Reporting requirements for such a commentary as the SEC-required MD&A vary across exchanges, but some are similar to the SEC requirements. The IASB issued an IFRS Practice Statement, "Management Commentary," in December 2010 to guide all companies reporting under IFRS.

Excerpt from **Consolidated Statements of Cash Flows**
For the years ended 31 December 2018, 2017, and 2016

	2018	2017	2016
CASH FLOWS FROM OPERATING ACTIVITIES:			
Net Income	$15,865,100	$2,019,200	$18,774,300
Adjustments to reconcile net income to net cash (used in) provided by operating activities:			
⋮	⋮	⋮	⋮
Gain on debt extinguishment	(2,345,000)	—	—
⋮	⋮	⋮	⋮
Total adjustments	(16,636,000)	38,842,400	19,815,800
Net cash (used in) provided by operating activities	(770,900)	40,861,600	38,590,100
⋮	⋮	⋮	⋮
CASH FLOWS FROM FINANCING ACTIVITIES:			
Payments for debt financing costs	(294,000)	(1,526,500)	(1,481,500)
⋮	⋮	⋮	⋮
Purchase of debt securities	(2,155,000)	—	(5,000,000)
⋮	⋮	⋮	⋮
Payments of unsecured debt	—	(31,402,960)	(1,356,000)

Excerpt from **NOTE 8: BONDS PAYABLE**

On December 12, 2014, the Company issued $25 million of unsecured bonds… Interest on the bonds is equal to Libor plus 4%, payable quarterly in arrears. … During the 4th quarter of 2018, the Company repurchased the unsecured bonds with a face value of $4.5 million and realized a $2.3 million gain.

1 The balance in bonds payable was reduced at redemption by:

A $2,155,000.

B $2,345,000.

C $4,500,000.

2 How much cash did the Company pay to redeem the bonds?

A $2,155,000

B $2,345,000

C $4,500,000

Solution to 1:

C is correct. The bonds payable is reduced at redemption by the carrying amount of the bonds redeemed. The cash paid to extinguish the bonds plus the gain on redemption equals the carrying amount of the bonds. The carrying amount of the bonds was $4,500,000. In this case, the carrying amount equals the face value. The company recognised a gain of $2,345,000 when it extinguished the debt of $4,500,000 by paying only $2,155,000.

Solution to 2:

A is correct. As shown in the Statement of Cash flow, the company paid $2,155,000 to redeem the bonds. The company recognised a gain of $2,345,000 when it extinguished the debt of $4,500,000 by paying only $2,155,000.

DEBT COVENANTS

6

d describe the role of debt covenants in protecting creditors

Borrowing agreements (the bond indenture) often include restrictions called covenants that protect creditors by restricting activities of the borrower. Debt covenants benefit borrowers to the extent that they lower the risk to the creditors and thus reduce the cost of borrowing. Affirmative covenants restrict the borrower's activities by requiring certain actions. For instance, covenants may require that the borrower maintain certain ratios above a specified amount or perform regular maintenance on real assets used as collateral. Negative covenants require that the borrower not take certain actions. These covenants may restrict the borrower's ability to invest, pay dividends, or make other operating and strategic decisions that might adversely affect the company's ability to pay interest and principal.

Common covenants include limitations on how borrowed monies can be used, maintenance of collateral pledged as security (if any), restrictions on future borrowings, and requirements that limit dividends. Covenants may also specify minimum acceptable levels of financial ratios, such as debt-to-equity ratio, current ratio, or interest coverage.

When a company violates a debt covenant, it is a breach of contract. Depending on the severity of the breach and the terms of the contract, lenders may choose to waive the covenant, be entitled to a penalty payment or higher interest rate, renegotiate, or call for payment of the debt. Bond contracts typically require that the decision to call for immediate repayment be made, on behalf of all the bondholders, by holders of some minimum percentage of the principal amount of the bond issue.

Example 7 illustrates common disclosures related to debt covenants included in financial statement disclosures (notes to the financial statements).

EXAMPLE 7

The following excerpts are from the 2017 Form 20-F filing of TORM plc, a tanker company which describes itself as one of the world's largest carriers of refined oil products. TORM plc was established in 2016 following the restructuring of TORM A/S.

Illustration of Debt Covenant Disclosures

The following excerpt is from the Risk Factors section of TORM's fiscal year 2017 Form 20-F.

> Our current debt facilities impose restrictions on our financial and operational flexibility. Our debt facilities impose, and any future debt facility may impose, covenants and other operating and financial restrictions on our ability to, among other things, pay dividends, charter-in vessels, incur additional debt, sell vessels or refrain from procuring the timely release of arrested vessels. Our debt facilities require us to maintain various financial ratios, including a specified

> minimum liquidity requirement, a minimum equity requirement and a collateral maintenance requirement. Our ability to comply with these restrictions and covenants is dependent on our future performance and our ability to operate our fleet and may be affected by events beyond our control, including fluctuating vessel values. We may therefore need to seek permission from our lenders in order to engage in certain corporate actions.
>
> Failure to comply with the covenants and financial and operational restrictions under our debt facilities may lead to an event of default under those agreements. An event of default may lead to an acceleration of the repayment of debt. In addition, any default or acceleration under our existing debt facilities or agreements governing our other existing or future indebtedness is likely to lead to an acceleration of the repayment of debt under any other debt instruments that contain cross-acceleration or cross-default provisions. If all or a part of our indebtedness is accelerated, we may not be able to repay that indebtedness or borrow sufficient funds to refinance that debt, which could have a material adverse effect on our future performance, results of operations, cash flows and financial position and could lead to bankruptcy or other insolvency proceedings.
>
> … As of December 31, 2017, we were in compliance with the financial covenants contained in our debt facilities.

The following excerpt is from the Liquidity and Capital Resources section in TORM's fiscal year 2017 Form 20-F.

> The DSF [Danish Ship Finance] Facility contains, among others, the following financial and other covenants:
>
> - Loan-to-value. If at any time the aggregate market value of the vessels and the value of any additional security is less than 133% of the loan amount less amounts on credit in the deposit accounts and reserve account and the value of any additional security, the borrower and guarantors shall, within 30 days of a written request, post additional security or prepay the loan to reduce the excess to zero.
> - Free Liquidity. Minimum unencumbered cash and cash equivalents … of the higher of $75 million and 5% of our total debt, of which $40 million is required to be unencumbered cash and cash equivalents.
> - Equity Ratio. The ratio of market value adjusted shareholders' equity to total market value adjusted assets shall be at least 25%.
> - Dividends. We are restricted from making any distributions, including payment of dividends and repayments of shareholders loans, except ….

1 Which of the covenants described in the above excerpts are affirmative covenants?

2 Based on the excerpt above, what are the potential consequences of breaching the loan covenants?

Solution to 1:

Examples of affirmative covenants in the above excerpts are from TORM's disclosure about the DSF Facility and include: the requirement for TORM to maintain a loan-to-value relationship such that the assets securing the loan (the vessels) are 133% of the loan amount; the requirement for TORM to maintain "free liquidity" (i.e., a minimum level of cash and cash equivalents); and the requirement that the equity ratio be at least 25%. These covenants require the issuer to do something. The dividend covenant requiring that TORM not take certain actions (i.e., not pay dividends unless certain conditions are met) is a negative covenant. In addition, the excerpt from Risk Factors describes other negative covenants in TORM's debt facilities including restrictions on chartering-in vessels, incurring additional debt, or selling vessels.

Solution to 2:

A breach of a loan covenant by TORM—an event of default—may result in the entire amount of its debt becoming due.

PRESENTATION AND DISCLOSURE OF LONG-TERM DEBT

7

e describe the financial statement presentation of and disclosures relating to debt

The non-current (long-term) liabilities section of the balance sheet usually includes a single line item of the total amount of a company's long-term debt due after one year, with the portion of long-term debt due in the next twelve months shown as a current liability. Notes to the financial statements provide more information on the types and nature of a company's debt. These note disclosures can be used to determine the amount and timing of future cash outflows. The notes generally include stated and effective interest rates, maturity dates, restrictions imposed by creditors (covenants), and collateral pledged (if any). The amount of scheduled debt repayments for the next five years also is shown in the notes.

Example 8 contains an excerpt from the 2017 Form 10-K of Johnson & Johnson (J&J), a US manufacturer of health care products.

EXAMPLE 8

Illustration of Long-Term Debt Disclosures

Exhibit 1 is an excerpt from Note 7 of Johnson & Johnson's 2017 annual report that illustrates financial statement disclosures for long-term debt, including type and nature of long-term debt, carrying amounts, effective interest rates, and required payments over the next five years. Johnson & Johnson reports its debt at amortised cost.

Exhibit 1 Johnson & Johnson Borrowings

The components of long-term debt are as follows:

(Dollars in Millions)	2017	Effective Rate %	2016	Effective Rate %
5.55% Debentures due 2017	$—	—	$1,000	5.55
1.125% Notes due 2017	—	—	699	1.15
5.15% Debentures due 2018	900	5.18	899	5.18
1.65% Notes due 2018	597	1.70	600	1.70

(continued)

Exhibit 1 (Continued)

4.75% Notes due 2019 (1B Euro 1.1947)[2]/(1B Euro 1.0449)[3]	1,192	5.83	1,041	5.83
1.875% Notes due 2019	496	1.93	499	1.93
0.89% Notes due 2019	300	1.75	299	1.20
1.125% Notes due 2019	699	1.13	699	1.13
3% Zero Coupon Convertible Subordinated Debentures due 2020	60	3.00	84	3.00
2.95% Debentures due 2020	547	3.15	546	3.15
[PORTIONS OMITTED]				
Subtotal	**32,174**	**3.19**[1]	**24,146**	**3.33**[1]
Less current portion	1,499		1,704	
Total long-term debt	**$30,675**		**$22,442**	

(1) Weighted average effective rate.
(2) Translation rate at December 31, 2017.
(3) Translation rate at January 1, 2017.

"Fair value of the long-term debt was estimated using market prices, which were corroborated by quoted broker prices and significant other observable inputs.

"The Company has access to substantial sources of funds at numerous banks worldwide. In September 2017, the Company secured a new 364-day Credit Facility. Total credit available to the Company approximates $10 billion, which expires on September 13, 2018. Interest charged on borrowings under the credit line agreements is based on either bids provided by banks, the prime rate or London Interbank Offered Rates (Libor), plus applicable margins. Commitment fees under the agreements are not material… Throughout 2017, the Company continued to have access to liquidity through the commercial paper market. Short-term borrowings and the current portion of long-term debt amounted to approximately $3.9 billion at the end of 2017, of which $2.3 billion was borrowed under the Commercial Paper Program, $1.5 billion is the current portion of the long term debt, and the remainder principally represents local borrowing by international subsidiaries. …"

Aggregate maturities of long-term obligations commencing in 2018 are (dollars in millions):

2018	2019	2020	2021	2022	After 2022
$1,499	2,752	1,105	1,797	2,189	22,832

Use the information in Exhibit 1 to answer the following questions:

1. Why are the effective interest rates unchanged from 2016 to 2017 for most of the borrowings listed?

2 Why does the carrying amount of the "1.125% Notes due 2019" remain the same in 2016 and 2017?

3 Why is the carrying amount of the "4.75% Notes due 2019" higher in 2017 than in 2016?

Solution to 1:

The effective rate typically refers to the market rate at which the bonds are issued and typically does not change from year to year.

Solution to 2:

The carrying amount of the "1.125% Notes due 2019" remains the same because the effective interest rate at which the debentures were issued (1.13%) is approximately the same as the coupon rate indicating that the notes were issued approximately at par. Thus, there would be no amortization of a premium or discount to affect the carrying amount of the notes, and assuming no repurchases, the carrying amount would not change.

Solution to 3:

The notes are denominated in Euros, with a face value of €1 billion. The dollar/euro translation exchange rate at the end of 2017 was higher than the exchange rate at the end of 2016 (1.1947 versus 1.0449). That increase explains part of the increase in carrying value. In addition, the effective interest rate of 5.83% is higher than the 4.75% coupon rate – implying that the notes were issued at a discount. Thus, the increase in the carrying amount of the notes also reflects the amortisation of the issuance discount.

In this reading, we focus on accounting for simple debt contracts. Debt contracts can take on additional features, which lead to more complexity. For instance, convertible debt and debt with warrants are more complex instruments that have both debt and equity features. Convertible debt gives the debt holder the option to exchange the debt for equity. Bonds issued with warrants give holders the right to purchase shares of the issuer's common stock at a specific price, similar to stock options. Issuance of bonds with warrants is more common by non-US companies. Example 9 provides an example of a financial statement disclosure of bonds with warrants.

EXAMPLE 9

Financial Statement Disclosure of Bonds with Warrants

The following describes a company's issuance of convertible bonds with warrants.

> On 1 February 2018, the Company issued convertible bonds with stock warrants due 2024 with an aggregate principal amount of RMB 30 billion (the "Bonds with Warrants"). The Bonds with Warrants with fixed interest rate of 0.8% per annum and interest payable annually, were issued at par value of RMB 100. Each lot of the Bonds with Warrants, comprising ten Bonds with Warrants are entitled to warrants (the "Warrants") to subscribe 50.5 shares of the Company during the 5 trading days prior to 3 March 2020 at an initial exercise price of RMB 19.68 per share, subject to adjustment for, amongst other things, cash dividends, subdivision or consolidation of shares,

> bonus issues, rights issues, capital distribution, change of control and other events which have a dilutive effect on the issued share capital of the Company.

If all warrants were exercised, how many shares would be subscribed for?

Solution:

1,515,000,000 shares would be subscribed for [aggregate principal amount divided by par value of a lot times shares subscribed per lot = (RMB 30,000,000,000/RMB 100 × 10) × 50.5 shares].

In addition to disclosures in the notes to the financial statements, an MD&A commonly provides other information about a company's capital resources, including debt financing and off-balance-sheet financing. In the MD&A, management often provides a qualitative discussion on any material trends, favorable or unfavorable, in capital resources and indicates any expected material changes in their mix and relative cost. Additional quantitative information is typically provided, including schedules summarising a company's contractual obligations (e.g., bond payables) and other commitments (e.g., lines of credit and guarantees) in total and over the next five years.

8 LEASES

f explain motivations for leasing assets instead of purchasing them

A **lease** is a contract that conveys the right to use an asset for a period of time in exchange for consideration. The party who uses the asset and pays the consideration is the **lessee**, and the party who owns the asset, grants the right to use the asset, and receives consideration is the **lessor**. For a contract to be a lease, it must

a identify a specific underlying asset,

b give the customer the right to obtain largely all of the economic benefits from the asset over the contract term, and

c give the customer, not the supplier, the ability to direct how and for what objective the underlying asset is used.

For example, a contract between a customer and a trucking company is a lease if the contract identifies a specific truck, allows the customer exclusive use of it during the contract term, and lets the customer direct its use. However, if the customer contracts with a trucking company to ship goods for a fee, the contract would not be a lease, because a specific truck is not identified nor does the customer obtain largely all of the economic benefits from the truck over the contract term.

Leasing is a way to obtain the benefits of the asset without purchasing it outright. From the perspective of a lessee, it is a form of financing that resembles acquiring an asset with a note payable. From the perspective of a lessor, a lease is a form of investment and can also be an effective selling strategy, because customers generally prefer to pay in installments.

8.1 Examples of Leases

Leasing is among the most prevalent forms of financing. In 2014, the International Accounting Standards Board found that more than 14,000 publicly listed companies were lessees and they owed over $3.3 trillion in future lease payments in aggregate.[7] The following exhibit illustrates several examples of these arrangements.

Lessee	Lease Disclosure Excerpt
Alibaba	"The Company entered into operating lease agreements primarily for shops and malls, offices, warehouses, and land."
Copa Airlines	"The Company leases some aircraft under long-term lease agreements with an average duration of 10 years. Other leased assets include real estate, airport and terminal facilities, sales offices, maintenance facilities, and general offices."
Facebook	"We have entered into various non-cancelable operating lease agreements for certain of our offices, data center, land, colocations, and equipment."
Standard Bank	"The group leases various offices, branch space, and ATM space."

Sources: Companies' 2020 and 2019 annual reports.

Lessors are often banks, although there are some large independent leasing companies, such as AerCap Holdings N.V., which describes itself as "the global leader in aircraft leasing, . . . leasing aircraft to customers in every major geographical region. As of December 31, 2019, [it] owned 939 aircraft."[8]

8.2 Advantages of Leasing

There are several advantages to leasing an asset compared with purchasing it:

- Less cash is needed up front. Leases typically require little, if any, down payment.
- Cost effectiveness: Leases are a form of secured borrowing; in the event of nonpayment, the lessor simply repossesses the leased asset. As a result, the effective interest rate for a lease is typically lower than what the lessee would pay on an unsecured loan or bond.
- Convenience and lower risks associated with asset ownership, such as obsolescence.

From the perspective of a lessor, leasing has advantages over selling outright, which include earning interest income over the lease term and increasing the market for its product by offering customers the ability to use or control an asset while paying smaller amounts over time.

8.3 Lease Classification as Finance or Operating

Leases can resemble either the purchase of an asset or a rental contract. For example, a 10-year lease of an automobile with a 10-year useful life for monthly payments that, in aggregate, are equal to the fair value of the automobile is effectively a debt-financed purchase of that automobile. In contrast, a 1-year lease of a machine with a 20-year useful life resembles a rental contract. A lease that resembles a purchase is classified as a **finance lease**. All other leases are **operating leases**.

---

[7] IFRS, "IASB Shines Light on Leases by Bringing Them onto the Balance Sheet" (13 January 2016). www.ifrs.org/news-and-events/2016/01/iasb-shines-light-on-leases-by-bringing-them-onto-the-balance-sheet.

[8] AerCap Holdings N.V. annual report for the fiscal year ended 31 December 31 2019 on Form 20-F.

More specifically, a lease is a finance lease if *any* of the following five criteria are met. These criteria are the same for the International Financial Reporting Standards (IFRS) and US generally accepted accounting principles (GAAP). If none of the criteria are met, the lease is an operating lease. The same criteria are used for both lessees and lessors in classifying the lease.

1. The lease transfers ownership of the underlying asset to the lessee.
2. The lessee has an option to purchase the underlying asset and is reasonably certain it will do so.
3. The lease term is for a major part of the asset's useful life.
4. The present value of the sum of the lease payments equals or exceeds substantially all of the fair value of the asset.
5. The underlying asset has no alternative use to the lessor.

EXAMPLE 10

Lease Identification and Classification

1. Company C enters a contract with Company D that requires Company C to pay ¥100 million at the end of each of the next two years to Company D for exclusive use of a specific machine over that time period. The present value of the payments is ¥186 million. At the end of the contract, Company C will return the machine to Company D. The contract does not contain a purchase option. The machine can be used in many applications by many types of customers. The remaining useful life of the machine is four years, and its fair value is ¥190 million. This contract is:
   - **A** not a lease.
   - **B** an operating lease.
   - **C** a finance lease.
2. If the fair value of the machine in Question 1 was ¥300 million, would the classification of the contract change?
   - **A** No
   - **B** Yes, from an operating lease to a finance lease
   - **C** Yes, from a finance lease to an operating lease

Solution to 1:

C is correct. This contract is a lease because a specific asset is identified, Company C will exclusively use it, and Company C will have the ability to direct its use. The contract is a finance lease because one of the five criteria is met: The present value of the lease payments equals substantially all of the fair value (186/190 = 98%).

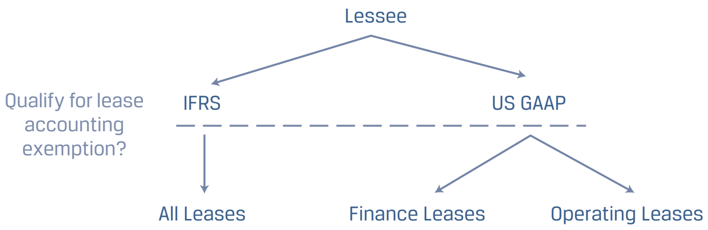
Solution to 2:

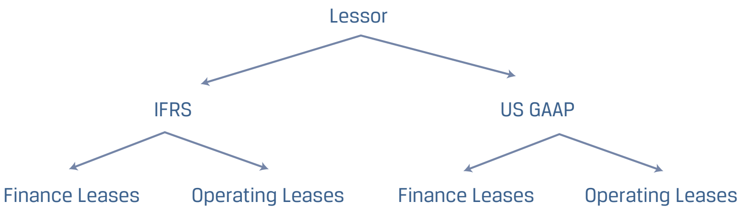
C is correct. This change would result in the lease not meeting any of the five criteria for a finance lease. If a lease does not meet any of the five criteria, it is an operating lease.

8.4 Financial Reporting of Leases

The financial reporting of a lease depends on whether the party is the lessee or lessor, whether the party reports with IFRS or US GAAP, and the classification of the lease as finance or operating. Additionally, for lessees, there are lease accounting exemptions for certain lease contracts: If its term is 12 months or less (IFRS and US GAAP) or it is for a “low-value asset,” up to $5,000 in sales price (IFRS only), then the lessee can elect to simply expense the lease payments on a straight-line basis. These exemptions are not available to lessors. The following diagram illustrates the different permutations for lease accounting.

Lease Classifications for Lessee and Lessor:





Fortunately, lessor accounting under both IFRS and US GAAP is substantially identical, and the differences in treatment for lessees are modest.

8.5 Lessee Accounting—IFRS

g explain the financial reporting of leases from a lessee’s perspective

Under IFRS, there is a single accounting model for both finance and operating leases for lessees. At lease inception, the lessee records a lease payable liability and a “right-of-use” (ROU) asset on its balance sheet, both equal to the present value of future lease payments. The discount rate used in the present value calculation is either the rate implicit in the lease or an estimated secured borrowing rate.

The lease liability is subsequently reduced by each lease payment using the effective interest method. Each lease payment is composed of interest expense, which is the product of the lease liability and the discount rate, and principal repayment, which is the difference between the interest expense and lease payment.

The ROU asset is subsequently amortized, often on a straight-line basis, over the lease term. So, although the lease liability and ROU asset begin with the same carrying value on the balance sheet, they typically diverge in subsequent periods because the principal repayment that reduces the lease liability and the amortization expense that reduces the ROU asset are calculated differently.

The following list shows how the transaction affects the financial statements:

1. The lease liability net of principal repayments and the ROU asset net of accumulated amortization are reported on the balance sheet.
2. Interest expense on the lease liability and the amortization expense related to the ROU asset are reported separately on the income statement.

3 The principal repayment component of the lease payment is reported as a cash outflow under financing activities on the statement of cash flows, and depending on the lessee's reporting policies, interest expense is reported under either operating or financing activities on the statement of cash flows.

EXAMPLE 11

Lessee Accounting—IFRS

Proton Enterprises, a hypothetical manufacturer based in Germany, is offered the following terms to lease a machine: five-year lease with an implied interest rate of 10% and an annual lease payment of EUR100,000 per year payable at the end of each year. The present value of the machinery is therefore EUR379,079 (in Microsoft Excel, the formula is =PV(10%,5,-100,000). The asset will be amortized over the five-year lease term on a straight-line basis. Proton reports under IFRS.

What would be the impact of this lease on Proton's:

1 balance sheet at the beginning of the year?

2 income statement during the following year?

3 statement of cash flows during the following year?

Solution to 1:

Proton would report a EUR379,079 lease liability and ROU asset.

Solution to 2:

Interest expense and amortization expense are reported on the income statement. In Year 2, interest expense is EUR31,699 and amortization expense is EUR 75,816, as illustrated in the following tables:

	Lease Payment	Interest Expense (10% × Lease Liability)	Principal Repayment (Payment – Interest)	Lease Liability
	FO.1	FO.2	FO.3	FO.4
Year 0				379,079
Year 1	100,000	37,908	62,092	316,987
Year 2	100,000	31,699	68,301	248,685
Year 3	100,000	24,869	75,131	173,554
Year 4	100,000	17,355	82,645	90,909
Year 5	100,000	9,091	90,909	0
Total	500,000	120,921	379,079	

	Amortization Expense	ROU Asset
	Straight-Line F.1	F.2
Year 0		379,079
Year 1	75,816	303,263
Year 2	75,816	227,447
Year 3	75,816	151,631
Year 4	75,816	75,816

	Amortization Expense	ROU Asset
Year 5	75,816	0
Total	379,079	

Note: Totals may not sum due to rounding.

Solution to 3:

Principal repayments are reported as a cash outflow under financing activities on the statement of cash flows, and depending on Proton's reporting policies, interest expense is reported under operating or financing activities on the statement of cash flows. From the previous tables, Year 2 principal repayment is EUR68,301 and interest expense is EUR31,699, for a total of EUR100,000.

8.6 Lessee Accounting—US GAAP

Under US GAAP, there are two accounting models for lessees: one for finance leases and another for operating leases. The finance lease accounting model is identical to the lessee accounting model for IFRS. The operating lease accounting model is different.

At operating lease inception, the lessee records a lease payable liability and a corresponding right-of-use asset on its balance sheet that are subsequently reduced by the principal repayment component of the lease payment and amortization, respectively, in the same manner that an IFRS lessee would.

The key difference between an operating lease and a finance lease is how the amortization of the ROU asset is calculated. For an operating lease, the lessee's ROU asset amortization expense is the lease payment minus the interest expense. The implication is that the total expense reported on the income statement (interest plus amortization) will equal the lease payment and that the lease liability and the ROU asset will always equal each other because the principal repayment and amortization are calculated in an identical manner.

The following list shows how the transaction appears on the financial statements:

a The lease liability net of principal repayments and the ROU asset net of accumulated amortization are reported on the balance sheet.

b Interest expense on the lease liability and the amortization expense related to the ROU asset are reported as a single line titled "lease expense" as an operating expense on the income statement. The interest and amortization components are *not* reported separately, nor are they grouped with other types of interest and amortization expense (e.g., interest on a bond, amortization of an intangible asset).

c The entire lease payment is reported as a cash outflow under operating activities on the statement of cash flows. The interest and principal repayment components are *not* reported separately.

EXAMPLE 12

Lessee Accounting—Operating Lease under US GAAP

If Proton Enterprises classified the lease of the machinery from Example 11 as an operating lease:

1 how would its financial statements differ, if at all?

2 how would the classification, all else equal, affect EBITDA margin, total asset turnover, and cash flow per share?

Solution to 1:

The first step is to construct the lease liability and ROU asset amortization tables under an operating lease scenario. The lease liability amortization is the same as the finance lease columns FO.1–FO.4 in Example 11.

	Amortization Expense	ROU Asset	Lease Expense
	(Lease Payment – Interest)		(Amortization + Interest)
	0.1	0.2	0.3
Year 0		379,079	
Year 1	62,092	316,987	100,000
Year 2	68,301	248,685	100,000
Year 3	75,131	173,554	100,000
Year 4	82,645	90,909	100,000
Year 5	90,909	0	100,000
Total	379,078		500,000

Now we can compare the financial statement impacts under both finance and operating lease scenarios.

Balance Sheet	Year 1	Year 2	Year 3	Year 4	Year 5
Finance lease:					
ROU asset, net: F.2	303,263	227,447	151,631	75,816	0
Lease liability, net: FO.4	316,987	248,685	173,554	90,909	0
Operating lease:					
ROU asset, net: O.2	316,987	248,685	173,554	90,909	0
Lease liability, net: FO.4	316,987	248,685	173,554	90,909	0

The ROU asset is lower in each period under a finance lease because the amortization expense is higher.

Income Statement	Year 1	Year 2	Year 3	Year 4	Year 5
Finance lease:					
Amortization: F.1	75,816	75,816	75,816	75,816	75,816
Interest: FO.2	37,908	31,699	24,869	17,355	9,091
Total	113,724	107,515	100,685	93,171	84,907

Income Statement	Year 1	Year 2	Year 3	Year 4	Year 5
Operating lease:					
Lease expense: O.3	100,000	100,000	100,000	100,000	100,000

Total expense is higher for a finance lease in Years 1–3 but lower in Years 4 and 5. The largest difference is classification; amortization and interest are presented separately for a finance lease, whereas operating lease expense is an operating expense.

Statement of Cash Flows	Year 1	Year 2	Year 3	Year 4	Year 5
Finance lease:					
Cash flow from operating activities	(37,908)	(31,699)	(24,869)	(17,355)	(9,091)
Cash flow from financing activities	(62,902)	(68,301)	(75,131)	(82,645)	(90,909)
Total	(100,000)	(100,000)	(100,000)	(100,000)	(100,000)
Operating lease:					
Cash flows from operating activities	(100,000)	(100,000)	(100,000)	(100,000)	(100,000)

The difference on the statement of cash flows is only in classification, because in both cases the total cash outflow is equal to the lease payment.

Solution to 2:

The following table shows how the classification affects the indicated financial ratios.

Ratio	Formula	Impact of Using an Operating Lease Instead of a Finance Lease
EBITDA margin	$\frac{\text{EBITDA}}{\text{Total revenues}}$	Lower: Lease expense is classified as an operating expense rather than interest and amortization.
Asset turnover	$\frac{\text{Total revenues}}{\text{Total assets}}$	Lower: Total assets are higher under an operating lease because the ROU asset is amortized at a slower pace in Years 1–3.
Cash flow per share	$\frac{\text{Cash flow from operations}}{\text{Shares outstanding}}$	Lower: Cash flow from operations is lower because the entire lease payment is included in operating activities versus solely interest expense for a finance lease.

8.7 Lessor Accounting

h explain the financial reporting of leases from a lessor's perspective

The accounting for lessors is substantially identical under IFRS and US GAAP. Under both accounting standards, lessors classify leases as finance or operating leases, which determines the financial reporting. Although lessors under US GAAP recognize finance leases as either "sales-type" or "direct financing," the distinction is immaterial from an analyst's perspective.

At finance lease inception, the lessor recognizes a lease receivable asset equal to the present value of future lease payments and de-recognizes the leased asset, simultaneously recognizing any difference as a gain or loss. The discount rate used in the present value calculation is the rate implicit in the lease.

The lease receivable is subsequently reduced by each lease payment using the effective interest method. Each lease payment is composed of interest income, which is the product of the lease receivable and the discount rate, and principal proceeds, which equals the difference between the interest income and cash receipt.

The following list shows how the transaction affects the financial statements:

1 Lease receivable net of principal proceeds is reported on the balance sheet.
2 Interest income is reported on the income statement. If leasing is a primary business activity for the entity, as it commonly is for financial institutions and independent leasing companies, it is reported as revenue.
3 The entire cash receipt is reported under operating activities on the statement of cash flows.

The accounting treatment for an operating lease is different in that, because the contract is essentially a rental agreement, the lessor keeps the leased asset on its books and recognizes lease revenue on a straight-line basis. Interest revenue is not recognized because the transaction is not considered a financing.

The following list shows how the transaction affects the financial statements:

1 The balance sheet is not affected. The lessor continues to recognize the leased asset at cost net of accumulated depreciation.
2 Lease revenue is recognized on a straight-line basis on the income statement. Depreciation expense continues to be recognized.
3 The entire cash receipt is reported under operating activities on the statement of cash flows. This is the same as a finance lease.

EXAMPLE 13

Lessor Accounting

Let's examine Proton's machine lease from Examples 11 and 12 from the perspective of the lessor. Assume that the carrying value of the asset immediately prior to the lease is EUR350,000, accumulated depreciation is zero, and the lessor elects to depreciate it on a straight-line basis over five years.

How are the lessor's financial statements affected by the classification of the lease as a finance or operating lease?

Solution

The difference on the balance sheet is material, because a finance lease requires the lessor to de-recognize the asset and recognize a lease receivable, whereas an operating lease lessor continues to recognize the asset and depreciate it over its useful life. In this case, where the present value of the lease payments is well above the carrying value of the asset, the finance lease classification results in a significant increase in assets.

Balance Sheet	Year 1	Year 2	Year 3	Year 4	Year 5
Finance lease:					
Lease receivable, net	316,987	248,685	173,554	90,909	0
Operating lease:					
Property, plant, and equipment, net	280,000	210,000	140,000	70,000	0

The difference on the income statement is also material, because a finance lease lessor recognizes interest revenue under the effective interest method whereas the operating lease lessor recognizes straight-line lease revenue.

Income Statement	Year 1	Year 2	Year 3	Year 4	Year 5
Finance lease:					
Interest revenue	37,908	31,699	24,869	17,355	9,091
Operating lease:					
Lease revenue	100,000	100,000	100,000	100,000	100,000

The statement of cash flows, however, is no different for the lessor under a finance or operating lease: The entire cash inflow from the lease payment is recognized under operating activities.

Statement of Cash Flows	Year 1	Year 2	Year 3	Year 4	Year 5
Finance lease:					
Cash flows from operating activities	100,000	100,000	100,000	100,000	100,000
Operating lease:					
Cash flows from operating activities	100,000	100,000	100,000	100,000	100,000

INTRODUCTION TO PENSIONS AND OTHER POST-EMPLOYMENT BENEFITS

9

i. compare the presentation and disclosure of defined contribution and defined benefit pension plans

Pensions and other post-employment benefits give rise to non-current liabilities reported by many companies. Companies may offer various types of benefits to their employees following retirement, such as pension plans, health care plans, medical insurance, and life insurance. Pension plans often are the most significant post-employment benefits provided to retired employees.

The accounting and reporting for pension plans depends on the type of pension plan offered. Two common types of pension plans are **defined contribution pension plans** and **defined benefit pension plans**. Under a defined-contribution plan, a company contributes an agreed-upon (defined) amount into the plan. The agreed-upon amount is the pension expense. The amount the company contributes to the plan is treated as an operating cash outflow. The only impact on assets and liabilities is a decrease in cash, although if some portion of the agreed-upon amount has not been paid by fiscal year-end, a liability would be recognised on the balance sheet. Because the amount of the contribution is defined and the company has no further obligation once the contribution has been made, accounting for a defined-contribution plan is fairly straightforward.

Accounting for a defined-benefit plan is more complicated. Under a defined-benefit plan, a company makes promises of future benefits to be paid to the employee during retirement. For example, a company could promise an employee annual pension payments equal to 70 percent of his final salary at retirement until death. Estimating the eventual amount of the obligation arising from that promise requires the company to make many assumptions, such as the employee's expected salary at retirement and the number of years the employee is expected to live beyond retirement. The company estimates the future amounts to be paid and discounts the future estimated amounts to a present value (using a rate reflective of a high-quality corporate bond yield) to determine the pension obligation. The discount rate used to determine the pension obligation significantly affects the amount of the pension obligation. The pension obligation is allocated over the employee's employment as part of pension expense.

Most defined-benefit pension plans are funded through a separate legal entity, typically a pension trust fund. A company makes payments into the pension fund, and retirees are paid from the fund. The payments that a company makes into the fund are invested until they are needed to pay the retirees. If the fair value of the fund's assets is higher than the present value of the estimated pension obligation, the plan has a surplus and the company's balance sheet will reflect a net pension asset.[9] Conversely, if the present value of the estimated pension obligation exceeds the fund's assets, the plan has a deficit and the company's balance sheet will reflect a net pension liability.[10] Thus, a company reports either a net pension asset or a net pension liability. Each period, the change in the net pension asset or liability is recognised either in profit or loss or in other comprehensive income.

Under IFRS, the change in the net pension asset or liability each period is viewed as having three general components. Two of the components of this change are recognised as pension expense in profit and loss: (1) employees' service costs, and (2) the net interest expense or income accrued on the beginning net pension asset or liability. The service cost during the period for an employee is the present value of the increase in the pension benefit earned by the employee as a result of providing one more year of service. The service cost also includes past service costs, which are changes in the present value of the estimated pension obligation related to employees' service in prior periods, such as might arise from changes in the plan. The net interest expense or income represents the change in value of the net defined benefit pension asset or liability and is calculated as the net pension asset or liability multiplied by the discount rate used in estimating the present value of the pension obligation. The third component of the change in the net pension asset or liability during a period —"remeasurements"— is recognised in other comprehensive income. Remeasurements are not amortised into profit or loss over time.

Remeasurements include (a) actuarial gains and losses and (b) the actual return on plan assets less any return included in the net interest expense or income. Actuarial gains and losses can occur when changes are made to the assumptions on which a company bases its estimated pension obligation (e.g., employee turnover, mortality rates, retirement ages, compensation increases). The actual return on plan assets includes interest, dividends and other income derived from the plan assets, including realized and unrealized gains or losses. The actual return typically differs from the amount included in the net interest expense or income, which is calculated using a rate reflective of a high-quality corporate bond yield; plan assets are typically allocated across various asset classes, including equity as well as bonds.

---

[9] The amount of any reported net pension asset is capped at the amount of any expected future economic benefits to the company from the plan; this cap is referred to as the asset ceiling.

[10] The description of accounting for pensions presented in this reading corresponds to the June 2011 version of IAS 19 *Employee Benefits*, which took effect on 1 January 2013. Both IFRS and US GAAP require companies to present the amount of net pension liability or asset on the balance sheet.

Under US GAAP, the change in net pension asset or liability each period is viewed as having five components, some of which are recognised in profit and loss in the period incurred and some of which are recognised in other comprehensive income and amortised into profit and loss over time. The three components recognised in profit and loss in the period incurred are (1) employees' service costs for the period, (2) interest expense accrued on the beginning pension obligation, and (3) expected return on plan assets, which is a reduction in the amount of expense recognised. The other two components are past service costs and actuarial gains and losses. Past service costs are recognised in other comprehensive income in the period in which they arise and then subsequently amortised into pension expense over the future service period of the employees covered by the plan. Actuarial gains and losses are typically also recognised in other comprehensive income in the period in which they occur and then amortised into pension expense over time. In effect, this treatment allows companies to "smooth" the effects on pension expense over time for these latter two components. US GAAP does permit companies to immediately recognize actuarial gains and losses in profit and loss.

Similar to other forms of employee compensation for a manufacturing company, the pension expense related to production employees is added to inventory and expensed through cost of sales (cost of goods sold). For employees not involved directly in the production process, the pension expense is included with salaries and other administrative expenses. Therefore, pension expense is not directly reported on the income statement. Rather, extensive disclosures are included in the notes to the financial statements.

Example 14 presents excerpts from the balance sheet and pension-related disclosures in BT Group plc's Annual Report for the year ended 31 March 2018.

EXAMPLE 14

BT Group plc: Excerpt from Balance Sheets

Below is an excerpt of BT Group plc's balance sheet from the annual report for the year ended 31 March 2018. BT reports under IFRS.

Non-current liabilities, GBP million	Mar. 31, 2018	Mar. 31, 2017	Mar. 31, 2016
Loans and other borrowings	11,994	10,081	11,025
Derivative financial instruments	787	869	863
Retirement benefit obligations	6,371	9,088	6,382
Other payables	1,326	1,298	1,106
Deferred tax liabilities	1,340	1,240	1,262
Provisions	452	536	565
Non-current liabilities	22,270	23,112	21,203

Pension-Related Disclosures

The following are excerpts of pension-related disclosures from BT Group plc's 2018 Annual Report.

Extract from Note 3 "Summary of Significant Accounting Policies"

Retirement benefits

The group's net obligation in respect of defined benefit pension plans is the present value of the defined benefit obligation less the fair value of the plan assets.

The calculation of the obligation is performed by a qualified actuary using the projected unit credit method and key actuarial assumptions at the balance sheet date.

The income statement expense is allocated between an operating charge and net finance income or expense. The operating charge reflects the increase in the defined benefit obligation resulting from the pension benefit earned by active employees in the current period, the costs of administering the plans and any past service costs/credits such as those arising from curtailments or settlements. The net finance income or expense reflects the interest on the net retirement benefit obligations recognised in the group balance sheet, based on the discount rate at the start of the year. Actuarial gains and losses are recognised in full in the period in which they occur and are presented in the group statement of comprehensive income.

The group also operates defined contribution pension plans and the income statement expense represents the contributions payable for the year.

Extract from Note 20 "Retirement Benefit Plans"
Information on Defined Benefit Pension Plans

£m	2018	2017	2016
Present value of liabilities	**57,327**	60,200	50,350
Fair value of plan assets	**50,956**	51,112	43,968

Use information in the excerpts to answer the following questions:

1. What type(s) of pension plans does BT have?
2. What proportion of BT's total non-current liabilities are related to its retirement benefit obligations?
3. Describe how BT's retirement benefit obligation is calculated.

Solution to 1:

Note 3 "Summary of Significant Accounting Policies" indicates that the company has both defined contribution and defined benefit pension plans.

Solution to 2:

Retirement benefit obligations represent 29%, 39%, and 30% of BT's total non-current liabilities for the years 2018, 2017, and 2016. Using 2018 to illustrate, £6,371/£22,270 = 29%. (£ million)

Solution to 3:

Note 3 "Summary of Significant Accounting Policies" indicates that BT's Retirement benefit obligation is calculated as the present value of the defined benefit obligation minus the fair value of the plan assets.

Using data from Note 20 "Retirement Benefit Plans" the retirement benefit obligation for each year can be calculated. Using 2018 to illustrate, £57,327 - £50,956 = £6,371 (£ million).

EVALUATING SOLVENCY: LEVERAGE AND COVERAGE RATIOS

10

j calculate and interpret leverage and coverage ratios

Solvency refers to a company's ability to meet its long-term debt obligations, including both principal and interest payments. In evaluating a company's solvency, ratio analyses can provide information about the relative amount of debt in the company's capital structure and the adequacy of earnings and cash flow to cover interest expense and other fixed charges (such as lease or rental payments) as they come due. Ratios are useful to evaluate a company's performance over time compared to the performance of other companies and industry norms. Ratio analysis has the advantage of allowing the comparison of companies regardless of their size and reporting currency.

The two primary types of solvency ratios are leverage ratios and coverage ratios. Leverage ratios focus on the balance sheet and measure the extent to which a company uses liabilities rather than equity to finance its assets. Coverage ratios focus on the income statement and cash flows and measure the ability of a company to cover its debt-related payments.

Exhibit 2 describes the two types of commonly used solvency ratios. The first three leverage ratios use total debt in the numerator.[11] The *debt-to-assets ratio* expresses the percentage of total assets financed with debt. Generally, the higher the ratio, the higher the financial risk and thus the weaker the solvency. The *debt-to-capital ratio* measures the percentage of a company's total capital (debt plus equity) financed through debt. The *debt-to-equity ratio* measures the amount of debt financing relative to equity financing. A *debt-to-equity ratio* of 1.0 indicates equal amounts of debt and equity, which is the same as a debt-to-capital ratio of 50 percent. Interpretations of these ratios are similar. Higher debt-to-capital or debt-to-equity ratios imply weaker solvency. A caveat must be made when comparing debt ratios of companies in different countries. Within certain countries, companies historically have obtained more capital from debt than equity financing, so debt ratios tend to be higher for companies in these countries.

Exhibit 2 Definitions of Commonly Used Solvency Ratios

Solvency Ratios	Numerator	Denominator
Leverage ratios		
Debt-to-assets ratio	Total debt[a]	Total assets
Debt-to-capital ratio	Total debt[a]	Total debt[a] + Total shareholders' equity

(continued)

---

[11] For calculations in this reading, total debt is the sum of interest-bearing short-term and long-term debt, excluding non-interest-bearing liabilities, such as accrued expenses, accounts payable, and deferred income taxes. This definition of total debt differs from other definitions that are more inclusive (e.g., all liabilities) or more restrictive (e.g., long-term debt only). If the use of different definitions of total debt materially changes conclusions about a company's solvency, the reasons for the discrepancies should be further investigated.

Exhibit 2 (Continued)

Solvency Ratios	Numerator	Denominator
Leverage ratios		
Debt-to-equity ratio	Total debt[a]	Total shareholders' equity
Financial leverage ratio	Average total assets	Average shareholders' equity
Coverage ratios		
Interest coverage ratio	EBIT[b]	Interest payments
Fixed charge coverage ratio	EBIT[b] + lease payments	Interest payments + lease payments

[a] In this reading, debt is defined as the sum of interest-bearing short-term and long-term debt.
[b] EBIT is earnings before interest and taxes.

The *financial leverage ratio* (also called the 'leverage ratio' or 'equity multiplier') measures the amount of total assets supported by one money unit of equity. For example, a value of 4 for this ratio means that each €1 of equity supports €4 of total assets. The higher the financial leverage ratio, the more leveraged the company in the sense of using debt and other liabilities to finance assets. This ratio often is defined in terms of average total assets and average total equity and plays an important role in the DuPont decomposition of return on equity.[12]

The *interest coverage ratio* measures the number of times a company's EBIT could cover its interest payments. A higher interest coverage ratio indicates stronger solvency, offering greater assurance that the company can service its debt from operating earnings. The *fixed charge coverage ratio* relates fixed financing charges, or obligations, to the cash flow generated by the company. It measures the number of times a company's earnings (before interest, taxes, and lease payments) can cover the company's interest and lease payments.

Example 15 demonstrates the use of solvency ratios in evaluating the creditworthiness of a company.

EXAMPLE 15

Evaluating Solvency Ratios

A credit analyst is evaluating and comparing the solvency of two companies—BT Group plc (BT) and Telefonica S A (Telefonica). The following data are gathered from the companies' fiscal 2017 annual reports (line item titles may vary between the two companies):

	BT Group plc (£ millions)		Telefonica S A (€ millions)	
	31-Mar-18	31-Mar-17	31-Dec-17	31-Dec-16
Short-term borrowings	2,281	2,632	9,414	14,749
Long-term debt	11,994	10,081	46,332	45,612
Total shareholders' equity	10,304	8,335	26,618	28,385

---

[12] The basic DuPont decomposition is: Return on Equity = Net income/Average shareholders' equity = (Sales/Average total assets) × (Net income/Sales) × (Average total assets/Average shareholders' equity).

	BT Group plc (£ millions)		Telefonica S A (€ millions)	
	31-Mar-18	31-Mar-17	31-Dec-17	31-Dec-16
Total assets	42,759	42,372	115,066	123,641
EBIT*	3,381	3,167	6,791	5,469
Interest expense	776	817	3,363	4,476

* Operating profit (or operating income) is used as a proxy for EBIT for both companies.

Use the above information to answer the following questions:

1 With regard to leverage ratios of BT and Telefonica:

- **A** What are each company's debt-to-assets, debt-to-capital, and debt-to-equity ratios for 2017 and 2016?
- **B** Comment on any changes in the calculated leverage ratios from year-to-year for each company.
- **C** Comment on the calculated leverage ratios of BT Group plc compared to Telefonica SA.

2 With regard to coverage ratios of BT and Telefonica:

- **A** What is each company's interest coverage ratio for 2017 and 2016?
- **B** Comment on any changes in the interest coverage ratio from year to year for each company.
- **C** Comment on the interest coverage ratio of BT Group plc compared to Telefonica SA.

Solution to 1:

A The debt-to-assets, debt-to-capital, and debt-to-equity ratios are as follows, with supporting calculations from each company's most recent year demonstrated below.

	BT Group plc		Telefonica S A	
	31-Mar-18	31-Mar-17	31-Dec-17	31-Dec-16
Debt-to-assets	33.4%	30.0%	48.4%	48.8%
Debt-to-capital	58.1%	60.4%	67.7%	68.0%
Debt-to-equity	1.39	1.53	2.09	2.13

	BT Group plc	Telefonica S A
	31-Mar-18	*31-Dec-17*
Debt-to-assets	*33.4% = (2,281 + 11,994)/42,759*	*48.4% = (9,414 + 46,332)/115,066*
Debt-to-capital	*58.1% = (2,281 + 11,994)/(2,281 + 11,994 + 10,304)*	*67.7% = (9,414 + 46,332)/(9,414 + 46,332 + 26,618)*
Debt-to-equity	*1.39 = (2,281 + 11,994)/10,304*	*2.09 = (9,414 + 46,332)/26,618*

B BT's debt-to-assets ratio increased, while its debt-to-capital and debt-to-equity ratios both decreased. The decrease in BT's debt-to-capital and debt-to-equity ratios resulted primarily from the company's increase in total equity and indicate stronger solvency. In addition, we observe that BT decreased its short-term borrowings and increased its long-term debt.

Telefonica's leverage ratios appear fairly similar, albeit slightly lower, for 2017 compared to 2016. Similar to BT, it appears that Telefonica shifted away from short borrowings to long-term debt in 2017.

C In both years, all three of BT's leverage ratios were lower than Telefonica's. Based on these ratios, this may imply higher solvency of BT relative to Telefonica.

Solution to 2:

A The interest coverage ratios are as follows, with supporting calculations from each company's most recent year demonstrated below.

	BT Group plc		Telefonica S A	
	31-Mar-18	**31-Mar-17**	**31-Dec-17**	**31-Dec-16**
Interest coverage ratio	4.36	3.88	2.02	1.22

	BT Group plc	**Telefonica S A**
	31-Mar-18	***31-Dec-17***
Interest coverage ratio	*4.36 = 3,381/776*	*2.02 = 6,791/3,363*

B Both companies' interest coverage ratios increased from 2017 to 2018, indicating an improvement in solvency, consistent with the conclusions drawn from the companies' ratios in question 1. Both companies have sufficient operating earnings to cover interest payments.

C BT's ability to cover interest payments is greater than Telefonica's, although both companies have sufficient operating earnings to service its interest payments. This comparison indicates that BT has greater financial strength than Telefonica, which is also consistent with the conclusions drawn from a comparison of the companies' ratios in question 1.

SUMMARY

Non-current liabilities arise from different sources of financing and different types of creditors. Bonds are a common source of financing from debt markets. Key points in accounting and reporting of non-current liabilities include the following:

- The sales proceeds of a bond issue are determined by discounting future cash payments using the market rate of interest at the time of issuance (effective interest rate). The reported interest expense on bonds is based on the effective interest rate.
- Future cash payments on bonds usually include periodic interest payments (made at the stated interest rate or coupon rate) and the principal amount at maturity.

- When the market rate of interest equals the coupon rate for the bonds, the bonds will sell at par (i.e., at a price equal to the face value). When the market rate of interest is higher than the bonds' coupon rate, the bonds will sell at a discount. When the market rate of interest is lower than the bonds' coupon rate, the bonds will sell at a premium.
- An issuer amortises any issuance discount or premium on bonds over the life of the bonds.
- If a company redeems bonds before maturity, it reports a gain or loss on debt extinguishment computed as the net carrying amount of the bonds (including bond issuance costs under IFRS) less the amount required to redeem the bonds.
- Debt covenants impose restrictions on borrowers, such as limitations on future borrowing or requirements to maintain a minimum debt-to-equity ratio.
- The carrying amount of bonds is typically the amortised historical cost, which can differ from their fair value.
- Companies are required to disclose the fair value of financial liabilities, including debt. Although permitted to do so, few companies opt to report debt at fair values on the balance sheet.
- A lease is a contract in which a lessor grants the lessee the exclusive right to use a specific underlying asset for a period of time in exchange for payments.
- Leasing is a common arrangement because it has several advantages over purchasing an asset outright: less upfront cash commitment, generally low interest rates, and lower risks associated with ownership such as obsolescence.
- Leases are classified as operating or finance leases. Finance leases resemble an asset purchase or sale while operating leases resemble a rental agreement.
- US GAAP and IFRS share the same accounting treatment for lessors but differ for lessees. IFRS has a single accounting model for both operating leases and finance lease lessees, while US GAAP has an accounting model for each.
- Lessees reporting under IFRS and finance lease lessees reporting under US GAAP
  - Recognize a lease liability and corresponding right-of-use asset on the balance sheet, equal to the present value of lease payments. The liability is subsequently reduced using the effective interest method and the right-of-use asset is amortized. Interest expense and amortization expense are shown separately on the income statement. The statement of cash flows shows the entire lease payment.
- Operating lease lessees reporting under US GAAP
  - Recognize a lease liability and corresponding right-of-use asset on the balance sheet, equal to the present value of lease payments. The liability is subsequently reduced using the effective interest method, but the amortization of the right-of-use asset is the lease payment less the interest expense. Interest expense and amortization expense are shown together as a single operating expense on the income statement.
- Finance lease lessors (IFRS and US GAAP)
  - Recognize a lease receivable asset equal to the present value of future lease payments and de-recognize the leased asset, simultaneously recognizing any difference as a gain or loss. The lease receivable is subsequently reduced by each lease payment using the effective interest method. Interest income is reported on the income statement, typically as revenue, and the entire cash receipt is reported under operating activities on the statement of cash flows.
- Operating lease lessors (IFRS and US GAAP)

  - The balance sheet is not affected: the lessor continues to recognize the underlying asset and depreciate it. Lease revenue is recognized on a straight-line basis on the income statement and the entire cash receipt is reported under operating activities on the statement of cash flows.
- Two types of pension plans are defined contribution plans and defined benefits plans. In a defined contribution plan, the amount of contribution into the plan is specified (i.e., defined) and the amount of pension that is ultimately paid by the plan (received by the retiree) depends on the performance of the plan's assets. In a defined benefit plan, the amount of pension that is ultimately paid by the plan (received by the retiree) is defined, usually according to a benefit formula.
- Under a defined contribution pension plan, the cash payment made into the plan is recognised as pension expense.
- Under both IFRS and US GAAP, companies must report the difference between the defined benefit pension obligation and the pension assets as an asset or liability on the balance sheet. An underfunded defined benefit pension plan is shown as a non-current liability.
- Under IFRS, the change in the defined benefit plan net asset or liability is recognised as a cost of the period, with two components of the change (service cost and net interest expense or income) recognised in profit and loss and one component (remeasurements) of the change recognised in other comprehensive income.
- Under US GAAP, the change in the defined benefit plan net asset or liability is also recognised as a cost of the period with three components of the change (current service costs, interest expense on the beginning pension obligation, and expected return on plan assets) recognised in profit and loss and two components (past service costs and actuarial gains and losses) typically recognised in other comprehensive income.
- Solvency refers to a company's ability to meet its long-term debt obligations.
- In evaluating solvency, leverage ratios focus on the balance sheet and measure the amount of debt financing relative to equity financing.
- In evaluating solvency, coverage ratios focus on the income statement and cash flows and measure the ability of a company to cover its interest payments.

PRACTICE PROBLEMS

1 A company issues €1 million of bonds at face value. When the bonds are issued, the company will record a:

A cash inflow from investing activities.

B cash inflow from financing activities.

C cash inflow from operating activities.

2 At the time of issue of 4.50% coupon bonds, the effective interest rate was 5.00%. The bonds were *most likely* issued at:

A par.

B a discount.

C a premium.

3 Oil Exploration LLC paid $45,000 in printing, legal fees, commissions, and other costs associated with its recent bond issue. It is *most likely* to record these costs on its financial statements as:

A an asset under US GAAP and reduction of the carrying value of the debt under IFRS.

B a liability under US GAAP and reduction of the carrying value of the debt under IFRS.

C a cash outflow from investing activities under both US GAAP and IFRS.

4 A company issues $1,000,000 face value of 10-year bonds on 1 January 2015 when the market interest rate on bonds of comparable risk and terms is 5%. The bonds pay 6% interest annually on 31 December. At the time of issue, the bonds payable reflected on the balance sheet is *closest* to:

A $926,399.

B $1,000,000.

C $1,077,217.

5 Midland Brands issues three-year bonds dated 1 January 2015 with a face value of $5,000,000. The market interest rate on bonds of comparable risk and term is 3%. If the bonds pay 2.5% annually on 31 December, bonds payable when issued are most likely reported as *closest* to:

A $4,929,285.

B $5,000,000.

C $5,071,401.

6 A firm issues a bond with a coupon rate of 5.00% when the market interest rate is 5.50% on bonds of comparable risk and terms. One year later, the market interest rate increases to 6.00%. Based on this information, the effective interest rate is:

A 5.00%.

B 5.50%.

C 6.00%.

7 On 1 January 2010, Elegant Fragrances Company issues £1,000,000 face value, five-year bonds with annual interest payments of £55,000 to be paid each 31 December. The market interest rate is 6.0 percent. Using the effective interest rate method of amortisation, Elegant Fragrances is *most likely* to record:

© 2021 CFA Institute. All rights reserved.

A an interest expense of £55,000 on its 2010 income statement.

B a liability of £982,674 on the 31 December 2010 balance sheet.

C a £58,736 cash outflow from operating activity on the 2010 statement of cash flows.

8 Consolidated Enterprises issues €10 million face value, five-year bonds with a coupon rate of 6.5 percent. At the time of issuance, the market interest rate is 6.0 percent. Using the effective interest rate method of amortisation, the carrying value after one year will be *closest* to:

A €10.17 million.

B €10.21 million.

C €10.28 million.

9 A company issues €10,000,000 face value of 10-year bonds dated 1 January 2015 when the market interest rate on bonds of comparable risk and terms is 6%. The bonds pay 7% interest annually on 31 December. Based on the effective interest rate method, the interest expense on 31 December 2015 is *closest* to:

A €644,161.

B €700,000.

C €751,521.

10 A company issues $30,000,000 face value of five-year bonds dated 1 January 2015 when the market interest rate on bonds of comparable risk and terms is 5%. The bonds pay 4% interest annually on 31 December. Based on the effective interest rate method, the carrying amount of the bonds on 31 December 2015 is *closest* to:

A $28,466,099.

B $28,800,000.

C $28,936,215.

11 Lesp Industries issues five-year bonds dated 1 January 2015 with a face value of $2,000, 000 and 3% coupon rate paid annually on 31 December. The market interest rate on bonds of comparable risk and term is 4%. The sales proceeds of the bonds are $1,910,964. Under the effective interest rate method, the interest expense in 2017 is *closest* to:

A $77,096.

B $77,780.

C $77,807.

12 For a bond issued at a premium, using the effective interest rate method, the:

A carrying amount increases each year.

B amortization of the premium increases each year.

C premium is evenly amortized over the life of the bond.

13 Comte Industries issues $3,000,000 worth of three-year bonds dated 1 January 2015. The bonds pay interest of 5.5% annually on 31 December. The market interest rate on bonds of comparable risk and term is 5%. The sales proceeds of the bonds are $3,040,849. Under the straight-line method, the interest expense in the first year is *closest* to:

A $150,000.

B $151,384.

C $152,042.

14 The management of Bank EZ repurchases its own bonds in the open market. They pay €6.5 million for bonds with a face value of €10.0 million and a carrying value of €9.8 million. The bank will *most likely* report:

A other comprehensive income of €3.3 million.

B other comprehensive income of €3.5 million.

C a gain of €3.3 million on the income statement.

15 A company redeems $1,000,000 face value bonds with a carrying value of $990,000. If the call price is 104 the company will:

A reduce bonds payable by $1,000,000.

B recognize a loss on the extinguishment of debt of $50,000.

C recognize a gain on the extinguishment of debt of $10,000.

16 Innovative Inventions, Inc. needs to raise €10 million. If the company chooses to issue zero-coupon bonds, its debt-to-equity ratio will *most likely*:

A rise as the maturity date approaches.

B decline as the maturity date approaches.

C remain constant throughout the life of the bond.

17 Fairmont Golf issued fixed rate debt when interest rates were 6 percent. Rates have since risen to 7 percent. Using only the carrying amount (based on historical cost) reported on the balance sheet to analyze the company's financial position would *most likely* cause an analyst to:

A overestimate Fairmont's economic liabilities.

B underestimate Fairmont's economic liabilities.

C underestimate Fairmont's interest coverage ratio.

18 Which of the following is an example of an affirmative debt covenant? The borrower is:

A prohibited from entering into mergers.

B prevented from issuing excessive additional debt.

C required to perform regular maintenance on equipment pledged as collateral.

19 Debt covenants are *least likely* to place restrictions on the issuer's ability to:

A pay dividends.

B issue additional debt.

C issue additional equity.

20 Regarding a company's debt obligations, which of the following is *most likely* presented on the balance sheet?

A Effective interest rate

B Maturity dates for debt obligations

C The portion of long-term debt due in the next 12 months

21 Beginning with fiscal year 2019, for leases with a term longer than one year, lessees report a right-to-use asset and a lease liability on the balance sheet:

A only for finance leases.

B only for operating leases.

C for both finance and operating leases.

22 For a lessor, the leased asset appears on the balance sheet and continues to be depreciated when the lease is classified as:

A a finance lease.

B a sales-type lease.

C an operating lease.

23 Under US GAAP, a lessor's reported revenues at lease inception will be *highest* if the lease is classified as:

A a sales-type lease.

B an operating lease.

C a direct financing lease.

24 Under both IFRS and US GAAP, a lessor in an operating lease recognizes:

A selling profit at lease inception.

B a lease asset comprising the lease receivable and relevant residual value at lease inception.

C lease receipts as income and related costs, including depreciation, as expenses over the lease term.

25 Compared with a finance lease, an operating lease:

A is similar to renting an asset.

B is equivalent to the purchase of an asset.

C term is for the majority of the economic life of the leased asset.

26 Under US GAAP, a lessee's accounting for a long-term finance lease after inception will include:

A recognizing a single lease expense.

B recording depreciation expense on the right-of-use asset.

C increasing the balance of the lease liability by a portion of the lease payment.

27 A company enters into a finance lease agreement to acquire the use of an asset for three years with lease payments of €19,000,000 starting next year. The leased asset has a fair market value of €49,000,000 and the present value of the lease payments is €47,250,188. Based on this information, the value of the lease liability reported on the company's balance sheet at lease inception is *closest* to:

A €47,250,188.

B €49,000,000.

C €57,000,000.

28 The following presents selected financial information for a company:

	$ Millions
Short-term borrowing	4,231
Current portion of long-term interest-bearing debt	29
Long-term interest-bearing debt	925
Average shareholders' equity	18,752
Average total assets	45,981

The financial leverage ratio is *closest* to:

A 0.113

B 0.277

C 2.452

29 An analyst evaluating three industrial companies calculates the following ratios:

	Company A	Company B	Company C
Debt-to-Equity	23.5%	22.5%	52.5%
Interest Coverage	15.6	49.5	45.5

The company with both the lowest financial leverage and the greatest ability to meet interest payments is:

A Company A.

B Company B.

C Company C.

30 An analyst evaluating a company's solvency gathers the following information:

	$ Millions
Short-term interest-bearing debt	1,258
Long-term interest-bearing debt	321
Total shareholder's equity	4,285
Total assets	8,750
EBIT	2,504
Interest payments	52

The company's debt-to-assets ratio is *closest* to:

A 0.18.

B 0.27.

C 0.37.

31 Penben Corporation has a defined benefit pension plan. At 31 December, its pension obligation is €10 million and pension assets are €9 million. Under either IFRS or US GAAP, the reporting on the balance sheet would be *closest* to which of the following?

A €10 million is shown as a liability, and €9 million appears as an asset.

B €1 million is shown as a net pension obligation.

C Pension assets and obligations are not required to be shown on the balance sheet but only disclosed in footnotes.

32 The following information is associated with a company that offers its employees a defined benefit plan:

Fair value of fund's assets	$1,500,000,000
Estimated pension obligations	$2,600,000,000
Present value of estimated pension obligations	$1,200,000,000

Based on this information, the company's balance sheet will present a net pension:

A asset of $300,000,000.

B asset of $1,400,000,000.

C liability of $1,100,000,000.

SOLUTIONS

1. B is correct. The company receives €1 million in cash from investors at the time the bonds are issued, which is recorded as a financing activity.
2. B is correct. The effective interest rate is greater than the coupon rate and the bonds will be issued at a discount.
3. A is correct. Under US GAAP, expenses incurred when issuing bonds are generally recorded as an asset and amortised to the related expense (legal, etc.) over the life of the bonds. Under IFRS, they are included in the measurement of the liability. The related cash flows are financing activities.
4. C is correct. The bonds will be issued at a premium because the coupon rate is higher than the market interest rate. The future cash outflows, the present value of the cash outflows, and the total present value are as follows:

Date	Interest Payment ($)	Present Value at Market Rate 5% ($)		Present Value at Market Rate 5% ($)	Total Present Value ($)
31 December 2015	60,000.00	57,142.86			
31 December 2016	60,000.00	54,421.77			
31 December 2017	60,000.00	51,830.26			
31 December 2018	60,000.00	49,362.15			
31 December 2019	60,000.00	47,011.57			
31 December 2020	60,000.00	44,772.92			
31 December 2021	60,000.00	42,640.88			
31 December 2022	60,000.00	40,610.36			
31 December 2023	60,000.00	38,676.53			
31 December 2024	60,000.00	36,834.80	1,000,000.00	613,913.25	
		463,304.10		613,913.25	1,077,217.35
					Sales Proceeds

The following illustrates the keystrokes for many financial calculators to calculate sales proceeds of $1,077,217.35:

Calculator Notation	Numerical Value for This Problem
N	10
% *i* or I/Y	5
FV	$1,000,000.00
PMT	$60,000.00
PV compute	X

Thus, the sales proceeds are reported on the balance sheet as an increase in long-term liability, bonds payable of $1,077,217.

5. A is correct. The bonds payable reported at issue is equal to the sales proceeds. The interest payments and future value of the bond must be discounted at the market interest rate of 3% to determine the sales proceeds.

Date	Interest Payment	Present Value at Market Rate (3%)	Face Value Payment	Present Value at Market Rate (3%)	Total Present Value
31 December 2015	$125,000.00	$121,359.22			
31 December 2016	$125,000.00	$117,824.49			
31 December 2017	$125,000.00	$114,392.71	$5,000,000.00	$4,575,708.30	
Total		$353,576.42		$4,575,708.30	$4,929,284.72

The following illustrates the keystrokes for many financial calculators to calculate sales proceeds of $4,929,284.72:

Calculator Notation	Numerical Value for This Problem
N	3
% *i* or I/Y	3.0
FV	$5,000,000.00
PMT	$125,000.00
PV compute	X

6 B is correct. The market interest rate at the time of issuance is the effective interest rate that the company incurs on the debt. The effective interest rate is the discount rate that equates the present value of the coupon payments and face value to their selling price. Consequently, the effective interest rate is 5.50%.

7 B is correct. The bonds will be issued at a discount because the market interest rate is higher than the stated rate. Discounting the future payments to their present value indicates that at the time of issue, the company will record £978,938 as both a liability and a cash inflow from financing activities. Interest expense in 2010 is £58,736 (£978,938 times 6.0 percent). During the year, the company will pay cash of £55,000 related to the interest payment, but interest expense on the income statement will also reflect £3,736 related to amortisation of the initial discount (£58,736 interest expense less the £55,000 interest payment). Thus, the value of the liability at 31 December 2010 will reflect the initial value (£978,938) plus the amortised discount (£3,736), for a total of £982,674. The cash outflow of £55,000 may be presented as either an operating or financing activity under IFRS.

8 A is correct. The coupon rate on the bonds is higher than the market rate, which indicates that the bonds will be issued at a premium. Taking the present value of each payment indicates an issue date value of €10,210,618. The interest expense is determined by multiplying the carrying amount at the beginning of the period (€10,210,618) by the market interest rate at the time of issue (6.0 percent) for an interest expense of €612,637. The value after one year will equal the beginning value less the amount of the premium amortised to date, which is the difference between the amount paid (€650,000) and the expense accrued (€612,637) or €37,363. €10,210,618 – €37,363 = €10,173,255 or €10.17 million.

9 A is correct. The future cash outflows, the present value of the cash outflows, and the total present value are as follows:

Date	Interest Payment (€)	Present Value at Market Rate 6% (€)		Present Value at Market Rate 6% (€)	Total Present Value (€)
31 December 2015	700,000.00	660,377.36			
31 December 2016	700,000.00	622,997.51			
31 December 2017	700,000.00	587,733.50			
31 December 2018	700,000.00	554,465.56			
31 December 2019	700,000.00	523,080.72			
31 December 2020	700,000.00	493,472.38			
31 December 2021	700,000.00	465,539.98			
31 December 2022	700,000.00	439,188.66			
31 December 2023	700,000.00	414,328.92			
31 December 2024	700,000.00	390,876.34	10,000,000.00	5,583,947.77	
		5,152,060.94		5,583,947.77	10,736,008.71
					Sales Proceeds

The following illustrates the keystrokes for many financial calculators to calculate sales proceeds of €10,736,008.71:

Calculator Notation	Numerical Value for This Problem
N	10
% *i* or I/Y	6
FV	$10,000,000.00
PMT	$700,000.00
PV compute	X

The interest expense is calculated by multiplying the carrying amount at the beginning of the year by the effective interest rate at issuance. As a result, the interest expense at 31 December 2015 is €644,161 (€10,736,008.71 × 6%).

10 C is correct. The future cash outflows, the present value of the cash outflows, and the total present value are as follows:

Date	Interest Payment ($)	Present Value at Market Rate 5% ($)		Present Value at Market Rate 5% ($)	Total Present Value ($)
31 December 2015	1,200,000	1,142,857.14			
31 December 2016	1,200,000	1,088,435.37			
31 December 2017	1,200,000	1,036,605.12			
31 December 2018	1,200,000	987,242.97			
31 December 2019	1,200,000	940,231.40	30,000,000	23,505,785.00	
		5,195,372.00		23,505,785.00	28,701,157.00
					Sales Proceeds

The following illustrates the keystrokes for many financial calculators to calculate sales proceeds of $28,701,157.00:

Calculator Notation	Numerical Value for This Problem
N	5
% *i* or I/Y	5
FV	$30,000,000.00
PMT	$1,200,000.00
PV compute	X

The following table illustrates interest expense, premium amortization, and carrying amount (amortized cost) for 2015.

Year	Carrying Amount (beginning of year)	Interest Expense (at effective interest rate of 5%)	Interest Payment (at coupon rate of 4%)	Amortization of Discount	Carrying Amount (end of year)
2015	$28,701,157.00	$1,435,057.85	$1,200,000.00	$235,057.85	$28,936,214.85

The carrying amount at the end of the year is found by adding the amortization of the discount to the carrying amount at the beginning of the year. As a result, the carrying amount on 31 December 2015 is $28,936,215.

Alternatively, the following illustrates the keystrokes for many financial calculators to calculate the carrying value at the end of first year of $28,936, 215:

Calculator Notation	Numerical Value for This Problem
N	4
% *i* or I/Y	5
FV	$30,000,000.00
PMT	$1,200,000.00
PV compute	X

11 B is correct. The interest expense for a given year is equal to the carrying amount at the beginning of the year times the effective interest of 4%. Under the effective interest rate method, the difference between the interest expense and the interest payment (based on the coupon rate and face value) is the discount amortized in the period, which increases the carrying amount annually. For 2017, the interest expense is the beginning carrying amount ($1,944,499) times the effective interest of 4%.

Year	Carrying Amount (beginning)	Interest Expense (at effective interest of 4%)	Interest Payment (at coupon rate of 3%)	Amortization of Discount	Carrying Amount (end of year)
2015	$1,910,964	$76,439	$60,000.00	$16,439	$1,927,403
2016	$1,927,403	$77,096	$60,000.00	$17,096	$1,944,499
2017	$1,944,499	$77,780	$60,000.00	$17,780	$1,962,279

12 B is correct. The amortization of the premium equals the interest payment minus the interest expense. The interest payment is constant and the interest expense decreases as the carrying amount decreases. As a result, the amortization of the premium increases each year.

13 B is correct. Under the straight-line method, the bond premium is amortized equally over the life of the bond. The annual interest payment is \$165,000 (\$3,000,000 × 5.5%) and annual amortization of the premium under the straight-line method is \$13,616 [(\$3,040,849 – \$3,000,000)/3)]. The interest expense is the interest payment less the amortization of the premium (\$165,000 – \$13,616 = \$151,384).

14 C is correct. A gain of €3.3 million (carrying amount less amount paid) will be reported on the income statement.

15 B is correct. If a company decides to redeem a bond before maturity, bonds payable is reduced by the carrying amount of the debt. The difference between the cash required to redeem the bonds and the carrying amount of the bonds is a gain or loss on the extinguishment of debt. Because the call price is 104 and the face value is \$1,000,000, the redemption cost is 104% of \$1,000,000 or \$1,040,000. The company's loss on redemption would be \$50,000 (\$990,000 carrying amount of debt minus \$1,040,000 cash paid to redeem the callable bonds).

16 A is correct. The value of the liability for zero-coupon bonds increases as the discount is amortised over time. Furthermore, the amortised interest will reduce earnings at an increasing rate over time as the value of the liability increases. Higher relative debt and lower relative equity (through retained earnings) will cause the debt-to-equity ratio to increase as the zero-coupon bonds approach maturity.

17 A is correct. When interest rates rise, bonds decline in value. Thus, the carrying amount of the bonds being carried on the balance sheet is higher than the market value. The company could repurchase the bonds for less than the carrying amount, so the economic liabilities are overestimated. Because the bonds are issued at a fixed rate, there is no effect on interest coverage.

18 C is correct. Affirmative covenants require certain actions of the borrower. Requiring the company to perform regular maintenance on equipment pledged as collateral is an example of an affirmative covenant because it requires the company to do something. Negative covenants require that the borrower not take certain actions. Prohibiting the borrower from entering into mergers and preventing the borrower from issuing excessive additional debt are examples of negative covenants.

19 C is correct. Covenants protect debtholders from excessive risk taking, typically by limiting the issuer's ability to use cash or by limiting the overall levels of debt relative to income and equity. Issuing additional equity would increase the company's ability to meet its obligations, so debtholders would not restrict that ability.

20 C is correct. The non-current liabilities section of the balance sheet usually includes a single line item of the total amount of a company's long-term debt due after 1 year, and the current liabilities section shows the portion of a company's long-term debt due in the next 12 months. Notes to the financial statements generally present the stated and effective interest rates and maturity dates for a company's debt obligations

21 C is correct. Beginning with fiscal year 2019, lessees report a right-of-use asset and a lease liability for all leases longer than one year. An exception under IFRS exists for leases when the underlying asset is of low value.

22 C is correct. When a lease is classified as an operating lease, the underlying asset remains on the lessor's balance sheet. The lessor will record a depreciation expense that reduces the asset's value over time.

23 A is correct. A sales-type lease treats the lease as a sale of the asset, and revenue is recorded at the time of sale equal to the value of the leased asset. Under a direct financing lease, only interest income is reported as earned. Under an operating lease, revenue from lease receipts is reported when collected.

24 C is correct. Lessor accounting for an operating lease under US GAAP is similar to that under IFRS: Over the lease term, the lessor recognizes lease receipts as income and recognizes related costs, including depreciation of the leased asset, as expenses. Under IFRS, at inception of a finance lease—not an operating lease—the lessor derecognizes the underlying leased asset and recognizes a lease asset comprising the lease receivable and relevant residual value. Further, an IFRS-reporting lessor will recognize selling profit at the beginning of all leases that are not classified as operating leases. In contrast, a US GAAP–reporting lessor will recognize selling profit only on sales-type leases at the beginning of the lease term.

25 A is correct. An operating lease is an agreement that allows the lessee to use an asset for a period of time. Thus, an operating lease is similar to renting an asset, whereas a finance lease is equivalent to the purchase of an asset by the lessee that is directly financed by the lessor.

26 B is correct. A lessee's accounting for a long-term finance lease under US GAAP and after lease inception includes recording depreciation expense on the right-of-use asset, recognizing interest expense on the lease liability, and reducing the balance of the lease liability for the portion of the lease payments that represents repayment of the lease liability. A lessee's accounting for an operating lease under US GAAP and after lease inception will recognize a single lease expense, which is a straight-line allocation of the cost of the lease over its term.

27 A is correct. Under the revised reporting standards under IFRS and US GAAP, a lessee must recognize an asset and a lease liability at inception of each of its leases (with an exception for short-term leases). The lessee reports a "right-of-use" (ROU) asset and a lease liability, calculated essentially as the present value of fixed lease payments, on its balance sheet. Thus, at lease inception, the company will record a lease liability on the balance sheet of €47,250,188.

28 C is correct. The financial leverage ratio is calculated as follows:

$$\frac{\text{Average total assets}}{\text{Average shareholder's equity}} = \frac{\$45{,}981\text{ million}}{\$18{,}752\text{ million}} = \$2.452\text{ million}$$

29 B is correct. Company B has the lowest debt-to-equity ratio, indicating the lowest financial leverage, and the highest interest coverage ratio, indicating the greatest number of times that EBIT covers interest payments.

30 A is correct because the debt-to-assets (total debt)/(total assets) ratio is (1,258 + 321)/(8,750) = 1,579/8,750 = 0.18

31 B is correct. The company will report a net pension obligation of €1 million equal to the pension obligation (€10 million) less the plan assets (€9 million).

32 A is correct. A company that offers a defined benefit plan makes payments into a pension fund and the retirees are paid from the fund. The payments that a company makes into the fund are invested until they are needed to pay retirees. If the fair value of the fund's assets is higher than the present value of the estimated pension obligation, the plan has a surplus and the company's balance sheet will reflect a net pension asset. Because the fair value of the fund's assets is $1,500,000,000 and the present value of estimated pension obligations is $1,200,000,000, the company will present a net pension asset of $300,000,000 on its balance sheet.

FINANCIAL STATEMENT ANALYSIS

STUDY SESSION

8

Financial Statement Analysis (4)

This study session introduces the concept of financial reporting quality. The session examines the financial reporting quality differences that may exist between companies and the means for identifying them. Warning signs of poor or low quality reporting are covered. The application of financial analysis techniques to evaluate a company's past and projected performance, assess credit risk, and screen for potential equity investments follows. Common adjustments to reported financials to facilitate cross-company comparisons conclude the session.

READING ASSIGNMENTS

Reading 25	Financial Reporting Quality by Jack T. Ciesielski, CPA, CFA, Elaine Henry, PhD, CFA, and Thomas I. Selling, PhD, CPA
Reading 26	Applications of Financial Statement Analysis by Thomas R. Robinson, PhD, CFA, J. Hennie van Greuning, DCom, CFA, Elaine Henry, PhD, CFA, and Michael A. Broihahn, CPA, CIA, CFA

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2021 CFA Institute. All rights reserved.

READING

25

Financial Reporting Quality

by Jack T. Ciesielski, CPA, CFA, Elaine Henry, PhD, CFA, and
Thomas I. Selling, PhD, CPA

Jack T. Ciesielski, CPA, CFA, is at R.G. Associates, Inc., former publisher of The Analyst's Accounting Observer (USA). Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). Thomas I. Selling, PhD, CPA, is at the Cox School of Business, Southern Methodist University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** compare and contrast financial reporting quality with the quality of reported results (including quality of earnings, cash flow, and balance sheet items);
☐	**b.** describe a spectrum for assessing financial reporting quality;
☐	**c.** explain the difference between conservative and aggressive accounting;
☐	**d.** describe motivations that might cause management to issue financial reports that are not high quality;
☐	**e.** describe conditions that are conducive to issuing low-quality, or even fraudulent, financial reports;
☐	**f.** describe mechanisms that discipline financial reporting quality and the potential limitations of those mechanisms;
☐	**g.** describe presentation choices, including non-GAAP measures, that could be used to influence an analyst's opinion;
☐	**h.** describe accounting methods (choices and estimates) that could be used to manage earnings, cash flow, and balance sheet items;
☐	**i.** describe accounting warning signs and methods for detecting manipulation of information in financial reports.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION & CONCEPTUAL OVERVIEW

a compare and contrast financial reporting quality with the quality of reported results (including quality of earnings, cash flow, and balance sheet items);

b describe a spectrum for assessing financial reporting quality;

Ideally, analysts would always have access to financial reports that are based on sound financial reporting standards, such as those from the International Accounting Standards Board (IASB) and the Financial Accounting Standards Board (FASB), and are free from manipulation. But, in practice, the quality of financial reports can vary greatly. High-quality financial reporting provides information that is useful to analysts in assessing a company's performance and prospects. Low-quality financial reporting contains inaccurate, misleading, or incomplete information.

Extreme lapses in financial reporting quality have given rise to high-profile scandals that resulted not only in investor losses but also in reduced confidence in the financial system. Financial statement users who were able to accurately assess financial reporting quality were better positioned to avoid losses. These lapses illustrate the challenges analysts face as well as the potential costs of failing to recognize practices that result in misleading or inaccurate financial reports.[1] Examples of misreporting can provide an analyst with insight into various signals that may indicate poor-quality financial reports.

This reading addresses *financial reporting quality*, which pertains to the quality of information in financial reports, including disclosures in notes. High-quality reporting provides decision-useful information, which is relevant and faithfully represents the economic reality of the company's activities during the reporting period as well as the company's financial condition at the end of the period. A separate but interrelated attribute of quality is *quality of reported results* or *earnings quality*, which pertains to the earnings and cash generated by the company's actual economic activities and the resulting financial condition. The term "earnings quality" is commonly used in practice and will be used broadly to encompass the quality of earnings, cash flow, and/or balance sheet items. High-quality earnings result from activities that a company will likely be able to sustain in the future and provide a sufficient return on the company's investment. The concepts of earnings quality and financial reporting quality are interrelated because a correct assessment of earnings quality is possible only when there is some basic level of financial reporting quality. Beyond this basic level, as the quality of reporting increases, the ability of financial statement users to correctly assess earnings quality and to develop expectations for future performance arguably also increases.

Sections 1–5 provide a conceptual overview of reporting quality. Sections 6 and 7 discuss motivations that might cause, and conditions that might enable, management to issue financial reports that are not high quality and mechanisms that aim to provide discipline to financial reporting quality. Sections 8–12 describe choices made by management that can affect financial reporting quality—presentation choices, accounting methods, and estimates—as well as warning signs of poor-quality financial reporting.

---

[1] In this reading, the examples of misleading or inaccurate financial reports occurred in prior years—*not* because there are no current examples of questionable financial reporting, but rather because it has been conclusively resolved that misreporting occurred in the historical examples.

1.1 Conceptual Overview

As indicated in the introduction, financial reporting quality and results or earnings quality are interrelated attributes of quality. Exhibit 1 illustrates this interrelationship and its implications.

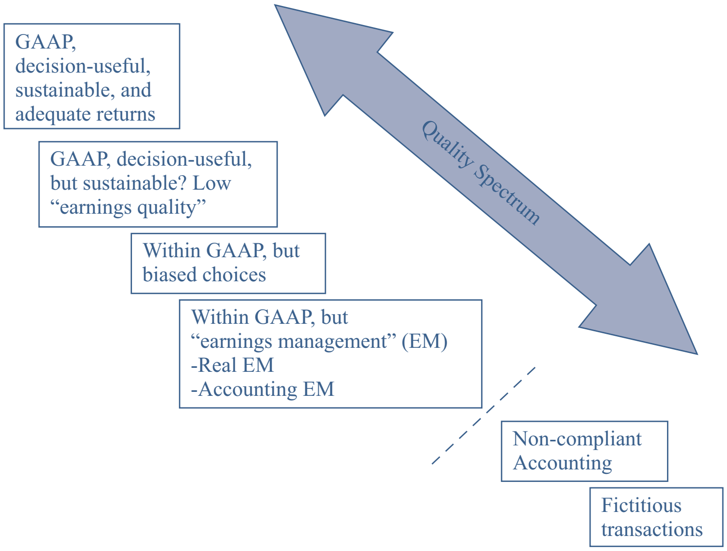
Exhibit 1 Relationships between Financial Reporting Quality and Earnings Quality

		Financial Reporting Quality	
		Low	**High**
Earnings (Results) Quality	High	LOW financial reporting quality impedes assessment of earnings quality and impedes valuation.	HIGH financial reporting quality enables assessment. HIGH earnings quality increases company value.
	Low		HIGH financial reporting quality enables assessment. LOW earnings quality decreases company value.

As can be seen in Exhibit 1, if financial reporting quality is low, the information provided is of little use in assessing the company's performance, and thus in making investment and other decisions.

Financial reporting quality varies across companies. High-quality reports contain information that is relevant, complete, neutral, and free from error. The lowest-quality reports contain information that is pure fabrication. Earnings (results) quality can range from high and sustainable to low and unsustainable. Providers of resources prefer high and sustainable earnings. Combining the two measures of quality—financial reporting and earnings—the overall quality of financial reports from a user perspective can be thought of as spanning a continuum from the highest to the lowest. Exhibit 2 presents a quality spectrum that provides a basis for evaluating better versus poorer quality reports. This spectrum ranges from reports that are of high financial reporting quality and reflect high and sustainable earnings quality to reports that are not useful because of poor financial reporting quality.

Exhibit 2 Quality Spectrum of Financial Reports


GAAP, decision-useful, sustainable, and adequate returns
GAAP, decision-useful, but sustainable? Low "earnings quality"
Within GAAP, but biased choices
Within GAAP, but "earnings management" (EM) -Real EM -Accounting EM
Non-compliant Accounting
Fictitious transactions
Quality Spectrum

2 GAAP, DECISION USEFUL FINANCIAL REPORTING

b describe a spectrum for assessing financial reporting quality;

At the top of the spectrum, labeled in Exhibit 2 as "GAAP, decision-useful, sustainable, and adequate returns," are high-quality reports that provide useful information about high-quality earnings.

- High-quality financial reports conform to the generally accepted accounting principles (GAAP) of the jurisdiction, such as International Financial Reporting Standards (IFRS), US GAAP, or other home-country GAAP. The exhibit uses the term GAAP to refer generically to the accounting standards accepted in a company's jurisdiction.
- In addition to conforming to GAAP, high-quality financial reports also embody the characteristics of decision-useful information such as those defined in the *Conceptual Framework*.[2] Recall that the fundamental characteristics of useful information are relevance and faithful representation. Relevant information is defined as information that can affect a decision and encompasses the notion of materiality. (Information is considered material if "omitting it or misstating it

---

[2] The characteristics of decision-useful information are identical under IFRS and US GAAP. In September 2010, the IASB adopted the *Conceptual Framework for Financial Reporting* in place of the *Framework for the Preparation and Presentation of Financial Statements* (1989). The *Conceptual Framework* represents the partial completion of a joint convergence project between the IASB and FASB on an updated framework. The *Conceptual Framework* (2010) contains two updated chapters: "The Objective of Financial Reporting" and "Qualitative Characteristics of Useful Financial Information." The remainder of the material in the *Conceptual Framework* is from the *Framework* (1989) and will be updated as the project is completed. Also in September 2010, the FASB issued Concepts Statement 8, "Conceptual Framework for Financial Reporting," to replace Concepts Statements 1 and 2.

could influence decisions that users make on the basis of the financial information of a specific reporting entity."[3]) Faithful representation of economic events is complete, neutral, and free from error.

The *Conceptual Framework* also enumerates enhancing characteristics of useful information: comparability, verifiability, timeliness, and understandability. Of course, the desirable characteristics for financial information require trade-offs. For example, financial reports must balance the aim of providing information that is produced quickly enough to be timely and thus relevant, and yet not so quickly that errors occur. Financial reports must balance the aim of providing information that is complete but not so exhaustive that immaterial information is included. High-quality information results when these and other tradeoffs are made in an unbiased, skillful manner.

- High-quality earnings indicate an adequate level of return on investment and derive from activities that a company will likely be able to sustain in the future. An adequate level of return on investment exceeds the cost of the investment and also equals or exceeds the expected return. Sustainable activities and sustainable earnings are those expected to recur in the future. Sustainable earnings that provide a high return on investment contribute to higher valuation of a company and its securities.

2.1 GAAP, Decision-Useful, but Sustainable?

The next level down in Exhibit 2, "GAAP, decision-useful, but sustainable?" refers to circumstances in which high-quality reporting provides useful information, but that information reflects results or earnings that are not sustainable (lower earnings quality). The earnings may not be sustainable because the company cannot expect earnings that generate the same level of return on investment in the future or because the earnings, although replicable, will not generate sufficient return on investment to sustain the company. Earnings quality is low in both cases. Reporting can be high quality even when the economic reality being depicted is not of high quality. For example, consider a company that generates a loss, or earnings that do not provide an adequate return on investment, or earnings that resulted from non-recurring activities. The relatively undesirable economic reality could nonetheless be depicted in financial reporting that provides high-quality, decision-useful information.

Exhibit 3 presents an excerpt from the fiscal year 2014 first-quarter results of Toyota Motor Corporation, a Japanese automobile company. As highlighted by a *Wall Street Journal* article,[4] the company sold fewer cars but reported an 88% increase in operating profits compared with the prior year, primarily because of the change in exchange rates. The weaker yen benefited Toyota both because the company manufactures more cars in Japan (compared with its competitors) and because the company sells a significant number of cars outside of Japan. Exchange rate weakening is a less sustainable source of profits than manufacturing and selling cars. In summary, this example is a case of high-quality financial reporting coupled with lower earnings quality.

---

[3] Text from conceptual frameworks referenced in Note 4.

[4] Back (2013).

Exhibit 3 Excerpt from Toyota Motor Corporation's Consolidated Financial Results for FY2014 First Quarter Ending 30 June 2013

Consolidated vehicle unit sales in Japan and overseas decreased by 37 thousand units, or 1.6%, to 2,232 thousand units in FY2014 first quarter (the three months ended June 30, 2013) compared with FY2013 first quarter (the three months ended June 30, 2012). Vehicle unit sales in Japan decreased by 51 thousand units, or 8.8%, to 526 thousand units in FY2014 first quarter compared with FY2013 first quarter. Meanwhile, overseas vehicle unit sales increased by 14 thousand units, or 0.8%, to 1,706 thousand units in FY2014 first quarter compared with FY2013 first quarter.

As for the results of operations, net revenues increased by 753.7 billion yen, or 13.7%, to 6,255.3 billion yen in FY2014 first quarter compared with FY2013 first quarter, and operating income increased by 310.2 billion yen, or 87.9%, to 663.3 billion yen in FY2014 first quarter compared with FY2013 first quarter. The factors contributing to an increase in operating income were the effects of changes in exchange rates of 260.0 billion yen, cost reduction efforts of 70.0 billion yen, marketing efforts of 30.0 billion yen and other factors of 10.2 billion yen. On the other hand, the factors contributing to a decrease in operating income were the increase in expenses and others of 60.0 billion yen.

3 BIASED ACCOUNTING CHOICES

b describe a spectrum for assessing financial reporting quality;

c explain the difference between conservative and aggressive accounting;

The next level down in the spectrum in Exhibit 2 is "Within GAAP, but biased choices." Biased choices result in financial reports that do not faithfully represent the economic substance of what is being reported. The problem with bias in financial reporting, as with other deficiencies in reporting quality, is that it impedes an investor's ability to correctly assess a company's past performance, to accurately forecast future performance, and thus to appropriately value the company.

Choices are deemed to be "aggressive" if they increase a company's reported performance and financial position in the period under review. The choice can increase the amount of revenues, earnings, and/or operating cash flow reported for the period, or decrease expenses, and/or reduce the level of debt reported on the balance sheet. Aggressive choices may lead to a reduction in the company's reported performance and in its financial position in later periods. In contrast, choices are deemed "conservative" if they decrease a company's performance and financial position in the reporting period. This can include lowering the reported revenues, earnings, and/or operating cash flow reported or increasing expenses, or recording a higher level of debt on the balance sheet. Conservative choices may lead to a rise in the company's reported performance and financial position in later periods.

Another type of bias is understatement of earnings volatility, so-called earnings "smoothing". Earnings smoothing can result from conservative choices to understate earnings in periods when a company's operations are performing well, building up (often hidden) reserves that allow aggressive choices in periods when its operations are struggling.

Biased choices can be made not only in the context of reported amounts but also in the context of how information is presented. For example, companies can disclose information transparently, which facilitates analysis, or they can disclose it in a manner that aims to obscure unfavorable and/or emphasize favorable information.

EXAMPLE 1

Quality of Financial Reports

PACCAR Inc. designs, manufactures, and distributes trucks and related aftermarket parts that are sold worldwide under the Kenworth, Peterbilt, and DAF nameplates. In 2013, the US SEC charged PACCAR for various accounting deficiencies that "clouded their financial reporting to investors in the midst of the financial crisis." The SEC complaint cites the company's 2009 segment reporting. Exhibit 4A presents an excerpt from the notes to PACCAR's financial statements, and Exhibit 4B presents an excerpt from the management's discussion and analysis (MD&A) of PACCAR's annual report.

Exhibit 4A Excerpt from Notes to PACCAR's 2009 Financial Statements

S. SEGMENT AND RELATED INFORMATION

PACCAR operates in two principal segments, Truck and Financial Services. The Truck segment includes the manufacture of trucks and the distribution of related aftermarket parts, both of which are sold through a network of independent dealers... The Financial Services segment is composed of finance and leasing products and services provided to truck customers and dealers … Included in All Other is PACCAR's industrial winch manufacturing business. Also within this category are other sales, income and expenses not attributable to a reportable segment, including a portion of corporate expense.

Business Segment Data ($ millions)

	2009	2008	2007
Income before Income Taxes			
Truck	$25.9	$1,156.5	$1,352.8
All other	42.2	6.0	32.0
	68.1	1,162.5	1,384.8
Financial services	84.6	216.9	284.1
Investment income	22.3	84.6	95.4
	$175.0	$1,464.0	$1,764.3

Exhibit 4B Excerpt from MD&A of PACCAR's 2009 Annual Report

Net sales and revenues and gross margins for truck units and aftermarket parts are provided below. The aftermarket parts gross margin includes direct revenues and costs, but excludes certain truck segment costs.

(continued)

Exhibit 4B (Continued)

	2009	2008	% Change
Net Sales and Revenues			
Trucks	$5,103.30	$11,281.30	–55
Aftermarket parts	1,890.70	2,266.10	–17
	$6,994.00	$13,547.40	–48
Gross Margin			
Trucks	–$46.6	$1,141.70	–104
Aftermarket parts	625.7	795.20	–21
	$579.1	$1,936.90	–70

1 Based on the segment data excerpted from the notes to the financial statements, was PACCAR's truck segment profitable in 2009?

2 Based on the data about the truck's gross margin presented in the MD&A, was PACCAR's truck segment profitable in 2009?

3 What is the main difference between the note presentation and the MD&A presentation?

4 The SEC complaint stated that "PACCAR failed to report the operating results of its aftermarket parts business separately from its truck sales business as required under segment reporting requirements, which are in place to ensure that investors gain the same insight into a company as its executives." Is the PACCAR situation an example of issues with financial reporting quality, earnings quality, or both?

Solution to 1:

Yes, the segment data presented in the note to the financial statements indicates that the Truck segment earned $25.9 million in 2009.

Solution to 2:

No, the segment data presented in the MD&A indicates that the Truck segment had a negative gross margin.

Solution to 3:

The main difference between the note presentation and the MD&A presentation is that the aftermarket parts business is combined with the trucks business in the notes but separated in the MD&A. Although the data are not exactly comparable in the two disclosures (because the note shows income before taxes and the MD&A shows gross profit), the two disclosures present a different picture of PACCAR's profits from truck sales.

Solution to 4:

The PACCAR situation appears to be an example of issues with both financial reporting quality and earnings quality. The substantial decrease in truck sales and the negative gross margin reflect poor earnings quality. The failure to disclose clear segment information is an instance of poor financial reporting quality.

While choices exist within GAAP for the presentation of a desired economic picture, non-GAAP reporting adds yet another dimension of management discretion. Non-GAAP reporting of financial metrics not in compliance with generally accepted accounting principles such as US GAAP and IFRS includes both financial metrics and operating metrics.[5] Non-GAAP financial metrics relate directly to the financial statements. A common non-GAAP financial metric is "non-GAAP earnings," which are created by companies "that adjust standards-compliant earnings to *exclude items required* by accounting standards or to *include items not permitted* by accounting standards" (Ciesielski and Henry, 2017). In contrast, non-GAAP operating metrics do not relate directly to the financial statements and include metrics that are typically industry-driven, such as subscriber numbers, active users, and occupancy rates.

Non-GAAP financial reporting has become increasingly common, presenting challenges to analysts. An important challenge is that non-GAAP financial reporting diminishes comparability across financial statements. The adjustments that companies make to create non-GAAP earnings, for example, are generally ad hoc and thus differ significantly. When evaluating non-GAAP metrics, investors must decide the extent to which specific adjustments should be incorporated into their analyses and forecasts.[6]

Another challenge arises from differences in terminology. Non-GAAP earnings are sometimes referred to as underlying earnings, adjusted earnings, recurring earnings, core earnings, or similar. Exhibit 5 provides an example from Jaguar Land Rover Automotive plc (JLR), a subsidiary of Tata Motors Ltd. The company prepares its financial reports under IFRS. The exhibit is an excerpt from JLR's 2016/17 annual report and uses the term "alternative performance measures". Exhibit 6 is from Tata Motors Ltd's Form 6-K filed with the US SEC, containing supplemental information regarding JLR and using the term "non-IFRS Financial Measures". The information in the two exhibits is essentially identical, but the terminology and formatting differ.

Exhibit 5

JLR's 2016/17 Annual Report: Footnote 3 [Excerpt]

3) ALTERNATIVE PERFORMANCE MEASURES

Many companies use alternative performance measures (APMs) to provide helpful additional information for users of their financial statements, telling a clearer story of how the business has performed over the period... These measures exclude certain items that are included in comparable statutory measures….

Reconciliations between these alternative performance measures and statutory reported measures are shown below.

EBIT AND EBITDA (£m)

Year ended 31 March	2017
EBITDA	**2,955**
Depreciation and amortisation	–1,656
Share of profit/(loss) of equity accounted investments	159
EBIT	**1,458**

(continued)

---

[5] The term "non-GAAP" refers generally to all metrics that are non-compliant with generally accepted accounting principles and thus includes "non-IFRS" metrics.

[6] A survey of non-GAAP earnings in the S&P 500 is presented in Ciesielski and Henry (2017). Some observers even recommend that investors shift their focus from a company's earnings to a company's "strategic assets" and the contribution of these assets to its competitive edge (Gu and Lev, 2017).

Exhibit 5 (Continued)

Year ended 31 March	2017
Foreign exchange (loss)/gain on derivatives	–11
Unrealised gain/(loss) on commodities	148
Foreign exchange loss on loans	–101
Finance income	33
Finance expense (net)	–68
Exceptional item	151
Profit before tax	**1,610**

Exhibit 6

Tata Motors Ltd. SEC Form 6-K [Excerpt]

Non-IFRS Financial Measures

This Report includes references to certain non-IFRS measures, including EBITDA, EBIT … [These measures] and related ratios should not be considered in isolation and are not measures of JLR's financial performance or liquidity under IFRS and should not be considered as an alternative to profit or loss for the period or any other performance measures derived in accordance with IFRS or as an alternative to cash flow from operating, investing or financing activities or any other measure of JLR's liquidity derived in accordance with IFRS. … In addition, EBITDA, EBIT… as defined, may not be comparable to other similarly titled measures used by other companies.

Exhibit 1 to Form 6-K Supplemental Information Regarding the Jaguar and Land Rover Business of Tata Motors Limited [Excerpt]

The reconciliation of JLR's EBIT and EBITDA to profit for the period line item is:

Fiscal year ended March 31, 2017	£m
Profit for the period	1,272
Add back taxation	338
Add/(less) back exceptional charge/(credit)	–151
Add back/(less) foreign exchange (gains)/loss – financing	101
Add back/(less) foreign exchange (gains)/loss – derivatives	11
Add back/(less) unrealized commodity losses/(gains) – unrealized derivatives	–148
Less finance income	–33
Add back finance expense (net)	68
EBIT	**1,458**
Add back depreciation and amortization	1,656

Exhibit 1 (Continued)

Fiscal year ended March 31, 2017	£m
Add/(less) back share of loss/(profit) from equity accounted investees	–159
EBITDA	**2,955**

Management emphasis on non-GAAP financial measures to deflect attention from less-than-desirable GAAP financial results is an example of an aggressive presentation choice. Since 2003, if a company uses a non-GAAP financial measure[7] in an SEC filing, it is required to display the most directly comparable GAAP measure with equal prominence and to provide a reconciliation between the non-GAAP measure and the equivalent GAAP measure. In other words, a company is not allowed to give more prominence to a non-GAAP financial measure in an SEC filing.

Similarly, the IFRS Practice Statement "Management Commentary," issued December 2010, requires disclosures when non-IFRS measures are included in financial reports:

> If information from the financial statements has been adjusted for inclusion in management commentary, that fact should be disclosed. If financial performance measures that are not required or defined by IFRSs are included within management commentary, those measures should be defined and explained, including an explanation of the relevance of the measure to users. When financial performance measures are derived or drawn from the financial statements, those measures should be reconciled to measures presented in the financial statements that have been prepared in accordance with IFRSs. (Page 17)

The reconciliation between as-reported measures (GAAP financial measures presented in the financial statements) and as-adjusted measures (non-GAAP financial measures presented in places other than the financial statements) can provide important information.

The European Securities and Markets Authority (ESMA) published guidelines in October 2015 *(ESMA Guidelines on Alternative Performance Measures)* covering such points as the definition of APMs, reconciliation to GAAP, explanation of the metrics' relevance, and consistency over time. We discuss ESMA in more detail later in this reading.

EXAMPLE 2

Presentation of Non-GAAP Financial Measures

Convatec Group PLC (Convatec), a global medical products manufacturer, raised $1.8 billion via an initial public offering (IPO) on the London Stock Exchange in 2016. The company had been purchased by private equity firms from Bristol-Myers Squibb in 2008 for $4.1 billion. Exhibit 7 presents excerpts from the company's regulatory filing at the London Stock Exchange announcing its full year 2016 results.

---

[7] Non-domestic private issuers can file financial statements prepared in accordance with IFRS without reconciliation to US GAAP. The SEC recognizes US GAAP and IFRS as GAAP.

Exhibit 7 Excerpt from ConvaTec's Press Release for Full Year 2016 Results

Headline: "Strong results, delivering on strategy"

CEO Review [Excerpt]

At constant currency, revenue grew 4% to $1,688 million and adjusted EBITDA was $508 million, up 6.5% at constant currency...

[Footnote] Constant currency growth 'CER' is calculated by restating 2016 results using 2015 foreign exchange rates for the relevant period.

Consolidated Statement of Profit or Loss for the year ended 31 December 2016 ($ m)

	2016	2015
Revenue	1,688.3	1,650.4
Cost of goods sold	–821.0	–799.9
Gross profit	867.3	850.5
Selling and distribution expenses	–357.0	–346.7
General and administrative expenses	–318.2	–233.1
Research and development expenses	–38.1	–40.3
Operating profit	154.0	230.4
Finance costs	–271.4	–303.6
Other expense, net	–8.4	–37.1
Loss before income taxes	–125.8	–110.3
Income tax (expense) benefit	–77.0	16.9
Net loss	–202.8	–93.4

Non-IFRS Financial Information [Excerpt]

This release contains certain financial measures that are not defined or recognised under IFRS. These measures are referred to as "Adjusted" measures... These measures are not measurements of financial performance or liquidity under IFRS and should not replace measures of liquidity or operating profit that are derived in accordance with IFRS.

Reconciliation to adjusted earnings [Excerpt]

2016	Reported	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Adjusted
Revenue	1,688.3	—	—	—	—	—	—	—	1,688.3
…									
Operating profit	**154.0**	**155.1**	**30.9**	**11.7**	**0.8**	**—**	**90.2**	**29.5**	**472.2**
…									
(Loss) profit before income taxes	–125.8	155.1	30.9	11.7	0.8	37.6	90.2	29.5	230.0

2016	Reported	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Adjusted
Income tax expense(h)	–77.0								–51.2
Net (loss) profit	–202.8								178.8

(a) Represents an adjustment to exclude (i) acquisition-related amortisation expense … (ii) accelerated depreciation …related to the closure of certain manufacturing facilities, and (iii) impairment charges and assets write offs related to property, plant and equipment and intangible assets ….

(b) Represents restructuring costs and other related costs …

(c) Represents remediation costs which include regulatory compliance costs related to FDA activities, IT enhancement costs, and professional service fees associated with activities that were undertaken in respect of the Group's compliance function and to strengthen its control environment within finance.

(d) Represents costs primarily related to (i) corporate development activities and (ii) a settlement of ordinary course multi-year patent-related litigations in 2015 ….

(e) Represents adjustments to exclude (i) loss on extinguishment of debt and write off of deferred financing fees … and (ii) foreign exchange related transactions.

(f) Represents an adjustment to exclude (i) share-based compensation expense … arising from pre-IPO employee equity grants and (ii) pre-IPO ownership structure related costs, including management fees to Nordic Capital and Avista (refer to Note 6 Related Party Transactions for further information).

(g) Represents IPO related costs, primary advisory fees.

(h) Adjusted income tax expense/benefit is income tax (expense) benefit net of tax adjustments.

Adjusted EBITDA [Excerpt]

Adjusted EBITDA is defined as Adjusted EBIT…further adjusted to exclude (i) software and R&D amortisation, (ii) depreciation and (iii) post-IPO share-based compensation.

The following table reconciles the Group's Adjusted EBIT to Adjusted EBITDA.

	2016 ($m)
Adjusted EBIT	472.2
Software and R&D amortization	6.7
Depreciation	27.9
Post-IPO share-based compensation	0.8
Adjusted EBITDA	507.6

1 Based on the information provided, explain the differences between the following two disclosures contained in Convatec's press release:

- **A** The CEO Review of 2016 results, at the beginning of the release, states that "revenue grew 4% to $1,688 million."
- **B** Convatec's Consolidated Statement of Profit or Loss shows 2016 revenues of $1,688.3 million and 2015 revenues of $1,650.4 million.

2 Based on the information provided, explain the differences between the following two disclosures contained in Convatecs's earnings release:

- **A** The CEO Review of 2016 results states that "adjusted EBITDA was $508 million, up 6.5% at constant currency."
- **B** Convatec's Consolidated Statement of Profit or Loss shows 2016 net loss of $202.8 million and 2015 net loss of $93.4 million.

Solution to 1:

The amount of revenue reported on the company's income statement conforms to International Financial Reporting Standards (IFRS). Using the amounts from the income statement, the company's total revenue increased by 2.3 % (= \$1,688.3/\$1,650.4 − 1). The revenue growth rate of 4% in the CEO review is a non-IFRS measure, calculated on a "constant currency" basis, which the footnote describes as a comparison using 2016 revenues restated at 2015 foreign exchange rates.

Solution to 2:

The amounts reported on the company's income statement conform to IFRS. Using amounts from the income statement, the company reported a loss in 2016 of \$202.8 million, which was more than twice as large a loss as the \$93.4 million loss reported in 2015. Also referring to the income statement, the company reported 2016 operating profit (referred to elsewhere as EBIT) of \$154.0 million, a decline of 33.2% from the \$230.4 million operating profit reported in 2016.

In contrast, the "Adjusted EBITDA" amount highlighted in the CEO Review is neither defined nor recognized under IFRS. It is a non-IFRS measure. To create the Adjusted EBITDA, the company first begins with EBIT (called Operating profit in excerpts II and III) of \$154.0 and creates Adjusted EBIT (\$472.2 million) by adding back 8 different expenses that IFRS requires the company to recognize. These adjustments are listed beneath the first tabular reconciliation in Items a through g. After developing Adjusted EBIT, the company creates Adjusted EBITDA (\$507.6 million) by adding back a further 3 different expenses that IFRS requires the company to recognize.

Overall, there are three key differences between Disclosures A and B: (1) Most importantly, disclosure A refers to a non-IFRS metric rather than an IFRS-compliant metric; (2) Disclosure A refers to operating profit, which was positive, rather than to net income, which was negative; and (3) Disclosure A highlights a positive economic outcome—i.e., an increase, on a currency adjusted basis. An analyst should be aware of the alternative means by which earnings announcements can paint a positive picture of companies' results.

Often, poor reporting quality occurs simultaneously with poor earnings quality; for example, aggressive accounting choices are made to obscure poor performance. It is also possible, of course, for poor reporting quality to occur with high-quality earnings. Although a company with good performance would not require aggressive accounting choices to obscure poor performance, it might nonetheless produce poor-quality reports for other reasons. A company with good performance might be unable to produce high-quality reports because of inadequate internal systems.

Another scenario in which poor reporting quality might occur simultaneously with high quality earnings is that a company with good performance might deliberately produce reports based on "conservative" rather than aggressive accounting choices—that is, choices that make current performance look worse. One motivation might be to avoid unwanted political attention. Another motivation could arise in a period in which management had already exceeded targets before the end of the period and thus made conservative accounting choices that would delay reporting profits until the following period (so-called "hidden reserves"). Similar motivations might also contribute to accounting choices that create the appearance that the trajectory of future results would appear more attractive. For example, a company might make choices to accelerate losses in the first year of an acquisition or the first year of a new CEO's tenure so that the trajectory of future results would appear more attractive.

Overall, *unbiased* financial reporting is the ideal. Some investors may prefer conservative choices rather than aggressive ones, however, because a positive surprise is easier to tolerate than a negative surprise. Biased reporting, whether conservative or aggressive, adversely affects a user's ability to assess a company.

The quality spectrum considers the more intuitive situation in which less-than-desired underlying economics are the central motivation for poor reporting quality. In addition, it is necessary to have some degree of reporting quality in order to evaluate earnings quality. Proceeding down the spectrum, therefore, the concepts of reporting quality and earnings quality become progressively less distinguishable.

3.1 Within GAAP, but "Earnings Management"

The next level down on the spectrum in Exhibit 2 is labeled "Within GAAP, but 'earnings management.'" The term "earnings management" is defined here as making intentional choices that create biased financial reports.[8] The distinction between earnings management and biased choices is subtle and, primarily, a matter of intent. Earnings management represents "deliberate actions to influence reported earnings and their interpretation" (Ronen and Yaari, 2008). Earnings can be "managed" upward (increased) by taking *real* actions, such as deferring research and development (R&D) expenses into the next reporting period. Alternatively, earnings can be increased by *accounting* choices, such as changing accounting estimates. For example, the amount of estimated product returns, bad debt expense, or asset impairment could be decreased to create higher earnings. Because it is difficult to determine intent, we include earnings management under the biased choices discussion.

DEPARTURES FROM GAAP

4

b describe a spectrum for assessing financial reporting quality;

The next levels down on the spectrum in Exhibit 2 mark departures from GAAP. Financial reporting that departs from GAAP can generally be considered low quality. In such situations, earnings quality is likely difficult or impossible to assess because comparisons with earlier periods and/or other entities cannot be made. An example of improper accounting was Enron (accounting issues revealed in 2001), whose inappropriate use of off-balance-sheet structures and other complex transactions resulted in vastly understated indebtedness as well as overstated profits and operating cash flow. Another notorious example of improper accounting was WorldCom (accounting issues discovered in 2002), a company that by improperly capitalizing certain expenditures dramatically understated its expenses and thus overstated its profits. More recently, New Century Financial (accounting issues revealed in 2007) issued billions of dollars of subprime mortgages and improperly reserved only minimal amounts for loan repurchase losses. Each of these companies subsequently filed for bankruptcy.

---

[8] Various definitions have appeared in academic research. Closest to the discussion here is Schipper (1989), which uses the term "earnings management" to mean "'disclosure management' in the sense of a purposeful intervention in the external financial reporting process, with the intent of obtaining some private gain (as opposed to, say, merely facilitating the neutral operation of the process)."

In the 1980s, Polly Peck International (PPI) reported currency losses, incurred in the normal course of operations, directly through equity rather than in its profit and loss statements. In the 1990s, Sunbeam improperly reported revenues from "bill-and-hold" sales and also manipulated the timing of expenses in an effort to falsely portray outstanding performance of its then-new chief executive.

At the bottom of the quality spectrum, fabricated reports portray fictitious events, either to fraudulently obtain investments by misrepresenting the company's performance and/or to obscure fraudulent misappropriation of the company's assets. Examples of fraudulent reporting are unfortunately easy to find, although they were not necessarily easy to identify at the time. In the 1970s, Equity Funding Corp. created fictitious revenues and even fictitious policyholders. In the 1980s, Crazy Eddie's reported fictitious inventory as well as fictitious revenues supported by fake invoices. In 2004, Parmalat reported fictitious bank balances.

EXAMPLE 3

Spectrum for Assessing Quality of Financial Reports

Jake Lake, a financial analyst, has identified several items in the financial reports of several (hypothetical) companies. Describe each of these items in the context of the financial reporting quality spectrum.

1. ABC Co.'s 2018 earnings totaled \$233 million, including a \$100 million gain from selling one of its less profitable divisions. ABC's earnings for the prior three years totaled \$120 million, \$107 million, and \$111 million. The company's financial reports are extremely clear and detailed, and the company's earnings announcement highlights the one-time nature of the \$100 million gain.
2. DEF Co. discloses that in 2018, it changed the depreciable life of its equipment from 3 years to 15 years. Equipment represents a substantial component of the company's assets. The company's disclosures indicate that the change is permissible under the accounting standards of its jurisdiction but provide only limited explanation of the change.
3. GHI Co.'s R&D expenditures for the past five years have been approximately 3% of sales. In 2018, the company significantly reduced its R&D expenditures. Without the reduction in R&D expenditures, the company would have reported a loss. No explanation is disclosed.

Solution to 1:

ABC's 2018 total earnings quality can be viewed as low because nearly half of the earnings are derived from a non-sustainable activity, namely the sale of a division. ABC's 2018 quality of earnings from continuing operations may be high because the amounts are fairly consistent from year to year, although an analyst would undertake further analysis to confirm earnings quality. In general, a user of financial reports should look beyond the bottom-line net income. The description provided suggests that the company's reporting quality is high; the reports are clear and detailed, and the one-time nature of the \$100 million gain is highlighted.

Solution to 2:

DEF's accounting choice appears to be within permissible accounting standards, but its effect is to substantially lower depreciation expense and thus to increase earnings for the year. The quality of reported earnings is questionable. Although the new level of earnings may be sustainable, similar increases in earnings for future periods might not be achievable, because increasing earnings solely by changing accounting estimates is likely not sustainable. In addition, the description provided suggests that the company's reporting quality is low because it offers only a limited explanation for the change.

Solution to 3:

GHI's operational choice to reduce its R&D may reflect real earnings management because the change enabled the company to avoid reporting a loss. In addition, the description provided suggests that the company's reporting quality is low because it does not offer an explanation for the change.

DIFFERENTIATE BETWEEN CONSERVATIVE AND AGGRESSIVE ACCOUNTING

5

c explain the difference between conservative and aggressive accounting;

This section returns to the implications of conservative and aggressive accounting choices. As mentioned earlier, *unbiased* financial reporting is the ideal. But some investors may prefer or be perceived to prefer conservative rather than aggressive accounting choices because a positive surprise is acceptable. In contrast, management may make, or be perceived to make, aggressive accounting choices because they increase the company's reported performance and financial position.

Aggressive accounting choices in the period under review may decrease the company's reported performance and financial position in later periods, which creates a sustainability issue. Conservative choices do not typically create a sustainability issue because they decrease the company's reported performance and financial position, and may increase them in later periods. In terms of establishing expectations for the future, however, financial reporting that is relevant and faithfully representative is the most useful.

A common presumption is that financial reports are typically biased upward, but that is not always the case. Although accounting standards ideally promote unbiased financial reporting, some accounting standards may specifically require a conservative treatment of a transaction or an event. Also, managers may choose to take a conservative approach when applying standards. It is important that an analyst consider the possibility of conservative choices and their effects.

At its most extreme, conservatism follows accounting practices that "anticipate no profit, but anticipate all losses" (Bliss, 1924). But in general, conservatism means that revenues may be recognized once a verifiable and legally enforceable receivable has been generated and that losses need not be recognized until it becomes "probable" that an actual loss will be incurred. Conservatism is not an absolute but is characterized by degrees, such as "the accountant's tendency to require a higher degree of verification to recognize good news as gains than to recognize bad news as losses" (Basu, 1997). From this perspective, "verification" (e.g., physical existence of inventories, evidence of costs incurred or to be incurred, or establishment of rights and obligations on legal grounds) drives the degree of conservatism. For recognition of revenues, a higher degree of verification would be required than for expenses.

5.1 Conservatism in Accounting Standards

The *Conceptual Framework* supports neutrality of information: "A neutral depiction is without bias in the selection or presentation of financial information."[9] Neutrality—lack of upward or downward bias—is considered a desirable characteristic of financial reporting. Conservatism directly conflicts with the characteristic of neutrality because the asymmetric nature of conservatism leads to bias in measuring assets and liabilities—and ultimately, earnings.

Despite efforts to support neutrality in financial reporting, many conservatively biased standards remain. Standards across jurisdictions may differ on the extent of conservatism embedded within them. An analyst should be aware of the implications of accounting standards for the financial reports.

An example is the different treatment by IFRS and US GAAP of the impairment of long-lived assets.[10] Both IFRS and US GAAP specify an impairment analysis protocol that begins with an assessment of whether recent events indicate that the economic benefit from an individual or group of long-lived assets may be less than its carrying amount(s). From that point on, however, the two regimes diverge:

- Under IFRS, if the "recoverable amount" (the higher of fair value less costs to sell and value in use) is less than the carrying amount, then an impairment charge will be recorded.
- Under US GAAP, an impairment charge will be recorded only when the sum of the undiscounted future cash flows expected to be derived from the asset(s) is less than the carrying amount(s). If the undiscounted future cash flows are less than the carrying amount, the asset is written down to fair value.

To illustrate the difference in application, assume that a factory is the unit of account eligible for impairment testing. Its carrying amount is $10,000,000; "fair value" and "recoverable amount" are both $6,000,000; and the undiscounted future net cash flows associated with the factory total $10,000,000. Under IFRS, an impairment charge of $4,000,000 would be recorded; but under US GAAP, no impairment charge would be recognized.

Thus, on its face, IFRS would be regarded as more conservative than US GAAP because impairment losses would normally be recognized earlier under IFRS than under US GAAP. But, taking the analysis one step further, such a broad generalization may not hold up. For example, if an asset is impaired under both IFRS and US GAAP and the asset's value in use exceeds its fair value, the impairment loss under US GAAP will be greater. Also, IFRS permits the recognition of recoveries of the recoverable amount in subsequent periods if evidence indicates that the recoverable amount has subsequently increased. In contrast, US GAAP prohibits the subsequent write-up of an asset after an impairment charge has been taken; it would recognize the asset's increased value only when the asset is ultimately sold.

Common examples of conservatism in accounting standards include the following:

- *Research costs*. Because the future benefit of research costs is uncertain at the time the costs are incurred, both US GAAP and IFRS require immediate expensing instead of capitalization.

---

[9] IASB and FASB, *The Conceptual Framework for Financial Reporting* (2010): QC 14.

[10] See IAS 36 and FASB ASC Section 360-10-35.

- *Litigation losses*. When it becomes "probable" that a cost will be incurred, both US GAAP and IFRS require expense recognition, even though a legal liability may not be incurred until a future date.
- *Insurance recoverables*. Generally, a company that receives payment on an insurance claim may not recognize a receivable until the insurance company acknowledges the validity of the claimed amount.

Watts (2003) reviews empirical studies of conservatism, and identifies four potential benefits of conservatism:

- Given asymmetrical information, conservatism may protect the contracting parties with less information and greater risk. This protection is necessary because the contracting party may be at a disadvantage. For example, corporations that access debt markets have limited liability, and lenders thus have limited recourse to recover their losses from shareholders. As another example, executives who receive earnings-based bonuses might not be subject to having those bonuses "clawed back" if earnings are subsequently discovered to be overstated.
- Conservatism reduces the possibility of litigation and, by extension, litigation costs. Rarely, if ever, is a company sued because it understated good news or overstated bad news.
- Conservative rules may protect the interests of regulators and politicians by reducing the possibility that fault will be found with them if companies overstate earnings or assets.
- In many tax jurisdictions, financial and tax reporting rules are linked. For example, in Germany and Japan, only deductions taken against reported income can be deducted against taxable income. Hence, companies can reduce the present value of their tax payments by electing conservative accounting policies for certain types of events.

Analysts should consider possible conservative and aggressive biases and their consequences when examining financial reports. Current-period financial reports may be unbiased, upward biased through aggressive accounting choices, downward biased through conservative accounting choices, or biased through a combination of conservative and aggressive accounting choices.

5.2 Bias in the Application of Accounting Standards

Any application of accounting standards, whether the standard itself is neutral or not, often requires significant amounts of judgment. Characterizing the application of an accounting standard as conservative or aggressive is more a matter of intent rather than definition.

Careful analysis of disclosures, facts, and circumstances contributes to making an accurate inference of intent. Management seeking to manipulate earnings may take a longer view by sacrificing short-term profitability in order to ensure higher profits in later periods. One example of biased accounting in the guise of conservatism is the so-called "big bath" restructuring charges. Both US GAAP and IFRS provide for accrual of future costs associated with restructurings, and these costs are often associated with and presented along with asset impairments. But in some instances, companies use the accounting provisions to estimate "big" losses in the period under review so that performance in future periods will appear better. Having observed numerous instances of manipulative practices in the late 1990s, in which US companies set up opportunities to report higher profits in future periods that were not connected with performance in those periods, the SEC staff issued rules that narrowed the circumstances under which costs can be categorized as part of a "non-recurring" restructuring event and enhanced the transparency surrounding restructuring charges and asset impairments.[11]

A similar manifestation of "big bath" accounting is often referred to as "cookie jar reserve accounting." Both US GAAP and IFRS require accruals of estimates of future non-payments of loans. In his 1998 speech "The 'Numbers Game,'" SEC chair Arthur Levitt expressed the general concern that corporations were overstating loans and other forms of loss allowances for the purpose of smoothing income over time.[12] In 2003, the SEC issued interpretive guidance that essentially requires a company to provide a separate section in management's discussion and analysis (MD&A) titled "Critical Accounting Estimates."[13] If the effects of subjective estimates and judgments of highly uncertain matters are material to stakeholders (investors, customers, suppliers, and other users of the financial statements), disclosures of their nature and exposure to uncertainty should be made in the MD&A. This requirement is in addition to required disclosures in the notes to the financial statements.

6

CONTEXT FOR ASSESSING FINANCIAL REPORTING QUALITY: MOTIVATIONS AND CONDITIONS CONDUCIVE TO ISSUING LOW QUALITY FINANCIAL REPORTS

d describe motivations that might cause management to issue financial reports that are not high quality;

e describe conditions that are conducive to issuing low-quality, or even fraudulent, financial reports;

In assessing financial reporting quality, it is useful to consider whether a company's managers may be motivated to issue financial reports that are not high quality. If motivation exists, an analyst should consider whether the reporting environment is conducive to managers' misreporting. It is important to consider mechanisms within the reporting environment that discipline financial reporting quality, such as the regulatory regime.

---

[11] SEC, "Restructuring and Impairment Charges," Staff Accounting Bulletin (SAB) No. 100 (1999): www.sec.gov/interps/account/sab100.htm.

[12] Arthur Levitt, "The 'Numbers Game,'" Remarks given at NYU Center for Law and Business (28 September 1998): www.sec.gov/news/speech/speecharchive/1998/spch220.txt.

[13] SEC, "Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations," Financial Reporting Release (FRR) No. 72 (2003): www.sec.gov/rules/interp/33-8350.htm.

6.1 Motivations

Managers may be motivated to issue financial reports that are not high quality to mask poor performance, such as loss of market share or lower profitability than competitors. Lewis (2012) stated, "A firm experiencing performance problems, particularly those it considers transient, may induce a response that inflates current earnings numbers in exchange for lower future earnings."

- Even when there is no need to mask poor performance, managers frequently have incentives to meet or beat market expectations as reflected in analysts' forecasts and/or management's own forecasts. Exceeding forecasts typically increases the stock price, if only temporarily. Additionally, exceeding forecasts can increase management compensation that is linked to increases in stock price or to reported earnings. Graham, Harvey, and Rajgopal (2005) found that the CFOs they surveyed view earnings as the most important financial metric to financial markets. Achieving (or exceeding) particular benchmarks, including prior-year earnings and analysts' forecasts, is very important. The authors examined a variety of motivations for why managers might "exercise accounting discretion to achieve some desirable earnings goal." Motivations to meet earnings benchmarks include equity market effects (for example, building credibility with market participants and positively affecting stock price) and trade effects (for example, enhancing reputation with customers and suppliers). Equity market effects are the most powerful incentives, but trade effects are important, particularly for smaller companies.
- Career concerns and incentive compensation may motivate accounting choices. For example, managers might be concerned that working for a company that performs poorly will limit their future career opportunities or that they will not receive a bonus based on exceeding a particular earnings target. In both cases, management might be motivated to make accounting choices to increase earnings. In a period of marginally poor performance, a manager might accelerate or inflate revenues and/or delay or under report expenses. Conversely, in a period of strong performance, a manager might delay revenue recognition or accelerate expense recognition to increase the probability of exceeding the next period's targets (i.e., to "bank" some earnings for the next period.) The surveyed managers indicated a greater concern with career implications of reported results than with incentive compensation implications.

Avoiding debt covenant violations can motivate managers to inflate earnings. Graham, Harvey, and Rajgopal's survey indicates that avoidance of bond covenant violation is important to highly leveraged and unprofitable companies but relatively unimportant overall.

6.2 Conditions Conducive to Issuing Low-Quality Financial Reports

As discussed, deviations from a neutral presentation of financial results could be driven by management choices or by a jurisdiction's financial reporting standards. Ultimately, a decision to issue low-quality, or even fraudulent, financial reports is made by an individual or individuals. Why individuals make such choices is not always immediately apparent. For example, why would the newly appointed CEO of Sunbeam, who already had a net worth of more than $100 million, commit accounting fraud by improperly reporting revenues from "bill-and-hold" sales and manipulating the timing of expenses, rather than admit to lower-than-expected financial results?

Typically, three conditions exist when low-quality financial reports are issued: opportunity, motivation, and rationalization. Opportunity can be the result of internal conditions, such as poor internal controls or an ineffective board of directors, or external conditions, such as accounting standards that provide scope for divergent choices or minimal consequences for an inappropriate choice. Motivation can result from pressure to meet some criteria for personal reasons, such as a bonus, or corporate reasons, such as concern about financing in the future. Rationalization is important because if an individual is concerned about a choice, he or she needs to be able to justify it to him- or herself.

Former Enron CFO Andrew Fastow, speaking at the 2013 Association of Certified Fraud Examiners Annual Fraud Conference, indicated that he knew at the time he was doing something wrong but followed procedure to justify his decision (Pavlo, 2013). He made sure to get management and board approval, as well as legal and accounting opinions, and to include appropriate disclosures. The incentive and corporate culture was to create earnings rather than focus on long-term value. Clearly, as reflected in his prison sentence, he did something that was not only wrong but illegal.

7

MECHANISMS THAT DISCIPLINE FINANCIAL REPORTING QUALITY

f describe mechanisms that discipline financial reporting quality and the potential limitations of those mechanisms;

Markets potentially discipline financial reporting quality. Companies and nations compete for capital, and the cost of capital is a function of perceived risk—including the risk that a company's financial statements will skew investors' expectations. Thus, in the absence of other conflicting economic incentives, a company seeking to minimize its long-term cost of capital should aim to provide high-quality financial reports. In addition to markets, other mechanisms that discipline financial reporting quality include market regulatory authorities, auditors, and private contracts.

7.1 Market Regulatory Authorities

Companies seeking to minimize the cost of capital should maximize reporting quality, but as discussed earlier, conflicting incentives often exist. For this reason, national regulations, and the regulators that establish and enforce rules, can play a significant role in financial reporting quality. Many of the world's securities regulators are members of the International Organization of Securities Commissions (IOSCO). IOSCO is recognized as the "global standard setter for the securities sector" although it does not actually set standards but rather establishes objectives and principles to guide securities and capital market regulation. IOSCO's membership includes more than 120 securities regulators and 80 other securities market participants, such as stock exchanges.[14]

One member of IOSCO is ESMA , an independent EU authority with a mission to "enhance the protection of investors and reinforce stable and well-functioning financial markets in the European Union."[15] ESMA organizes financial reporting enforcement activities through a forum consisting of European enforcers from European Economic Area countries. Direct supervision and enforcement activities are performed at the national level. For example, the Financial Conduct Authority (FCA) is the IOSCO member with primary responsibility for securities regulation in the United Kingdom. ESMA reported that European enforcers examined the interim and/or annual financial statements of 1,141 issuers in 2017, which in turn led to enforcement actions for 328 issuers with the following outcomes: 12 required reissuances of financial statements, 71 public corrective notes, and 245 required corrections in future financial statements.[16]

---

[14] Visit www.iosco.org for more information.

[15] Text from ESMA's mission statement on their website: www.esma.europa.eu.

Another member of IOSCO is the US regulatory authority, the Securities and Exchange Commission. The SEC is responsible for overseeing approximately 9,100 US public companies (along with investment advisers, broker/dealers, securities exchanges, and other entities) and reviews the disclosures of these companies at least once every three years with the aim of improving information available to investors and potentially uncovering possible violations of securities laws.[17] In 2017, the SEC reported that it had filed 754 total and 446 standalone enforcement actions, about 20% of which concerned issuer reporting/accounting and auditing.[18]

Examples of regulatory bodies in Asia include the Financial Services Agency in Japan, the China Securities Regulatory Commission, and the Securities and Exchange Board of India. Examples of regulatory bodies in South America include the Comisión Nacional de Valores in Argentina, Comissão de Valores Mobiliários in Brazil, and Superintendencia de Valores y Seguros in Chile. A full list of IOSCO members can be found on the organization's website.

Typical features of a regulatory regime that most directly affect financial reporting quality include the following:

- *Registration requirements*. Market regulators typically require publicly traded companies to register securities before offering the securities for sale to the public. A registration document typically contains current financial statements, other relevant information about the risks and prospects of the company issuing the securities, and information about the securities being offered.
- *Disclosure requirements*. Market regulators typically require publicly traded companies to make public periodic reports, including financial reports and management comments. Standard-setting bodies, such as the IASB and FASB, are typically private sector, self-regulated organizations with board members who are experienced accountants, auditors, users of financial statements, and academics. Regulatory authorities, such as the Accounting and Corporate Regulatory Authority in Singapore, the Securities and Exchange Commission in the United States, the Securities and Exchange Commission in Brazil, and the Financial Reporting Council in the United Kingdom, have the legal authority to enforce financial reporting requirements and exert other controls over entities that participate in the capital markets within their jurisdiction. In other words, *generally*, standard-setting bodies set the standards, and regulatory authorities recognize and enforce those standards. Without the recognition of standards by regulatory authorities, the private-sector standard-setting bodies would have no authority. Regulators often retain the legal authority to establish financial reporting standards in their jurisdiction and can overrule the private-sector standard-setting bodies.

---

[16] ESMA, "Enforcement and Regulatory Activities of Accounting Enforcers in 2017," ESMA32-63-424, European Securities and Markets Authority (03 April 2018): www.esma.europa.eu.

[17] SEC, "FY2013 Congressional Justification," Securities and Exchange Commission (February 2012): www.sec.gov/about/secfy13congbudgjust.pdf.

[18] SEC, Securities and Exchange Commission Division of Enforcement Annual Report, "A Look Back at Fiscal Year 2017" www.sec.gov/report.

- *Auditing requirements*. Market regulators typically require companies' financial statements to be accompanied by an audit opinion attesting that the financial statements conform to the relevant set of accounting standards. Some regulators, such as the SEC in the United States, require an additional audit opinion attesting to the effectiveness of the company's internal controls over financial reporting.
- *Management commentaries*. Regulations typically require publicly traded companies' financial reports to include statements by management. For example, the FCA in the United Kingdom requires a management report containing "(1) a fair review of the issuer's business; and (2) a description of the principal risks and uncertainties facing the issuer" (Disclosure Guidance and Transparency Rules sourcebook.)
- *Responsibility statements*. Regulations typically require a statement from the person or persons responsible for the company's filings. Such statements require the responsible individuals to explicitly acknowledge responsibility and to attest to the correctness of the financial reports. Some regulators, such as the SEC in the United States, require formal certifications that carry specific legal penalties for false certifications.
- *Regulatory review of filings*. Regulators typically undertake a review process to ensure that the rules have been followed. The review process typically covers all initial registrations and a sample of subsequent periodic financial reports.
- *Enforcement mechanisms*. Regulators are granted various powers to enforce the securities market rules. Such powers can include assessing fines, suspending or permanently barring market participants, and bringing criminal prosecutions. Public announcements of disciplinary actions are also a type of enforcement mechanism.

In summary, market regulatory authorities play a central role in encouraging high-quality financial reporting.

7.2 Auditors

As noted, regulatory authorities typically require that publicly traded companies' financial statements be audited by an independent auditor. Private companies also obtain audit opinions for their financial statements, either voluntarily or because audit reports are required by an outside party, such as providers of debt or equity capital.

Audit opinions provide financial statement users with some assurance that the information complies with the relevant set of accounting standards and presents the company's information fairly. Exhibits 8, 9, 10, and 11 provide excerpts from the independent auditors' reports for GlaxoSmithKline plc, Alibaba Group Holding Limited, Apple Inc., and Tata Motors Limited, respectively. For each company, the auditor issued an unqualified opinion on the financial statements, indicating that the financial statements present fairly the company's performance in accordance with relevant standards. (Note: The term "unqualified opinion" means that the opinion did not include any qualifications or exceptions; the term is synonymous with the less formal term "clean opinion." Unqualified opinions are the most common.) Other items in the audit reports reflect the specific requirements of the company's regulatory regime. For example, the audit report for GlaxoSmithKline spans nine pages and includes opinions on the company's financial statements as well as the Strategic Report and the Directors' Report. This audit report also includes disclosures about "Key audit matters," in accordance with International Standards on Auditing (ISAs) issued by the International Auditing and Assurance Standards Board (IAASB) in 2015 and effective for periods ending on or after December 15, 2016.

The excerpts for Alibaba, Apple, and Tata Motors show the auditors' opinions on the companies' financial statements and additionally the SEC-required opinions on the effectiveness of the companies' internal controls because these companies are listed in the United States. For Alibaba, a single report includes both unqualified opinions: (i) the financial statements present fairly the financial position, results of operations, and cash flows… in conformity with US GAAP; and (ii) the company maintained effective control over financial reporting. For Apple, the first report includes the unqualified opinion on the financial statements, and the second report includes the unqualified opinion on the company's effective internal controls. For Tata Motors, the first report includes the unqualified opinion that the financial statements present the company's position and results fairly in accordance with IFRS. (The SEC permits non-US companies to report using US GAAP, IFRS as issued by the IASB, or home-country GAAP.) However, the second report includes an *adverse* opinion on the effectiveness of the company's internal controls: "In our opinion, because of the effect of the material weakness … the company has not maintained effective internal control." The report explains that the material weakness involved a third party's inappropriate access to the company's systems. The report further states that although the material weakness resulted in ineffective internal controls, it did not affect the audit opinion on the financial statements. Elsewhere in Tata Motors' annual report (not shown in the excerpt), the company discloses that the weakness did not result in a financial misstatement and that it has undertaken remedial measures.

Exhibit 8 Excerpts from Audit Opinion of PricewaterhouseCoopers LLP from the 2017 Annual Report *(pages 149–157)* of GlaxoSmithKline plc

In our opinion, GlaxoSmithKline plc's Group financial statements (the "financial statements"):

- give a true and fair view of the state of the Group's affairs as at 31 December 2017 and of its profit and cash flows for the year then ended;
- have been properly prepared in accordance with International Financial Reporting Standards ("IFRSs") as adopted by the European Union; and
- have been prepared in accordance with the requirements of the Companies Act 2006 and Article 4 of the IAS Regulation.

…

In our opinion, the Group financial statements have been properly prepared in accordance with IFRSs as issued by the IASB.

…

Key audit matters

Key audit matters are those matters that, in the auditors' professional judgement, were of most significance in the audit of the financial statements of the current period and include the most significant assessed risks of material misstatement (whether or not due to fraud) identified by the auditors, including those which had the greatest effect on: the overall audit strategy; the allocation of resources in the audit; and directing the efforts of the engagement team. These matters, and any comments we make on the results of our procedures thereon, were addressed in the context of our audit of the financial statements as a whole, and in forming our opinion thereon, and we do not provide a separate opinion on these matters. This is not a complete list of all risks identified by our audit.

…

(continued)

Exhibit 8 (Continued)

In our opinion, based on the work undertaken in the course of the audit, the information given in the Strategic Report and Directors' Report for the year ended 31 December 2017 is consistent with the financial statements and has been prepared in accordance with applicable legal requirements.

Exhibit 9 Excerpts from Audit Opinion of PricewaterhouseCoopers Hong Kong, SAR from the Annual Report ***(SEC Form 20-F, Pages F-2 and F-3)*** **of Alibaba Group Holding Limited for the year ended March 31, 2018**

In our opinion, the consolidated financial statements referred to above present fairly, in all material respects, the financial position of the Company as of March 31, 2017 and 2018, and the results of their operations and their cash flows for each of the three years in the period ended March 31, 2018 in conformity with accounting principles generally accepted in the United States of America. Also in our opinion, the Company maintained, in all material respects, effective internal control over financial reporting as of March 31, 2018, based on criteria established in Internal Control — Integrated Framework (2013) issued by the COSO.

Exhibit 10 Excerpt from Audit Opinion of Ernst & Young from the Annual Report (***SEC Form 10-K, pages 70 and 71*****) of Apple Inc. for the year ended September 30, 2017**

[From the Financial Statement Opinion]

We have audited the accompanying consolidated balance sheets of Apple Inc. as of September 30, 2017 and September 24, 2016, and the related consolidated statements of operations, comprehensive income, shareholders' equity and cash flows for each of the three years in the period ended September 30, 2017.

....

In our opinion, the financial statements referred to above present fairly, in all material respects, the consolidated financial position of Apple Inc. at September 30, 2017 and September 24, 2016, and the consolidated results of its operations and its cash flows for each of the three years in the period ended September 30, 2017, in conformity with U.S. generally accepted accounting principles.

...

We also have audited, in accordance with the standards of the Public Company Accounting Oversight Board (United States), Apple Inc.'s internal control over financial reporting as of September 30, 2017, based on criteria established in Internal Control – Integrated Framework issued by the Committee of Sponsoring Organizations of the Treadway Commission (2013 framework) and our report dated November 3, 2017 expressed an unqualified opinion thereon.

[From the Internal Controls Opinion]

We have audited Apple Inc.'s internal control over financial reporting as of September 30, 2017, based on criteria established in Internal Control – Integrated Framework issued by the Committee of Sponsoring Organizations of the Treadway Commission (2013 framework) ("the COSO criteria").

Exhibit 10 (Continued)

…

In our opinion, Apple Inc. maintained, in all material respects, effective internal control over financial reporting as of September 30, 2017, based on the COSO criteria.

We also have audited, in accordance with the standards of the Public Company Accounting Oversight Board (United States), the 2017 consolidated financial statements of Apple Inc. and our report dated November 3, 2017 expressed an unqualified opinion thereon.

Exhibit 11 Excerpt from Audit Opinion of KPMG Mumbai, India from the Annual Report (*SEC Form 20-F, pages F2 to F4*) of Tata Motors Limited for the year ended March 31, 2018

Opinion on the Consolidated Financial Statements

We have audited the accompanying consolidated balance sheet of Tata Motors Limited and its subsidiaries (the "Company") as of March 31, 2018, the related consolidated income statement, statement of comprehensive income, statement of cash flows, and statement of changes in equity for the year ended March 31, 2018, and the related notes and financial statement schedule 1 (collectively, the consolidated financial statements).

In our opinion, the consolidated financial statements present fairly, in all material respects, the financial position of the Company as of March 31, 2018, and the results of its operations and its cash flows for the year ended March 31, 2018, in conformity with the International Financial Reporting Standards as issued by the International Accounting Standards Board ("IFRS").

We also have audited, in accordance with the standards of the Public Company Accounting Oversight Board (United States) (PCAOB), the Company's internal control over financial reporting as of March 31, 2018, based on criteria established in *Internal Control – Integrated Framework (2013)* issued by the Committee of Sponsoring Organizations of the Treadway Commission, and our report dated July, 31, 2018 expressed an adverse opinion on the effectiveness of the Company's internal control over financial reporting.

…

Opinion on Internal Control Over Financial Reporting

We have audited Tata Motors Limited's and subsidiaries' (the Company) internal control over financial reporting as of March 31, 2018, based on criteria established in Internal Control – Integrated Framework (2013) issued by the Committee of Sponsoring Organizations of the Treadway Commission. In our opinion, because of the effect of the material weakness described below, on the achievement of the objectives of the control criteria, the Company has not maintained effective internal control over financial reporting as of March 31, 2018, based on criteria established in *Internal Control – Integrated Framework (2013)* issued by the Committee of Sponsoring Organizations of the Treadway Commission.

…

A material weakness is a deficiency, or a combination of deficiencies, in internal control over financial reporting, such that there is a reasonable possibility that a material misstatement of the company's annual or interim financial statements will not be prevented or detected on a timely basis. A material weakness related to inappropriate system access restrictions at a third party logistics provider

(continued)

Exhibit 11 (Continued)

has been identified and included in management's assessment. The material weakness was considered in determining the nature, timing, and extent of audit tests applied in our audit of the 2018 consolidated financial statements, and this report does not affect our report on those consolidated financial statements.

Although audit opinions provide discipline for financial reporting quality, inherent limitations exist. First, an audit opinion is based on a review of information prepared by the company. If a company deliberately intends to deceive its auditor, a review of information might not uncover misstatements. Second, an audit is based on sampling, and the sample might not reveal misstatements. Third, an "expectations gap" may exist between the auditor's role and the public's expectation of auditors. An audit is not typically intended to detect fraud; it is intended to provide assurance that the financial reports are fairly presented. Finally, the company being audited pays the audit fees, often established through a competitive process. This situation could provide an auditor with an incentive to show leniency to the company being audited, particularly if the auditor's firm provides additional services to the company.

7.3 Private Contracting

Aspects of private contracts, such as loan agreements or investment contracts, can serve as mechanisms to discipline financial reporting quality. Many parties that have a contractual arrangement with a company have an incentive to monitor that company's performance and to ensure that the company's financial reports are high quality. For example, loan agreements often contain loan covenants, which create specifically tailored financial reporting requirements that are legally binding for the issuer. As noted earlier, avoidance of debt covenant violation is a potential motivation for managers to inflate earnings. As another example, an investment contract could contain provisions giving investors the option to recover all or part of their investment if certain financial triggers occur. Such provisions could motivate the investee's managers to manipulate reported results to avoid the financial triggers.

Because the financial reports prepared by the investees or borrowers directly affect the contractual outcomes—potentially creating a motivation for misreporting—investors and lenders are motivated to monitor financial reports and to ensure that they are high quality.

EXAMPLE 4

Financial Reporting Manipulation: Motivations and Disciplining Mechanisms

For each of the following two scenarios, identify (1) factors that might motivate the company's managers to manipulate reported financial amounts and (2) applicable mechanisms that could discipline financial reporting quality.

1 ABC Co. is a private company. Bank NTBig has made a loan to ABC Co. ABC is required to maintain a minimum 2.0 interest coverage ratio. In its most recent financial reports, ABC reported earnings before interest and taxes of $1,200 and interest expense of $600. In the report's notes, the

company discloses that it changed the estimated useful life of its property, plant, and equipment during the year. Depreciation was approximately $150 lower as a result of this change in estimate.

2 DEF Co. is a publicly traded company. For the most recent quarter, the average of analysts' forecasts for earnings per share was $2.50. In its quarterly earnings announcement, DEF reported net income of $3,458,780. The number of common shares outstanding was 1,378,000. DEF's main product is a hardware device that includes a free two-year service contract in the selling price. Based on management estimates, the company allocates a portion of revenues to the hardware device, which it recognizes immediately, and a portion to the service contract, which it defers and recognizes over the two years of the contract. Based on the disclosures, a higher percentage of revenue was allocated to hardware than in the past, with an estimated after-tax impact on net income of $27,000.

Solution to 1:

The need to maintain a minimum interest coverage ratio of 2.0 might motivate ABC's managers to manipulate reported financial amounts. The company's coverage ratio based on the reported amounts is exactly equal to 2.0. If ABC's managers had not changed the estimated useful life of the property, plant, and equipment, the coverage ratio would have fallen below the required level.

EBIT, as reported	$1,200
Impact on depreciation expense of changed assumptions about useful life	150
EBIT, as adjusted	$1,050
Interest expense	$600
Coverage ratio, as reported	2.00
Coverage ratio, as adjusted	1.75

The potential disciplining mechanisms include the auditors, who will assess the reasonableness of the depreciable lives estimates. In addition, the lenders will carefully scrutinize the change in estimate because the company only barely achieved the minimum coverage ratio and would not have achieved the minimum without the change in accounting estimate.

Solution to 2:

The desire to meet or exceed the average of analysts' forecasts for earnings per share might motivate DEF Co.'s managers to manipulate reported financial amounts. As illustrated in the following calculations, the impact of allocating a greater portion of revenue to hardware enabled the company to exceed analysts' earnings per share forecasts by $0.01.

Net income, as reported	$3,458,780
Impact on gross profit of changed revenue recognition, net of tax	27,000
Net income, as adjusted	$3,431,780
Weighted average number of shares	1,378,000

(continued)

Earnings per share, as reported	$2.51
Earnings per share, as adjusted	$2.49

The potential disciplining mechanisms include the auditors, market regulators, financial analysts, and financial journalists.

8

DETECTION OF FINANCIAL REPORTING QUALITY ISSUES: INTRODUCTION & PRESENTATION CHOICES

g describe presentation choices, including non-GAAP measures, that could be used to influence an analyst's opinion;

h describe accounting methods (choices and estimates) that could be used to manage earnings, cash flow, and balance sheet items;

Choices in the application of accounting standards abound, which is perhaps one reason why accounting literature and texts are so voluminous. Compounding the complexity, measurement often depends on estimates of economic phenomena. Two estimates might be justifiable, but they may have significantly different effects on the company's financial statements. As discussed earlier, the choice of a particular estimate may depend on the motivations of the reporting company's managers. With many choices available, and the inherent flexibility of estimates in the accounting process, managers have many tools for managing and meeting analysts' expectations through financial reporting.

An understanding of the choices that companies make in financial reporting is fundamental to evaluating the overall quality—both financial reporting and earnings quality—of the reports produced. Choices exist both in how information is presented (financial reporting quality) and in how financial results are calculated (earnings quality). Choices in presentation (financial reporting quality) may be fairly transparent to investors. Choices in the calculation of financial results (earnings quality), however, are more difficult to discern because they can be deeply embedded in the construction of reported financial results.

The availability of accounting choices enables managers to affect the reporting of financial results. Some choices increase performance and financial position in the reporting period (aggressive choices), and others increase them in later periods (conservative choices). A manager that wants to increase performance and financial position in the reporting period could:

- Recognize revenue prematurely;
- Use non-recurring transactions to increase profits;
- Defer expenses to later periods;
- Measure and report assets at higher values; and/or
- Measure and report liabilities at lower values.

A manager that wants to increase performance and financial position in a later period could:

- Defer current income to a later period (save income for a "rainy day"); and/or
- Recognize future expenses in a current period, setting the table for improving future performance.

The following sections describe some of the potential choices for how information is presented and how accounting elements [assets, liabilities, owners' equity, revenue and gains (income), and expenses and losses] are recognized, measured, and reported. In addition to choices within GAAP, companies may prepare fraudulent reports. For example, these reports may include non-existent revenue or assets. These sections conclude with some of the warning signs that can indicate poor-quality financial reports.

8.1 Presentation Choices

The technology boom of the 1990s and the internet bubble of the early 2000s featured companies, popular with investors, that often shared the same characteristic: They could not generate enough current earnings to justify their stock prices using the traditional price-to-earnings ratio (P/E) approaches to valuation. Many investors chose to explain these apparent anomalies by rationalizing that the old focus on profits and traditional valuation approaches no longer applied to such companies. Strange new metrics for determining operating performance emerged. Website operators spoke of the "eyeballs" they had captured in a quarter, or the "stickiness" of their websites for web surfers' visits. Various versions of "pro forma earnings"—that is, "non-GAAP earnings measures"—became a financial reporting staple of the era.

Many technology companies were accomplished practitioners of pro forma reporting, but they were not the first to use it. In the early 1990s, downsizing of large companies was a commonplace event, and massive restructuring charges obscured the operating performance at many established companies. For example, as it learned to cope in a world that embraced the personal computer rather than mainframe computing, International Business Machines (IBM) reported massive restructuring charges in 1991, 1992, and 1993: $3.7 billion, $11.6 billion, and $8.9 billion, respectively. IBM was not alone. Sears incurred $2.7 billion of restructuring charges in 1993, and AT&T reported restructuring charges of $7.7 billion in 1995. These events were not isolated; restructuring charges were a standard quarterly reporting event. To counter perceptions that their operations were floundering, and supposedly to assist investors in evaluating operating performance, companies often sanitized earnings releases by excluding restructuring charges in pro forma measures of financial performance.

Accounting principles for reporting business combinations also played a role in boosting the popularity of pro forma earnings. Before 2001, acquisitions of one company by another often resulted in goodwill amortization charges that made subsequent earnings reports look weak. Complicating matters, there were two accounting methods for recording acquisitions: pooling-of-interests and purchase methods. The now-extinct pooling-of-interests treatment was difficult for companies to achieve because of the many restrictive criteria for its use, but it was greatly desired because it did not result in goodwill amortization charges. In the technology boom period, acquisitions were common and many were reported as purchases, with consequential goodwill amortization dragging down earnings for as long as 40 years under the then-existing rules. Acquisitive companies reporting under purchase accounting standards perceived themselves to be at a reporting disadvantage compared with companies able to apply pooling-of-interests. They responded by presenting earnings adjusted for the exclusion of amortization of intangible assets and goodwill.

Because investors try to make intercompany comparisons on a consistent basis, earnings before interest, taxes, depreciation, and amortization has become an extremely popular performance measure. EBITDA is widely viewed as eliminating noisy reporting signals. That noise may be introduced by different accounting methods among companies for depreciation, amortization of intangible assets, and restructuring charges.

Companies may construct and report their own version of EBITDA, sometimes referring to it as "adjusted EBITDA," by adding to the list of items to exclude from net income. Items that analysts might encounter include the following:

- Rental payments for operating leases, resulting in EBITDAR (earnings before interest, taxes, depreciation, amortization, and rentals);
- Equity-based compensation, usually justified on the grounds that it is a non-cash expense;
- Acquisition-related charges;
- Impairment charges for goodwill or other intangible assets;
- Impairment charges for long-lived assets;
- Litigation costs; and
- Loss/gain on debt extinguishments.

Among other incentives for the spread of non-GAAP earnings measures are loan covenants. Lenders may make demands on a borrowing company that require achieving and maintaining performance criteria that use GAAP net income as a starting point but arrive at a measure suitable to the lender. The company may use this measure as its preferred non-GAAP metric in earnings releases, and also when describing its liquidity or solvency situation in the management commentary (called management discussion and analysis in the United States).

As mentioned earlier, if a company uses a non-GAAP financial measure in an SEC filing, it must display the most directly comparable GAAP measure with equal prominence and provide a reconciliation between the two. Management must explain why it believes that the non-GAAP financial measure provides useful information regarding the company's financial condition and operations. Management must also disclose additional purposes, if material, for which it uses the non-GAAP financial measures.

Similarly, IFRS requires a definition and explanation of any non-IFRS measures included in financial reports, including why the measure is potentially relevant to users. Management must provide reconciliations of non-IFRS measures with IFRS measures. There is a concern that management may use non-GAAP measures to distract attention from GAAP measures.

The SEC intended that the definition of non-GAAP financial measures would capture all measures with the effect of depicting either:

- a measure of performance that differs from that presented in the financial statements, such as income or loss before taxes or net income or loss, as calculated in accordance with GAAP; or
- a measure of liquidity that differs from cash flow or cash flow from operations computed in accordance with GAAP.[19]

The SEC prohibits the exclusion of charges or liabilities requiring cash settlement from any non-GAAP liquidity measures, other than EBIT and EBITDA. Also prohibited is the calculation of a non-GAAP performance measure intended to eliminate or smooth items tagged as non-recurring, infrequent, or unusual when such items are very likely to occur again. The SEC views the period within two years of either before or after the reporting date as the relevant time frame for considering whether a charge or gain is a recurring item. Example 5 describes a case of misuse and misreporting of non-GAAP measures.

---

[19] SEC, "Final Rule: Conditions for Use of Non-GAAP Financial Measures," Securities and Exchange Commission (www.sec.gov/rules/final/33-8176.htm).

EXAMPLE 5

Misuse and Misreporting of Non-GAAP Measures

Groupon is an online discount merchant. In the company's initial S-1 registration statement in 2011, then-CEO Andrew Mason gave prospective investors an up-front warning in a section entitled "We don't measure ourselves in conventional ways", which described Groupon's adjusted consolidated segment operating income (adjusted CSOI) measure. Exhibit 12 provides excerpts from a section entitled "Non-GAAP Financial Measures," which offered a more detailed explanation. Exhibit 13, also from the initial registration statement, shows a reconciliation of CSOI to the most comparable US GAAP measure. In its review, the SEC took the position that online marketing expenses were a recurring cost of business. Groupon responded that the marketing costs were similar to acquisition costs, not recurring costs, and that "we'll ramp down marketing just as fast as we ramped it up, reducing the customer acquisition part of our marketing expenses" as time passes.[20]

Eventually, and after much negative publicity, Groupon changed its non-GAAP measure. Exhibit 14 shows an excerpt from the final prospectus filed in November, after the SEC's review. Use the three exhibits to answer the questions that follow.

Exhibit 12 Groupon's "Non-GAAP Financial Measures"

Disclosures from June S-1 Filing

Adjusted CSOI is operating income of our two segments, North America and International, adjusted for online marketing expense, acquisition-related costs and stock-based compensation expense. Online marketing expense primarily represents the cost to acquire new subscribers and is dictated by the amount of growth we wish to pursue. Acquisition-related costs are non-recurring non-cash items related to certain of our acquisitions. Stock-based compensation expense is a non-cash item. We consider Adjusted CSOI to be an important measure of the performance of our business as it excludes expenses that are non-cash or otherwise not indicative of future operating expenses. We believe it is important to view Adjusted CSOI as a complement to our entire consolidated statements of operations.

Our use of Adjusted CSOI has limitations as an analytical tool, and you should not consider this measure in isolation or as a substitute for analysis of our results as reported under GAAP. Some of these limitations are:

- Adjusted CSOI does not reflect the significant cash investments that we currently are making to acquire new subscribers;
- Adjusted CSOI does not reflect the potentially dilutive impact of issuing equity-based compensation to our management team and employees or in connection with acquisitions;
- Adjusted CSOI does not reflect any interest expense or the cash requirements necessary to service interest or principal payments on any indebtedness that we may incur;
- Adjusted CSOI does not reflect any foreign exchange gains and losses;

(continued)

---

[20] Correspondence between Groupon and SEC, filed in EDGAR on 16 September 2011.

Exhibit 12 (Continued)

- Adjusted CSOI does not reflect any tax payments that we might make, which would represent a reduction in cash available to us;
- Adjusted CSOI does not reflect changes in, or cash requirements for, our working capital needs; and
- Other companies, including companies in our industry, may calculate Adjusted CSOI differently or may use other financial measures to evaluate their profitability, which reduces the usefulness of it as a comparative measure.

Because of these limitations, Adjusted CSOI should not be considered as a measure of discretionary cash available to us to invest in the growth of our business. When evaluating our performance, you should consider Adjusted CSOI alongside other financial performance measures, including various cash flow metrics, net loss and our other GAAP results.

Exhibit 13 Groupon's "Adjusted CSOI"

Excerpt from June S-1 Filing

The following is a reconciliation of CSOI to the most comparable US GAAP measure, "loss from operations," for the years ended December 31, 2008, 2009, and 2010 and the three months ended March 31, 2010 and 2011:

	Year Ended December 31,			Three Months Ended March 31,	
(in $ thousands)	**2008**	**2009**	**2010**	**2010**	**2011**
(Loss) Income from operations	(1,632)	(1,077)	(420,344)	8,571	(117,148)
Adjustments:					
Online marketing	162	4,446	241,546	3,904	179,903
Stock-based compensation	24	115	36,168	116	18,864
Acquisition-related	—	—	203,183	—	—
Total adjustments	186	4,561	480,897	4,020	198,767
Adjusted CSOI	(1,446)	3,484	60,553	12,591	81,619

Exhibit 14 Groupon's "CSOI"

Excerpt from Revised S-1 Filing

The following is a reconciliation of CSOI to the most comparable US GAAP measure, "loss from operations," for the years ended December 31, 2008, 2009, and 2010 and the nine months ended September 30, 2010 and 2011:

Exhibit 14 (Continued)

	Year Ended December 31,			Nine Months Ended September 30,	
(in $ thousands)	**2008**	**2009**	**2010**	**2010**	**2011**
Loss from operations	(1,632)	(1,077)	(420,344)	(84,215)	(218,414)
Adjustments:					
Stock-based compensation	24	115	36,168	8,739	60,922
Acquisition-related	—	—	203,183	37,844	(4,793)
Total adjustments	24	115	239,351	46,583	56,129
CSOI	(1,608)	(962)	(180,993)	(37,632)	(162,285)

1. What cautions did Groupon include along with its description of the “Adjusted CSOI” metric?
2. Groupon excludes “online marketing” from “Adjusted CSOI.” How does the exclusion of this expense compare with the SEC’s limits on non-GAAP performance measures?
3. In the first quarter of 2011, what was the effect of excluding online marketing expenses on the calculation of “Adjusted CSOI”?
4. For 2010, how did results under the revised non-GAAP metric compare with the originally reported metric?

Solution to 1:

Groupon cautioned that the “Adjusted CSOI” metric should not be considered in isolation, should not be considered as a substitute for analysis using GAAP results, and “should not be considered a measure of discretionary cash flow.” The company lists numerous limitations, primarily citing items that adjusted CSOI did not reflect.

Solution to 2:

The SEC specifies that non-GAAP measures should not eliminate items tagged as non-recurring, infrequent, or unusual when such items may be very likely to occur again. Because the online marketing expense occurred in every period reported and is likely to occur again, exclusion of this item appears contrary to SEC requirements.

Solution to 3:

As shown in Exhibit 13, in the first quarter of 2011, the exclusion of the online marketing expense was enough to swing the company from a net loss under US GAAP reporting to a profit—at least, a profit as defined by adjusted CSOI. Using adjusted CSOI as a performance measure, the company showed results that were 35% higher for the first *quarter* of 2011 compared with the entire previous *year*.

Solution to 4:

As shown in Exhibit 14, the revised metric is now called "CSOI" and no longer refers to "Adjusted CSOI." For 2010, results under the revised non-GAAP metric, which includes online marketing costs, shows a loss of $180,993,000 instead of a profit of $60,553,000.

In the case described in Example 5, Groupon changed its reporting and corrected the non-GAAP metric that the SEC had identified as misleading. In other cases, the SEC has pursued enforcement actions against companies for reporting misleading non-GAAP information. One such action was brought in 2009 against SafeNet Inc., where the SEC charged the company with improperly classifying ordinary operating expenses as non-recurring. This related to the integration of an acquired company and exclusion of the expenses from non-GAAP earnings in order to exceed earnings targets. A second action was brought by the SEC in 2017 against MDC Partners Inc. ("MDCA") for improper reconciliation of a non-GAAP measure and for improperly displaying the non-GAAP measure with greater prominence in its earnings releases. The case was brought after the company agreed to follow the rules but then failed to do so, as evidenced by the remark in the SEC's action: "Despite agreeing to comply with non-GAAP financial measure disclosure rules in December 2012 correspondence with the [SEC's] Division of Corporation Finance, MDCA continued to violate those rules for six quarters …" Exhibit 15 presents the headline and sub-headings for one of MDC Partners' earnings announcements that was the subject of the enforcement action.

Exhibit 15 MDC Partners Inc. Press Release [Excerpt]

SEC Form 8-K filed 24 April 2014

This excerpt shows the headline, sub-heads, and lead sentence of the company's press release announcing periodic earnings.

MDC PARTNERS INC. REPORTS RECORD RESULTS FOR THE THREE MONTHS ENDED MARCH 31, 2014

ORGANIC REVENUE GROWTH OF 8.3%, EBITDA GROWTH OF 18.1% AND 90 BASIS POINTS OF MARGIN IMPROVEMENT

FREE CASH FLOW GROWTH OF 34.0%

INCREASED 2014 GUIDANCE IMPLIES YEAR-OVER-YEAR EBITDA GROWTH OF +13.5% TO +16.1%, MARGIN IMPROVEMENT OF 60 TO 70 BASIS POINTS, AND FREE CASH FLOW GROWTH OF +15.8% TO +20.2%

FIRST QUARTER HIGHLIGHTS:

- Revenue increased to $292.6 million from $265.6 million, an increase of 10.1%
- Organic revenue increased 8.3%
- EBITDA increased to $36.4 million from $30.8 million, an increase of 18.1%
- EBITDA margin increased 90 basis points to 12.5% from 11.6%
- Free Cash Flow increased to $20.6 million from $15.4 million, an increase of 34.0%
- Net New Business wins totaled $24.4 million

Exhibit 15 (Continued)

NEW YORK, NY (April 24, 2014) – MDC Partners Inc. (NASDAQ: MDCA; TSX: MDZ.A) today announced financial results for the three months ended March 31, 2014.

…

In general, management may choose to construct non-GAAP financial measures not only to help investors better understand the company's performance but also to paint a more flattering picture of its performance. In some cases, management may attempt to present non-GAAP measures in a way that diverts attention from the standards-compliant financial information that it is required to present.

ACCOUNTING CHOICES AND ESTIMATES AND HOW ACCOUNTING CHOICES AND ESTIMATES AFFECT EARNINGS AND BALANCE SHEETS

9

h describe accounting methods (choices and estimates) that could be used to manage earnings, cash flow, and balance sheet items;

Choices do not necessarily involve complex accounting standards. Something as simple as the shipping terms for goods delivered to customers can have a profound effect on the timing of revenue. On the last day of the first quarter, suppose a company ships $10,000 of goods to a customer on the terms "free on board (FOB) shipping point," arriving the next day. This shipping term means that the customer takes title to the goods, and bears the risk of loss, at the time the goods leave the seller's loading dock. Barring any issues with collectability of the receivable, or a likelihood of a return, the seller would be able to recognize revenue on the sale along with the associated profit. That revenue and profit would be recognized in the first quarter of the year. Change the point at which the goods' title transfers to the customer to "FOB destination" and the revenue pattern will be completely different. Under these terms, the title—and risk of loss—transfers to the customer when the goods arrive at their destination, which is the customer's address. The seller cannot recognize the sale and profit until the shipment arrives the following day, which is the start of a new accounting period.

A simple change in shipping terms can make the difference between revenue and profits in the reporting period or postponing them until the next period. Shipping terms can also influence management behavior. To "make the numbers," managers might push product out the door prematurely under FOB shipping point arrangements in order to reflect as much revenue as possible in the reporting period. Alternatively, in the case of an over-abundance of orders, the company could run the risk of exceeding analysts' consensus estimates by a large margin. Management might be uncomfortable with this situation because investors might extrapolate too much from one reporting period in which expectations were exceeded. Management might want to prevent investors from becoming too optimistic and, if possible, delay revenue recognition until the next quarter. This result could be accomplished by fulfilling customer orders by initiating delivery on the last day of the quarter, with shipping terms set as FOB destination. By doing so, title would transfer in the next accounting period. Another possibility in this scenario is that if the customers insisted on FOB shipping point terms, the selling company could simply delay shipment until after the close of the quarter.

This illustration also highlights a difficult distinction for investors to make. A company may use accounting as a tool to aggressively promote earnings growth—as in the example with the premature shipment of goods with FOB shipping point terms—but it may be aggressively managing the business flow by slacking off on shipping goods when business is "too good," as in the second example. In either case, a desired management outcome is obtained by a simple change in shipping terms. Yet, many investors might be inclined to say that the second example is a conservative kind of earnings management and accept it, even though it artificially masks the actual economic activity that occurred at the time.

9.1 How Accounting Choices and Estimates Affect Earnings and Balance Sheets

Assumptions about inventory cost flows provide another example of how accounting choices can affect financial reporting. Companies may assume that their purchases of inventory items are sold to customers on a first-in-first-out (FIFO) basis, with the result that the remaining inventory reflects the most recent costs. Alternatively, they may assume that their purchases of inventory items are sold to customers on a weighted-average cost basis. Example 6 makes the point that merely choosing a cost flow assumption can affect profitability.

EXAMPLE 6

Effect of Cost Flow Assumption

A company starts operations with no inventory at the beginning of a fiscal year and makes purchases of a good for resale five times during the period at increasing prices. Each purchase is for the same number of units of the good. The purchases, and the cost of goods available for sale, appear in the following table. Notice that the price per unit has increased by 140% by the end of the period.

	Units	Price	Cost
Purchase 1	5	$100	$500
Purchase 2	5	150	750
Purchase 3	5	180	900
Purchase 4	5	200	1,000
Purchase 5	5	240	1,200
Cost of goods available for sale			$4,350

During the period, the company sells, at $250 each, all of the goods purchased except for five of them. Although the ending inventory consists of five units, the cost attached to those units can vary greatly.

1 What are the ending inventory and cost of goods sold if the company uses the FIFO method of inventory costing?

2 What are the ending inventory and cost of goods sold if the company uses the weighted-average method of inventory costing?

3 Compare cost of goods sold and gross profit calculated under the two methods.

Solution to 1:

The ending inventory and cost of goods sold if the company uses the FIFO method of inventory costing are $1,200 and $3,150.

Solution to 2:

The ending inventory and cost of goods sold if the company uses the weighted-average method of inventory costing are $870 and $3,480.

Solution to 3:

The following table shows how the choice of inventory costing methods—FIFO versus weighted average—affects the cost of goods sold and gross profit.

Cost Flow Assumption	FIFO	Weighted Average
Cost of goods available for sale	$4,350	$4,350
Ending inventory (5 units)	(1,200)	(870)
Cost of goods sold	$3,150	$3,480
Sales	$5,000	$5,000
Cost of goods sold	3,150	3,480
Gross profit	$1,850	$1,520
Gross profit margin	37.0%	30.4%

Note: Average inventory cost is calculated as Cost of goods available for sale/Units purchased = $4,350/25 = $174. There are five units in ending inventory, yielding an inventory value of $870.

Depending on which cost flow assumption the company uses, the end-of-period inventory is either $870 (under the weighted-average method) or $1,200 (under FIFO). The choice of method results in a difference of $330 in gross profit and 6.6% in gross profit margin.

The previous example is simplified and extreme for purposes of illustration clarity, but the point is important: Management's choice among acceptable inventory assumptions and methods affects profit. The selection of an inventory costing method is a policy decision, and companies cannot arbitrarily switch from one method to another. The selection does matter to profitability, however, and it also matters to the balance sheet.

In periods of changing prices, the FIFO cost assumption will provide a more current picture of ending inventory value, because the most recent purchases will remain in inventory. The balance sheet will be more relevant to investors. Under the weighted-average cost assumption, however, the balance sheet will display a blend of old and new costs. During inflationary periods, the value of the inventory will be understated: The company will not be able to replenish its inventory at the value shown. At the same time, the weighted-average inventory cost method ensures that the more current costs are shown in cost of sales, making the income statement more relevant than under the FIFO assumption. Trade-offs exist, and investors should be aware of how accounting choices affect financial reports. High-quality financial reporting provides users with sufficient information to assess the effects of accounting choices.

Estimates abound in financial reporting because of the use of accrual accounting, which attempts to show the effects of all economic events on a company during a particular period. Accrual accounting stands in contrast to cash basis accounting, which shows only the cash transactions conducted by a company. Although a high degree of certainty exists with reporting only cash transactions, much information is hidden. For instance, a company with growing revenues that makes the majority of its sales on credit would be understating its revenues for each period if it reported only cash transactions. On an accrual basis, revenues reflect all transactions that occurred, whether they transacted on a cash basis or credit-extended basis. Estimates enter the process because some facts related to events occurring in a particular period might not yet be known. Estimates can be well grounded in reality and applied to present a complete picture of the events affecting a company, or they can be management tools for achieving a desired financial picture.

To illustrate how estimates can affect financial reporting, consider sales made on credit. A company sells $1,000,000 of merchandise on credit and records the sale just before year end. Under accrual accounting, that amount is included in revenues and accounts receivable. The company's managers know from experience that they will never collect every dollar of the accounts receivable. Past experience is that, on average, only 97% of accounts receivable is collected. The company would estimate an amount of the uncollectible accounts at the time the sales occur and record an uncollectible accounts expense of $30,000, lowering earnings. The other side of the entry would be to establish an allowance for uncollectible accounts of $30,000. This allowance would be a contra asset account, presented as an offset to accounts receivable. The accounts receivable, net of the allowance for uncollectible accounts, would be stated at $970,000, which is the amount of cash the company ultimately expects to receive. If cash-basis accounting had been used, no revenues or accounts receivable would have been reported even though sales of merchandise had occurred. Accrual accounting, which contains estimates about future events, provides a much fuller picture of what transpired in the period than pure cash-basis accounting.

Yet, accrual accounting poses temptations to managers to manage the numbers, rather than to manage the business. Suppose a company's managers realize that the company will not meet analysts' consensus estimates in a particular quarter, and further, their bonus pay is dependent on reaching specified earnings targets. By offering special payment terms, or discounts, the managers may induce customers to take delivery of products that they would normally not order, so they could ship the products on FOB shipping point terms and recognize the revenues in the current quarter. They could even be so bold as to ship the goods under those terms even if the customer did not order them, in the hope that the customer would keep them or, at worst, return them in the next accounting period. Their aim would be to move the product off the company's property with FOB shipping point terms.

To further improve earnings in order to meet the consensus estimates, the company's managers might revise their estimate of the uncollectible accounts. The company's collection history shows a typical non-collection rate of 3% of sales, but the managers might rationalize the use of a 2% non-collection rate. This change will reduce the allowance for uncollectible accounts and the expense reported for the period. The managers might be able to justify the reduction on the grounds that the sales occurred in a part of the country that was experiencing an improved economic outlook, or that the company's collection history had been biased by the inclusion of a prolonged period of economic downturn. Whatever the justification, it would be hard to prove that the new estimate was completely right or wrong until time had passed. Because proof of the reliability of estimates is rarely available at the time the estimate is recorded, managers have a readily available means for manipulating earnings at their discretion.

ConAgra Foods Inc. provides an example of how the allowance for uncollectible accounts may be manipulated to manage earnings.[21] A subsidiary, called United Agri-Products (UAP), engaged in several improper accounting practices, one of them being the understatement of uncollectible accounts expense for several years. Exhibit 16 presents an excerpt from the SEC's Accounting and Auditing Enforcement Release.

Exhibit 16 SEC's Accounting and Auditing Enforcement Release Regarding United Agri-Products

> … Generally, UAP's policy required that accounts which were past due between 90 days and one year should be reserved at 50%, and accounts over one year past due were to be reserved at 100%.
>
> … In FY 1999 and continuing through FY 2000, UAP had substantial bad debt problems. In FY 2000, certain former UAP senior executives were informed that UAP needed to record an additional $50 million of bad debt expense. Certain former UAP senior executives were aware that in FY 1999 the size of the bad debt at certain IOCs had been substantial enough that it could have negatively impacted those IOC's ability to achieve PBT (profits before taxes) targets. In addition, just prior to the end of UAP's FY 2000, the former UAP COO (chief operating officer), in the presence of other UAP employees, ordered that UAP's bad debt reserve be reduced by $7 million in order to assist the Company in meeting its PBT target for the fiscal year.
>
> … At the end of FY 2000, former UAP senior executives reported financial results to ConAgra which they knew, or were reckless in not knowing, overstated UAP's income before income taxes because UAP had failed to record sufficient bad debt expense. The misconduct with respect to bad debt expense caused ConAgra to overstate its reported income before income taxes by $7 million, or 1.13%, in FY 2000. At the Agricultural Products' segment level, the misconduct caused that segment's reported operating profit to be overstated by 5.05%.

Deferred-tax assets provide a similar example of choices in estimates affecting the earnings outcome. Deferred-tax assets may arise when a company reports a net operating loss under tax accounting rules. A company may record a deferred-tax asset based on the expectation that losses in the reporting period will offset expected future profits and reduce the company's future income tax liability. Accounting standards require that the deferred tax asset be reduced by a "valuation allowance" to account for the possibility that the company will be unable to generate enough profit to use all of the available tax benefits.[22]

Assume a company loses €1 billion in 2012, generating a net operating loss of the same amount for tax purposes. The company's income tax rate is 25%, and it will be able to apply the net operating loss to its taxable income for the next 10 years. The net operating loss results in a deferred tax asset with a nominal value of €250 million (25% × €1,000,000,000). Initial recognition would result in a deferred tax asset of €250 million and a credit to deferred tax expense of €250 million. The company must address the question of whether or not the €250 million will ever be completely applied to future income. It may be experiencing increased competition and other circumstances that resulted in the €1 billion loss, and it may be unreasonable to assume it will have taxable income against which to apply the loss. In fact, the company's managers might believe it is reasonable to assume only that it will survive for five years, and with marginal profitability. The €250 million deferred tax asset is thus overstated if no valuation allowance is recorded to offset it.

---

[21] Accounting and Auditing Enforcement Release No. 2542, "SEC v. James Charles Blue, Randy Cook, and Victor Campbell,) United States District Court for the District of Colorado, Civ. Action No. 07-CV-00095 REB-MEH (17 January 2007).

[22] See Accounting Standards Codification 740-10-30-16 to 25, "Establishment of a Valuation Allowance for Deferred Tax Assets."

The managers believe that only €100 million of the net operating losses will actually be applied to the company's taxable income. That belief implies that only €25 million of the tax benefits will ever be realized. The deferred tax assets reported on the balance sheet should not exceed this amount. The company should record a valuation allowance of €225 million, which would offset the deferred tax asset balance of €250 million, resulting in a net deferred tax asset balance of €25 million. There would also be a €225 million credit to the deferred tax provision. It is important to understand that the valuation allowance should be revised whenever facts and circumstances change.

The ultimate value of the deferred tax asset is driven by management's outlook for the future—and that outlook may be influenced by other factors. If the company needs to stay in compliance with debt covenants and needs every euro of value that can be justified by the outlook, its managers may take a more optimistic view of the future and keep the valuation allowance artificially low (in other words, the net deferred tax asset high).

PowerLinx Inc. provides an example of how over-optimism about the realizability of a deferred tax asset can lead to misstated financial reports. PowerLinx was a maker of security video cameras, underwater cameras, and accessories. Aside from fraudulently reporting 90% of its fiscal year 2000 revenue, PowerLinx had problems with valuation of its deferred tax assets. Exhibit 17 provides an excerpt from the SEC's Accounting and Auditing Enforcement Release with emphasis added.[23]

Exhibit 17 SEC's Accounting and Auditing Enforcement Release Regarding PowerLinx

PowerLinx improperly recorded on its fiscal year 2000 balance sheet a deferred tax asset of $1,439,322 without any valuation allowance. The tax asset was material, representing *almost forty percent of PowerLinx's total assets* of $3,841,944. PowerLinx also recorded deferred tax assets of $180,613, $72,907, and $44,921, respectively, in its financial statements for the first three quarters of 2000.

PowerLinx did not have a proper basis for recording the deferred tax assets. The company had accumulated significant losses in 2000 and had no historical operating basis from which to conclude that it would be profitable in future years. Underwater camera sales had declined significantly and the company had devoted most of its resources to developing its SecureView product. The sole basis for PowerLinx's "expectation" of future profitability was the purported $9 million backlog of SecureView orders, which management assumed would generate taxable income; however, this purported backlog, which predated Bauer's hiring, did not reflect actual demand for SecureView cameras and, consequently, was not a reasonable or reliable indicator of future profitability.

Another example of misstated financial results caused by improper reflection of the realizability of a deferred tax asset occurred with Hampton Roads Bankshares Inc. ("HRBS"), a commercial bank with deteriorating loan portfolio quality and commensurate losses in the years following the financial crisis. The company reported a deferred tax asset related to its loan losses; however, it did not establish a valuation allowance against its deferred tax asset. This decision was based on dubious projections indicating that the company would earn the necessary future taxable income "to fully utilize the [deferred tax asset] DTA over the applicable carry-forward period."[24] Over time, it became clear that the earnings projections were not realistic, and ultimately the company restated its financial results to include a valuation allowance against almost the entire deferred tax asset. Exhibit 18 presents an excerpt from the company's amended Form 10-Q/A containing the restatement.

---

[23] Accounting and Auditing Enforcement Release No. 2448, "In the Matter of Douglas R. Bauer, Respondent," SEC (27 June 2006): www.sec.gov/litigation/admin/2006/34-54049.pdf.

EXHIBIT 18 Hampton Roads Bankshares, Inc. Form 10-Q/A filed August 13, 2010 [Excerpt from footnotes]

NOTE B – RESTATEMENT OF CONSOLIDATED FINANCIAL STATEMENTS

Subsequent to filing the Company's annual report on Form 10-K for the year ended December 31, 2009 and its Form 10-Q for the three months ended March 31, 2010 the Company determined that a valuation allowance on its deferred tax assets should be recognized as of December 31, 2009. The Company decided to establish a valuation allowance against the deferred tax asset because it is uncertain when it will realize this asset.

Accordingly, the December 31, 2009 consolidated balance sheet and the March 31, 2010 consolidated financial statements have been restated to account for this determination. The effect of this change in the consolidated financial statements was as follows (in thousands, except per share amounts).

Consolidated Balance Sheet at March 31, 2010

	As Reported	Adjustment	As Restated
Deferred tax assets, net	$70,323	$(70,323)	—
Total assets	3,016,470	(70,323)	$2,946,147
Retained earnings deficit	(158,621)	(70,323)	(228,944)
Total shareholder's equity	156,509	(70,323)	86,186
Total Liabilities and shareholders' equity	3,016,470	(70,323)	2,946,147

Consolidated Balance Sheet at December 31, 2009

	As Reported	Adjustment	As Restated
Deferred tax assets, net	$56,380	$(55,983)	$397
Total assets	2,975,559	(55,983)	2,919,576
Retained earnings deficit	(132,465)	(55,983)	(188,488)

(continued)

---

[24] Accounting and Auditing Enforcement Release No. 3600, "In the Matter of Hampton Roads Bankshares Inc., Respondent," SEC (5 December 2014) https://www.sec.gov/litigation/admin/2014/34-73750.pdf.

(Continued)

	As Reported	Adjustment	As Restated
Total shareholder's equity	180,996	(55,983)	125,013
Total Liabilities and shareholders' equity	2,975,559	(55,983)	2,919,576

Another example of how choices and estimates can affect reported results lies in the selection of a depreciation method for allocating the cost of long-lived assets to accounting periods subsequent to their acquisition. A company's managers may choose to depreciate long-lived assets (1) on a straight-line basis, with each year bearing the same amount of depreciation expense; (2) using an accelerated method, with greater depreciation expense recognition in the earlier part of an asset's life; or (3) using an activity-based depreciation method, which allocates depreciation expense based on units of use or production. Depreciation expense is affected by another set of choices and estimates regarding the salvage value of the assets being depreciated. A salvage value of zero will always increase depreciation expense under any method compared with the choice of a non-zero salvage value.

Assume a company invests $1,000,000 in manufacturing equipment and expects it to have a useful economic life of 10 years. During its expected life, the equipment will produce 400,000 units of product, or $2.50 depreciation expense per unit produced. When it is disposed of at the end of its expected life, the company's managers expect to realize no value for the equipment. The following table shows the differences in the three alternative methods of depreciation: straight-line, accelerated on a double-declining balance basis, and units-of-production method, with no salvage value assumed at the end of the equipment's life.

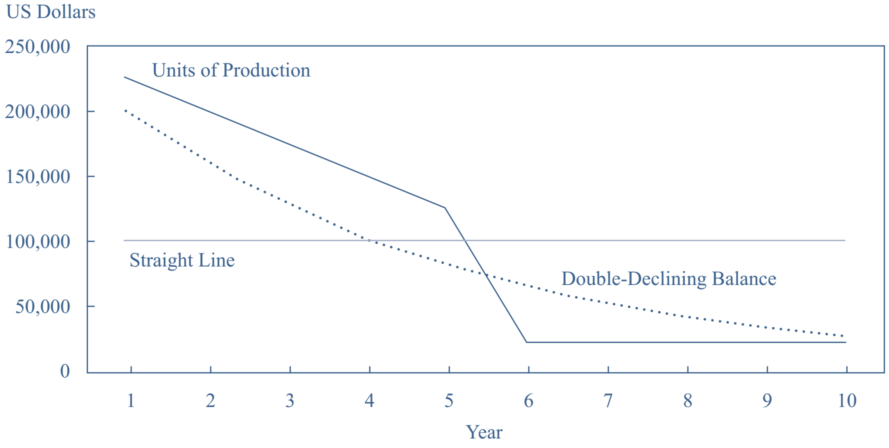
	Straight-Line Method	Double-Declining Balance Method			Units-of-Production Method		
Year	Depreciation Expense	Balance	Declining Balance Rate[1]	Depreciation Expense	Units Produced	Depreciation Rate/Unit	Depreciation Expense
1	$100,000	$1,000,000	20%	$200,000	90,000	$2.50	$225,000
2	100,000	800,000	20%	160,000	80,000	$2.50	200,000
3	100,000	640,000	20%	128,000	70,000	$2.50	175,000
4	100,000	512,000	20%	102,400	60,000	$2.50	150,000
5	100,000	409,600	20%	81,920	50,000	$2.50	125,000
6	100,000	327,680	20%	65,536	10,000	$2.50	25,000
7	100,000	262,144	20%	52,429	10,000	$2.50	25,000
8	100,000	209,715	20%	41,943	10,000	$2.50	25,000
9	100,000	167,772	20%	33,554	10,000	$2.50	25,000
10	100,000	134,218	20%	26,844	10,000	$2.50	25,000
Total	$1,000,000			$892,626	400,000		$1,000,000

[1] Declining balance rate of 20% calculated as 10-year life being equivalent to 10% annual depreciation rate, multiplied by 2 = 20%.

The straight-line method allocates the cost of the equipment evenly to all 10 years of the equipment's life. The double-declining balance method will have a higher allocation of cost to the earlier years of the equipment's life. As its name implies, the depreciation expense will decline in each succeeding year because it is based on a fixed rate applied to a declining balance. The rate used was double the straight-line rate, but it could have been any other rate that the company's managers believed was representative of the way the actual equipment depreciation occurred. Notice that the double-declining balance method also results in an incomplete depreciation of the machine at the end of 10 years; a balance of $107,374 (= $1,000,000 – $892,626) remains at the end of the expected life, which will result in a loss upon the retirement of the equipment if the company's expectation of zero salvage value turns out to be correct. Some companies may choose to depreciate the equipment to its expected salvage value, zero in this case, in its final year of use. Some companies may use a policy of switching to straight-line depreciation after the mid-life of its depreciable assets in order to fully depreciate them. That particular pattern is coincidentally displayed in the units-of-production example, in which the equipment is used most heavily in the earliest part of its useful life, and then levels off to much less utilization in the second half of the expected life.

Exhibit 19 shows the different expense allocation patterns of the methods over the same life. Each will affect earnings differently.

Exhibit 19 Expense Allocation Patterns of Different Depreciation Methods



The company's managers could justify any of these methods. Each might fairly represent the way the equipment will be consumed over its expected economic life, which is a subjective estimate itself. The choices of methods and lives can profoundly affect reported income. These choices are not proven right or wrong until far into the future—but managers must estimate their effects in the present.

Exhibit 20 shows the effects of the three different methods on operating profit and operating profit margins, assuming that the production output of the equipment generates revenues of $500,000 each year and $200,000 of cash operating expenses are incurred, leaving $300,000 of operating profit before depreciation expense.

Exhibit 20 Effects of Depreciation Methods on Operating Profit

	Straight Line		
Year	**Depreciation**	**Operating Profit**	**Operating Profit Margin**
1	$100,000	$200,000	40.0%
2	100,000	200,000	40.0%

(continued)

Exhibit 20 (Continued)

	Straight Line		
Year	**Depreciation**	**Operating Profit**	**Operating Profit Margin**
3	100,000	200,000	40.0%
4	100,000	200,000	40.0%
5	100,000	200,000	40.0%
6	100,000	200,000	40.0%
7	100,000	200,000	40.0%
8	100,000	200,000	40.0%
9	100,000	200,000	40.0%
10	100,000	200,000	40.0%

	Double Declining Balance		
Year	**Depreciation**	**Operating Profit**	**Operating Profit Margin**
1	$200,000	$100,000	20.0%
2	160,000	140,000	28.0%
3	128,000	172,000	34.4%
4	102,400	197,600	39.5%
5	81,920	218,080	43.6%
6	65,536	234,464	46.9%
7	52,429	247,571	49.5%
8	41,943	258,057	51.6%
9	33,554	266,446	53.3%
10	134,218*	165,782	33.2%

	Units of Production		
Year	**Depreciation**	**Operating Profit**	**Operating Profit Margin**
1	$225,000	$75,000	15.0%
2	200,000	100,000	20.0%
3	175,000	125,000	25.0%
4	150,000	150,000	30.0%
5	125,000	175,000	35.0%
6	25,000	275,000	55.0%
7	25,000	275,000	55.0%
8	25,000	275,000	55.0%
9	25,000	275,000	55.0%
10	25,000	275,000	55.0%

* Includes $107,374 of undepreciated basis, treated as depreciation expense in final year of service.

The straight-line method shows consistent operating profit margins, and the other two methods show varying degrees of increasing operating profit margins as the depreciation expense decreases over time.

The example above shows the differences among alternative methods, but even more depreciation expense variation is possible by changing estimated lives and assumptions about salvage value. For instance, change the expected life assumption to 5 years from 10 and add an expectation that the equipment will have a 10% salvage value at the end of its expected life. Exhibit 21 shows the revised depreciation calculations. Notice that under the double-declining balance method, the depreciation rate is applied to the gross cost, unlike the other two methods. The straight-line method and the units-of-production method subtract the salvage value from the cost before depreciation expense is calculated. Also note that the assumption about the usage of the equipment is revised so that it is depreciated only to its salvage value of $100,000 by the end of its estimated life. The total depreciation under each method is $900,000.

Exhibit 21 Depreciation Calculations for Each Method in Changed Scenario

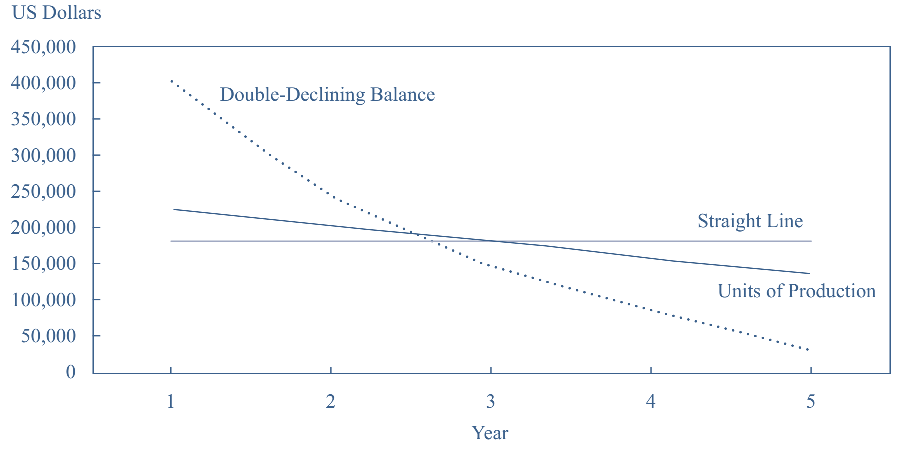
	Straight-Line Method	Double-Declining Balance Method			Units-of-Production Method		
Year	Depreciation Expense	Balance	Declining Balance Rate[1]	Depreciation Expense	Units Produced	Depreciation Rate/Unit	Depreciation Expense
1	**$180,000**	$1,000,000	40%	**$400,000**	100,000	$2.25	**$225,000**
2	**180,000**	600,000	40%	**240,000**	90,000	$2.25	**202,500**
3	**180,000**	360,000	40%	**144,000**	80,000	$2.25	**180,000**
4	**180,000**	216,000	40%	**86,400**	70,000	$2.25	**157,500**
5	**180,000**	129,600	40%	**29,600**[2]	60,000	$2.25	**135,000**
Total	**$900,000**			**$900,000**	400,000		**$900,000**

[1] Declining balance rate of 40% calculated as 5-year life being equivalent to 20% annual depreciation rate, multiplied by 2 = 40%.

[2] Depreciation calculated as $29,600 instead of 40% × $129,600. Rote application of the declining-balance rate would have resulted in $51,840 of expense, which would have depreciated the asset below salvage value.

Exhibit 22 shows the different expense allocation patterns of the methods over the five-year expected life, and assuming a 10% salvage value. Although each method is distinctly different in the timing of the cost allocation over time, the variation is less pronounced than over the longer life used in the previous example.

Exhibit 22 Expense Allocation Patterns of Depreciation Methods in Changed Scenario



One of the clearest examples of how choices affect both the balance sheet and income statement can be found in capitalization practices. In classifying a payment made, management must determine whether the payment will benefit only the current period—making it an expense—or whether it will benefit future periods, leading to classification as a cost to be capitalized as an asset. This management judgment embodies an implicit forecast of how the item acquired by the payment will be used, or not used, in the future.

That judgment can be biased by the powerful effect a capitalization policy can have on earnings. Every amount capitalized on the balance sheet as a building, an item of inventory, a deferred cost, or any "other asset" is an amount that does not get recognized as an expense in the reporting period.

A real-life example can be found in the case of WorldCom Inc., a telecom concern that grew rapidly in the late 1990s. Much of WorldCom's financial reporting was eventually found to be fraudulent. An important part of the misreporting centered on its treatment of what is known in the telecom industry as "line costs". These are the costs of carrying a voice call or data transmission from its starting point to its ending point, and they represented WorldCom's largest expense. WorldCom's chief financial officer decided to capitalize such costs instead of treating them as an operating expense. As a consequence, from the second quarter of 1999 through the first quarter of 2002, WorldCom increased its operating income by $7 billion. In three of the five quarters in which the improper line cost capitalization took place, WorldCom would have recognized pre-tax losses instead of profits.[25]

Similarly, acquisitions are an area in which managers must exercise judgment. An allocation of the purchase price must be made to all of the different assets acquired based on their fair values, and those fair values are not always objectively verifiable. Management may have to make its own estimate of fair values for assets acquired, and it may be biased towards a low estimate for the values of depreciable assets in order to depress future depreciation expense. Another benefit to keeping depreciable asset values low is that the amount of the purchase price that cannot be allocated to specific assets is classified as goodwill, which is neither depreciated nor amortized in future reporting periods.

---

[25] See Report of Investigation by the Special Investigative Committee of the Board of Directors of WorldCom, Inc., by Dennis R. Beresford, Nicholas deB. Katzenbach, & C.B. Rogers, Jr.PP 9-11: www.sec.gov/Archives/edgar/data/723527/000093176303001862/dex991.htm.

Goodwill reporting has choices of its own. Although goodwill has no effect on future earnings when unimpaired, annual testing of its fair value may reveal that the excess of price paid over the fair value of assets may not be recoverable, which should lead to a write-down of goodwill. The estimation process for the fair value of goodwill may depend heavily on projections of future performance. Those projections may be biased upward in order to avoid a goodwill write-down.

HOW CHOICES THAT AFFECT THE CASH FLOW STATEMENT

10

h describe accounting methods (choices and estimates) that could be used to manage earnings, cash flow, and balance sheet items;

The cash flow statement consists of three sections: the operating section, which shows the cash generated or used by operations; the investing section, which shows cash used for investments or provided by their disposal; and the financing section, which shows transactions attributable to financing activities.

The operating section is closely scrutinized by investors. Many of them consider it a reality check on reported earnings, on the grounds that earnings attributable to accrual accounting only and unsupported by actual cash flows may indicate earnings manipulation. Such investors believe that amounts shown for cash generated by operations is more insulated from managerial manipulation than the income statement. Cash generated by operations can be managed to an extent, however.

The operating section of the cash flow statement can be shown either under the direct method or the indirect method. Under the direct method, "entities are encouraged to report major classes of gross cash receipts and gross cash payments and their arithmetic sum—the net cash flow from operating activities."[26] In practice, companies rarely use the direct method. Instead, they use the indirect method, which shows a reconciliation of net income to cash provided by operations. The reconciliation shows the non-cash items affecting net income along with changes in working capital accounts affecting cash from operations. Exhibit 23 provides an example of the indirect presentation method.

Exhibit 23 Indirect Presentation Method

Cash Flows from Operating Activities ($ millions)	2018
Net income	$3,000
Adjustments to reconcile net income to net cash provided by operating activities:	
Provision for doubtful receivables	10
Provision for depreciation and amortization	1,000
Goodwill impairment charges	35
Share-based compensation expense	100
Provision for deferred income taxes	200
Changes in assets and liabilities:	

(continued)

---

[26] Accounting Standards Codification Section 230-10-45-25, "Reporting Operating, Investing, and Financing Activities." The direct method and indirect method are similar in IFRS, as addressed in IAS 7, Paragraph 18.

Exhibit 23 (Continued)

Cash Flows from Operating Activities ($ millions)	2018
Trade, notes and financing receivables related to sales	(2,000)
Inventories	(1,500)
Accounts payable	1,200
Accrued income taxes payable/receivable	(80)
Retirement benefits	90
Other	(250)
Net cash provided by operating activities	$1,805

Whether the indirect method or direct method is used, simple choices exist for managers to improve the *appearance* of cash flow provided by operations without actually improving it. One such choice is in the area of accounts payable management, shaded in Exhibit 23. Assume that the accounts payable balance is $5,200 million at the end of the period, an increase of $1,200 million from its previous year-end balance of $4,000 million. The $1,200 million increase in accounts payable matched increased expenses and/or assets but did not require cash. If the company's managers had further delayed paying creditors $500 million until the day *after* the balance sheet date, they could have increased the cash provided by operating activities by $500 million. If the managers believe that cash generated from operations is a metric of focus for investors, they can impress them with artificially strong cash flow by simply stretching the accounts payable credit period.

What might alert investors to such machinations? They need to examine the composition of the operations section of the cash flow statement—if they do not, then *nothing* will ever alert them. Studying changes in the working capital can reveal unusual patterns that may indicate manipulation of the cash provided by operations.

Another practice that might lead an investor to question the quality of cash provided by operations is to compare a company's cash generation with an industry-wide level or with the cash operating performance of one or more similar competitors. Cash generation performance can be measured several ways. One way is to compare the relationship between cash generated by operations and net income. Cash generated by operations in excess of net income signifies better quality of earnings, whereas a chronic excess of net income over cash generated by operations should be a cause for concern; it may signal the use of accounting methods to simply raise net income instead of depicting financial reality. Another way to measure cash generation performance is to compare cash generated by operations with debt service, capital expenditures, and dividends (if any). When there is a wide variance between the company's cash generation performance and that of its benchmarks, investors should seek an explanation and carefully examine the changes in working capital accounts.

Because investors may focus on cash from operations as an important metric, managers may resort to managing the working capital accounts as described in order to present the most favorable picture. But there are other ways to do this. A company may misclassify operating uses of cash into either the investing or financing sections of the cash flow statement, which enhances the appearance of cash generated by operating activities.

Dynegy Inc. provides an example of manipulation of cash from operations through clever construction of contracts and assistance from an unconsolidated special purpose entity named ABG Gas Supply LLC (ABG). In April 2001, Dynegy entered into a contract for the purchase of natural gas from ABG. According to the contract, Dynegy would purchase gas at *below-market* rates from ABG for nine months and sell it at the current market rate. The nine-month term coincided with Dynegy's 2001 year-end and would result in gains backed by cash flows. Dynegy also agreed to buy gas at *above-market* rates from ABG for the following 51 months and sell it at the current market rate. The contract was reported at its fair value at the end of fiscal year 2001. It had no effect on net income for the year. The earlier portion of the contract resulted in a gain, supported by $300 million of cash flow, but the latter portion of the contract resulted in non-cash losses that offset the profit. The mark-to-market rules required the recognition of both gains and losses from all parts of the contract, and hence the net effect on earnings was zero.

In April 2002, a *Wall Street Journal* article exposed the chicanery, thanks to leaked documents. The SEC required Dynegy to restate the cash flow statement by reclassifying $300 million from the operating section of the cash flow statement to the financing section, on the grounds that Dynegy had used ABG as a conduit to effectively borrow $300 million from Citigroup. The bank had extended credit to ABG, which it used to finance its losses on the contract (Lee, 2012).

Another area of flexibility in cash flow reporting is found in the area of interest capitalization, which creates differences between total interest payments and total interest costs.[27] Assume a company incurs total interest cost of $30,000, composed of $3,000 of discount amortization and $27,000 of interest payments. Of the $30,000, two-thirds of it ($20,000) is expensed; the remaining third ($10,000) is capitalized as plant assets. If the company uses the same interest expense/capitalization proportions to allocate the interest payments between operating and investing activities, then it will report $18,000 (2/3 × $27,000) as an operating outflow and $9,000 (1/3 × $27,000) as an investing outflow. The company might also choose to offset the entire $3,000 of non-cash discount amortization against the $20,000 treated as expense, resulting in an operating outflow as low as $17,000, or as much as $20,000 if it allocated all of the non-cash discount amortization to interest capitalized as investing activities. Similarly, the investing outflow could be as much as $10,000 or as little as $7,000, depending on the treatment of the non-cash discount amortization. There are choices within the choices, all in areas where investors believe choices do not even exist. Nurnberg and Largay (1998) note that companies apparently favor the method that reports the lowest operating outflow, presumably to maximize reported cash from operations.

Investors and analysts need to be aware that presentation choices permitted in IAS 7, "Statement of Cash Flows," offer flexibility in classification of certain items in the cash flow statement. This flexibility can drastically change the results in the operating section of the cash flow statement. An excerpt from IAS 7, Paragraphs 33 and 34, provides the background:

> 33. Interest paid and interest and dividends received are usually classified as operating cash flows for a financial institution. However, there is no consensus on the classification of these cash flows for other entities. Interest paid and interest and dividends received may be classified as operating cash flows because they enter into the determination of profit or loss. *Alternatively, interest paid and interest and dividends received may be classified as financing cash flows and investing cash flows respectively, because they are costs of obtaining financial resources or returns on investments.*
>
> 34. Dividends paid may be classified as a financing cash flow because they are a cost of obtaining financial resources. *Alternatively, dividends paid may be classified as a component of cash flows from operating activities in order to assist users to determine the ability of an entity to pay dividends out of operating cash flows.* [Emphasis added.]

---

[27] See Nurnberg and Largay (1998) and Nurnberg (2006).

By allowing a choice of operating or financing for the placement of interest and dividends received or paid, IAS 7 gives a company's managers the opportunities to select the presentation that gives the best-looking picture of operating performance. An example is Norse Energy Corp. ASA, a Norwegian gas explorer and producer, which changed its classifications of interest paid and interest received in 2007 (Gordon, Henry, Jorgensen, and Linthicum, 2017). Interest paid was switched to financing instead of decreasing cash generated from operations. Norse Energy also switched its classification of interest received to investing from operating cash flow. The net effect of these changes was to report positive, rather than negative, operating cash flows in both 2007 and 2008. With these simple changes, the company could also change the perception of its operations. The cash flow statement formerly presented the appearance of a company with operations that used more cash than it generated, and it possibly raised questions about the sustainability of operations. After the revision, the operating section of the cash flow statement depicted a much more viable operation.

Exhibit 24 shows the net effect of the reclassifications on Norse Energy's cash flows.

Exhibit 24 Reclassification of Cash Flows (amounts in $ millions)

	As Reported (following 2007 reclassification)		Adjustments, If No Reclassification*		Pro-forma (if no reclassification)	
	2008	2007	2008	2007	2008	2007
Operating	**$5.30**	**$2.80**	**($13.70)**	**($14.40)**	**($8.40)**	**($11.60)**
Investing	$0.90	($56.80)	($9.00)	($3.50)	($8.10)	($60.30)
Financing	($16.60)	$34.50	$22.70	$17.90	$6.10	$52.40
Total	($10.40)	($19.50)	$0	$0	($10.40)	($19.50)

* The adjustments reverse the addition of interest received to investing and instead add it to operating. The adjustments also reverse the deduction of interest paid from financing and instead subtract it from operating.

11 CHOICES THAT AFFECT FINANCIAL REPORTING

h describe accounting methods (choices and estimates) that could be used to manage earnings, cash flow, and balance sheet items;

Exhibit 25 summarizes some of the areas where choices can be made that affect financial reports.

Exhibit 25 Areas Where Choices and Estimates Affect Financial Reporting

Area of Choice/ Estimate	Analyst Concerns
Revenue recognition	■ How is revenue recognized: upon shipment or upon delivery of goods? ■ Is the company engaging in “channel stuffing”—the practice of overloading a distribution channel with more product than it is normally capable of selling? This can be accomplished by inducing customers to buy more through unusual discounts, the threat of near-term price increases, or both—or simply by shipping goods that were not ordered. These transactions may be corrected in a subsequent period and may result in restated results. Are accounts receivable relative to revenues abnormally high for relative to the company’s history or to its peers? If so, channel stuffing may have occurred. ■ Is there unusual activity in the allowance for sales returns relative to past history? ■ Does the company’s days sales outstanding show any collection issues that might indicate shipment of unneeded or unwanted goods to customers? ■ Does the company engage in “bill-and-hold” transactions? This is when a customer purchases goods but requests that they remain with the seller until a later date. This kind of transaction makes it possible for a seller to manufacture fictitious sales by declaring end-of-period inventory as “sold but held,” with a minimum of effort and phony documentation. ■ Does the company use rebates as part of its marketing approach? If so, how significantly do the estimates of rebate fulfillment affect net revenues, and have any unusual breaks with history occurred? ■ Does the company separate its revenue arrangements into multiple deliverables of goods or services? This area is one of great revenue recognition flexibility, and also one that provides little visibility to investors. They simply cannot examine a company’s arrangements and decide for themselves whether revenue has been properly allocated to different components of a contract. If a company uses multiple deliverable arrangements with its customers as a routine matter, investors might be more sensitive to revenue reporting risks. In seeking a comfort level, they might ask the following questions: Does the company explain adequately how it determines the different allocations of deliverables and how revenue is recognized on each one? Do deferred revenues result? If not, does it seem reasonable that there are no deferred revenues for this kind of arrangement? Are there unusual trends in revenues and receivables, particularly with regard to cash conversion? If an investor is not satisfied with the answers to these questions, he or she might be more comfortable with other investment choices.

(continued)

Exhibit 25 (Continued)

Area of Choice/ Estimate	Analyst Concerns
Long-lived assets: Depreciation policies	■ Do the estimated life spans of the associated assets make sense, or are they unusually low compared with others in the same industry? ■ Have there been changes in depreciable lives that have a positive effect on current earnings? ■ Do recent asset write-downs indicate that company policy on asset lives might need to be reconsidered?
Intangibles: Capitalization policies	■ Does the company capitalize expenditures related to intangibles, such as software? Does its balance sheet show any R&D capitalized as a result of acquisitions? Or, if the company is an IFRS filer, has it capitalized any internally generated development costs? ■ How do the company's capitalization policies compare with the competition? ■ Are amortization policies reasonable?
Allowance for doubtful accounts/ loan loss reserves	■ Are additions to such allowances lower or higher than in the past? ■ Does the collection experience justify any difference from historical provisioning? ■ Is there a possibility that any lowering of the allowance may be the result of industry difficulties along with the difficulty of meeting earnings expectations?
Inventory cost methods	■ Does the company use a costing method that produces fair reporting results in view of its environment? How do its inventory methods compare with others in its industry? Are there differences that will make comparisons uneven if there are unusual changes in inflation? ■ Does the company use reserves for obsolescence in its inventory valuation? If so, are they subject to unusual fluctuations that might indicate adjusting them to arrive at a specified earnings result? ■ If a company reports under US GAAP and uses last-in-first-out (LIFO) inventory accounting, does LIFO liquidation (the assumed sale of old, lower-cost layers of inventory) occur through inventory reduction programs? This inventory reduction may generate earnings without supporting cash flow, and management may intentionally reduce the layers to produce specific earnings benefits.

Exhibit 25 (Continued)

Area of Choice/ Estimate	Analyst Concerns
Tax asset valuation accounts	■ Tax assets, if present, must be stated at the value at which management expects to realize them, and an allowance must be set up to restate tax assets to the level expected to eventually be converted into cash. Determining the allowance involves an estimate of future operations and tax payments. Does the amount of the valuation allowance seem reasonable, overly optimistic, or overly pessimistic? ■ Are there contradictions between the management commentary and the allowance level, or the tax note and the allowance level? There cannot be an optimistic management commentary and a fully reserved tax asset, or vice versa. One of them has to be wrong. ■ Look for changes in the tax asset valuation account. It may be 100% reserved at first, and then "optimism" increases whenever an earnings boost is needed. Lowering the reserve decreases tax expense and increases net income.
Goodwill	■ Companies must annually assess goodwill balances for impairment on a qualitative basis. If further testing appears necessary, it is based on estimates of the fair value of the reporting units (US GAAP issuers), or cash-generating units (IFRS issuers), associated with goodwill balances. The tests are based on subjective estimates, including future cash flows and the employment of discount rates. ■ Do the disclosures on goodwill testing suggest that the exercise was skewed to avoid impairment charges?
Warranty reserves	■ Have additions to the reserves been reduced, perhaps to make earnings targets? Examine the trend in the charges of actual costs against the reserves: Do they support or contradict the warranty provisioning activity? Do the actual costs charged against the reserve give the analyst any indication about the quality of the products sold?
Related-party transactions	■ Is the company engaged in transactions that disproportionately benefit members of management? Does one company have control over another's destiny through supply contracts or other dealings? ■ Do extensive dealings take place with *non-public* companies that are under management control? If so, those companies could absorb losses (through supply arrangements that are unfavorable to them, for example) in order to make the public company's performance look good. This scenario may provide opportunities for an owner to cash out.

The most important lesson is that choices exist among accounting methods and estimates, and an analyst needs a working knowledge of them in order to understand whether management may have made choices to achieve a desired result.

12

WARNING SIGNS

i. describe accounting warning signs and methods for detecting manipulation of information in financial reports.

The choices management makes to achieve desired results leave a trail, like tracks in sand or snow. The evidence, or warning signs, of information manipulation in financial reports is directly linked to the basic means of manipulation: biased revenue recognition and biased expense recognition. The bias may relate to timing and/or location of recognition. An example of the timing issue is that a company may choose to defer expenses by capitalizing them. Regarding location, it may recognize a loss in other comprehensive income or directly through equity, rather than through the profit and loss statement. The alert investor or analyst should do the following to find warning signs.

1) Pay attention to revenue.

The single largest number on the income statement is revenue, and revenue recognition is a recurring source of accounting manipulation and even outright fraud. Answering the question, "Is revenue higher or lower than the previous comparable period?" is not sufficient. Many analytical procedures can be routinely performed to identify warning signals of malfeasance:

- *Examine the accounting policies note for a company's revenue recognition policies.*
  - Consider whether the policies make it easier to prematurely recognize revenue, such as recognizing revenue immediately upon shipment of goods, or if the company uses bill-and-hold arrangements whereby a sale is recognized before goods are actually shipped to the customer.
  - Barter transactions may exist, which can be difficult to value properly.
  - Rebate programs involve many estimates, including forecasts of the amount of rebates that will ultimately be incurred, which can have significant effects on revenue recognition.
  - Multiple-deliverable arrangements of goods and services are common, but clarity about the timing of revenue recognition for each item or service delivered is necessary for the investor to be comfortable with the reporting of revenues.

  Although none of these decisions necessarily violates accounting standards, each involves significant judgement and warrants close attention if other warning signs are present.
- *Look at revenue relationships.* Compare a company's revenue growth with its primary competitors or its industry peer group.
  - If a company's revenue growth is out of line with its competitors, its industry, or the economy, the investor or analyst needs to understand the reasons for the outperformance. It may be a result of superior management or products and services, but not all management is superior, nor are the products and services of their companies. Revenue quality might be suspect, and the investor should take additional analytical steps.
  - Compare accounts receivable with revenues over several years.

    - Examine the trend to determine whether receivables are increasing as a percentage of total revenues. If so, a company might be engaging in channel-stuffing activities, or worse, recording fictitious sales transactions.
    - Calculate receivables turnover for several years:
        - Examine the trend for unusual changes and seek an explanation if they exist.
        - Compare a company's days sales outstanding (DSO) or receivables turnover with that of relevant competitors or an industry peer group and determine whether the company is an outlier.

  An increase in DSO or decrease in receivables turnover could suggest that some revenues are recorded prematurely or are even fictitious, or that the allowance for doubtful accounts is insufficient.

  - Examine asset turnover. If a company's managers make poor asset allocation choices, revenues may not be sufficient to justify the investment. Be particularly alert when asset allocation choices involve acquisitions of entire companies. If post-acquisition revenue generation is weak, managers might reach for revenue growth anywhere it can be found. That reach for growth might result in accounting abuses.

    Revenues, divided by total assets, indicate the productivity of assets in generating revenues. If the company's asset turnover is continually declining, or lagging the asset turnover of competitors or the industry, it may portend future asset write-downs, particularly in the goodwill balances of acquisitive companies.

2) Pay attention to signals from inventories.

Although inventory is not a component of every company's asset base, its presence creates an opportunity for accounting manipulation.

- *Look at inventory relationships.* Because revenues involve items sold from inventory, the kind of examination an investor should perform on inventory is similar to that for revenues.
  - Compare growth in inventories with competitors and industry benchmarks. If a company's inventory growth is out of line with its peers, without any concurrent sales growth, then it may be simply the result of poor inventory management—an operational inefficiency that might affect an investor's view of a company. It may also signal obsolescence problems in the company's inventory that have not yet been recognized through markdowns to the inventory's net realizable value. Reported gross and net profits could be overstated because of overstated inventory.
  - Calculate the inventory turnover ratio. This ratio is the cost of sales divided by the average ending inventory. Declining inventory turnover could also suggest obsolescence problems that should be recognized.
  - Companies reporting under US GAAP may use LIFO inventory cost flow assumptions. When this assumption is part of the accounting policies, and a company operates in an inflationary environment, investors should note whether old, low-cost inventory costs have been passed through current earnings and artificially improved gross, operating, and net profits.

3) Pay attention to capitalization policies and deferred costs.

In a study of enforcement actions over a five-year period, the SEC found that improper revenue recognition was the most prevalent accounting issue.[28] Suppression of expenses was the next most prevalent problem noted. As the earlier discussion of WorldCom showed, improper capitalization practices can result in a significant misstatement of financial results.

- *Examine the company's accounting policy note for its capitalization policy for long-term assets, including interest costs, and for its handling of other deferred costs.* Compare the company's policy with the industry practice. If the company is the only one capitalizing certain costs while other industry participants treat them as expenses, a red flag is raised. If an outlier company of this type is encountered, it would be useful to cross-check such a company's asset turnover and profitability margins with others in its industry. An investor might expect such a company to be more profitable than its competitors, but the investor might also have lower confidence in the quality of the reported numbers.

4) Pay attention to the relationship of cash flow and net income.

Net income propels stock prices, but cash flow pays bills. Management can manipulate either one, but sooner or later, net income must be realized in cash if a company is to remain viable. When net income is higher than cash provided by operations, one possibility is that aggressive accrual accounting policies have shifted current expenses to later periods. Increasing earnings in the presence of declining cash generated by operations might signal accounting irregularities.

- *Construct a time series of cash generated by operations divided by net income.* If the ratio is consistently below 1.0 or has declined repeatedly, there may be problems in the company's accrual accounts.

5) Other potential warnings signs.

Other areas that might suggest the need for further analysis include the following:

- *Depreciation methods and useful lives.* As discussed earlier, selection of depreciation methods and useful lives can greatly influence profitability. An investor should compare a company's policies with those of its peers to determine whether it is particularly lenient in its effects on earnings. Investors should likewise compare the length of depreciable lives used by a company with those used by its peers.
- *Fourth-quarter surprises.* An investor should be suspicious of possible earnings management if a company routinely disappoints investors with poor earnings or overachieves in the fourth quarter of the year when no seasonality exists in the business. The company may be over- or under-reporting profits in the first three quarters of the year.
- *Presence of related-party transactions.* Related-party transactions often arise when a company's founders are still very active in managing the company, with much of their wealth tied to the company's fortunes. They may be more biased

---

[28] SEC, "Report Pursuant to Section 704 of the Sarbanes–Oxley Act of 2002" (www.sec.gov/news/studies/sox704report.pdf): 5–6.

in their view of a company's performance because it relates directly to their own wealth and reputations, and they may be able to transact business with the company in ways that may not be detected. For instance, they may purchase unsellable inventory from the company for disposal in another company of their own in order to avoid markdowns.

- *Non-operating income or one-time sales included in revenue.* To disguise weakening revenue growth, or just to enhance revenue growth, a company might classify non-operating income items into revenues or fail to clarify the nature of revenues. In the first quarter of 1997, Sunbeam Corporation included one-time disposal of product lines in sales without indicating that such non-recurring sales were included in revenues. This inclusion gave investors a false impression of the company's sustainable revenue-generating capability.
- *Classification of expenses as "non-recurring."* To make operating performance look more attractive, managers might carve out "special items" in the income statement. Particularly when these items appear period after period, equity investors might find their interests best served by not accepting the carve-out of serial "special items" and instead focusing on the net income line in evaluating performance over long periods.
- *Gross/operating margins out of line with competitors or industry.* This disparity is an ambivalent warning sign. It might signal superior management ability. But it might also signal the presence of accounting manipulations to add a veneer of superior management ability to the company's reputation. Only the compilation and examination of other warning signals will enable an investor or analyst to decide which signal is being given.

Warning signals are just that: signals, not indisputable declarations of accounting manipulation guilt. Investors and analysts need to evaluate them cohesively, not on an isolated basis. When an investor finds a number of these signals, the investee company should be viewed with caution or even discarded in favor of alternatives.

Furthermore, as discussed earlier, context is important in judging the value of warning signals. A few examples of facts and circumstances to be aware of are as follows.

- *Younger companies with an unblemished record of meeting growth projections.* It is plausible, especially for a younger company with new and popular product offerings, to generate above-average returns for a period of time. But, as demand dissipates, products mature, and competitors challenge for market share, management may seek to extend its recent record of rapid growth in sales and profitability by unconventional means. At this point, the "earnings games" begin: aggressive estimates, drawing down "cookie jar" reserves, selling assets for accounting gains, taking on excess leverage, or entering into financial transactions with no apparent business purpose other than financial statement "window dressing."
- *Management has adopted a minimalist approach to disclosure.* Confidence in accounting quality depends on disclosure. If management does not seem to take seriously its obligation to provide information, one needs to be concerned. For example, for a large company, management might claim that it has only one reportable segment, or its commentary might be similar from period to period. A plausible explanation for minimalist disclosure policies could be that management is protecting investors' interests by withholding valuable information from competitors. But, this is not necessarily the case. For example, after Sony Corporation acquired CBS Records and Columbia Pictures, it incurred substantial losses for a number of years. Yet, Sony chose to hide its negative trends and doubtful future prospects by aggregating the results within a much larger "Entertainment Division." In 1998, after Sony ultimately wrote off much

of the goodwill associated with these ill-fated acquisitions, the SEC sanctioned Sony and its CFO for failing to separately discuss them in MD&A in a balanced manner.[29]

- *Management fixation on earnings reports.* Beware of companies whose management appears to be fixated on reported earnings, sometimes to the detriment of attending to real drivers of value. Indicators of excessive earnings fixation include the aggressive use of non-GAAP measures of performance, special items, or non-recurring charges. Another indicator of earnings fixation is highly decentralized operations in which division managers' compensation packages are heavily weighted toward the attainment of reported earnings or non-GAAP measures of performance.

Company Culture

A company's culture is an intangible that investors should bear in mind when they are evaluating financial statements for the possibility of accounting manipulation. A management's highly competitive mentality may serve investors well when the company conducts business (assuming that actions taken are not unethical, illegal, or harmfully myopic), but that kind of thinking should not extend to communications with the owners of the company: the shareholders. That mentality can lead to the kind of accounting gamesmanship seen in the early part of the century. In examining financial statements for warning signs of manipulation, the investor should consider whether that mindset exists in the preparation of the financial statements.

One notable example of the mindset comes from one of the most recognized corporate names in the world, General Electric. In the mid-1980s, GE acquired Kidder Peabody, and it was ultimately determined that much of the earnings that Kidder had reported were bogus. As a consequence, GE would announce within two days that it would take a non-cash write-off of $350 million. Here is how former CEO/Chair Jack Welsh described the ensuing meeting with senior management in his memoir, *Straight from the Gut*:

> "The response of our business leaders to the crisis was typical of the GE *culture* [emphasis added]. Even though the books had closed on the quarter, many immediately offered to pitch in to cover the Kidder gap. Some said they could find an extra $10 million, $20 million, and even $30 million from their businesses to offset the surprise. Though it was too late, their willingness to help was a dramatic contrast to the excuses I had been hearing from the Kidder people." (p. 225)

It appears that the corporate governance apparatus fostered a GE culture that extended the concept of teamwork to the point of "sharing" profits to win one for the team as a whole, which is incompatible with the concept of neutral financial reporting. Although research is not conclusive on this question, it may also be worth considering that predisposition to earnings manipulation is more likely to be present when the CEO and board chair are one and the same, or when the audit committee of the board essentially serves at the pleasure of the CEO and lacks financial reporting sophistication. Finally, one could discuss whether the financial reporting environment today would reward or penalize a CEO who openly endorsed a view that he could legitimately exercise financial reporting discretion—albeit within limits—for the purpose of artificially smoothing earnings.

---

[29] Accounting and Auditing Enforcement Release No. 1061, "In the Matter of Sony Corporation and Sumio Sano, Respondents," SEC (5 August 1998).

Restructuring and/or impairment charges.

At times, a company's stock price has been observed to rise after it recognized a "big bath" charge to reported earnings. The conventional wisdom explaining the stock price rise is that accounting recognition signals something positive: that management is now ready to part with the lagging portion of a company, so as to redirect its attention and talents to more-profitable activities. Consequently, the earnings charge should be disregarded for being solely related to past events.

The analyst should also consider, however, that the events leading ultimately to the big bath on the financial statements did not happen overnight, even though the accounting for those events occurs at a subsequent point. Management may want to communicate that the accounting adjustments reflect the company's new path, but the restructuring charge also indicates that the old path of reported earnings was not real. In particular, expenses reported in prior years were very likely understated—even assuming that no improper financial statement manipulation had occurred. To extrapolate historical earnings trends, an analyst should consider making pro forma analytical adjustments to prior years' earnings to reflect a reasonable division of the latest period's restructuring and impairment charges.

Management has a merger and acquisition orientation.

Tyco International Ltd. acquired more than 700 companies from 1996 to 2002. Even assuming the best of intentions regarding financial reporting, a growth-at-any-cost corporate culture poses a severe challenge to operational and financial reporting controls. In Tyco's case, the SEC found that it consistently and fraudulently understated assets acquired (lowering future depreciation and amortization charges) and overstated liabilities assumed (avoiding expense recognition and potentially increasing earnings in future periods).[30]

SUMMARY

Financial reporting quality varies across companies. The ability to assess the quality of a company's financial reporting is an important skill for analysts. Indications of low-quality financial reporting can prompt an analyst to maintain heightened skepticism when reading a company's reports, to review disclosures critically when undertaking financial statement analysis, and to incorporate appropriate adjustments in assessments of past performance and forecasts of future performance.

- Financial reporting quality can be thought of as spanning a continuum from the highest (containing information that is relevant, correct, complete, and unbiased) to the lowest (containing information that is not just biased or incomplete but possibly pure fabrication).
- *Reporting quality*, the focus of this reading, pertains to the information disclosed. High-quality reporting represents the economic reality of the company's activities during the reporting period and the company's financial condition at the end of the period.

---

[30] Accounting and Auditing Enforcement Release No. 2414, "SEC Brings Settled Charges Against Tyco International Ltd. Alleging Billion Dollar Accounting Fraud," SEC (17 April 2006): www.sec.gov/litigation/litreleases/2006/lr19657.htm.

- *Results quality* (commonly referred to as earnings quality) pertains to the earnings and cash generated by the company's actual economic activities and the resulting financial condition, relative to expectations of current and future financial performance. Quality earnings are regarded as being sustainable, providing a sound platform for forecasts.
- An aspect of financial reporting quality is the degree to which accounting choices are conservative or aggressive. "Aggressive" typically refers to choices that aim to enhance the company's reported performance and financial position by inflating the amount of revenues, earnings, and/or operating cash flow reported in the period; or by decreasing expenses for the period and/or the amount of debt reported on the balance sheet.
- Conservatism in financial reports can result from either (1) accounting standards that specifically require a conservative treatment of a transaction or an event or (2) judgments made by managers when applying accounting standards that result in conservative results.
- Managers may be motivated to issue less-than-high-quality financial reports in order to mask poor performance, to boost the stock price, to increase personal compensation, and/or to avoid violation of debt covenants.
- Conditions that are conducive to the issuance of low-quality financial reports include a cultural environment that result in fewer or less transparent financial disclosures, book/tax conformity that shifts emphasis toward legal compliance and away from fair presentation, and limited capital markets regulation.
- Mechanisms that discipline financial reporting quality include the free market and incentives for companies to minimize cost of capital, auditors, contract provisions specifically tailored to penalize misreporting, and enforcement by regulatory entities.
- Pro forma earnings (also commonly referred to as non-GAAP or non-IFRS earnings) adjust earnings as reported on the income statement. Pro forma earnings that exclude negative items are a hallmark of aggressive presentation choices.
- Companies are required to make additional disclosures when presenting any non-GAAP or non-IFRS metric.
- Managers' considerable flexibility in choosing their companies' accounting policies and in formulating estimates provides opportunities for aggressive accounting.
- Examples of accounting choices that affect earnings and balance sheets include inventory cost flow assumptions, estimates of uncollectible accounts receivable, estimated realizability of deferred tax assets, depreciation method, estimated salvage value of depreciable assets, and estimated useful life of depreciable assets.
- Cash from operations is a metric of interest to investors that can be enhanced by operating choices, such as stretching accounts payable, and potentially by classification choices.

REFERENCES

Back, Aaron. 2013. "Toyota, What a Difference the Yen Makes." *Wall Street Journal* (4 August 2013).

Basu, Sudipta. 1997. "The Conservatism Principle and the Asymmetric Timeliness of Earnings." *Journal of Accounting and Economics*, vol. 24, no. 1 (December):3–37.

Bliss, James Harris. 1924. *Management through Accounts*. New York: Ronald Press Company.

Ciesielski, Jack T, and Elaine Henry. 2017. "Accounting's Tower of Babel: Key Considerations in Assessing Non-GAAP Earnings." *Financial Analysts Journal*, vol. 73, no. 2:34–50.

Dichev, Ilia, John Graham, Campbell Harvey, and Shivaram Rajgopal. 2013. "Earnings Quality: Evidence from the Field." *Journal of Accounting and Economics*, vol. 56, issues 2–3:1–33.

Dichev, Ilia, John Graham, Campbell Harvey, and Shiva Rajgopal. 2016. "The Misrepresentation of Earnings." *Financial Analysts Journal*, vol. 72, issue 1:22–35.

Ernst & Young. 2013. *Navigating Today's Complex Business Risks*. Europe, Middle East, India and Africa Fraud Survey 2013 (May): www.ey.com/Publication/vwLUAssets/Navigating_todays_complex_business_risks/$FILE/Navigating_todays_complex_business_risks.pdf.

Ernst & Young Global Limited. 2016. *Corporate misconduct—individual consequences*. 14th Global Fraud Survey. www.ey.com/gl/en/services/assurance/fraud-investigation---dispute-services/ey-global-fraud-survey-2016.

Gordon, Elizabeth, Elaine Henry, Bjorn Jorgensen, and Cheryl Linthicum. 2017. "Flexibility in Cash-Flow Classification under IFRS: Determinants and Consequences." *Review of Accounting Studies*, vol. 22, no. 2:839–872.

Graham, John, Campbell Harvey, and Shiva Rajgopal. 2005. "The Economic Implications of Corporate Financial Reporting." *Journal of Accounting and Economics*, vol. 40, no. 1 (December):3–73.

Lewis, Craig M. 2012. "Risk Modeling at the SEC: The Accounting Quality Model" Speech, the Financial Executives International Committee on Finance and Information Technology (13 December): www.sec.gov/news/speech/2012/spch121312cml.htm.

Nurnberg, H. 2006. "Perspectives on the Cash Flow Statement under FASB Statement No. 95." Center for Excellence in Accounting and Security Analysis Occasional Paper Series. Columbia Business School.

Nurnberg, H., and J. Largay. 1998. "Interest Payments in the Cash Flow Statement." *Accounting Horizons*, vol. 12, no. 4 (December):407–418.

Pavlo, Walter. 2013. "Fmr Enron CFO Andrew Fastow Speaks At ACFE Annual Conference," *Forbes* (26 June): www.forbes.com/sites/walterpavlo/2013/06/26/fmr-enron-cfo-andrew-fastow-speaks-at-acfe-annual-conference/.

Ronen, Joshua, and Varda Yaari. 2008. *Earnings Management: Emerging Insights in Theory, Practice, and Research*. New York: Springer.

Schipper, Katherine. 1989. "Commentary on Earnings Management." *Accounting Horizons*, vol. 3, no. 4 (December):91–102.

Watts, Ross. 2003. "Conservatism in Accounting Part I: Explanations and Implications." *Accounting Horizons*, vol. 17, no. 3 (September):207–221.

PRACTICE PROBLEMS

1 In contrast to earnings quality, financial reporting quality *most likely* pertains to:

A sustainable earnings.
B relevant information.
C adequate return on investment.

2 The information provided by a low-quality financial report will *most likely*:

A decrease company value.
B indicate earnings are not sustainable.
C impede the assessment of earnings quality.

3 To properly assess a company's past performance, an analyst requires:

A high earnings quality.
B high financial reporting quality.
C both high earnings quality and high financial reporting quality.

4 Low quality earnings *most likely* reflect:

A low-quality financial reporting.
B company activities which are unsustainable.
C information that does not faithfully represent company activities.

5 Earnings that result from non-recurring activities *most likely* indicate:

A lower-quality earnings.
B biased accounting choices.
C lower-quality financial reporting.

6 Which attribute of financial reports would *most likely* be evaluated as optimal in the financial reporting spectrum?

A Conservative accounting choices
B Sustainable and adequate returns
C Emphasized pro forma earnings measures

7 Financial reports of the lowest level of quality reflect:

A fictitious events.
B biased accounting choices.
C accounting that is non-compliant with GAAP.

8 When earnings are increased by deferring research and development (R&D) investments until the next reporting period, this choice is considered:

A non-compliant accounting.
B earnings management as a result of a real action.
C earnings management as a result of an accounting choice.

9 A high-quality financial report may reflect:

A earnings smoothing.
B low earnings quality.
C understatement of asset impairment.

© 2019 CFA Institute. All rights reserved.

10 If a particular accounting choice is considered aggressive in nature, then the financial performance for the reporting period would *most likely*:

A be neutral.

B exhibit an upward bias.

C exhibit a downward bias.

11 Which of the following is *most likely* to reflect conservative accounting choices?

A Decreased reported earnings in later periods

B Increased reported earnings in the period under review

C Increased debt reported on the balance sheet at the end of the current period

12 Which of the following is *most likely* to be considered a potential benefit of accounting conservatism?

A A reduction in litigation costs

B Less biased financial reporting

C An increase in current period reported performance

13 Which of the following statements *most likely* describes a situation that would motivate a manager to issue low-quality financial reports?

A The manager's compensation is tied to stock price performance.

B The manager has increased the market share of products significantly.

C The manager has brought the company's profitability to a level higher than competitors.

14 Which of the following concerns would *most likely* motivate a manager to make conservative accounting choices?

A Attention to future career opportunities

B Expected weakening in the business environment

C Debt covenant violation risk in the current period

15 Which of the following conditions *best* explains why a company's manager would obtain legal, accounting, and board level approval prior to issuing low-quality financial reports?

A Motivation

B Opportunity

C Rationalization

16 A company is experiencing a period of strong financial performance. In order to increase the likelihood of exceeding analysts' earnings forecasts in the next reporting period, the company would *most likely* undertake accounting choices for the period under review that:

A inflate reported revenue.

B delay expense recognition.

C accelerate expense recognition.

17 Which of the following situations represents a motivation, rather than an opportunity, to issue low-quality financial reports?

A Poor internal controls

B Search for a personal bonus

C Inattentive board of directors

18 Which of the following situations will *most likely* motivate managers to inflate reported earnings?

- **A** Possibility of bond covenant violation
- **B** Earnings in excess of analysts' forecasts
- **C** Earnings that are greater than the previous year

19 Which of the following *best* describes an opportunity for management to issue low-quality financial reports?

- **A** Ineffective board of directors
- **B** Pressure to achieve some performance level
- **C** Corporate concerns about financing in the future

20 An audit opinion of a company's financial reports is *most likely* intended to:

- **A** detect fraud.
- **B** reveal misstatements.
- **C** assure that financial information is presented fairly.

21 If a company uses a non-GAAP financial measure in an SEC filing, then the company must:

- **A** give more prominence to the non-GAAP measure if it is used in earnings releases.
- **B** provide a reconciliation of the non-GAAP measure and equivalent GAAP measure.
- **C** exclude charges requiring cash settlement from any non-GAAP liquidity measures.

22 A company wishing to increase earnings in the reporting period may choose to:

- **A** decrease the useful life of depreciable assets.
- **B** lower estimates of uncollectible accounts receivables.
- **C** classify a purchase as an expense rather than a capital expenditure.

23 Bias in revenue recognition would *least likely* be suspected if:

- **A** the firm engages in barter transactions.
- **B** reported revenue is higher than the previous quarter.
- **C** revenue is recognized before goods are shipped to customers.

24 Which technique *most likely* increases the cash flow provided by operations?

- **A** Stretching the accounts payable credit period
- **B** Applying all non-cash discount amortization against interest capitalized
- **C** Shifting classification of interest paid from financing to operating cash flows

25 Which of the following is an indication that a company may be recognizing revenue prematurely? Relative to its competitors, the company's:

- **A** asset turnover is decreasing.
- **B** receivables turnover is increasing.
- **C** days sales outstanding is increasing.

26 Which of the following would *most likely* signal that a company may be using aggressive accrual accounting policies to shift current expenses to later periods? Over the last five-year period, the ratio of cash flow to net income has:

- **A** increased each year.
- **B** decreased each year.
- **C** fluctuated from year to year.

27 An analyst reviewing a firm with a large reported restructuring charge to earnings should:

A view expenses reported in prior years as overstated.

B disregard it because it is solely related to past events.

C consider making pro forma adjustments to prior years' earnings.

SOLUTIONS

1 B is correct. Financial reporting quality pertains to the quality of information in financial reports. High-quality financial reporting provides decision-useful information, which is relevant and faithfully represents the economic reality of the company's activities. Earnings of high quality are sustainable and provide an adequate level of return. Highest-quality financial reports reflect both high financial reporting quality and high earnings quality.

2 C is correct. Financial reporting quality pertains to the quality of the information contained in financial reports. High-quality financial reports provide decision-useful information that faithfully represents the economic reality of the company. Low-quality financial reports impede assessment of earnings quality. Financial reporting quality is distinguishable from earnings quality, which pertains to the earnings and cash generated by the company's actual economic activities and the resulting financial condition. Low-quality earnings are not sustainable and decrease company value.

3 B is correct. Financial reporting quality pertains to the quality of the information contained in financial reports. If financial reporting quality is low, the information provided is of little use in assessing the company's performance. Financial reporting quality is distinguishable from earnings quality, which pertains to the earnings and cash generated by the company's actual economic activities and the resulting financial condition.

4 B is correct. Earnings quality pertains to the earnings and cash generated by the company's actual economic activities and the resulting financial condition. Low-quality earnings are likely not sustainable over time because the company does not expect to generate the same level of earnings in the future or because earnings will not generate sufficient return on investment to sustain the company. Earnings that are not sustainable decrease company value. Earnings quality is distinguishable from financial reporting quality, which pertains to the quality of the information contained in financial reports.

5 A is correct. Earnings that result from non-recurring activities are unsustainable. Unsustainable earnings are an example of lower-quality earnings. Recognizing earnings that result from non-recurring activities is neither a biased accounting choice nor indicative of lower quality financial reporting because it faithfully represents economic events.

6 B is correct. At the top of the quality spectrum of financial reports are reports that conform to GAAP, are decision useful, and have earnings that are sustainable and offer adequate returns. In other words, these reports have both high financial reporting quality and high earnings quality.

7 A is correct. Financial reports span a quality continuum from high to low based on decision-usefulness and earnings quality (see Exhibit 2 of the reading). The lowest-quality reports portray fictitious events, which may misrepresent the company's performance and/or obscure fraudulent misappropriation of the company's assets.

8 B is correct. Deferring research and development (R&D) investments into the next reporting period is an example of earnings management by taking a *real* action.

9 B is correct. High-quality financial reports offer useful information, meaning information that is relevant and faithfully represents actual performance. Although low earnings quality may not be desirable, if the reported earnings

are representative of actual performance, they are consistent with high-quality financial reporting. Highest-quality financial reports reflect both high financial reporting quality and high earnings quality.

10 B is correct. Aggressive accounting choices aim to enhance the company's reported performance by inflating the amount of revenues, earnings, and/or operating cash flow reported in the period. Consequently, the financial performance for that period would most likely exhibit an upward bias.

11 C is correct. Accounting choices are considered conservative if they decrease the company's reported performance and financial position in the period under review. Conservative choices may increase the amount of debt reported on the balance sheet. They may decrease the revenues, earnings, and/or operating cash flow reported for the period and increase those amounts in later periods.

12 A is correct. Conservatism reduces the possibility of litigation and, by extension, litigation costs. Rarely, if ever, is a company sued because it understated good news or overstated bad news. Accounting conservatism is a type of bias in financial reporting that decreases a company's reported performance. Conservatism directly conflicts with the characteristic of neutrality.

13 A is correct. Managers often have incentives to meet or beat market expectations, particularly if management compensation is linked to increases in stock prices or to reported earnings.

14 B is correct. Managers may be motivated to understate earnings in the reporting period and increase the probability of meeting or exceeding the next period's earnings target.

15 C is correct. Typically, conditions of opportunity, motivation, and rationalization exist when individuals issue low-quality financial reports. Rationalization occurs when an individual is concerned about a choice and needs to be able to justify it to herself or himself. If the manager is concerned about a choice in a financial report, she or he may ask for other opinions to convince herself or himself that it is okay.

16 C is correct. In a period of strong financial performance, managers may pursue accounting choices that increase the probability of exceeding earnings forecasts for the next period. By accelerating expense recognition or delaying revenue recognition, managers may inflate earnings in the next period and increase the likelihood of exceeding targets.

17 B is correct. Motivation can result from pressure to meet some criteria for personal reasons, such as a bonus, or corporate reasons, such as concern about future financing. Poor internal controls and an inattentive board of directors offer opportunities to issue low-quality financial reports.

18 A is correct. The possibility of bond covenant violations may motivate managers to inflate earnings in the reporting period. In so doing, the company may be able to avoid the consequences associated with violating bond covenants.

19 A is correct. Opportunities to issue low-quality financial reports include internal conditions, such as an ineffective board of directors, and external conditions, such as accounting standards that provide scope for divergent choices. Pressure to achieve a certain level of performance and corporate concerns about future financing are examples of motivations to issue low-quality financial reports. Typically, three conditions exist when low-quality financial reports are issued: opportunity, motivation, and rationalization.

20 C is correct. An audit is intended to provide assurance that the company's financial reports are presented fairly, thus providing discipline regarding financial reporting quality. Regulatory agencies usually require that the financial

statements of publicly traded companies be audited by an independent auditor to provide assurance that the financial statements conform to accounting standards. Privately held companies may also choose to obtain audit opinions either voluntarily or because an outside party requires it. An audit is not typically intended to detect fraud. An audit is based on sampling and it is possible that the sample might not reveal misstatements.

21 B is correct. If a company uses a non-GAAP financial measure in an SEC filing, it is required to provide the most directly comparable GAAP measure with equivalent prominence in the filing. In addition, the company is required to provide a reconciliation between the non-GAAP measure and the equivalent GAAP measure. Similarly, IFRS require that any non-IFRS measures included in financial reports must be defined and their potential relevance explained. The non-IFRS measures must be reconciled with IFRS measures.

22 B is correct. If a company wants to increase reported earnings, the company's managers may reduce the allowance for uncollected accounts and the related expense reported for the period. Decreasing the useful life of depreciable assets would increase depreciation expense and decrease earnings in the reporting period. Classifying a purchase as an expense, rather than capital expenditure, would decrease earnings in the reporting period. The use of accrual accounting may result in estimates in financial reports, because all facts associated with events may not be known at the time of recognition. These estimates can be grounded in reality or managed by the company to present a desired financial picture.

23 B is correct. Bias in revenue recognition can lead to manipulation of information presented in financial reports. Addressing the question as to whether revenue is higher or lower than the previous period is not sufficient to determine if there is bias in revenue recognition. Additional analytical procedures must be performed to identify warning signals of accounting malfeasance. Barter transactions are difficult to value properly and may result in bias in revenue recognition. Policies that make it easier to prematurely recognize revenue, such as before goods are shipped to customers, may be a warning sign of accounting malfeasance.

24 A is correct. Managers can temporarily show a higher cash flow from operations by stretching the accounts payable credit period. In other words, the managers delay payments until the next accounting period. Applying all non-cash discount amortization against interest capitalized causes reported interest expenses and operating cash outflow to be higher, resulting in a lower cash flow provided by operations. Shifting the classification of interest paid from financing to operating cash flows lowers the cash flow provided by operations.

25 C is correct. If a company's days sales outstanding (DSO) is increasing relative to competitors, this may be a signal that revenues are being recorded prematurely or are even fictitious. There are numerous analytical procedures that can be performed to provide evidence of manipulation of information in financial reporting. These warning signs are often linked to bias associated with revenue recognition and expense recognition policies.

26 B is correct. If the ratio of cash flow to net income for a company is consistently below 1 or has declined repeatedly over time, this may be a signal of manipulation of information in financial reports through aggressive accrual accounting policies. When net income is consistently higher than cash provided by operations, one possible explanation is that the company may be using aggressive accrual accounting policies to shift current expenses to later periods.

27 C is correct. To extrapolate historical earnings trends, an analyst should consider making pro forma analytical adjustments of prior years' earnings to reflect in those prior years a reasonable share of the current period's restructuring and impairment charges.

READING

26

Applications of Financial Statement Analysis

by Thomas R. Robinson, PhD, CFA, J. Hennie van Greuning, DCom, CFA, Elaine Henry, PhD, CFA, and Michael A. Broihahn, CPA, CIA, CFA

Thomas R. Robinson, PhD, CFA, is at AACSB International (USA). J. Hennie van Greuning, DCom, CFA, is at BIBD (Brunei). Elaine Henry, PhD, CFA, is at Stevens Institute of Technology (USA). Michael A. Broihahn, CPA, CIA, CFA, is at Barry University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** evaluate a company's past financial performance and explain how a company's strategy is reflected in past financial performance;
☐	**b.** demonstrate how to forecast a company's future net income and cash flow;
☐	**c.** describe the role of financial statement analysis in assessing the credit quality of a potential debt investment;
☐	**d.** describe the use of financial statement analysis in screening for potential equity investments;
☐	**e.** explain appropriate analyst adjustments to a company's financial statements to facilitate comparison with another company.

Note: Changes in accounting standards as well as new rulings and/or pronouncements issued after the publication of the readings on financial reporting and analysis may cause some of the information in these readings to become dated. Candidates are *not* responsible for anything that occurs after the readings were published. In addition, candidates are expected to be familiar with the analytical frameworks contained in the readings, as well as the implications of alternative accounting methods for financial analysis and valuation discussed in the readings. Candidates are also responsible for the content of accounting standards, but not for the actual reference numbers. Finally, candidates should be aware that certain ratios may be defined and calculated differently. When alternative ratio definitions exist and no specific definition is given, candidates should use the ratio definitions emphasized in the readings.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION & EVALUATING PAST FINANCIAL PERFORMANCE

a evaluate a company's past financial performance and explain how a company's strategy is reflected in past financial performance;

This reading presents several important applications of financial statement analysis. Among the issues we will address are the following:

- What are the key questions to address in evaluating a company's past financial performance?
- How can an analyst approach forecasting a company's future net income and cash flow?
- How can financial statement analysis be used to evaluate the credit quality of a potential fixed-income investment?
- How can financial statement analysis be used to screen for potential equity investments?
- How can differences in accounting methods affect financial ratio comparisons between companies, and what are some adjustments analysts make to reported financials to facilitate comparability among companies?

The reading "Financial Statement Analysis: An Introduction" described a framework for conducting financial statement analysis. Consistent with that framework, prior to undertaking any analysis, an analyst should explore the purpose and context of the analysis. The purpose and context guide further decisions about the approach, the tools, the data sources, and the format in which to report results of the analysis, and also suggest which aspects of the analysis are most important. Having identified the purpose and context, the analyst should then be able to formulate the key questions that the analysis must address. The questions will suggest the data the analyst needs to collect to objectively address the questions. The analyst then processes and analyzes the data to answer these questions. Conclusions and decisions based on the analysis are communicated in a format appropriate to the context, and follow-up is undertaken as required. Although this reading will not formally present applications as a series of steps, the process just described is generally applicable.

Section 1.1 of this reading describes the use of financial statement analysis to evaluate a company's past financial performance, and Sections 2 and 3 describe basic approaches to projecting a company's future financial performance. Section 4 presents the use of financial statement analysis in assessing the credit quality of a potential debt investment. Section 5 concludes the survey of applications by describing the use of financial statement analysis in screening for potential equity investments. Analysts often encounter situations in which they must make adjustments to a company's reported financial results to increase their accuracy or comparability with the financials of other companies. Sections 6–8 illustrate several common types of analyst adjustments. Section 9 presents a summary, and practice problems in the CFA Institute multiple-choice format conclude the reading.

1.1 Application: Evaluating Past Financial Performance

Analysts examine a company's past financial performance for a number of reasons. Cross-sectional analysis of financial performance facilitates understanding of the comparability of companies for a market-based valuation.[1] Analysis of a company's historical performance over time can provide a basis for a forward-looking analysis of the company. Both cross-sectional and trend analysis can provide information for evaluating the quality and performance of a company's management.

An evaluation of a company's past performance addresses not only *what* happened (i.e., how the company performed) but also *why* it happened—the causes behind the performance and how the performance reflects the company's strategy. Evaluative judgments assess whether the performance is better or worse than a relevant benchmark, such as the company's own historical performance, a competitor's performance, or market expectations. Some key analytical questions include the following:

- How and why have corporate measures of profitability, efficiency, liquidity, and solvency changed over the periods being analyzed?
- How do the level and trend in a company's profitability, efficiency, liquidity, and solvency compare with the corresponding results of other companies in the same industry? What factors explain any differences?
- What aspects of performance are critical for a company to successfully compete in its industry, and how did the company perform relative to those critical performance aspects?
- What are the company's business model and strategy, and how did they influence the company's performance as reflected in, for example, its sales growth, efficiency, and profitability?

Data available to answer these questions include the company's (and its competitors') financial statements, materials from the company's investor relations department, corporate press releases, and non-financial-statement regulatory filings, such as proxies. Useful data also include industry information (e.g., from industry surveys, trade publications, and government sources), consumer information (e.g., from consumer satisfaction surveys), and information that is gathered by the analyst firsthand (e.g., through on-site visits). Processing the data typically involves creating common-size financial statements, calculating financial ratios, and reviewing or calculating industry-specific metrics. Example 1 illustrates the effects of strategy on performance and the use of basic economic reasoning in interpreting results.

EXAMPLE 1

A Change in Products Reflected in Financial Performance

Apple Inc. is a company that has evolved and adapted over time. In its 1994 Prospectus filed with the US SEC, Apple identified itself as "one of the world's leading personal computer technology companies." At that time, most of its revenue was generated by computer sales. In the prospectus, however, Apple stated, "The Company's strategy is to expand its market share in the personal computing industry while developing and expanding into new related business such as Personal Interactive Electronics and Apple Business Systems." Over time, products other than computers became significant generators of revenue and profit.

---

[1] Pinto et al. (2010) describe market-based valuation as using price multiples—ratios of a stock's market price to some measure of value per share (e.g., price-to-earnings ratios). Although the valuation method may be used independently of an analysis of a company's past financial performance, such an analysis may provide reasons for differences in companies' price multiples.

In 2005, an article in *Barron's* said, "In the last year, the iPod has become Apple's best-selling product, bringing in a third of revenues for the Cupertino, Calif. firm . . . Little noticed by these iPod zealots, however is a looming threat . . . Wireless phone companies are teaming up with the music industry to make most mobile phones into music players" (*Barron's* 27 June 2005, p. 19). The threat noted by *Barron's* was not unnoticed or ignored by Apple.

In June 2007, Apple itself entered the mobile phone market with the launch of the original iPhone, followed in June 2008 by the second-generation iPhone 3G (a handheld device combining the features of a mobile phone, an iPod, and an internet connection device). Soon after, the company launched the iTunes App Store, which allows users to download third-party applications onto their iPhones. As noted in a 2009 *Business Week* article, Apple "is the world's largest music distributor, having passed Wal-Mart Stores in early 2008. Apple sells around 90% of song downloads and 75% of digital music players in the United States" (*Business Week*, 28 September 2009, p. 34). Product innovations continue as evidenced by the introduction of the iPad in January 2010.

In analyzing the historical performance of Apple in 2018, an analyst might refer to the information presented in Exhibit 1, which shows sales, profitability, sales by product line and product mix.

Exhibit 1 (dollars in millions)

Sales and Profitability	**2017**	**2016**	**2015**	**2014**	**2013**	**2012**	**2011**	**2010**
Sales	229,234	215,639	233,715	182,795	170,910	156,608	108,249	65,225
Cost of goods sold	141,048	131,376	140,089	112,258	106,606	87,846	64,431	39,541
Gross profit	88,186	84,263	93,626	70,537	64,304	68,762	43,818	25,684
Gross margin	38.5%	39.1%	40.1%	38.6%	37.6%	43.9%	40.5%	39.4%
Net sales by product								
Mac	25,850	22,831	25,471	24,079	21,483	23,221	21,783	17,480
iPhone and related	141,319	136,700	155,041	101,991	91,279	78,692	45,998	25,177
iPad and related	19,222	20,628	23,227	30,283	31,980	30,945	19,168	4,957
Services	29,980	24,348	19,909	18,063	16,051	12,890	9,373	10,110
Other (includes iPod)	12,863	11,132	10,067	8,379	10,117	10,760	11,927	7,501
Total	229,234	215,639	233,715	182,795	170,910	156,508	108,249	65,225
Net sales % by product								
Mac	11.3%	10.6%	10.9%	13.2%	12.6%	14.8%	20.1%	26.8%
iPhone and related	61.6%	63.4%	66.3%	55.8%	53.4%	50.3%	42.5%	38.6%

Exhibit 1 (Continued)

Sales and Profitability	2017	2016	2015	2014	2013	2012	2011	2010
iPad and related	8.4%	9.6%	9.9%	16.6%	18.7%	19.8%	17.7%	7.6%
Services	13.1%	11.3%	8.5%	9.9%	9.4%	8.2%	8.7%	15.5%
Other (includes iPod)	5.6%	5.2%	4.3%	4.6%	5.9%	6.9%	11.0%	11.5%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Source: Apple 10-K filings.

Using the information provided, address the following:

1. How have sales and gross margin changed over time?
2. How has the company's product mix changed since the introduction of the iPad in 2010, and what might this change suggest for an analyst in evaluating Apple's profitability over time and its ability to maintain that profitability?

Solution to 1:

Since 2010 total sales have increased from $65 billion to $229 billion. This represents an annualized growth rate of almost 20%. There was only one year that did not have sales growth in dollars (2016). Gross margin has ranged from 37.6% to 43.9%. Gross margin increased from 2010, when the iPad was introduced, through 2012, when it reached its peak. Gross margin then declined in 2013 and trended upward through 2015. There were modest declines in gross margin after 2015.

Solution to 2:

When the iPad was introduced in 2010 it received a significant share of the product mix, rising to 17.7% in 2011, the first full year after introduction. The iPad's product mix share approached 20% share in 2012 and then declined slightly for two years before a larger decline down to a relatively stable product mix share of around 9%. This could be explained by reaching fairly widespread adoption. The iPhone also gained significant product mix share, rising steadily from 38.6% in 2010 to 66.3% in 2015. Share declined slightly since 2015 but still remains the largest of Apple's product segments at more than 60%. Sales of their original product, the Mac, have declined from more than 25% of sales to around 10%. Services have changed significantly but have shown a steady increase in recent years, most likely due to Apple's music and other media subscription plans. Initially a blockbuster product, the iPod is now included in "other," and this is the largest driver of the decline in that category over time.

Apple had a history of introducing new products every few years, but in recent years the company has not created new product categories. Instead the company has periodically introduced new models of iPads and iPhones. The recent decline in margins is attributable in part to the lack of new products and services and highlights the importance of product innovation to Apple in maintaining historically healthy margins.

In calculating and interpreting financial statement ratios, an analyst needs to be aware of the potential impact on the financial statements and related ratios of companies reporting under different accounting standards, such as international financial reporting standards (IFRS), US generally accepted accounting principles (US GAAP), or other home-country GAAP. Furthermore, even within a given set of accounting standards, companies still have discretion to choose among acceptable methods. A company also may make different assumptions and estimates even when applying the same method as another company. Therefore, making selected adjustments to a company's financial statement data may be useful to facilitate comparisons with other companies or with the industry overall. Examples of such analyst adjustments will be discussed in Sections 6–8.

Non-US companies that use any acceptable body of accounting standards (other than IFRS or US GAAP) and file with the US SEC (because their shares or depositary receipts based on their shares trade in the United States) are required to reconcile their net income and shareholders' equity accounts to US GAAP. Note that in 2007, the SEC eliminated the reconciliation requirement for non-US companies using IFRS and filing with the SEC, however companies may still voluntarily provide this information for comparison purposes.

In general, because the reconciliation data are no longer required by the SEC, we cannot always determine whether differences in net income, equity, and thus ROE also exist between IFRS and the companies' home-country GAAP (including US GAAP).

Comparison of the levels and trends in a company's performance provide information about *how* the company performed. The company's management presents its view about causes underlying its performance in the management commentary or management discussion and analysis (MD&A) section of its annual report and during periodic conference calls with analysts and investors. To gain additional understanding of the causes underlying a company's performance, an analyst can review industry information or seek information from additional sources, such as consumer surveys.

The results of an analysis of past performance provide a basis for reaching conclusions and making recommendations. For example, an analysis undertaken as the basis for a forward-looking study might conclude that a company's future performance is or is not likely to reflect continuation of recent historical trends. As another example, an analysis to support a market-based valuation of a company might focus on whether the company's profitability and growth outlook, which is better (worse) than the peer group median, justifies its relatively high (low) valuation. This analysis would consider market multiples, such as price-to-earnings ratio (P/E), price-to-book ratio, and total invested capital to EBITDA (earnings before interest, taxes, depreciation, and amortization).[2] As another example, an analysis undertaken as part of an evaluation of the management of two companies might result in conclusions about whether one company has grown as fast as another company, or as fast as the industry overall, and whether each company has maintained profitability while growing.

2

APPLICATION: PROJECTING FUTURE FINANCIAL PERFORMANCE AS AN INPUT TO MARKET BASED VALUATION

b demonstrate how to forecast a company's future net income and cash flow;

---

[2] **Total invested capital** is the sum of market value of common equity, book value of preferred equity, and face value of debt.

Projections of future financial performance are used in determining the value of a company or its equity component. Projections of future financial performance are also used in credit analysis—particularly in project finance or acquisition finance—to determine whether a company's cash flows will be adequate to pay the interest and principal on its debt and to evaluate whether a company will likely remain in compliance with its financial covenants.

Sources of data for analysts' projections include some or all of the following: the company's projections, the company's previous financial statements, industry structure and outlook, and macroeconomic forecasts.

Evaluating a company's past performance may provide a basis for forward-looking analyses. An evaluation of a company's business and economic environment and its history may persuade the analyst that historical information constitutes a valid basis for such analyses and that the analyst's projections may be based on the continuance of past trends, perhaps with some adjustments. Alternatively, in the case of a major acquisition or divestiture, for a start-up company, or for a company operating in a volatile industry, past performance may be less relevant to future performance.

Projections of a company's near-term performance may be used as an input to market-based valuation or relative valuation (i.e., valuation based on price multiples). Such projections may involve projecting next year's sales and using the common-size income statement to project major expense items or particular margins on sales (e.g., gross profit margin or operating profit margin). These calculations will then lead to the development of an income measure for a valuation calculation, such as net income, earnings per share (EPS) or EBITDA. More complex projections of a company's future performance involve developing a more detailed analysis of the components of performance for multiple periods—for example, projections of sales and gross margin by product line, projection of operating expenses based on historical patterns, and projection of interest expense based on requisite debt funding, interest rates, and applicable taxes. Furthermore, a projection should include sensitivity analyses applied to the major assumptions.

2.1 Projecting Performance: An Input to Market-Based Valuation

One application of financial statement analysis involves projecting a company's near-term performance as an input to market-based valuation. For example, an analyst might project a company's sales and profit margin to estimate EPS and then apply a projected P/E to establish a target price for the company's stock.

Analysts often take a top-down approach to projecting a company's sales.[3] First, industry sales are projected on the basis of their historical relationship with some macroeconomic indicator, such as growth in real gross domestic product (GDP). In researching the automobile industry, for example, the analyst may find that the industry's annual domestic unit car sales (number of cars sold in domestic markets) bears a relationship to annual changes in real GDP. Regression analysis is often used to establish the parameters of such relationships. Other factors in projecting sales may include consumer income or tastes, technological developments, and the availability of substitute products or services. After industry sales are projected, a company's market share is projected. Company-level market share projections may be based on historical market share and a forward-looking assessment of the company's competitive position. The company's sales are then estimated as its projected market share multiplied by projected total industry sales.

---

[3] The discussion in this paragraph is indebted to Benninga and Sarig (1997).

After developing a sales forecast for a company, an analyst can choose among various methods for forecasting income and cash flow. An analyst must decide on the level of detail to consider in developing forecasts. For example, separate forecasts may be made for individual expense items or for more aggregated expense items, such as total operating expenses. Rather than stating a forecast in terms of expenses, the forecast might be stated in terms of a forecasted profit margin (gross, operating, or net). The net profit margin, in contrast to the gross or operating profit margins, is affected by financial leverage and tax rates, which are subject to managerial and legal/regulatory revisions; therefore, historical data may sometimes be more relevant for projecting gross or operating profit margins than for projecting net profit margins. Whatever the margin used, the forecasted amount of profit for a given period is the product of the forecasted amount of sales and the forecast of the selected profit margin.

As Example 2 illustrates, for relatively mature companies operating in non-volatile product markets, historical information on operating profit margins can provide a useful starting point for forecasting future operating profits (at least over short forecasting horizons). Historical operating profit margins are typically less reliable for projecting future margins for a new or relatively volatile business or one with significant fixed costs (which can magnify the volatility of operating margins).

EXAMPLE 2

Using Historical Operating Profit Margins to Forecast Operating Profit

One approach to projecting operating profit is to determine a company's average operating profit margin over the previous several years and apply that margin to a forecast of the company's sales. Use the following information on three companies to answer Questions 1 and 2 below:

- Johnson & Johnson (JNJ). This US health care conglomerate, founded in 1887, had 2017 sales of around $76.5 billion from its three main businesses: pharmaceuticals, medical devices and diagnostics, and consumer products.
- BHP Billiton (BHP). This company, with group headquarters in Australia and secondary headquarters in London, is the world's largest natural resources company, reporting revenue of approximately US$38.3 billion for the fiscal year ended June 2017. The company mines, processes, and markets coal, copper, nickel, iron, bauxite, and silver and also has substantial petroleum operations.
- Baidu. This Chinese company, which was established in 2000 and went public on NASDAQ in 2005, is the leading Chinese language search engine. The company's revenues for 2017 were 84.8 billion renminbi (RMB), an increase of 20 percent from 2016 and almost 4 times revenue in 2012.

1. For each of the three companies, state and justify whether the suggested forecasting method (applying the average operating profit over the previous several years to a forecast of sales) would be a reasonable starting point for projecting future operating profit.
2. Assume that the 2017 forecast of sales was perfect and, therefore, equal to the realized sales by the company in 2017. Compare the forecast of 2017 operating profit, using an average of the previous five years' operating profit margins, with the actual 2017 operating profit reported by the company given the following additional information:

- JNJ: For the five years prior to 2017, JNJ's average operating profit margin was approximately 25.6 percent. The company's actual operating profit for 2017 was $18.2 billion.
- BHP: For the four years prior to the year ending June 2017, BHP's average operating profit margin was approximately 24.0 percent. The company's actual operating profit for the year ended June 2017 was US$11.8 billion.
- Baidu: Over the four years prior to 2017, Baidu's average operating profit margin was approximately 28.5 percent. The company's actual operating profit for 2017 was RMB15.7 billion.

Using the additional information given, state and justify whether actual results support the usefulness of the stable operating margin assumption.

Solution to 1:

JNJ. Because JNJ is an established company with diversified operations in relatively stable businesses, the suggested approach to projecting the company's operating profit would be a reasonable starting point.

BHP. Because commodity prices tend to be volatile and the mining industry is relatively capital intensive, the suggested approach to projecting BHP's operating profit would probably not be a useful starting point.

Baidu. Compared to the other two companies, Baidu has a more limited operating history and remains in a period of rapid growth. These aspects about the company suggests that the broad approach to projecting operating profit would not be a useful starting point for Baidu.

Solution to 2:

JNJ. JNJ's actual operating profit margin for 2017 was 23.8 percent ($18.2 billion divided by sales of $76.5 billion), which is a little less than company's five-year average operating profit margin of approximately 25.6 percent.

BHP. BHP's actual operating profit margin for the year ended June 2017 was 30.8 percent ($11.8 billion divided by sales of $38.3 billion). If the company's average profit margin of 24.0 percent had been applied to perfectly forecasted sales, the forecasted operating profit would have been approximately US$9.2 billion, around 22 percent lower than actual operating profit.

Baidu. Baidu's actual operating profit margin for 2017 was 18.5 percent (RMB15.7 billion divided by sales of RMB84.8 billion). If the average profit margin of 28.5 percent had been applied to perfectly forecasted sales, the forecasted operating profit would have been approximately RMB24.2 billion, or around 54 percent higher than Baidu's actual operating profit.

Although prior years' profit margins can provide a useful starting point in projections for companies with relatively stable business, the underlying data should, nonetheless, be examined to identify items that are not likely to occur again in the following year(s). Such non-recurring (i.e., transitory) items should be removed from computations of any profit amount or profit margin that will be used in projections. Example 3 illustrates this principle.

EXAMPLE 3

Issues in Forecasting

Following are excerpts from the 2017 annual report of Textron, a global aircraft, defense and industrial company.

Textron Consolidated Statements of Operations for each of the years in the three-year period ended December 31

(In millions, except per share data)	2017	2016	2015
Revenues			
Manufacturing revenues	$14,129	$13,710	$13,340
Finance revenues	69	78	83
Total revenues	14,198	13,788	13,423
Costs, expenses and other			
Cost of sales	11,795	11,311	10,979
Selling and administrative expense	1,337	1,304	1,304
Interest expense	174	174	169
Special charges	130	123	—
Total costs, expenses and other	13,436	12,912	12,452
Income from continuing operations before income taxes	762	876	971
Income tax expense	456	33	273
Income from continuing operations	306	843	698
Income (loss) from discontinued operations, net of income taxes*	1	119	(1)
Net income	307	962	697

Footnotes:
2017 Note 12 Special Charges

(Continued)

> In 2016, we initiated a plan to restructure and realign our businesses by implementing headcount reductions, facility consolidations and other actions in order to improve overall operating efficiency across Textron. Under this plan, Textron Systems discontinued production of its sensor-fuzed weapon product within its Weapons and Sensors operating unit, we combined our Jacobsen business with the Textron Specialized Vehicles business by consolidating facilities and general and administrative functions, and we reduced headcount at Textron Aviation, as well as other businesses and corporate functions. In December 2017, we decided to take additional restructuring actions to further consolidate operating facilities and streamline product lines, primarily within the Bell, Textron Systems and Industrial segments, which resulted in additional special charges of $45 million in the fourth quarter of 2017. We recorded total special charges of $213 million since the inception of the 2016 plan, which included $97 million of severance costs, $84 million of asset impairments and $32 million in contract terminations and other costs. Of these amounts, $83 million was incurred at Textron Systems, $63 million at Textron Aviation, $38 million at Industrial, $28 million at Bell and $1 million at Corporate. The total headcount reduction under this plan is expected to be approximately 2,100 positions, representing 5% of our workforce. In connection with the acquisition of Arctic Cat, as discussed in Note 2, we initiated a restructuring plan in the first quarter of 2017 to integrate this business into our Textron Specialized Vehicles business within the Industrial segment and reduce operating redundancies and maximize efficiencies. Under the Arctic Cat plan, we recorded restructuring charges of $28 million in 2017, which included $19 million of severance costs, largely related to change-of-control provisions, and $9 million of contract termination and other costs. In addition, we recorded $12 million of acquisition-related integration and transaction costs in 2017.

2016 Financial Statement General Footnote

> *Income from discontinued operations, net of income taxes for the year ended December 31, 2016 primarily includes the settlement of a U.S. federal income tax audit. See Note 13 to the Consolidated Financial Statements for additional information.

2016 Note 13 Income Taxes

> The provision for income taxes for 2016 included a benefit of $319 million to reflect the settlement with the U.S. Internal Revenue Service Office of Appeals for our 1998 to 2008 tax years, which resulted in a $206 million benefit attributable to continuing operations and $113 million attributable to discontinued operations.

Source: Textron annual reports.

Discussion:

Results of discontinued operations and restructuring charges should generally not be included when assessing past performance or when forecasting future net income. For purposes of evaluating the company's ongoing operating and net profit margins the special charges related to restructuring and the special tax benefit related to discontinued operations should be removed. For example, the company's operating margin for 2017 including special charges would be 5.4% (\$762 million/\$14,198 million). Excluding special charges, the operating margin would be 6.3% (\$762 million + \$130 million)/\$14,198 million. Similarly, the net profit margin would be determined by eliminating the income from discontinued operations, particularly for 2016.

In general, when earnings projections are used as a foundation for market-based valuations, an analyst will make appropriate allowance for transitory components of past earnings. Occasionally, an analyst will observe that a company takes special charges virtually every year. In such cases, they are not transitory and should not be removed in evaluating past and future margins.

3

PROJECTING MULTIPLE-PERIOD PERFORMANCE

b demonstrate how to forecast a company's future net income and cash flow;

Projections of future financial performance over multiple periods are needed in valuation models that estimate the value of a company or its equity by discounting future cash flows. The value of a company or its equity developed in this way can then be compared with its current market price as a basis for investment decisions.

Projections of future performance are also used for credit analysis. These projections are important in assessing a borrower's ability to repay interest and principal of debt obligations. Investment recommendations depend on the needs and objectives of the client and on an evaluation of the risk of the investment relative to its expected return—both of which are a function of the terms of the debt obligation itself as well as financial market conditions. Terms of the debt obligation include amount, interest rate, maturity, financial covenants, and collateral.

Example 4 presents an elementary illustration of net income and cash flow forecasting to illustrate a format for analysis and some basic principles. In Example 4, assumptions are shown first; then, the period-by-period abbreviated financial statement resulting from the assumptions is shown.

Depending on the use of the forecast, an analyst may choose to compute further, more specific cash flow metrics. For example, free cash flow to equity, which is used in discounted cash flow approaches to equity valuation, can be estimated as net income adjusted for noncash items, minus investment in net working capital and in net fixed assets, plus net borrowing.

EXAMPLE 4

Basic Example of Financial Forecasting

Assume a company is formed with \$100 of equity capital, all of which is immediately invested in working capital. Assumptions are as follows:

Variable	Assumption
First-year sales	$100
Sales growth	10% per year
Cost of goods sold/Sales	20%
Operating expense/Sales	70%
Interest income rate	5%
Tax rate	30%
Working capital as percent of sales	90%
Dividends	Non-dividend paying

Based on this information, forecast the company's net income and cash flow for five years.

Solution:

Exhibit 2 shows the net income forecasts in Line 7 and cash flow forecasts ("Change in cash") in Line 18.

Exhibit 2 Basic Financial Forecasting

	Time Period					
	0	1	2	3	4	5
(1) Sales		100.0	110.0	121.0	133.1	146.4
(2) Cost of goods sold		(20.0)	(22.0)	(24.2)	(26.6)	(29.3)
(3) Operating expenses		(70.0)	(77.0)	(84.7)	(93.2)	(102.5)
(4) Interest income		0.0	0.9	0.8	0.8	0.7
(5) Income before tax		10.0	11.9	12.9	14.1	15.3
(6) Taxes		(3.0)	(3.6)	(3.9)	(4.2)	(4.6)
(7) Net income		7.0	8.3	9.0	9.9	10.7
(8) Cash/Borrowing	0.0	17.0	16.3	15.4	14.4	13.1
(9) Working capital (non-cash)	100.0	90.0	99.0	108.9	119.8	131.8
(10) Total assets	100.0	107.0	115.3	124.3	134.2	144.9
(11) Liabilities	0.0	0.0	0.0	0.0	0.0	0.0
(12) Equity	100.0	107.0	115.3	124.3	134.2	144.9
(13) Total liabilities + Equity	100.0	107.0	115.3	124.3	134.2	144.9
(14) Net income		7.0	8.3	9.0	9.9	10.7
(15) Plus: Non-cash items		0.0	0.0	0.0	0.0	0.0
(16) Less: Investment in working capital		−10.0	9.0	9.9	10.9	12.0
(17) Less: Investment in fixed capital		0.0	0.0	0.0	0.0	0.0
(18) Change in cash		17.0	−0.7	−0.9	−1.0	−1.3
(19) Beginning cash		0.0	17.0	16.3	15.4	14.4
(20) Ending cash		17.0	16.3	15.4	14.4	13.1

Exhibit 2 indicates that at time 0, the company is formed with $100 of equity capital (Line 12). All of the company's capital is assumed to be immediately invested in working capital (Line 9). In future periods, because it is assumed that no dividends are paid, book equity increases each year by the amount of net income (Line 14). Future periods' required working capital (Line 9) is assumed to be 90 percent of annual sales (Line 1). Sales are assumed to be $100 in the first period and to grow at a constant rate of 10 percent per year (Line 1). The cost of goods sold is assumed to be constant at 20 percent of sales (Line 2), so the gross profit margin is 80 percent. Operating expenses are assumed to be 70 percent of sales each year (Line 3). Interest income (Line 4) is calculated as 5 percent of the beginning balance of cash/borrowing or the ending balance of the previous period (Line 8) and is an income item when there is a cash balance, as in this example. (If available cash is inadequate to cover required cash outflows, the shortfall is presumed to be covered by borrowing. This borrowing would be shown as a negative balance on Line 8 and an associated interest expense on Line 4. Alternatively, a forecast can be presented with separate lines for cash and borrowing.) Taxes of 30 percent are deducted to obtain net income (Line 7).

To calculate each period's cash flow, begin with net income (Line 7 = Line 14), add back any noncash items, such as depreciation (Line 15), deduct investment in working capital in the period or change in working capital over the period (Line 16), and deduct investment in fixed capital in the period (Line 17).[4] In this simple example, we are assuming that the company does not invest in any fixed capital (long-term assets) but, rather, rents furnished office space. Therefore, there is no depreciation and noncash items are zero. Each period's change in cash (Line 18) is added to the beginning cash balance (Line 19) to obtain the ending cash balance (Line 20 = Line 8).

Example 4 is simplified to demonstrate some principles of forecasting. In practice, each aspect of a forecast presents a range of challenges. Sales forecasts may be very detailed, with separate forecasts for each year of each product line, each geographical, and/or each business segment. Sales forecasts may be based on past results (for relatively stable businesses), management forecasts, industry studies, and/or macroeconomic forecasts. Similarly, gross profit margins may be based on past results or forecasted relationships and may be detailed. Expenses other than cost of goods sold may be broken down into more detailed line items, each of which may be forecasted on the basis of its relationship with sales (if variable) or on the basis of its historical levels. Working capital requirements may be estimated as a proportion of the amount of sales (as in Example 4) or the change in sales or as a compilation of specific forecasts for inventory, receivables, and payables. Most forecasts will involve some investment in fixed assets, in which case, depreciation amounts affect taxable income and net income but not cash flow. Example 4 makes the simplifying assumption that interest is paid on the beginning-of-year cash balance.

Example 4 develops a series of point estimates for future net income and cash flow. In practice, forecasting generally includes an analysis of the risk in forecasts—in this case, an assessment of the impact on income and cash flow if the realized values of variables differ significantly from the assumptions used in the base case or if actual sales are much different from forecasts. Quantifying the risk in forecasts requires an analysis of the economics of the company's businesses and expense structure and the potential impact of events affecting the company, the industry, and the economy in general. When that investigation is completed, the analyst can use scenario analysis or Monte Carlo simulation to assess risk. Scenario analysis involves specifying assumptions that differ from those used as the base-case assumptions. In Example 4, the projections of net income and cash flow could be recast in a more pessimistic scenario, with assumptions changed to reflect slower sales growth and higher costs. A Monte Carlo simulation involves specifying probability distributions of values for variables and random sampling from those distributions. In the analysis in Example 4, the projections would be repeatedly recast with the selected values for the drivers of net income and cash flow, thus permitting the analyst to evaluate a range of possible results and the probability of simulating the possible actual outcomes.

---

[4] Working capital represents funds that must be invested in the daily operations of a business to, for example, carry inventory and accounts receivable. The term "investment" in this context means "addition to" or "increase in." The "investment in fixed capital" is also referred to as "capital expenditure" ("capex").

An understanding of financial statements and ratios can enable an analyst to make more detailed projections of income statement, balance sheet, and cash flow statement items. For example, an analyst may collect information on normal inventory and receivables turnover and use this information to forecast accounts receivable, inventory, and cash flows based on sales projections rather than use a composite working capital investment assumption, as in Example 4.

As the analyst makes detailed forecasts, he or she must ensure that the forecasts are consistent with each other. For instance, in Example 5, the analyst's forecast concerning days of sales outstanding (which is an estimate of the average time to collect payment from sales made on credit) should flow from a model of the company that yields a forecast of the change in the average accounts receivable balance. Otherwise, predicted days of sales outstanding and accounts receivable will not be mutually consistent.

EXAMPLE 5

Consistency of Forecasts

Brown Corporation, a hypothetical company, had an average days-of-sales-outstanding (DSO) period of 19 days in 2017. An analyst thinks that Brown's DSO will decline in 2018 (because of expected improvements in the company's collections department) to match the industry average of 15 days. Total sales (all on credit) in 2017 were $300 million, and Brown expects total sales (all on credit) to increase to $320 million in 2018. To achieve the lower DSO, the change in the average accounts receivable balance from 2017 to 2018 that must occur is *closest* to:

A –$3.51 million.

B –$2.46 million.

C $2.46 million.

D $3.51 million.

Solution:

B is correct. The first step is to calculate accounts receivable turnover from the DSO collection period. Receivable turnover equals 365/19 (DSO) = 19.2 for 2017 and 365/15 = 24.3 in 2018. Next, the analyst uses the fact that the average accounts receivable balance equals sales/receivable turnover to conclude that for 2017, average accounts receivable was $300,000,000/19.2 = $15,625,000 and for 2018, it must equal $320,000,000/24.3 = $13,168,724. The difference is a reduction in receivables of $2,456,276.

The next section illustrates the application of financial statement analysis to credit risk analysis.

4 APPLICATION: ASSESSING CREDIT RISK

c describe the role of financial statement analysis in assessing the credit quality of a potential debt investment;

Credit risk is the risk of loss caused by a counterparty's or debtor's failure to make a promised payment. For example, credit risk with respect to a bond is the risk that the obligor (the issuer of the bond) will not be able to pay interest and/or principal according to the terms of the bond indenture (contract). **Credit analysis** is the evaluation of credit risk. Credit analysis may relate to the credit risk of an obligor in a particular transaction or to an obligor's overall creditworthiness.

In assessing an obligor's overall creditworthiness, one general approach is credit scoring, a statistical analysis of the determinants of credit default. Credit analysis for specific types of debt (e.g., acquisition financing and other highly leveraged financing) typically involves projections of period-by-period cash flows.

Whatever the techniques adopted, the analytical focus of credit analysis is on debt-paying ability. Unlike payments to equity investors, payments to debt investors are limited by the agreed contractual interest. If a company experiences financial success, its debt becomes less risky but its success does not increase the amount of payments to its debtholders. In contrast, if a company experiences financial distress, it may be unable to pay interest and principal on its debt obligations. Thus, credit analysis has a special concern with the sensitivity of debt-paying ability to adverse events and economic conditions—cases in which the creditor's promised returns may be most at risk. Because those returns are generally paid in cash, credit analysis usually focuses on cash flow rather than accrual income. Typically, credit analysts use return measures related to operating cash flow because it represents cash generated internally, which is available to pay creditors.

These themes are reflected in Example 6, which illustrates Moody's application of four quantitative factors in the credit analysis of the aerospace and defense industry.[5] These factors include

1. scale,
2. business profile,
3. leverage and coverage, and
4. financial policy.

"Scale" relates to a company's sensitivity to adverse events, adverse economic conditions, and other factors—such as market leadership, purchasing power with suppliers, and access to capital markets—that may affect debt-paying ability. "Business profile" represents a company's competitive position, stability of revenues, product and geographic diversity, growth prospects, and the stability and volatility of cash flows. "Leverage and coverage" reflects a company's "financial flexibility" and viability. Finally, "financial policy" relates to a company's financial risk tolerance and its capital structure.

---

[5] The information in this paragraph and in Example 7 are based upon the "Rating Methodology: Aerospace and Defense Industry" (Moody's, 2018).

EXAMPLE 6

Moody's Evaluation of Quantifiable Rating Factors for the Aerospace and Defense Industry

Moody's considers four broad rating factors for the aerospace and defense industry: scale; business profile, leverage and coverage; and financial policy. A company's ratings for each of these factors are weighted and aggregated in determining the overall credit rating assigned. The broad factors, the sub-factors, and weightings are as follows:

Broad Factor	Sub-factors	Sub-factor Weighting (%)	Broad Factor Weighting (%)
Scale	Total revenue	10	25
	Operating profit	15	
Business profile	Competitive position	10	20
	Expected revenue stability	10	
Leverage and coverage	Debt/EBITDA	10	35
	Retained cash flow[a]/Net debt	15	
	EBIT/Interest	10	
Financial policy	Financial policy	20	20
Total		100	100

[a] Retained cash flow is defined by Moody's as cash flow before working capital and after dividends.

Why might the leverage and coverage factor be weighted higher compared to the other rating factors?

Solution:

The level of debt relative to earnings and cash flow is a critical factor in assessing creditworthiness. Higher levels of debt for a company typically result in a higher risk in meeting interest and principal payments on its debt obligations.

A point to note regarding Example 6 is that the rating factors and the metrics used to represent each can vary by industry group.

Analyses of a company's historical and projected financial statements are an integral part of the credit evaluation process. Moody's and other rating agencies compute a variety of ratios in assessing creditworthiness. A comparison of a company's ratios with the ratios of its peers is informative in evaluating relative creditworthiness, as demonstrated in Example 7.

EXAMPLE 7

Peer Comparison of Ratios

A credit analyst is assessing the efficiency and leverage of two aerospace and defense companies based on certain sub-factors identified by Moody's in Example 7. The analyst collects the information from the companies' annual reports and calculates the following ratios:

	Company 1	Company 2
Debt/EBITDA	9.3	4.1
Retained cash flow/Net debt	2.6%	9.6%
EBIT/Interest	5.7	8.2

Based solely on the data given, which company is more likely to be assigned a higher credit rating, and why?

Solution:

The ratio comparisons are all in favor of Company 2, which has a lower level of debt relative to EBITDA, higher retained cash flow to net debt, and higher interest coverage. Based only on the data given, Company 2 is likely to be assigned a higher credit rating.

In calculating credit ratios, such as those presented in Example 8, analysts typically make certain adjustments to reported financial statements. We describe some common adjustments later in the reading.

Financial statement analysis, especially financial ratio analysis, can also be an important tool in selecting equity investments, as discussed in the next section.

5

SCREENING FOR POTENTIAL EQUITY INVESTMENTS

d describe the use of financial statement analysis in screening for potential equity investments;

Ratios constructed from financial statement data and market data are often used to screen for potential equity investments. **Screening** is the application of a set of criteria to reduce a set of potential investments to a smaller set having certain desired characteristics. Criteria involving financial ratios generally involve comparing one or more ratios with some pre-specified target or cutoff values.

A security selection approach incorporating financial ratios may be applied whether the investor uses top-down analysis or bottom-up analysis. **Top-down analysis** involves identifying attractive geographical segments and/or industry segments, from which the investor chooses the most attractive investments. **Bottom-up analysis** involves selection of specific investments from all companies within a specified investment universe. Regardless of the direction, screening for potential equity investments aims to identify companies that meet specific criteria. An analysis of this type may be used as the basis for directly forming a portfolio, or it may be undertaken as a preliminary part of a more thorough analysis of potential investment targets.

Fundamental to this type of analysis are decisions about which metrics to use as screens, how many metrics to include, what values of those metrics to use as cutoff points, and what weighting to give each metric. Metrics may include not only financial ratios but also characteristics such as market capitalization or membership as a component security in a specified index. Exhibit 3 presents a hypothetical example of a simple stock screen based on the following criteria: a valuation ratio (P/E) less than a specified value, a solvency ratio measuring financial leverage (total liabilities/total assets) not exceeding a specified value, positive operating margin, and dividend yield (dividends per share divided by price per share) greater than a specified value. Exhibit 3 shows the results of applying the screen in August 2018 to a set of 6,406 US companies with market capitalization greater than $100 million, which compose a hypothetical equity manager's investment universe.

Exhibit 3 Example of a Stock Screen

	Stocks Meeting Criterion	
Criterion	**Number**	**Percent of Total**
Market Capitalization > $100 million	4,357	68.01%
P/E < 15	1,104	17.23%
Total liabilities/Total Assets ≤ 0.9	61	0.95%
Operating Income/Sales > 0	3,509	54.78%
Dividend yield > 0.5%	2,391	37.32%
Meeting all five criteria simultaneously	17	0.27%

Source for data: http://google.com/finance/.

Several points about the screen in Exhibit 3 are consistent with many screens used in practice:

- Some criteria serve as checks on the results from applying other criteria. In this hypothetical example, the second criterion selects stocks that appear relatively cheaply valued. The stocks might be cheap for a good reason, however, such as poor profitability or excessive financial leverage. So, the requirement for net income to be positive serves as a check on profitability, and the limitation on financial leverage serves as a check on financial risk. Of course, financial ratios or other statistics cannot generally control for exposure to certain other types of risk (e.g., risk related to regulatory developments or technological innovation).
- If all the criteria were completely independent of each other, the set of stocks meeting all five criteria would be 2, equal to 6,406 times 0.023 percent—the product of the fraction of stocks satisfying the five criteria individually (i.e., $0.6801 \times 0.1723 \times 0.0095 \times 0.5478 \times 0.3732 = 0.000228$, or 0.023 percent). As the screen illustrates, criteria are often not independent, and the result is that more securities pass the screening than if criteria were independent. In this example, 17 of the securities pass all five screens simultaneously. For an example of the lack of independence, we note that dividend-paying status is probably positively correlated with the ability to generate positive operating margin. If stocks that pass one test tend to also pass another, few are eliminated after the application of the second test.
- The results of screens can sometimes be relatively concentrated in a subset of the sectors represented in the benchmark. The financial leverage criterion in Exhibit 3 would exclude banking stocks, for example. What constitutes a high or low value of a measure of a financial characteristic can be sensitive to the industry in which a company operates.

Screens can be used by both **growth investors** (focused on investing in high-earnings-growth companies), **value investors** (focused on paying a relatively low share price in relation to earnings or assets per share), and **market-oriented investors** (an intermediate grouping of investors whose investment disciplines cannot be clearly categorized as value or growth). Growth screens would typically feature criteria related to earnings growth and/or momentum. Value screens, as a rule, feature criteria setting upper limits for the value of one or more valuation ratios. Market-oriented screens would not strongly emphasize valuation or growth criteria. The use of screens involving financial ratios may be most common among value investors.

An analyst may want to evaluate how a portfolio based on a particular screen would have performed historically. For this purpose, the analyst uses a process known as "back-testing." **Back-testing** applies the portfolio selection rules to historical data and calculates what returns would have been earned if a particular strategy had been used. The relevance of back-testing to investment success in practice, however, may be limited. Haugen and Baker (1996) described some of these limitations:

- Survivorship bias: If the database used in back-testing eliminates companies that cease to exist because of a bankruptcy or merger, then the remaining companies collectively will appear to have performed better.
- Look-ahead bias: If a database includes financial data updated for restatements (where companies have restated previously issued financial statements to correct errors or reflect changes in accounting principles), then there is a mismatch between what investors would have actually known at the time of the investment decision and the information used in the back-testing.
- Data-snooping bias: If researchers build a model on the basis of previous researchers' findings, then use the same database to test that model, they are not actually testing the model's predictive ability. When each step is backward looking, the same rules may or may not produce similar results in the future. The predictive ability of the model's rules can validly be tested only by using future data. One academic study has argued that the apparent ability of value strategies to generate excess returns is largely explainable as the result of collective data snooping (Conrad, Cooper, and Kaul, 2003).

EXAMPLE 8

Ratio-Based Screening for Potential Equity Investments

Below are two alternative strategies under consideration by an investment firm:

> *Strategy A*: Invest in stocks that are components of a global equity index, have a ROE above the median ROE of all stocks in the index, and have a P/E less than the median P/E.
>
> *Strategy B*: Invest in stocks that are components of a broad-based US equity index, have a ratio of price to operating cash flow in the lowest quartile of companies in the index, and have shown increases in sales for at least the past three years.

Both strategies were developed with the use of back-testing.

1 How would you characterize the two strategies?

2 What concerns might you have about using such strategies?

Solution to 1:

Strategy A appears to aim for global diversification and combines a requirement for high relative profitability with a traditional measure of value (low P/E). Strategy B focuses on both large and small companies in a single market and apparently aims to identify companies that are growing and have a lower price multiple based on cash flow from operations.

Solution to 2:

The use of *any* approach to investment decisions depends on the objectives and risk profile of the investor. With that crucial consideration in mind, we note that ratio-based benchmarks may be an efficient way to screen for potential equity investments. In screening, however, many questions arise.

First, unintentional selections can be made if criteria are not specified carefully. For example, Strategy A might unintentionally select a loss-making company with negative shareholders' equity because negative net income divided by negative shareholders' equity arithmetically results in a positive ROE. Strategy B might unintentionally select a company with negative operating cash flow because price to operating cash flow will be negative and thus very low in the ranking. In both cases, the analyst can add additional screening criteria to avoid unintentional selections; these additional criteria could include requiring positive shareholders' equity in Strategy A and requiring positive operating cash flow in Strategy B.

Second, the inputs to ratio analysis are derived from financial statements, and companies may differ in the financial standards they apply (e.g., IFRS versus US GAAP), the specific accounting method(s) they choose within those allowed by the reporting standards, and/or the estimates made in applying an accounting method.

Third, back-testing may not provide a reliable indication of future performance because of survivorship bias, look-ahead bias, or data-snooping bias. Also, as suggested by finance theory and by common sense, the past is not necessarily indicative of the future.

Fourth, implementation decisions can dramatically affect returns. For example, decisions about frequency and timing of portfolio re-evaluation and changes affect transaction costs and taxes paid out of the portfolio.

FRAMEWORK FOR ANALYST ADJUSTMENTS & ADJUSTMENTS TO INVESTMENTS & ADJUSTMENTS TO INVENTORY

6

e explain appropriate analyst adjustments to a company's financial statements to facilitate comparison with another company.

When comparing companies that use different accounting methods or estimate key accounting inputs in different ways, analysts frequently adjust a company's financials. In this section, we first provide a framework for considering potential analyst adjustments to facilitate such comparisons and then provide examples of such adjustments. In practice, required adjustments vary widely. The examples presented here are not intended to be comprehensive but, rather, to illustrate the use of adjustments to facilitate a meaningful comparison.

6.1 A Framework for Analyst Adjustments

In this discussion of potential analyst adjustments to a company's financial statements, we use a framework focused on the *balance sheet*. Because the financial statements are interrelated, however, adjustments to items reported on one statement may also be reflected in adjustments to items on another financial statement. For example, an analyst adjustment to inventory on the balance sheet affects cost of goods sold on the income statement (and thus also affects net income and, subsequently, the retained earnings account on the balance sheet).

Regardless of the particular order in which an analyst considers the items that may require adjustment for comparability, the following aspects are appropriate:

- *Importance* (*materiality*). Is an adjustment to this item likely to affect the conclusions? In other words, does it matter? For example, in an industry where companies require minimal inventory, does it matter that two companies use different inventory accounting methods?
- *Body of standards*. Is there a difference in the body of standards being used (US GAAP versus IFRS)? If so, in which areas is the difference likely to affect a comparison?
- *Methods*. Is there a difference in accounting methods used by the companies being compared?
- *Estimates*. Is there a difference in important estimates used by the companies being compared?

The following sections illustrate analyst adjustments—first, those relating to the asset side of the balance sheet and then those relating to the liability side.

6.2 Analyst Adjustments Related to Investments

Accounting for investments in the debt and equity securities of other companies (other than investments accounted for under the equity method and investments in consolidated subsidiaries) depends on management's intention (i.e., whether to actively trade the securities, make them available for sale, or in the case of debt securities, hold them to maturity). When securities are classified as "financial assets measured at fair value through profit or loss" (similar to "trading" securities in US GAAP), unrealized gains and losses are reported in the income statement. When securities are classified as "financial assets measured at fair value through other comprehensive income" (similar to "available-for-sale" securities in US GAAP), unrealized gains and losses are not reported in the income statement and, instead, are recognized in equity. If two otherwise comparable companies have significant differences in the classification of investments, analyst adjustments may be useful to facilitate comparison.

6.3 Analyst Adjustments Related to Inventory

With inventory, adjustments may be required for different accounting methods. As described in previous readings, a company's decision about inventory method will affect the value of inventory shown on the balance sheet as well as the value of inventory that is sold (cost of goods sold). If a company not reporting under IFRS[6] uses LIFO (last-in, first-out) and another uses FIFO (first-in, first-out), comparability of the financial results of the two companies will suffer. Companies that use the LIFO method, must also, however, disclose the value of their inventory under the FIFO method. To recast inventory values for a company using LIFO reporting on a FIFO basis, the analyst adds the ending balance of the LIFO reserve to the ending value of inventory under LIFO accounting. To adjust cost of goods sold to a FIFO basis, the analyst subtracts the change in the LIFO reserve from the reported cost of goods sold under LIFO accounting. Example 9 illustrates the use of a disclosure of the value of inventory under the FIFO method to make a more consistent comparison of the current ratios of two companies reporting in different methods.

---

[6] IAS No. 2 does not permit the use of LIFO.

EXAMPLE 9

Adjustment for a Company Using LIFO Accounting for Inventories

An analyst is comparing the financial performance of LP Technology Corporation (LP Tech), a hypothetical company, with the financial performance of a similar company that uses IFRS for reporting. The company reporting under IFRS uses the FIFO method of inventory accounting. Therefore, the analyst converts LP Tech's results to a comparable basis. Exhibit 4 provides balance sheet information on LP Tech.

Exhibit 4 Data for LP Technology Corporation

	30 June	
	2018	**2017**
Total current assets	820.2	749.7
Total current liabilities	218.1	198.5

NOTE 6. INVENTORIES

Inventories consist of the following ($ millions):

	30 June	
	2018	**2017**
Raw materials	$30.7	$29.5
Work in process	109.1	90.8
Finished goods	63.8	65.1
	$203.6	$185.4

If the first-in, first-out method of inventory had been used instead of the LIFO method, inventories would have been $331.8 and $305.8 million higher as of June 30, 2018 and 2017, respectively.

1 Based on the information in Exhibit 4, calculate LP Tech's current ratio under FIFO and LIFO for 2017 and 2018.

2 LP Tech makes the following disclosure in the risk section of its MD&A. Assuming an effective tax rate of 35 percent, estimate the impact on LPTC's tax liability.

> **"We value most of our inventory using the LIFO method, which could be repealed resulting in adverse effects on our cash flows and financial condition.**
>
> The cost of our inventories is primarily determined using the Last-In First-Out ("LIFO") method. Under the LIFO inventory valuation method, changes in the cost of raw materials and production activities are recognized in cost of sales in the current period even though these materials and other costs may have been incurred at significantly different values due to the length of time of our production cycle. Generally in a period of rising prices, LIFO recognizes higher costs of goods sold, which both reduces current income and assigns a lower value to the year-end inventory. Recent proposals have been initiated aimed at repealing the election to use the LIFO method for income tax purposes. According to these proposals, generally taxpayers that currently use the LIFO method would be required to revalue their LIFO inventory to its first-in, first-out ("FIFO") value. As of June 30, 2018, if the FIFO method of inventory had been used instead of the LIFO method, our inventories would have been about $332 million higher. This increase in inventory would result in a one time increase in taxable income which would be taken into account ratably over the first taxable year and the following several taxable years. The repeal of LIFO could result in a substantial tax liability which could adversely impact our cash flows and financial condition."

3 LP Tech reported cash flow from operations of $115.2 million for the year ended 30 June 2018. In comparison with the company's operating cash flow, how significant is the additional potential tax liability?

Solution to 1:

The calculations of LP Tech's current ratio (current assets divided by current liabilities) are as follows:

	2018	2017
I. Current ratio (unadjusted)		
Total current assets	$820.2	$749.7
Total current liabilities	$218.1	$198.5
Current ratio (unadjusted)	3.8	3.8
II. Current ratio (adjusted)		
Total current assets	$820.2	$749.7
Adjust inventory to FIFO, add:	331.8	305.8
Total current assets (adjusted)	$1,152	$1,056
Total current liabilities	218.1	198.5
Current ratio (adjusted)	5.3	5.3

To adjust the LIFO inventory to FIFO, add the excess amounts of FIFO cost over LIFO cost to LIFO inventory and increase current assets by an equal amount. The effect of adjusting inventory on the current ratio is to increase the current ratio from 3.8 to 5.3 in both 2017 and 2018. LP Tech has greater liquidity according to the adjusted current ratio.

Solution to 2:

Assuming an effective tax rate of 35 percent, we find the total increase in LP Tech's tax liability to be $116.1 million (0.35 × $331.8 million).

Solution to 3:

The additional tax liability would be greater than the entire amount of the company's cash flow from operations of $115.2 million; the additional tax liability would be apportioned, however, over several years.

In summary, the information disclosed by companies that use LIFO allows an analyst to calculate the value of the company's inventory as if the company were using the FIFO method. If the LIFO method is used for a substantial part of a company's inventory and the LIFO reserve is large relative to reported inventory, however, the adjustment to a FIFO basis can be important for comparison of the LIFO-reporting company with a company that uses the FIFO method of inventory valuation. Example 10 illustrates a case in which such an adjustment would have a major impact on an analyst's conclusions.

EXAMPLE 10

Analyst Adjustment to Inventory Value for Comparability in a Current Ratio Comparison

Company A reports under IFRS and uses the FIFO method of inventory accounting. Company B reports under US GAAP and uses the LIFO method. Exhibit 5 gives data pertaining to current assets, LIFO reserves, and current liabilities of these companies.

Exhibit 5 Data for Companies Accounting for Inventory on Different Bases

	Company A (FIFO)	Company B (LIFO)
Current assets (includes inventory)	$ 300,000	$ 80,000
LIFO reserve	NA	$ 20,000
Current liabilities	$ 150,000	$ 45,000

NA = not applicable.

Based on the data given in Exhibit 5, compare the liquidity of the two companies as measured by the current ratio.

Solution:

Company A's current ratio is 2.0. Based on unadjusted balance sheet data, Company B's current ratio is 1.78. Company A's higher current ratio indicates that Company A appears to be more liquid than Company B; however, the use of unadjusted data for Company B is not appropriate for making comparisons with Company A.

After adjusting Company B's inventory to a comparable basis (i.e., to a FIFO basis), the conclusion changes. The following table summarizes the results when Company B's inventory is left on a LIFO basis and when it is placed on a FIFO basis for comparability with Company A.

		Company B	
	Company A (FIFO)	**Unadjusted (LIFO basis)**	**Adjusted (FIFO basis)**
Current assets (includes inventory)	$ 300,000	$ 80,000	$ 100,000
Current liabilities	$ 150,000	$ 45,000	$ 45,000
Current ratio	2.00	1.78	2.22

When both companies' inventories are stated on a FIFO basis, Company B appears to be the more liquid, as indicated by its current ratio of 2.22 versus Company A's ratio of 2.00.

The adjustment to place Company B's inventory on a FIFO basis was significant because Company B was assumed to use LIFO for its entire inventory and its inventory reserve was $20,000/$80,000 = 0.25, or 25 percent of its reported inventory.

As mentioned earlier, an analyst can also adjust the cost of goods sold for a company using LIFO to a FIFO basis by subtracting the change in the amount of the LIFO reserve from cost of goods sold. Such an adjustment would be appropriate for making profitability comparisons with a company reporting on a FIFO basis and is important to make when the impact of the adjustment would be material.

7

ADJUSTMENTS RELATED TO PROPERTY, PLANT, AND EQUIPMENT

e explain appropriate analyst adjustments to a company's financial statements to facilitate comparison with another company.

Management generally has considerable discretion in determination of depreciation expense. Depreciation expense affects the values of reported net income and reported net fixed assets. Analysts often consider management's choices related to depreciation as a qualitative factor in evaluating the quality of a company's financial reporting, and in some cases, analysts may adjust reported depreciation expense for a specific analytical purpose.

The amount of depreciation expense depends on both the accounting method and the estimates used in the calculations. Companies can use the straight-line method, an accelerated method, or a usage method to depreciate fixed assets (other than land). The straight-line method reports an equal amount of depreciation expense each period, and the expense is computed as the depreciable cost divided by the estimated useful life of the asset (when acquired, an asset's depreciable cost is calculated as its total cost minus its estimated salvage value). Accelerated methods depreciate the asset more quickly; they apportion a greater amount of the depreciable cost to depreciation expense in the earlier periods. Usage-based methods depreciate an asset in proportion to its usage. In addition to selecting a depreciation method, companies must estimate an asset's salvage value and useful life to compute depreciation.

Disclosures required for depreciation often do not facilitate specific adjustments, so comparisons of companies concerning their decisions in depreciating assets are often qualitative and general. The accounts that are associated with depreciation include the balance sheet accounts for gross property, plant, and equipment (PPE) and accumulated depreciation; the income statement amount for depreciation expense; and the statement of cash flows disclosure of capital expenditure (capex) and asset disposals. The relationships among these items can reveal various pieces of information. Note, however, that PPE typically includes a mix of assets with different depreciable lives and salvage values, so the items in the following list reflect general relationships in the total pool of assets.

- Accumulated depreciation divided by gross PPE, from the balance sheet, suggests how much of the useful life of the company's overall asset base has passed.
- Accumulated depreciation divided by depreciation expense suggests how many years' worth of depreciation expense have already been recognized (i.e., the average age of the asset base).
- Net PPE (net of accumulated depreciation) divided by depreciation expense is an approximate indicator of how many years of useful life remain for the company's overall asset base.
- Gross PPE divided by depreciation expense suggests the average life of the assets at installation.
- Capex divided by the sum of gross PPE plus capex can suggest what percentage of the asset base is being renewed through new capital investment.
- Capex in relation to asset disposal provides information on growth of the asset base.

As Example 11 shows, these relationships can be evaluated for companies in an industry to suggest differences in their strategies for asset utilization or areas for further investigation.

EXAMPLE 11

Differences in Depreciation

An analyst is evaluating the financial statements of two companies in the same industry. The companies have similar strategies with respect to the use of equipment in manufacturing their products. The following information is provided (amounts in millions):

	Company A	Company B
Net PPE	$1,200	$750
Depreciation expense	$120	$50

1. Based on the information given, estimate the average remaining useful lives of the asset bases of Company A and Company B.
2. Suppose that, based on a physical inspection of the companies' plants and other industry information, the analyst believes that the actual remaining useful lives of Company A's and Company B's assets are roughly equal at 10 years. Based only on the facts given, what might the analyst conclude about Company B's reported net income?

Solution to 1:

The estimated average remaining useful life of Company A's asset base is 10 years (calculated as net PPE divided by depreciation expense, or $1,200/$120 = 10 years). For Company B, the average remaining useful life of the asset base appears to be far longer, 15 years ($750/$50).

Solution to 2:

If 10 years were used to calculate Company B's depreciation expense, the expense would be $75 million (i.e., $25 million higher than reported) and higher depreciation expense would decrease net income. The analyst might conclude that Company B's reported net income reflects relatively more aggressive accounting estimates than estimates reflected in Company A's reported net income.

8

ADJUSTMENTS RELATED TO GOODWILL

e explain appropriate analyst adjustments to a company's financial statements to facilitate comparison with another company.

Goodwill arises when one company purchases another for a price that exceeds the fair value of the net identifiable assets acquired. Net identifiable assets include current assets, fixed assets, and certain intangible assets that have value and meet recognition criteria under accounting standards. A broad range of intangible assets might require valuation in the context of a business combination—for example, brands, technology, and customer lists. Goodwill is recorded as an asset and essentially represents the difference between the purchase price and the net identifiable assets. For example, assume ParentCo purchases TargetCo for a purchase price of $400 million and the fair value of TargetCo's identifiable assets is $300 million (which includes the fair values of current assets, fixed assets, and a recognized brand). ParentCo will record total assets of $400 million consisting of $300 million in identifiable assets (including the fair value of the brand) and $100 million of goodwill. The goodwill is tested annually for impairment and if the value of the goodwill is determined to be impaired, ParentCo will then reduce the amount of the asset and report a write-off resulting from impairment.

One of the conceptual difficulties with goodwill arises in comparative financial statement analysis. Consider, for example, two hypothetical US companies, one of which has grown by making an acquisition and the other of which has grown internally. Assume that the economic value of the two companies is identical: Each has an identically valuable branded product, well-trained workforce, and proprietary technology. The company that has grown by acquisition will have recorded the transaction to acquire the target company and its underlying net assets on the basis of the total consideration paid for the acquisition. The company that has grown internally will have done so by incurring expenditures for advertising, staff training, and research, all of which are expensed as incurred under US GAAP. Given the immediate expensing, the value of the internally generated assets is not capitalized onto the balance sheet and is thus not directly reflected on the company's balance sheet (revenues, income, and cash flows should reflect the benefits derived from the investment in the intangible assets). Ratios based on asset values and/or income, including profitability ratios (such as ROA) and market value to book value (MV/BV),[7] will generally differ for the two companies because of differences in the accounting values of assets and income related to acquired intangibles and goodwill, although, by assumption, the economic value of the companies is identical.

---

[7] MV/BV equals the total market value of the stock (the market capitalization) divided by total stockholders' equity. It is also referred to as the price-to-book ratio because it can also be calculated as price per share divided by stockholders' equity per share.

EXAMPLE 12

Ratio Comparisons for Goodwill

Miano Marseglia is an analyst who is evaluating the relative valuation of two securities brokerage companies: TD Ameritrade Holding Corporation (AMTD) and the Charles Schwab Corporation (SCHW). As one part of an overall analysis, Marseglia would like to see how the two companies compare with each other and with the industry based on market value to book value. Because both companies are large players in the industry, Marseglia expects them to sell at a higher MV/BV than the financial services sector median of 2.2. He collects the following data on the two companies.

	SCHW	AMTD
Market capitalization on 30 August 2018 (market price per share times the number of shares outstanding)	$68,620	$33,247
Total shareholders' equity (as of 30 June 2018 for both companies)	$20,097	$7,936
Goodwill	$1,227	$4,198
Other intangible assets	$93	$1,363

Marseglia computes the MV/BV for the companies as follows:

SCHW $68,620/$20,097 = 3.4

AMTD $33,247/$7,936 = 4.2

As expected, each company appears to be selling at a premium to the sector median MV/BV of 2.2. The companies have similar MV/BVs (i.e., they are somewhat equally valued relative to the book value of shareholders' equity). Marseglia is concerned, however, because he notes that AMTD has significant amounts of goodwill and acquired intangible assets. He wonders what the relative value would be if the MV/BV were computed after adjusting book value, first, to remove goodwill and, second, to remove all intangible assets. Book value reduced by all intangible assets (including goodwill) is known as "tangible book value."

1 Compute the MV/BV adjusted for goodwill and the price/tangible book value for each company.

2 Which company appears to be a better value based *solely* on this data? (Note that the MV/BV is only one part of a broader analysis. Much more evidence related to the valuations and the comparability of the companies would be required to reach a conclusion about whether one company is a better value.)

Solution to 1:

	($ millions)	
	SCHW	AMTD
Total stockholders' equity	$20,097	$7,936
Less: Goodwill	$1,227	$4,198
Book value, adjusted	$18,870	$3,738
Adjusted MV/BV	**3.6**	**8.9**

	($ millions)	
	SCHW	AMTD
Total stockholders' equity	$20,097	$7,936
Less: Goodwill	$1,227	$4,198
Less: Other intangible assets	$93	$1,363
Tangible book value	$18,777	$2,375
MV/tangible book value	**3.7**	**14.0**

Solution to 2:

After adjusting for goodwill, SCHW appears to be selling for a much lower price relative to book value than does AMTD (3.6 versus 8.9) after adjusting for goodwill. The difference is more extreme after adjusting for other intangibles.

SUMMARY

This reading described selected applications of financial statement analysis, including the evaluation of past financial performance, the projection of future financial performance, the assessment of credit risk, and the screening of potential equity investments. In addition, the reading introduced analyst adjustments to reported financials. In all cases, the analyst needs to have a good understanding of the financial reporting standards under which the financial statements were prepared. Because standards evolve over time, analysts must stay current in order to make good investment decisions.

The main points in the reading are as follows:

- Evaluating a company's historical performance addresses not only what happened but also the causes behind the company's performance and how the performance reflects the company's strategy.
- The projection of a company's future net income and cash flow often begins with a top-down sales forecast in which the analyst forecasts industry sales and the company's market share. By projecting profit margins or expenses and the level of investment in working and fixed capital needed to support projected sales, the analyst can forecast net income and cash flow.
- Projections of future performance are needed for discounted cash flow valuation of equity and are often needed in credit analysis to assess a borrower's ability to repay interest and principal of a debt obligation.
- Credit analysis uses financial statement analysis to evaluate credit-relevant factors, including tolerance for leverage, operational stability, and margin stability.
- When ratios constructed from financial statement data and market data are used to screen for potential equity investments, fundamental decisions include which metrics to use as screens, how many metrics to include, what values of those metrics to use as cutoff points, and what weighting to give each metric.
- Analyst adjustments to a company's reported financial statements are sometimes necessary (e.g., when comparing companies that use different accounting methods or assumptions). Adjustments can include those related to investments; inventory; property, plant, and equipment; and goodwill.

REFERENCES

Benninga, Simon Z., and Oded H. Sarig. 1997. *Corporate Finance: A Valuation Approach*. New York: McGraw-Hill Publishing.

Conrad, J., M. Cooper, and G. Kaul. 2003. "Value versus Glamour." *Journal of Finance*, vol. 58, no. 5:1969–1996.

Haugen, R.A., and N.L. Baker. 1996. "Commonality in the Determinants of Expected Stock Returns." *Journal of Financial Economics*, vol. 41, no. 3:401–439.

PRACTICE PROBLEMS

1 Projecting profit margins into the future on the basis of past results would be *most* reliable when the company:

A is in the commodities business.

B operates in a single business segment.

C is a large, diversified company operating in mature industries.

2 Galambos Corporation had an average receivables collection period of 19 days in 2003. Galambos has stated that it wants to decrease its collection period in 2004 to match the industry average of 15 days. Credit sales in 2003 were \$300 million, and analysts expect credit sales to increase to \$400 million in 2004. To achieve the company's goal of decreasing the collection period, the change in the average accounts receivable balance from 2003 to 2004 that must occur is *closest* to:

A –\$420,000.

B \$420,000.

C \$836,000.

3 Credit analysts are likely to consider which of the following in making a rating recommendation?

A Business risk but not financial risk

B Financial risk but not business risk

C Both business risk and financial risk

4 When screening for potential equity investments based on return on equity, to control risk, an analyst would be *most likely* to include a criterion that requires:

A positive net income.

B negative net income.

C negative shareholders' equity.

5 One concern when screening for stocks with low price-to-earnings ratios is that companies with low P/Es may be financially weak. What criterion might an analyst include to avoid inadvertently selecting weak companies?

A Net income less than zero

B Debt-to-total assets ratio below a certain cutoff point

C Current-year sales growth lower than prior-year sales growth

6 When a database eliminates companies that cease to exist because of a merger or bankruptcy, this can result in:

A look-ahead bias.

B back-testing bias.

C survivorship bias.

7 In a comprehensive financial analysis, financial statements should be:

A used as reported without adjustment.

B adjusted after completing ratio analysis.

C adjusted for differences in accounting standards, such as international financial reporting standards and US generally accepted accounting principles.

© 2019 CFA Institute. All rights reserved.

8 When comparing a US company that uses the last in, first out (LIFO) method of inventory with companies that prepare their financial statements under international financial reporting standards (IFRS), analysts should be aware that according to IFRS, the LIFO method of inventory:

A is never acceptable.

B is always acceptable.

C is acceptable when applied to finished goods inventory only.

9 An analyst is evaluating the balance sheet of a US company that uses last in, first out (LIFO) accounting for inventory. The analyst collects the following data:

	31 Dec 05	31 Dec 06
Inventory reported on balance sheet	$500,000	$600,000
LIFO reserve	$ 50,000	$70,000
Average tax rate	30%	30%

After adjusting the amounts to convert to the first in, first out (FIFO) method, inventory at 31 December 2006 would be closest to:

A $600,000.

B $620,000.

C $670,000.

10 An analyst gathered the following data for a company ($ millions):

	31 Dec 2000	31 Dec 2001
Gross investment in fixed assets	$2.8	$2.8
Accumulated depreciation	$1.2	$1.6

The average age and average depreciable life of the company's fixed assets at the end of 2001 are *closest* to:

	Average Age	Average Depreciable Life
A	1.75 years	7 years
B	1.75 years	14 years
C	4.00 years	7 years

11 To compute tangible book value, an analyst would:

A add goodwill to stockholders' equity.

B add all intangible assets to stockholders' equity.

C subtract all intangible assets from stockholders' equity.

SOLUTIONS

1 C is correct. For a large, diversified company, margin changes in different business segments may offset each other. Furthermore, margins are most likely to be stable in mature industries.

2 C is correct. Accounts receivable turnover is equal to 365/19 (collection period in days) = 19.2 for 2003 and needs to equal 365/15 = 24.3 in 2004 for Galambos to meet its goal. Sales/turnover equals the accounts receivable balance. For 2003, \$300,000,000/19.2 = \$15,625,000, and for 2004, \$400,000,000/24.3 = \$16,460,905. The difference of \$835,905 is the increase in receivables needed for Galambos to achieve its goal.

3 C is correct. Credit analysts consider both business risk and financial risk.

4 A is correct. Requiring that net income be positive would eliminate companies that report a positive return on equity only because both net income and shareholders' equity are negative.

5 B is correct. A lower value of debt/total assets indicates greater financial strength. Requiring that a company's debt/total assets be below a certain cutoff point would allow the analyst to screen out highly leveraged and, therefore, potentially financially weak companies.

6 C is correct. Survivorship bias exists when companies that merge or go bankrupt are dropped from the database and only surviving companies remain. Look-ahead bias involves using updated financial information in back-testing that would not have been available at the time the decision was made. Back-testing involves testing models in prior periods and is not, itself, a bias.

7 C is correct. Financial statements should be adjusted for differences in accounting standards (as well as accounting and operating choices). These adjustments should be made prior to common-size and ratio analysis.

8 A is correct. LIFO is not permitted under IFRS.

9 C is correct. To convert LIFO inventory to FIFO inventory, the entire LIFO reserve must be added back: \$600,000 + \$70,000 = \$670,000.

10 C is correct. The company made no additions to or deletions from the fixed asset account during the year, so depreciation expense is equal to the difference in accumulated depreciation at the beginning of the year and the end of the year, or \$0.4 million. Average age is equal to accumulated depreciation/depreciation expense, or \$1.6/\$0.4 = 4 years. Average depreciable life is equal to ending gross investment/depreciation expense = \$2.8/\$0.4 = 7 years.

11 C is correct. Tangible book value removes all intangible assets, including goodwill, from the balance sheet.

Corporate Issuers

STUDY SESSION

Study Session 9	Corporate Issuers (1)
Study Session 10	Corporate Issuers (2)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to evaluate a company's corporate governance; to demonstrate methods used to make capital investment; to evaluate the management of working capital; to estimate a company's cost of capital; and to evaluate a company's operating and financial leverage.

Some academic studies have shown that well governed companies may perform better in financial terms. Increasingly, investment approaches that consider environmental, social, and governance factors, known as ESG, are being adopted. In addition to good governance practices, management decisions regarding investment and financing also play a central role in corporate profitability and performance. To remain in business as a going concern and to increase shareholder value over time, a company's management must consistently identify and invest in profitable long-term capital projects relative to cost of capital (financing) while making optimal use of leverage and working capital in day to day operations.

© 2021 CFA Institute. All rights reserved.

CORPORATE ISSUERS

STUDY SESSION

9

Corporate Issuers (1)

This study session provides an introduction to corporate governance and investing and financing decisions. An overview of corporate governance is presented along with a framework for understanding and analyzing corporate governance and stakeholder management. The growing impact of environmental and social considerations in investing is also highlighted. A discussion of the capital allocation process and the assessment of capital investment opportunities using NPV and IRR methods follows. The session ends with coverage of the short and long-term options companies may have to finance themselves and techniques for assessing a company's short-term liquidity management.

READING ASSIGNMENTS

Reading 27	Introduction to Corporate Governance and Other ESG Considerations by Assem Safieddine, PhD, Young Lee, CFA, Donna F. Anderson, CFA, and Deborah Kidd, CFA and Hardik Sanjay Shah, CFA
Reading 28	Uses of Capital by John D. Stowe, PhD, CFA, and Jacques R. Gagné, FSA, CFA, CIPM
Reading 29	Sources of Capital by Edgar A. Norton, Jr., PhD, CFA, Kenneth L. Parkinson, MBA, CCM, and Pamela Peterson Drake, PhD, CFA

© 2021 CFA Institute. All rights reserved.

READING

27

Introduction to Corporate Governance and Other ESG Considerations

by Assem Safieddine, PhD, Young Lee, CFA, Donna F. Anderson, CFA, Deborah S. Kidd, CFA, and Hardik Sanjay Shah, CFA

Assem Safieddine, PhD, is at Suliman Olayan Business School, American University of Beirut (Lebanon). Young Lee, CFA, is at MacKay Shields LLC (USA), MacKay Shields Europe Investment Management Ltd. (Ireland), and MacKay Shields UK LLP (United Kingdom). Donna F. Anderson, CFA (USA). Deborah S. Kidd, CFA, is at CFA Institute (USA). Hardik Sanjay Shah, CFA, is at GMO LLC (Singapore).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe corporate governance;
☐	**b.** describe a company's stakeholder groups, and compare interests of stakeholder groups;
☐	**c.** describe principal–agent and other relationships in corporate governance and the conflicts that may arise in these relationships;
☐	**d.** describe stakeholder management;
☐	**e.** describe mechanisms to manage stakeholder relationships and mitigate associated risks;
☐	**f.** describe functions and responsibilities of a company's board of directors and its committees;
☐	**g.** describe market and non-market factors that can affect stakeholder relationships and corporate governance;
☐	**h.** identify potential risks of poor corporate governance and stakeholder management, and identify benefits from effective corporate governance and stakeholder management;
☐	**i.** describe factors relevant to the analysis of corporate governance and stakeholder management;
☐	**j.** describe environmental and social considerations in investment analysis;
☐	**k.** describe how environmental, social, and governance factors may be used in investment analysis.

© 2021 CFA Institute. All rights reserved.

1 INTRODUCTION AND OVERVIEW OF CORPORATE GOVERNANCE

a describe corporate governance

Weak corporate governance is a common thread found in many company failures. Lack of proper oversight by the board of directors, inadequate protection for minority shareholders, and incentives at companies that promote excessive risk taking are just a few of the examples that can be problematic for a company. Poor corporate governance practices resulted in several high-profile accounting scandals and corporate bankruptcies over the past several decades and have been cited as significantly contributing to the 2008–2009 global financial crisis.

In response to these failures, regulations have been introduced to promote stronger governance practices to protect financial markets and investors. Academics, policy makers, and other groups have published numerous works discussing the benefits of good corporate governance and identifying core corporate governance principles believed to be essential to ensuring continuous, well-functioning capital markets and the stability of the financial system.

The investment community has also demonstrated a greater appreciation for the importance of good corporate governance. The assessment of a company's corporate governance structure and controls, including consideration of conflicts of interest and transparency of operations, has been an essential factor in investment analysis. More data availability and demands for better governance have increased the weight of corporate governance in the investment decision-making process. In addition, investors have become more attentive to environmental and social issues related to a company's operations. Collectively, these concepts often are referred to as environmental, social, and governance (**ESG**).

Section 1 of this reading provides an overview of corporate governance, including its underlying principles and theories. Section 2 discusses the various stakeholders of a company and conflicts of interest that exist among stakeholder groups. Section 3 describes the principal–agent and other relationships in corporate governance and the conflicts that may arise in these relationships. Section 4 describes stakeholder management, reflecting how companies manage their relationships with stakeholders. Section 5 focuses on mechanisms to manage stakeholder relationships and mitigate associated risks. Section 6 focuses on the role of the board of directors and its committees as overseers of the company. Section 7 explores certain key factors that affect stakeholder relationships and corporate governance. Section 8 highlights the risks and benefits that underlie a corporate governance structure. Section 9 synthesizes corporate governance concepts that should be considered by investment professionals. Finally, Section 10 discusses the growing use of environmental, social, and governance factors in investment analysis and portfolio construction processes.

1.1 Corporate Governance Overview

Corporate governance can be defined as "the system of internal controls and procedures by which individual companies are managed. It provides a framework that defines the rights, roles, and responsibilities of various groups within an organization. At its core, corporate governance is the arrangement of checks, balances, and incentives a company needs in order to minimize and manage the conflicting interests between insiders and external shareowners."[1]

Corporate governance practices differ among countries and jurisdictions, and even within countries different corporate governance systems may co-exist. The corporate governance systems adopted in most of the world typically reflect the influences of either *shareholder theory* or *stakeholder theory* to a varying extent, as well as historical, cultural, legal, political, and other influences specific to a region.

Shareholder theory takes the view that the most important responsibility of a company's managers is to maximize shareholder returns. Stakeholder theory broadens a company's focus beyond the interests of only its shareholders to its customers, suppliers, employees, and others who have an interest in the company. The approach to corporate governance in a given country typically places greater emphasis on one of the two theories but can also exhibit a combination of the two. Notwithstanding the system of corporate governance used, nearly all companies depend on contributions from a number of stakeholders for their long-term success. The company's strategy is set by the board of directors, which also oversees management. In turn, the company's strategy is executed by its managers; financial capital to fund the company's activities and operations is supplied by shareholders, creditors, and suppliers; human capital is provided by employees; and demand for goods and services comes from customers. Other stakeholders include governments and regulators, which seek to protect the interests and well-being of their citizens. Certain external forces, such as the legal environment and competition, affect the way a company operates and the relationships among its stakeholders.

Two reports issued during the 1990s, the Cadbury Report and the Principles of Corporate Governance, were particularly influential in shaping the global corporate governance landscape. In 1991, the Committee on the Financial Aspects of Corporate Governance was established in the United Kingdom by the Financial Reporting Council, the London Stock Exchange, and the accountancy profession to examine corporate governance. In the following year, the report of the committee—commonly referred to as the Cadbury Report, after its chairperson—defined corporate governance simply as "the system by which companies are directed and controlled." The report focused on the responsibilities of a company's board of directors, shareholders, and auditors, with shareholders implicitly identified as the primary stakeholders. In 1999, the Organisation for Economic Co-Operation and Development (OECD) produced the *Principles of Corporate Governance*, which expanded the scope of corporate governance to consider the interests of other stakeholders—notably, employees, creditors, and suppliers. According to the OECD, "Corporate governance includes a set of relationships between a company's management, its board, its shareholders, and other stakeholders." The *Principles of Corporate Governance*, which has been revised numerous times since its initial publication, also discusses potential positive outcomes of good corporate governance practices (including financial market stability and economic growth) and includes standards and guidelines designed to evaluate and improve the corporate governance framework throughout the world.

There is evidence that some movement toward global convergence of corporate governance systems is underway. One trend is the increased acceptance and adoption of corporate governance regulations with similar principles from one jurisdiction to another. For example, several countries implemented regulations similar to those of the US Sarbanes–Oxley Act of 2002 (SOX) in response to corporate and accounting scandals of the early 2000s. Although these regulations are not identical, they share the same objective of improving internal controls and restoring investor confidence in financial disclosures. Another trend is initiatives by international agencies to build greater consensus on important corporate governance principles. The *Principles of Corporate Governance*, for example, has been ratified by more than 30 member countries, representing a broad range of corporate governance models. The *Principles of Corporate Governance* do not mandate, or even promote, the adoption of a single corporate governance regime; rather, the principles were designed to serve as a framework that can be adopted by any number of corporate governance systems.

---

[1] CFA Institute Centre for Financial Market Integrity, *The Corporate Governance of Listed Companies: A Manual for Investors*, 2nd ed. (Charlottesville, VA: CFA Institute, 2009).

EXAMPLE 1

Corporate Governance Overview

Which statement regarding corporate governance is *most* accurate?

A Most countries have similar corporate governance regulations.

B A single definition of corporate governance is widely accepted in practice.

C Both shareholder theory and stakeholder theory consider the needs of a company's shareholders.

Solution:

C is correct. Both shareholder and stakeholder theories consider the needs of shareholders, with the latter extending to a broader group of stakeholders. A is incorrect because corporate governance regulations differ across countries, although there is a trend toward convergence. B is incorrect because a universally accepted definition of corporate governance remains elusive.

2

STAKEHOLDER GROUPS

b describe a company's stakeholder groups, and compare interests of stakeholder groups

A corporate governance system is likely to be influenced by several stakeholder groups. These groups do not necessarily share similar goals or needs; in fact, the interests of any one group may conflict with the interests of another group. The varying influences of these groups are important considerations for investment professionals when analyzing a corporate governance system. This section provides an overview of a corporation's primary stakeholder groups, followed by a discussion of principal–agent considerations and the conflicts that may arise among the groups.

2.1 Stakeholder Groups

The primary stakeholder groups of a corporation consist of shareholders, creditors, managers (or executives), other employees, board of directors, customers, suppliers, affected individuals and community groups, and governments/regulators. The interests of each of these groups are discussed in the following sections.

2.1.1 *Shareholders*

Shareholders own shares of stock in a corporation and are entitled to certain rights, such as the right to receive dividends and to vote on certain corporate issues.[2] In terms of capital structure, shareholders are the most junior class of capital providers; in case of a company bankruptcy, shareholders receive proceeds only after all creditors' claims are paid. Shareholder interests are, therefore, typically focused on growth in corporate profitability that maximizes the value of a company's equity.

As a company grows in size and its operations and structure become more complex, most individual shareholders have limited or no involvement in the company's activities. Shareholders maintain control over the company through their power to elect the board of directors and vote for specified resolutions. The board of directors is expected to represent shareholders—protecting their interests, appointing senior management, providing strategic direction, and monitoring company and management performance. In publicly traded companies that have dispersed ownership, the voting power in general meetings is distributed among a large number of shareholders. But in some companies, a particular shareholder or block of shareholders may hold a percentage of shares that gives them sufficient voting power to control the election of the board of directors and to influence the approval or blockage of a company resolution; these shareholders are known as **controlling shareholders**. In contrast, non-controlling shareholders (**minority shareholders**) hold a much smaller proportion of a company's outstanding shares, resulting in a more limited ability to exercise control in voting activities.

2.1.2 *Creditors*

Creditors, most commonly bondholders and banks, are a company's lenders and the providers of debt financing. Creditors do not hold voting power (unlike common shareholders) and typically have limited influence over a company's operations. Creditors may protect themselves and exert some control over a company by using covenants, which restrict activities of the borrower. In return for capital provided, creditors expect to receive interest and principal payments. These payments are pre-determined from the terms of a debt contract and are typically not contingent on the company's performance. Creditors usually do not participate in a company's superior performance beyond receiving promised interest and principal payments. The company's ability to generate cash flows, mainly through its operations, is the primary source of payments for creditors. Consequently, creditors generally prefer stability in company operations and performance, which contrasts with the interests of shareholders, who generally are inclined to tolerate higher risks in return for higher return potential from strong company performance.

2.1.3 *Managers and Employees*

Senior executives and other high-level managers are normally compensated through salary, bonuses, equity-based remuneration (or compensation),[3] and certain perquisites. As a result, managers may be motivated to maximize the value of their total remuneration while also protecting their employment positions. Lower-level employees typically seek fair remuneration, good working conditions, access to promotions, career development opportunities, training and development, job security, and a safe and healthy work environment. Overall, remuneration packages that are designed to prevent employees from chasing short-term profits are likely to be most effective and aligned with long-term shareholder interests.

---

[2] https://themoderncorporation.wordpress.com/company-law-memo/

[3] The terms "remuneration" and "compensation" are typically interchangeable, with compensation generally used in North America and remuneration generally used outside North America. In this reading, unless specifically identified with North America, we primarily use "remuneration."

As with shareholders and creditors, managers and employees have a significant interest in the company's viability. Managers and employees tend to benefit if the company performs well and are among the most adversely affected stakeholders if a company's financial position weakens. Despite some similarities, the interests of managers and employees and other stakeholders can conflict. For example, a company may be presented with a takeover offer that is attractive to shareholders but would jeopardize the interests of managers in preserving their employment at the company.

2.1.4 *Board of Directors*

A company's board of directors is elected by shareholders to protect shareholders' interests, provide strategic direction, and monitor company and management performance. A board is typically structured as either *one-tier* or *two-tier*.

A one-tier structure consists of a single board of directors, composed of executive and non-executive directors. Executive (sometimes called "internal") directors are employees, typically senior managers, of the company. Non-executive (sometimes called "external") directors are not employees of the company. Countries in which one-tier boards are common include the United States, the United Kingdom, and India. A two-tier structure consists of two separate boards: (1) a *supervisory board*, which is primarily composed of non-executive directors, and (2) a *management (executive) board*, which is composed of executive directors. The supervisory board oversees the management board. Two-tier boards are common in such countries as Germany, the Netherlands, Finland, and China.

In this reading, unless specified otherwise, the term "board" refers to the single board of directors in a one-tier structure and the supervisory board in a two-tier structure. Directors, both internal and external, are typically experienced individuals who are focused on fulfilling their responsibilities toward shareholders and the company while maintaining a good reputation in the business community. Directors are also typically concerned with their exposure to liability for breach of duty. Directors can mitigate this exposure, *inter alia*, by exercising appropriate levels of control over the company's operations and its management or by obtaining independent access to company documentation. A company's board of directors is discussed in more detail in Section 6 of this reading.

2.1.5 *Customers*

Customers expect a company's products or services to satisfy their needs and provide appropriate benefits given the price paid, as well as to meet applicable standards of safety. Depending on the type of product or service and the duration of their relationship with the company, customers may desire ongoing support, product guarantees, and after-sale service. Companies are concerned with customer satisfaction given its potential correlation with sales revenue and profit. They are also increasingly concerned about brand boycotts due to adverse environmental and social impacts and controversies caused by their products and services. Compared with other stakeholder groups, customers tend to be less concerned with, and affected by, a company's financial performance. However, customers, particularly those who are dependent on the goods or services provided by the company, typically have an interest in a company's stability.

2.1.6 *Suppliers*

A company's suppliers have a primary interest in being paid as contracted or agreed on, and in a timely manner, for products or services delivered to the company. Suppliers often seek to build long-term relationships with companies for the benefit of both parties and aim for these relationships to be fair and transparent. Suppliers, like creditors, are concerned with a company's ability to generate sufficient cash flows to meet its financial obligations.

2.1.7 *Governments/Regulators*

Governments and regulators seek to protect the interests of the general public and ensure the well-being of their nations' economies. Because corporations have a significant effect on a nation's economic output, capital flows, employment, and social welfare, among other factors, regulators have an interest in ensuring that corporations behave in a manner that is consistent with applicable laws. Moreover, as the collector of tax revenues from companies and their employees, a government can also be considered one of the company's major stakeholders.

STAKEHOLDERS IN NON-PROFIT ORGANIZATIONS

The stakeholders of a non-profit organization tend to differ from those of for-profit companies. Non-profit organizations do not have shareholders. Their stakeholders most commonly include board directors or trustees, employees, regulators, society, patrons of the organization, donors, and volunteers. The stakeholders of non-profit organizations are generally focused on ensuring that the organization is serving the intended cause and that donated funds are used as promised.

EXAMPLE 2

Stakeholder Groups

Which stakeholders would *most likely* realize the greatest benefit from a significant increase in the market value of the company?

A Creditors

B Customers

C Shareholders

Solution:

C is correct. Shareholders own shares of stock in the company, and their wealth is directly related to the market value of the company. A is incorrect because creditors are usually not entitled to any additional cash flows (beyond interest and debt repayment) if the company's value increases. B is incorrect because customers may have an interest in the company's stability and long-term viability, but they do not benefit directly from an increase in a company's value.

PRINCIPAL–AGENT AND OTHER RELATIONSHIPS IN CORPORATE GOVERNANCE

3

c describe principal–agent and other relationships in corporate governance and the conflicts that may arise in these relationships

A **principal–agent relationship** (also known as an agency relationship) is created when a principal hires an agent to perform a particular task or service. The principal–agent relationship involves obligations, trust, and expectations of loyalty; the agent is expected to act in the best interests of the principal. In a company, agency theory stipulates that principal–agent relationships often lead to conflicts—for example, when managers do not act in the best interests of shareholders.[4] Examples of principal–agent relationships and potential conflicts between the principal and agent are discussed in the following sections. Other conflicts among stakeholder groups not involving principal–agent relationships are also discussed.

3.1 Shareholder and Manager/Director Relationships

In shareholder-owned companies, shareholders typically grant directors and managers the responsibility to make most corporate decisions. According to traditional shareholder theory (discussed earlier), the central duty of directors and managers is to act in the best interests of shareholders. In certain circumstances, however, managers may seek to maximize their personal benefits (e.g., remuneration and perquisites) to the detriment of shareholders' interests.

Shareholder interests can also diverge from those of managers and directors with respect to risk tolerance. In some cases, shareholders with diversified investment portfolios may have a relatively high risk tolerance because the risk undertaken by a specific company can be diversified across the shareholders' investments. Managers and directors, however, are typically more risk averse in their corporate decision making so they can better protect their employment status. Such behavior may differ from the company's value creation objective. In addition, compared with shareholders, managers typically have greater access to information about the business and are more knowledgeable about its operations. Such "information asymmetry" (that is, unequal access to information) makes it easier for managers to make strategic decisions that are not necessarily in the best interest of shareholders and weakens the ability of shareholders to exercise control. Another conflict of interest might arise between shareholders and directors when the board is influenced by insiders. In this case, the ability of the board to properly perform its monitoring and control role may be hindered. Finally, a conflict between the two groups may occur if directors favor certain influential shareholders over other shareholders.

3.2 Controlling and Minority Shareholder Relationships

In companies in which a particular shareholder holds a controlling stake, conflicts of interest may arise among the controlling and minority shareholders. In such ownership structures, the opinions of minority shareholders are often outweighed or overshadowed by the influence of the controlling shareholders. Minority shareholders often have limited or no control over management and limited or no voice in director appointments or in major transactions that could have a direct effect on the value of their shares. For instance, in companies that adopt **straight voting** (that is, one vote for each share owned), controlling shareholders clearly wield the most influence in board of director elections, leaving minority shareholders with much less representation on the board.

---

[4] Agency theory considers the problems that can arise in a business relationship when one person delegates decision-making authority to another. The traditional view in the investment community is that directors and managers are agents of shareholders. More recently, however, many legal experts have argued that in several countries, corporations are separate "legal persons"; thus, directors and managers are agents of the corporations rather than shareholders (or a subset of shareholders). See https://themoderncorporation.wordpress.com/company-law-memo.

The decisions made by controlling shareholders, or their board representatives, could also have an effect on corporate performance and, consequently, on minority shareholders' wealth. Takeover transactions are notable situations in which controlling shareholders typically have greater influence than do minority shareholders with regard to the consideration received and other deal terms. A historical example of note occurred in 2007 when Qtel, Qatar's largest telecommunications company, executed a deal with a consortium of the shareholders of Wataniya, Kuwait's telecommunications company, to acquire the consortium's shares in Wataniya (representing a 51% stake in the target). The consortium of Wataniya's shareholders sold their shares to Qtel at a premium of 48% on the stock price to the exclusion of minority shareholders.

Related-party transactions are another example for which controlling shareholders may place their interests ahead of minority shareholders' interests. Such a situation could occur when a controlling shareholder maintains a financial interest in a transaction between the company and a third party and that transaction conflicts with the company's best interests. Consider, for example, a controlling shareholder that arranges a deal between the company and a third-party supplier that is owned by the shareholder's spouse whereby the supplier provides the company with inventory at above-market prices. Such a transaction would benefit the controlling shareholder and the spouse's interests but could harm the profitability of the company and the interests of minority shareholders.

Lastly, an equity structure with multiple share classes in which one class is non-voting or has limited voting rights creates a divergence between the ownership and control rights of different classes of shareholders. Under a multiple-class structure (traditionally called a *dual-class structure* when there are two share classes), the company's founders, executives, and other key insiders control the company by virtue of ownership of a share class with superior voting powers. The multiple-class structure enables controlling shareholders to mitigate dilution of their voting power when new shares are issued. Examples of companies that have adopted multiple-class stock structures are Alibaba and Facebook (each with two share classes).

3.3 Manager and Board Relationships

Given that a board of directors typically relies on management to operate the company, the board's monitoring role can be compromised in the event of limited information provided to the board. This conflict is particularly pronounced for non-executive directors who are typically not involved in the day-to-day operations of a company.

3.4 Shareholder versus Creditor Interests

Shareholders typically seek growth in corporate profitability because of the residual nature of equity returns. However, the pre-determined returns of debt obligations normally prevent creditors from receiving any cash flows beyond principal and interest payments but do expose creditors to default risk in case of extremely poor corporate performance. From an investment perspective, shareholders would likely prefer riskier projects with a strong likelihood of higher return potential, whereas creditors would likely prefer stable performance and lower-risk activities. A divergence in risk tolerance regarding the company's investments thus exists between shareholders and creditors.

Creditors may also find their interests jeopardized when the company attempts to increase its borrowings to a level that would increase default risk. If the company's operations and investments fail to generate sufficient returns required to repay the increased interest and debt obligations, creditors will be increasingly exposed to default risk. The distribution of excessive dividends to shareholders might also conflict with creditors' interests if it impairs the company's ability to pay interest and principal.

3.5 Other Stakeholder Conflicts

In a corporation, interests can also conflict among other stakeholders. Some of these situations are as follows:

- *Conflict between customers and shareholders*: For example, a company decides to charge a high price for its products or reduces product safety features to reduce costs.
- *Conflict between customers and suppliers*: A company offers overly lenient credit terms to its customers, whereby the company's ability to repay suppliers on time may be affected.
- *Conflict between shareholders and governments or regulators*: Examples of such conflicts may include a company adopting accounting and reporting practices that reduce its tax burden, thus potentially benefiting shareholders, or a bank's shareholders preferring a lower equity capital base while regulators prefer a higher capital position. This last conflict is fairly common in the banking industry and has been increasingly in focus since the global financial crisis of 2008–2009.

EXAMPLE 3

Stakeholder Relationships

A controlling shareholder of XYZ Company owns 55% of XYZ's shares, and the remaining shares are spread among a large group of shareholders. In this situation, conflicts of interest are *most likely* to arise between:

A shareholders and regulators.

B the controlling shareholder and managers.

C the controlling shareholder and minority shareholders.

Solution:

C is correct. In this ownership structure, the controlling shareholder's power is likely more influential than that of minority shareholders. Thus, the controlling shareholder may be able to exploit its position to the detriment of the interests of the remaining shareholders. Choices A and B are incorrect because the ownership structure in and of itself is unlikely to create material conflicts between shareholders and regulators or shareholders and managers.

4 OVERVIEW AND MECHANISMS OF STAKEHOLDER MANAGEMENT

d describe stakeholder management

e describe mechanisms to manage stakeholder relationships and mitigate associated risks

Because interests among stakeholder groups differ, companies often adopt mechanisms to more efficiently manage stakeholder relationships. **Stakeholder management** involves identifying, prioritizing, and understanding the interests of stakeholder groups, and, on that basis, managing the company's relationships with these groups.

4.1 Overview of Stakeholder Management

Effective communication and active engagement with the various stakeholders form the basis of stakeholder management. Although the practices underlying stakeholder management may vary, companies typically seek to balance the interests of their various stakeholders and thus limit the effect of conflicts.

To help balance these interests, corporate governance and stakeholder management frameworks reflect a legal, contractual, organizational, and governmental infrastructure that defines the rights, responsibilities, and powers of each group. The *legal infrastructure* defines the framework for rights established by law and the availability or ease of legal recourse for any violation of these rights. The *contractual infrastructure* is shaped by the contractual arrangements entered into by the company and its stakeholders that help define and secure the rights of both parties. The *organizational infrastructure* refers to internal systems, governance procedures, and practices adopted and controlled by the company in managing its stakeholder relationships. Lastly, the *governmental infrastructure* refers to regulations imposed on companies.

The corporate governance systems in such countries as France, Germany, and Japan focus on a broader range of stakeholders relative to the more shareholder-driven Anglo-American systems. Globally, there is a growing movement among regulators and practitioners to more effectively balance the interests of all stakeholders. For instance, the Companies Act 2006 in the United Kingdom introduced "enlightened shareholder value," which requires directors to consider the interests of all stakeholders—not just shareholders. Several regulators, such as those in the United Kingdom and Japan, have adopted stewardship codes that encourage more active engagement of institutional investors with companies.

EXAMPLE 4

Stakeholder Management

The component of stakeholder management in which a corporation has the *most* control is:

A legal infrastructure.

B contractual infrastructure.

C governmental infrastructure.

Solution:

B is correct. A corporation's contractual infrastructure refers to the contractual arrangements between the corporation and stakeholders. As such, the corporation has control over these arrangements. A is incorrect because the legal infrastructure is established by law, which is outside the corporation's own control. Similarly, C is incorrect because a corporation's governmental structure is largely imposed by regulators.

4.2 Mechanisms of Stakeholder Management

Stakeholder management and governance practices attempt to manage the interests of all stakeholders. As mentioned earlier, a prescribed or standard set of rights and practices does not exist across all companies, and the principles vary across countries and jurisdictions. Still, there are some common control elements and governance mechanisms among companies.

4.2.1 *General Meetings*

Corporate laws grant shareholders certain powers and controls. The participation of shareholders in general meetings, also known as general assemblies, and the exercise of their voting rights are among the most influential tools available. General meetings enable shareholders to participate in discussions and to vote on major corporate matters and transactions that are not delegated to the board of directors.

Companies are ordinarily required to hold an annual general meeting (AGM) within a certain period following the end of their fiscal year. The main purpose of those meetings is to present shareholders with the annual audited financial statements of the company, provide an overview of the company's performance and activities, and address shareholder questions. Shareholders also elect the directors at the AGM and, in some countries, may be required to approve the financial statements, discharge directors of their duties, appoint external auditors, or vote on the remuneration of the board and/or top management. Extraordinary general meetings can be called by the company or by shareholders throughout the year when significant resolutions requiring shareholder approval are proposed. These resolutions might relate to proposed material corporate changes, such as amendments to the company's bylaws or rights attached to a class of shares, mergers and acquisitions, or the sale of significant corporate assets or businesses.

All shareholders typically have the right to attend, speak at, and vote at general meetings. Regulations, particularly corporate laws, specify conditions for inviting shareholders to general meetings and circulating information to shareholders. These conditions vary across regulations but generally aim at ensuring the participation of a large number of shareholders in general meetings without imposing excessive restrictions on the ability of the company to hold a meeting. By engaging in general meetings, shareholders can exercise their voting rights on major corporate issues and better monitor the performance of the board and senior management. General meetings and the underlying voting procedures are among the most widely adopted practices by companies in mitigating agency problems and their associated risks.

Some resolutions, such as the approval of financial statements and the election of directors and auditors, are considered ordinary at general meetings because they require only a simple majority of votes to be passed. Decisions that are more material in nature may require a supermajority vote, such as two-thirds or 75% of votes, to be passed. Such special resolutions may include amendments to bylaws, voting on a merger or takeover transaction, or waiving pre-emptive rights. Depending on the ownership structure, supermajority requirements may make it harder for majority shareholders to influence corporate decisions at the expense of minority shareholders.

Proxy voting is a process that enables shareholders who are unable to attend a meeting to authorize another individual (for example, another shareholder or director) to vote on their behalf. Proxy voting is the most common form of investor participation in general meetings. Although most resolutions at most companies pass without controversy, sometimes minority shareholders attempt to strengthen their influence at companies via proxy voting. Several shareholders sometimes use this process to collectively vote their shares in favor of or in opposition to a certain resolution. **Cumulative voting** (as opposed to straight voting) enables shareholders to accumulate and vote all their shares for a single candidate in an election involving more than one director. This voting process raises the likelihood that minority shareholders are represented by at least one director on the board, but it may not be compatible with majority voting standards for director elections in which share ownership is widely dispersed. In terms of worldwide practice, the existence of cumulative voting varies; for example, it is mandated in Spain but not allowed in several countries, such as Germany, Japan, Singapore, and Turkey.

Minority shareholders are often granted rights to protect their interests in acquisitions. For example, companies in European Union member states are required to adopt sell-out rights. These rights allow minority shareholders who have voted against a merger offer to force a bidder with more than 90% of the target's voting rights to buy their shares at a fair price upon the deal's approval.

EXAMPLE 5

General Meetings

Which of the statements about extraordinary general meetings (EGMs) of shareholders is true?

A The appointment of external auditors occurs during the EGM.

B A corporation provides an overview of corporate performance at the EGM.

C An amendment to a corporation's bylaws typically occurs during the EGM.

Solution:

C is correct. An amendment to corporate bylaws would normally take place during an EGM, which covers significant changes to a company, such as bylaw amendments. A and B are incorrect because the appointment of external auditors and a corporate performance overview would typically take place during the AGM.

4.2.2 *Board of Director Mechanisms*

In companies with complex ownership structures and operations, it is impractical for shareholders to be involved in strategy formulation and day-to-day activities. Shareholders thus elect a board of directors to provide broad oversight of the company. Shareholders monitor the board's performance through exercise of voting power and participation in general meetings. The board, in turn, appoints the top management of the company. The board is accountable primarily to shareholders and is responsible for the proper governance of the company; in this regard, the board is the link between shareholders and managers. The board guides managers on the company's strategic direction, oversees and monitors management's actions in implementing the strategy, and evaluates and rewards or disciplines management performance. The board also supervises the company's audit, control, and risk management functions and ensures the adoption of proper governance systems and compliance with all applicable laws and regulations. In Section 6, we provide more detail regarding the functions and responsibilities of the board of directors and its committees.

4.2.3 *The Audit Function*

The audit function is an integral component of any governance structure. The function represents the systems, controls, and policies/procedures in place to examine the company's operations and financial records. Internal audits are conducted by an independent internal audit function or department. External auditors are independent from the company and conduct an annual audit of the company's financial records to provide reasonable and independent assurance of the accuracy of financial statements and their fair representation of the financial position of the company. External auditors are typically recommended by an audit committee (which we discuss later in the reading) for appointment by shareholders or, in some jurisdictions, by the board. The board of directors is generally required to receive and review the financial statements and auditors' reports and confirm their accuracy before they are presented to shareholders for approval at the AGM. Senior management of publicly traded companies is also required to review and provide assurance of the effectiveness of the internal control systems to the board of directors or to shareholders. Overall, a company's audit function limits insiders' discretion with regard to the use of company resources and to its financial reporting. The audit function is also designed to mitigate incidents of fraud or misstatements of accounting and financial information.

4.2.4 *Reporting and Transparency*

Shareholders have access to a range of financial and non-financial information concerning the company, typically through annual reports, proxy statements, disclosures on the company's website, the investor relations department, and other means of communication (e.g., social media). This information may relate to the company's operations, its strategic direction or objectives, audited financial statements, governance structure, ownership structure, remuneration policies, related-party transactions, and risk factors. Such information is essential for shareholders to

- reduce the extent of information asymmetry between shareholders and managers;
- assess the performance of the company and of its directors and managers;
- make informed decisions in valuing the company and deciding to purchase, sell, or transfer shares; and
- vote on key corporate matters or changes.

4.2.5 *Policies on Related-Party Transactions*

The development and implementation of policies for related-party transactions and other conflicts of interest is an increasingly common practice among companies. These policies establish the procedures for mitigating, managing, and disclosing such cases. Typically, directors and managers are required to disclose any actual or potential, or direct or indirect, conflict of interest they have with the company, as well as any material interests in a transaction that may affect the company. Often, these policies require such transactions or matters to be voted on by the board (or shareholders) excluding the director (or shareholder) holding the interest. The adoption of these policies and procedures aims at ensuring that related-party transactions are handled fairly and that they do not advance the interests of the related party at the expense of the interests of the company or other shareholders.

4.2.6 *Remuneration Policies*

Executive remuneration plans have gained significant attention in the investment world, with a primary goal of aligning the interests of managers with those of shareholders. For this purpose, incentive plans increasingly include a variable component—typically profit sharing, stocks, or stock options—that is contingent on corporate or stock price performance. However, the granting of stock-based remuneration does not serve its purpose if managers can improve their personal gains at the expense of the company while limiting their exposure to weak stock performance. As a result, companies are increasingly designing incentive plans that discourage either "short-termism" or excessive risk taking by managers. Some incentive plans include granting shares, rather than options, to managers and restricting their vesting or sale for several years or until retirement. A long-term incentive plan delays the payment of remuneration, either partially or in total, until company strategic objectives (typically performance targets) are met.

Regulators across the world are also increasingly focused on remuneration policies. In some cases, regulators require companies to base remuneration specifically on long-term performance measures. A number of regulators are requiring companies, including many in the financial industry, to adopt clawback provisions. These provisions allow a company to recover previously paid remuneration if certain events, such as financial restatements, misconduct, breach of the law, or risk management deficiencies, are uncovered.

4.2.7 *Say on Pay*

Given the role of remuneration plans in aligning the interests of executives with those of shareholders, regulators and companies are increasingly seeking shareholder views on pay. The concept of **say on pay** enables shareholders to vote on executive remuneration matters. Say on pay was first introduced in the United Kingdom in the early 2000s. In an early example of shareholder rejection, in 2003 the shareholders of GlaxoSmithKline rejected the company's remuneration report because of opposition to the proposed executive pay. This was the first such rejection by the shareholders of a large UK-based company. Shortly thereafter, the practice of say on pay spread to other parts of the world and was implemented in the United States as part of the Dodd–Frank Act in 2011. In 2018, there were a number of instances in which shareholders voted against remuneration reports. In the United Kingdom, for example, Inmarsat plc's remuneration report was rejected by 58% of voters over concerns that executive compensation was not aligned with company performance. In the United States, 52% of Walt Disney Co.'s shareholders voted against what they believed was an excessive executive compensation package.

The scope and effect of say on pay varies across countries and companies. Some countries, such as Canada, have a non-mandatory and advisory (non-binding) say on pay system in which shareholders signal, rather than impose, their views on proposed remuneration. In other countries, such as the United States, France, and South Africa, say on pay is mandatory but non-binding. In these countries, the board is required to enable shareholders to vote on remuneration plans or packages, but the board does not have to abide by the result of the vote. Conversely, countries in which shareholder votes on say on pay are binding include the Netherlands, the United Kingdom, and China.

Say on pay has been subject to criticism because of the fact that shareholders often have limited involvement in a company's strategy and operations. These opponents argue that the board is better suited to determine remuneration matters. Conversely, by allowing shareholders to express their views on remuneration-related matters, companies can limit the discretion of directors and managers in granting themselves excessive or inadequate remuneration. This approach could thus reduce a critical agency conflict in stakeholder management by better aligning the interests of principals and agents.

MECHANISMS TO MITIGATE ASSOCIATED STAKEHOLDER RISKS

5

e describe mechanisms to manage stakeholder relationships and mitigate associated risks

The rights of creditors are established by laws and according to contracts executed with the company. Laws vary by jurisdiction but commonly contain provisions to protect creditors' interests and provide legal recourse. One such provision is an **indenture**, which is a legal contract that describes the structure of a bond, the obligations of the issuer, and the rights of the bondholders. To limit creditors' risk during the term of a bond (or loan), debtholders may choose to impose **covenants** within indentures or contracts. Covenants are the terms and conditions of lending agreements, enabling creditors to specify the actions an issuer is obligated to perform or prohibited from performing. Affirmative covenants require bond issuers to perform certain actions, such as maintaining adequate levels of insurance. Restrictive covenants require bond issuers to not perform certain actions, such as allowing the liquidity level to fall below a minimum coverage ratio. **Collaterals** are another tool often used by creditors to guarantee repayment, representing assets or financial guarantees that are above and beyond an issuer's promise to repay its obligations.

To further protect their rights, creditors usually require the company to provide periodic information (including financial statements) to ensure that covenants are not violated and thus potential default risk is not increased. Because it is usually impractical and costly for individual bondholders to fully scrutinize a bond issue, companies often hire a financial institution to act as a trustee and monitor the issue on behalf of a class of bondholders. In some countries, credit committees, particularly for unsecured bondholders, are established once a company files for bankruptcy. Such committees are expected to represent bondholders throughout the bankruptcy proceedings and protect bondholder interests in any restructuring or liquidation.

5.1 Employee Laws and Contracts

Employee rights are primarily secured through labor laws, which define the standards for employees' rights and responsibilities and cover such matters as labor hours, pension and retirement plans, hiring and firing, and vacation and leave. In most countries, employees have the right to create unions. Unions seek to influence certain matters affecting employees. In the European Union, companies meeting specific size and geographic criteria are required to establish European Works Councils that are composed of employees who meet with management at least annually. Although not a common practice in the United States and many other parts of the world, employees are sometimes represented on the board of directors—or supervisory boards—of companies meeting certain size or ownership criteria (e.g., in Germany, Austria, and Luxembourg). In Japan, the employee model stresses reaching a consensus between management and employees in decision making.

At the individual level, employment contracts specify an employee's various rights and responsibilities. Employment contracts typically do not cover every situation between employees and employers, and thus there is some area of discretion in the employment relationship. Human resources policies also help companies manage their relationships with employees. Effective human resources policies seek to attract and recruit high-quality employees while providing remuneration, training/development, and career growth to improve employee retention. Some companies have employee stock ownership plans (ESOPs) to help retain and motivate employees. As part of an ESOP, a company establishes a fund consisting of cash and/or company shares. The shares, which have designated vesting periods, are granted to employees.

Codes of ethics and business conduct also serve an important role in the relationship between employees and the company. Such codes establish the company's values and the standards of ethical and legal behavior that employees must follow. Companies typically assign a compliance or corporate governance officer (or a board committee) to implement these codes, receive violation reports, and resolve ethical matters. In addition, whistleblower protection policies are another mechanism to ensure misconduct can be safely reported so as to mitigate associated stakeholder risks.

By managing its relationships with its employees, a company seeks to comply with employees' rights and mitigate legal or reputational risks in violation of these rights. Employee relationship management also helps ensure that employees are fulfilling their responsibilities toward the company and are qualified and motivated to act in the company's best interests.

5.2 Contractual Agreements with Customers and Suppliers

Both customers and suppliers enter into contractual agreements with a company that specify the products and services underlying the relationship, the prices or fees and the payment terms, the rights and responsibilities of each party, the after-sale relationship, and any guarantees. Contracts also specify actions to be taken and recourse available if either party breaches the terms of the contract.

5.3 Laws and Regulations

As part of their public service roles, governments and regulatory authorities develop laws that companies must follow and monitor companies' compliance with these laws. Such laws may address or protect the rights of a specific group, such as consumers or the environment. Some industries or sectors whose services, products, or operations are more likely to endanger the public or specific stakeholders' interests are usually subject to a more rigorous regulatory framework. Examples of these industries are banks, food manufacturers, and health care companies.

Many regulatory authorities have also adopted corporate governance codes that consist of guiding principles for publicly traded companies. Publicly traded companies, in turn, are generally required to annually publish corporate governance reports describing their governance structure and explain any deviations from guiding principles. Companies normally seek to adopt internal governance and compliance procedures and adhere to the relevant financial reporting and transparency requirements imposed by regulators.

EXAMPLE 6

Stakeholder Relationships

Which of the following is **not** typically used to protect creditors' rights?

A Proxy voting

B Collateral to secure debt obligations

C The imposition of a covenant to limit a company's debt level

Solution:

A is correct. Proxy voting is a practice adopted by shareholders, not creditors. B and C are incorrect because both collateral and covenants are used by creditors to help mitigate the default risk of a company.

COMPANY BOARDS AND COMMITTEES

6

f describe functions and responsibilities of a company's board of directors and its committees

As discussed earlier in the reading, the board of directors is a central component of a company's governance structure. The board serves as the link between shareholders and managers and acts as the shareholders' monitoring tool within the company. As the relevance of corporate governance has grown within the investment field, the responsibilities of the board of directors have also increased in importance.

6.1 Composition of the Board of Directors

The structure and composition of a board of directors vary by company and geography. There is no single or optimal structure, and the number of directors may differ depending on the company size, structure, and complexity of operations. Most corporate governance codes require that the board include a diverse mix of expertise, backgrounds, and competencies. Such qualifications may include specialized knowledge of the company's industry as well as experience in certain functions, such as strategy, finance/audit, risk management, human resources, or legal. Moreover, many companies seek age, gender, and racial diversity in board composition.

Boards with one-tier structures comprise a mix of executive and non-executive directors. Executive (or internal) directors are employed by the company and are typically members of senior management. Non-executive (or external) directors provide objective decision making, monitoring, and performance assessment. Additionally, non-executive directors can serve an important role in challenging management and using past expertise in strategy and board issues. An *independent director* is a specific type of non-executive director that does not have a material relationship with the company with regard to employment, ownership, or remuneration.

In two-tier structures, the supervisory and management boards are independent from each other. Regulators generally prohibit members of the management board from serving on the supervisory board or limit the number of individuals serving on both boards. Employee representatives are typically elected by the company's employees and could make up half of the supervisory board in large companies.

In many countries, the chief executive officer (CEO) and chairperson roles are increasingly separated. In the United States, many companies have historically had "CEO duality," in which the CEO also serves as chairperson of the board. Nevertheless, the percentage of companies separating the two roles, particularly in the United States, has increased considerably since the global financial crisis of 2008–2009. The appointment of a lead independent director is an alternative that is sometimes practiced by boards of companies with CEO duality. The lead independent director generally has the authority to request and oversee meetings of all independent directors. Duality is not applicable in two-tier structures that prohibit the members of the management board from serving on the supervisory board. In these models, the chairperson of the supervisory board is typically external and the CEO usually chairs the management board.

STAGGERED BOARDS

The general practice for boards is that elections occur simultaneously and for specified terms (three years, for example). Some companies, however, have **staggered boards**, whereby directors are typically divided into three classes that are elected separately in consecutive years—that is, one class every year. Because shareholders would need several years to replace a full board, this election process limits their ability to effect a major change of control at the company. The positive aspect of a staggered board, though, is that it provides continuous implementation of strategy and oversight without constantly being reassessed by new board members thereby bringing short-termism into company strategy. This staggered board model was historically prevalent in the United States but has been generally replaced by regular board election terms. In contrast, the practice is common in Australia.

6.2 Functions and Responsibilities of the Board

As mentioned earlier, a company's board of directors is elected by shareholders to act on their behalf. In fulfilling their functions, directors have a responsibility to consider the interests of all stakeholders. The duties of directors are mandated by law in many countries but vary across jurisdictions. Two widely established elements of directors' responsibilities are the *duty of care* and the *duty of loyalty*. According to the OECD's *Principles of Corporate Governance*, duty of care "requires board members to act on a fully informed basis, in good faith, with due diligence and care." The OECD further notes that duty of loyalty "is the duty of the board member to act in the interest of the company and shareholders. The duty of loyalty should prevent individual board members from acting in their own interest, or the interest of another individual or group, at the expense of the company and all shareholders."

A board of directors does not typically engage in the company's day-to-day activities; rather, those activities are delegated to management. The board guides and approves the company's strategic direction, taking into consideration the company's risk profile. It delegates the implementation of the company's strategy to senior management, oversees the execution of the strategy, and establishes milestones to monitor the progress in reaching the objectives. The board also reviews corporate performance and determines relevant courses of action accordingly. In doing so, the board can monitor and evaluate management's performance and determine whether senior executive remuneration is aligned with the long-term interests of the company. The board is also responsible for selecting, appointing, and terminating the employment of senior managers (or the management board in case of a two-tier structure). One of the board's main responsibilities is to ensure leadership continuity through succession planning for the CEO and other key executives.

The board plays a central role in ensuring the effectiveness of the company's audit and control systems. It sets the overall structure of these systems and oversees their implementation, including oversight of the financial reporting practices and review of the financial statements for fairness and accuracy. The board also oversees reports by internal audit, the audit committee, and the external auditors and proposes and follows up on remedial actions. The board has the ultimate responsibility to ensure that the company adopts and implements proper corporate governance principles and complies with all applicable internal and external laws and regulations, including ethical standards.

In addition, the board typically ensures that the company has an appropriate enterprise risk management system in place, whereby risks are identified, mitigated, assessed, and managed appropriately. The board monitors the effectiveness of these systems through regular reviews and reports received from both management and the company's risk function. The board also has the responsibility to review any proposals for corporate transactions or changes, such as major capital acquisitions, divestures, mergers, and acquisitions, before they are referred to shareholders for approval, if applicable.

6.3 Board of Directors Committees

A company's board of directors typically establishes committees that focus on specific functions. Such committees, in turn, provide recommendations that are reported to the board on a regular basis. Despite the delegation of responsibilities to committees, the overall board remains ultimately responsible to shareholders and is not discharged of its liabilities to shareholders. Although board committees may vary by organization, some of the most common committees are discussed in the following sections.

6.3.1 *Audit Committee*

The audit committee is perhaps the most common board committee among companies worldwide. The audit committee plays a key role in overseeing the audit and control systems at the company and ensuring their effectiveness. In this regard, the committee monitors the financial reporting process, including the application of accounting policies; ensures the integrity of financial statements; supervises the internal audit function and ensures its independence and competence; and presents an annual audit plan to the board and monitors its implementation by the internal audit function. The audit committee is also responsible for recommending the appointment of an independent external auditor, periodically assessing the independence of the external auditor, and proposing the auditor's remuneration. Both internal and external auditors report their findings to the audit committee, which in turn proposes remedial action for highlighted issues or matters.

6.3.2 *Governance Committee*

The primary role of the governance (or corporate governance) committee is to ensure that the company adopts good corporate governance practices. In doing so, the committee develops and oversees the implementation of the corporate governance code, the charters of the board and its committees, and the company's code of ethics and conflict of interest policy. The governance committee reviews these policies on a regular basis to incorporate any regulatory requirements or relevant developments in the field. Most importantly, the committee monitors the implementation of the governance policies and standards as well as the compliance with applicable laws and regulations throughout the company. Remedial actions are recommended if any flaws or breaches of laws or regulations are identified. In some companies, the governance committee may be responsible for overseeing an annual evaluation of the board to ensure that its structure and activities are aligned with the governance principles.

6.3.3 *Remuneration or Compensation Committee*

The remuneration (or compensation) committee of the board specializes in remuneration matters. This committee develops and proposes remuneration policies for the directors and key executives and presents them for approval by the board or shareholders. The committee may also be involved in handling the contracts of managers and directors as well as in setting performance criteria and evaluating the performance of managers. The responsibilities of the remuneration committee may extend to establishing human resources policies for the company when remuneration matters are involved. In some companies, the remuneration committee also sets and oversees the implementation of employee benefit plans, including insurance, pension, severance benefits, and retirement plans (including monitoring the investment performance of benefit plan funds).

6.3.4 *Nomination Committee*

The nomination committee identifies candidates who are qualified to serve as directors and recommends their nomination for election by shareholders. The committee also establishes the nomination procedures and policies, including the criteria for board directors, the search process, and the identification of qualified candidates for director positions. Director independence, including what constitutes an independent director, is also a function of the nomination committee. Through these roles, the nomination committee can help ensure that the board's composition is well balanced and aligned with the company's governance principles.

6.3.5 *Risk Committee*

The risk committee assists the board in determining the risk policy, profile, and appetite of the company. Accordingly, the committee oversees establishing enterprise risk management plans and monitors their implementation. It also supervises the risk management functions in the company, receives regular reports, and reports on its findings and recommendations to the board.

6.3.6 *Investment Committee*

The investment committee of the board reviews material investment opportunities proposed by management and considers their viability for the company. Such opportunities may include large projects, acquisitions, and expansion plans, as well as divestures or major asset disposals. The committee often challenges, when necessary, management assumptions underlying investment prospects, monitors the performance of investments, and reports its findings to the board. The committee also is typically responsible for establishing and revising the investment strategy and policies of the company.

The specific board committees discussed in previous sections are the most commonly used, although the composition and number of committees may vary depending on the jurisdiction or on company-specific factors (e.g., company size, industry, complexity of operations, or regulatory requirements). An audit committee, for instance, is a regulatory requirement in a large number of jurisdictions. For banks and other financial institutions, a risk committee is strongly recommended by some regulators and required by others. In Brazil, the Central Bank of Brazil requires financial institutions to establish a remuneration committee at the board level. Some companies choose to combine two or more committees into one—for example, a nomination and remuneration (or compensation) committee or an audit and risk committee. Companies may also find it valuable to establish committees with other specializations, such as a compliance committee, an ethics committee, a human resources committee, or a health/environmental/safety committee.

The composition of a board committee is normally aligned with its scope of responsibilities. For instance, many regulators request that executive (internal) directors do not rule on matters underlying conflicts of interest or on matters requiring an unbiased judgment (such as audit, remuneration, or related-party transaction matters). As such, a large number of rules, including those adopted by the London Stock Exchange and the New York Stock Exchange, require that the audit and the compensation committees be composed of independent directors only. Countries with less-stringent regulations may require the audit committee to be composed of external (non-executive) directors, the majority of which should be independent (including the chairperson).

EXAMPLE 7

Responsibilities of Board Committees

A primary responsibility of a board's audit committee does **not** include the:

A proper application of accounting policies.

B adoption of proper corporate governance.

C recommendation of remuneration for the external auditor(s).

Solution:

B is correct. The adoption of proper corporate governance is the responsibility of a corporation's governance committee. Both A and C are incorrect because proper application of accounting policies and the remuneration of external auditors fall under the domain of the audit committee.

7 RELEVANT FACTORS IN ANALYZING CORPORATE GOVERNANCE AND STAKEHOLDER MANAGEMENT

g describe market and non-market factors that can affect stakeholder relationships and corporate governance

This section explores ways in which certain factors, both market and non-market related, can affect stakeholder relationships and corporate governance. For this section, market factors include those that relate to capital markets, whereas non-market factors do not.

7.1 Market Factors

This section focuses on shareholder engagement, shareholder activism, and competitive forces, all of which are influential market factors for a company. Shareholder engagement involves a company's interactions with its shareholders to better understand how it manages its material risks and opportunities. Shareholder activism, on the other hand, involves more aggressive and sometimes public efforts by large shareholder(s) to create a significant change in a corporation's behavior. Meanwhile, competitive dynamics can help align managerial interests with those of its stakeholders.

7.1.1 *Shareholder Engagement*

The engagement of companies with shareholders—called **shareholder engagement**—has traditionally involved certain events, such as the annual shareholder meeting and analyst calls, the scope of which was limited to financial and strategic matters. There is a growing trend, however, for greater engagement between companies and their shareholders beyond these venues and topics. Companies have increasingly recognized the benefits that frequent, year-round engagement with shareholders can provide, such as building support against short-term activist investors, countering negative recommendations from proxy advisory firms, and receiving greater support for management's position.

7.1.2 *Shareholder Activism*

Shareholder activism refers to strategies used by shareholders to attempt to compel a company to act in a desired manner. Although shareholder activism can focus on a range of issues, including those involving social or political considerations, the primary motivation of activist shareholders is to increase shareholder value. Activist shareholders often pressure management through such tactics as initiating proxy battles (fights), proposing shareholder resolutions, and publicly raising awareness on issues of contention.

Shareholder activists may pursue additional tactics, such as shareholder derivative lawsuits, which are legal proceedings initiated by one or more shareholders against board directors, management, and/or controlling shareholders of the company. The theory behind this type of lawsuit is that the plaintiff shareholder is deemed to be acting on behalf of the company in place of its directors and officers who have failed to adequately act for the benefit of the company and its shareholders. In many countries, however, the law restricts shareholders from pursuing legal action via the court system—in some cases, by imposing thresholds that enable only shareholders with interests above a minimum amount to pursue legal actions or by denying legal action altogether.

Hedge funds are among the most predominant shareholder activists. Compared with most traditional institutional investors, the fee structure of hedge funds often provides a significant stake in the financial success of any activist campaign. Furthermore, unlike regulated investment entities that are typically subject to restrictions on their investments (e.g., limitations on leverage or ownership of distressed or illiquid securities), hedge funds are loosely regulated and can thus pursue a greater range of activist opportunities.

7.1.3 *Competition and Takeovers*

Metrics that measure a company's success, such as market share or earnings growth, provide information that is useful for shareholders to judge the performance of a company's management team or board of directors and compare such performance with that of competitors. Senior managers risk their employment status in the event of underperformance, and directors, in turn, can be voted out by shareholders.

The traditional view of the market for corporate control (often known as the takeover market) is one in which shareholders of a company hire and fire management to achieve better resource utilization. Corporate takeovers can be pursued in several different ways. One mechanism is the **proxy contest** (or proxy fight). In a proxy contest, shareholders are persuaded to vote for a group seeking a controlling position on a company's board of directors. Managerial teams can also be displaced through a **tender offer**, which involves shareholders selling their interests directly to the group seeking to gain control. A contest for corporate control may attract arbitrageurs and takeover specialists, who facilitate transfers of control by accumulating long positions from existing shareholders in the target company and later selling the positions to the highest bidder. Finally, a **hostile takeover** is an attempt by one entity to acquire a company without the consent of the company's management.

Preservation of their employment status serves as an incentive for board members and managers to focus on shareholder wealth maximization. This threat of removal, however, can also have negative implications for a company's corporate governance practices if the company chooses to adopt anti-takeover measures, such as a staggered board or a shareholder rights plan (also known as a poison pill) to reduce the likelihood of an unwanted takeover. Staggering director elections can dilute the value of shareholder voting rights by extending the term that each director serves and eliminating the ability of shareholders to replace the entire board at any given election. Shareholder rights plans enable shareholders to buy additional shares at a discount if another shareholder purchases a certain percentage of the company's shares. These plans are designed to increase the cost to any bidder seeking to take over a company.

EXAMPLE 8

Shareholder Activism

Which of the following is true of shareholder activism?

A Shareholder activists rarely include hedge funds.

B Regulators play a prominent role in shareholder activism.

C A primary goal of shareholder activism is to increase shareholder value.

Solution:

C is correct. Although the subject of shareholder activism may involve social and political issues, activist shareholders' primary motivation is to increase shareholder value. A is incorrect because hedge funds commonly serve as shareholder activists. B is incorrect because regulators play a prominent role in standard setting, not shareholder activism.

7.2 Non-Market Factors

This section focuses on certain non-market factors, such as a company's legal environment, the role of the media, and the corporate governance industry, that can have an effect on stakeholder relationships and corporate governance.

7.2.1 *Legal Environment*

The legal environment in which a company operates can significantly influence the rights and remedies of stakeholders. Countries that have a common law system (such as the United Kingdom, the United States, India, and Canada) are generally considered to offer superior protection of the interests of shareholders and creditors relative to those that have adopted a civil law system (such as France, Germany, Italy, and Japan). The key difference between the two systems lies in the ability of a judge to create laws. In civil law systems, laws are created primarily through statutes and codes enacted by the legislature. The role of judges is generally limited to rigidly applying the statutes and codes to the specific case brought before the court. In contrast, in common law systems, laws are created both from statutes enacted by the legislature and by judges through judicial opinions. In common law systems, shareholders and creditors have the ability to appeal to a judge to rule against management actions and decisions that are not expressly forbidden by statute or code, whereas in civil law systems, this option is generally not possible.

Regardless of a country's legal system, creditors are generally more successful in seeking remedies in court to enforce their rights than shareholders are because shareholder disputes often involve complex legal theories, such as whether a manager or director breached a duty owed to shareholders. In contrast, disputes involving creditors, such as whether the terms of an indenture or other debt contract have been breached, are more straightforward and therefore more easily determinable by a court.

7.2.2 *The Media*

The media can affect corporate governance and influence stakeholder relationships through its ability to spread information quickly and shape public opinion. As an example, negative media attention can adversely affect the reputation or public perception of a company or its managers and directors. Senior management's concern over reputational risk can thus reduce the cost of monitoring management activities by stakeholders. Media attention can motivate politicians and regulators to introduce corporate governance reforms or enforce laws that protect stakeholders and society at large. This influence was evident in the aftermath of the 2008–2009 global financial crisis, when significant media attention was a factor in the adoption of new laws and regulations designed to address perceived deficiencies in corporate governance.

Social media has become a powerful tool that stakeholders have increasingly used to protect their interests or enhance their influence on corporate matters. Prior to the advent of social media, companies typically had an advantage in distributing information because of their considerable resources as well as relationships with traditional media organizations. Through social media, stakeholders can instantly broadcast information with little cost or effort and are thus better able to compete with company management in influencing public sentiment.

7.2.3 *The Corporate Governance Industry*

An important catalyst for the rise of the corporate governance industry occurred in 2003, when the Securities and Exchange Commission (SEC) required US-registered mutual funds to disclose their proxy voting records annually. The same rule also required US mutual funds to adopt policies and procedures designed to reasonably ensure that proxies would be voted in the best interests of their clients. In the years following the SEC's 2003 mutual fund rule, institutional investors have, to varying degrees, committed additional resources to monitor and vote proxies for the large number of companies in which they invest.

With the increased importance and relevance of corporate governance among investors, the demand for external corporate governance services has grown considerably. In particular, some institutional investors have retained proxy voting firms to assist with corporate governance monitoring and proxy voting. In response to this demand, an industry that provides corporate governance services, including governance ratings and proxy advice, has developed. Because the corporate governance industry is relatively concentrated, these vendors have considerable influence in corporate governance practices, and in turn, corporations are generally compelled to pay attention to ratings and recommendations produced by the corporate governance industry.

RISKS AND BENEFITS OF CORPORATE GOVERNANCE AND STAKEHOLDER MANAGEMENT

8

h identify potential risks of poor corporate governance and stakeholder management, and identify benefits from effective corporate governance and stakeholder management

As illustrated thus far, good corporate governance and stakeholder management can have a meaningfully positive effect on a company. A company will likely not meet the expectations of all stakeholders if one group is able to extract private benefits at the expense of another group. Depending on their nature and magnitude, unmanaged conflicts of interest and weak control over a company's operations may expose the company to various risks, such as legal, regulatory, reputational, or default risks. By adopting effective guidelines for managing the interests of stakeholder groups and instituting adequate levels of control, corporate governance can be reflected in better company relationships, superior levels of efficiency in operations, and improved financial performance.

8.1 Risks of Poor Governance and Stakeholder Management

Weaknesses in stakeholder management mechanisms or the adoption of poor governance structures can create various risks for a company and its stakeholders. A weak control environment can encourage misconduct and hinder the ability of the company to identify and manage risks.

8.1.1 *Weak Control Systems*

In a company with weak control systems or inefficient monitoring tools, such as poor audit procedures or insufficient scrutiny by the board, one stakeholder group may benefit at the expense of the company or other stakeholders. This could consequently have an adverse effect on the company's resources, performance, and value. The audit deficiencies at Enron, for instance, prevented the uncovering of the acts of fraud, erroneous accounting records, and other related issues that led to one of the largest corporate bankruptcies in US history.

8.1.2 *Ineffective Decision Making*

When the quality and quantity of information available to managers are superior to those available to the board or shareholders, in the absence of sufficient monitoring tools, managers have an opportunity to make decisions that benefit themselves relative to the company or shareholders. Without proper scrutiny, such practices might go unnoticed. Deficient decisions could include managing the company with a lower risk profile relative to shareholders' tolerance, thus avoiding investment opportunities that could create value for the company. Conversely, manager overconfidence may result in poor investment decisions without proper examination of their effect on the company or on shareholders' wealth.

Remuneration policies for management could also have significant implications for the company. Outsized remuneration packages for executives could have an adverse effect on shareholders' wealth, constitute a burden on corporate performance, and affect the interests of other stakeholders, such as employees, customers, or creditors. Remuneration policies that are not carefully designed may also encourage managers to seek immediate personal gains by taking excessive risks or focusing on creating short-term performance or stock price increases. Related-party transactions that underlie unfair terms for the company are another example of activities that are not aligned with the objective of value creation and that could be facilitated by a poor governance system.

8.1.3 *Legal, Regulatory, and Reputational Risks*

Compliance weaknesses in the implementation of regulatory requirements or lack of proper reporting practices may expose the company to legal, regulatory, or reputational risks. In such cases, the company may become subject to investigation by government or regulatory authorities for violation of applicable laws. A company could also be exposed to lawsuits filed by shareholders, employees, creditors, or other parties for breach of contractual agreements or company bylaws or for violation of stakeholders' legal rights. In today's markets, information flows rapidly. Improperly managed conflicts of interest or governance failures could bring reputational damage to the company, and its associated costs could be significant. Such risks are particularly acute for publicly listed companies subject to scrutiny by investors, analysts, and other market participants.

8.1.4 *Default and Bankruptcy Risks*

Poor corporate governance, including weak management of creditors' interests, can affect the company's financial position and may hinder its ability to honor its debt obligations. To the extent that the deterioration of corporate performance results in a debt default, the company may be exposed to bankruptcy risk if creditors choose to take legal action. The adverse consequences of corporate failures are not limited to the company's shareholders; they extend to other stakeholders, such as managers and employees and even society and the environment.

8.2 Benefits of Effective Governance and Stakeholder Management

The development of good governance practices can play a vital role in aligning the interests of managers and the board of directors with those of shareholders, while balancing the interests of the company's stakeholders. A good governance structure can be reflected in operational efficiency, improved control processes, better financial performance, and lower levels of risk.

8.2.1 *Operational Efficiency*

As part of a good governance structure, an organization clarifies the delegation of responsibilities and reporting lines across the company and ensures that all employees have a clear understanding of their respective duties. When balanced with adequate internal control mechanisms, the governance structure can ensure that corporate decisions and activities are properly monitored and controlled to mitigate risk and help improve the operational efficiency of the company.

8.2.2 *Improved Control*

Governance practices seek to institute more effective control exercised at all corporate levels, from shareholders to the board of directors and management. These practices can help identify and manage risk at early stages that can otherwise hinder corporate performance and/or damage reputation. Control can be enhanced by the proper functioning of a company's audit committee and the effectiveness of its audit systems. By adopting procedures for monitoring compliance with internal policies and external regulations and for reporting any violations, the company can better mitigate regulatory or legal risks and their associated costs. Additionally, the adoption of formal procedures with regard to conflicts of interest and related-party transactions allows the company to ensure fairness in its relationships with those parties.

8.2.3 *Better Operating and Financial Performance*

Good governance and stakeholder management can help a company improve its operating performance and reduce the costs associated with weak control systems. The costs of poor investments, legal proceedings against the company, and excessive perquisites are just a few examples that could be mitigated by well-functioning governance systems. Enhanced corporate governance could also allow the company to improve its decision-making process and respond faster to market factors. Proper remuneration policies are another governance tool that can motivate managers to make decisions with the objective of creating corporate value.

8.2.4 *Lower Default Risk and Cost of Debt*

Good corporate governance can lower business and investment risk. Governance arrangements that manage conflicts of interest with creditors, and that help protect creditor rights, can reduce a company's cost of debt and default risk. Default risk can also often be mitigated by proper functioning of audit systems, improved transparency (e.g., reporting of earnings), and the control of information asymmetries between the company and its capital providers. With regard to credit risk, corporate governance mechanisms are increasingly relevant criteria among credit rating agencies when assessing a company's creditworthiness.

EXAMPLE 9

Benefits of Corporate Governance

Which of the following is **not** a benefit of an effective corporate governance structure?

A Operating performance can be improved.

B A corporation's cost of debt can be reduced.

C Corporate decisions and activities require less control.

Solution:

C is correct. A benefit of an effective corporate governance structure is to enable adequate scrutiny and control over operations. B is incorrect because an effective governance structure can reduce investors' perceived credit risk of a corporation, thus potentially lowering the corporation's cost of debt. A is incorrect because operating efficiency may indeed be a benefit of an effective corporate governance structure.

9

FACTORS RELEVANT TO CORPORATE GOVERNANCE AND STAKEHOLDER MANAGEMENT ANALYSIS

i. describe factors relevant to the analysis of corporate governance and stakeholder management

In the past, analysts may have considered corporate governance and stakeholder management issues to be only peripherally related to traditional fundamental analysis. Generally, these issues were seen as obscure and unlikely to be material drivers of performance. Following a number of governance failures since the early 2000s, the global financial crisis, and the rise of shareholder activism around the world, there is little doubt that governance and stakeholder issues have become increasingly important topics for analysts.

Some key questions that analysts may consider when assessing a company's corporate governance or stakeholder management system are as follows:

- What is the company's ownership and voting structure among shareholders?
- Who represents shareholders on this company's board?
- What are the main drivers of the management team's remuneration and incentive structure?
- Who are the significant investors in the company?
- How robust are the shareholder rights at the company, including relative to peers?
- How effectively is the company managing long-term risks, such as securing access to necessary resources, managing human capital, exhibiting integrity and leadership, and strengthening the long-term sustainability of the enterprise?

A qualitative analysis of these issues—typically provided by a company's proxy statements, annual reports, and sustainability reports, if available—can provide important insights about the quality of management and sources of potential risk.

9.1 Economic Ownership and Voting Control

Generally speaking, corporations with publicly traded equity have a voting structure that involves one vote for each share. That is, any shareholder's voting power is equal to the percentage of the company's outstanding shares owned by that shareholder. When there are exceptions to this norm and economic ownership becomes separated from control, investors can face significant potential risks.

In a small number of markets, dual-class structures are allowed under the local regulatory framework or exchange rules, which is the most common way that voting power is decoupled from ownership. In these cases, common shares may be divided into two classes, one of which has superior voting rights to the other. A common arrangement is when a share class (for example, Class A) carries one vote per share and is publicly traded, whereas another share class (for example, Class B) carries several votes per share and is held exclusively by company insiders or family members. This structure is used by Facebook, for example. In this way, the founders and/or insiders of a company may continue to control board elections, strategic decisions, and all other significant voting matters for a long period—even once their ownership level declines to less than 50% of the company's shares.

Another mechanism used to separate voting control from economic ownership is when one class of stock (held by insiders) elects a majority of the board; outside shareholders who hold a different share class would then be entitled to elect only a minority of the board. Technically, each share carries equal voting rights, but with this structure, the insiders retain substantial power over the affairs of the corporation because they control a majority of the board. Alibaba's partnership structure is one example of this type of control.

Proponents of dual-share systems, such as those just mentioned, argue that the systems promote company stability and enable management to make long-term strategic investments, insulated from the short-term pressures of outside investors. Critics of these structures believe they create conflicts of interest between the providers of capital and the management of the business.

It is virtually impossible for outside investors to dismantle dual-class structures because of the inherent design of their unequal voting rights. Therefore, these structures can remain in place even through generational changes within a founding family. Investors with long time horizons may want to consider the motivations of the controlling stockholders, generational dynamics, succession planning, and the relationship between the board and management. In addition, there may be potential valuation implications because dual-class companies tend to trade at a discount to their peers.

9.2 Board of Directors Representation

In most markets, investors have access to basic biographical information about the non-executive members of corporate boards. Analysts can assess the available information to determine whether the experience and skill sets on the board match the current and future needs of the company.

In particular, questions regarding directors' independence, tenure, experience, and diversity may bring useful investment insights. For example, if the board has multiple directors engaging in related-party transactions with the company, investors may have cause for concern about any conflicts of interest that arise. The issue of board tenure can also be a useful tool for investors. Directors with long periods of service to a company clearly offer valuable experience and expertise, but if the board composition is dominated by such long-tenured members, it may have a negative effect on the board's diversity and adaptability.

An example in which board composition had a significant effect on company performance occurred at a European pharmaceutical company. At one point in its history, the company had become overleveraged and faced significant financial distress. In response, non-executive directors with banking and turnaround experience were added to the board. With the help of these directors, the company recovered. Seven years later, the most promising product in the company's pharmaceutical portfolio began to cause serious side effects in its patient population. The situation required both a meaningful understanding of the medical issues involved and a rapid response from the company. However, the board was still composed of directors with financial expertise rather than medical training. The company struggled with its response to the crisis, and its stock price fell sharply. This situation was one in which the board's composition had been ideal for a certain point in the company's history, but ultimately the directors failed to refresh the board's membership as the needs of the business changed.

9.3 Remuneration and Company Performance

The availability and quality of information about executive remuneration plans varies widely across markets. In those markets where such disclosures are available, analysts can assess the elements of the remuneration program to determine whether they support or conflict with the key drivers of performance for the company.

Generally speaking, current executive remuneration programs consist of a base salary, a short-term bonus usually delivered in the form of cash, and a multi-year incentive plan delivered in one or more forms of equity (options, time-vested shares, and/or performance-vested shares). Often, these short-term and long-term plans are contingent on achieving financial or operational objectives, and often these objectives are disclosed. In these situations, an analyst can assess whether the primary drivers of the remuneration plan are the same factors that, in the analyst's view, drive overall company results.

Assessment of the suitability of a remuneration plan for a particular company is a subjective exercise and is highly dependent on industry and geographic norms. But there are some warning signs that may warrant particular scrutiny:

- **Plans offering little alignment with shareholders.** As an example, if a plan offers only cash-based payouts and no equity, there may be a misalignment of incentives between executives and investors unless management already owns a significant stake in the company.
- **Plans exhibiting little variation in results over multiple years.** If an award is described as performance-based but still pays in full every year regardless of the company's results, investors may have concerns about the rigor of the performance hurdles underlying the awards.
- **Plans with excessive payouts relative to comparable companies with comparable performance.** Investors may want to understand the cause of the anomaly and whether it is a temporary issue or a result of flawed remuneration plan design.
- **Plans that may have specific strategic implications.** As an example, some remuneration plans contain payouts tied to specific milestones, such as regulatory approval of a product, completion of an acquisition, or achievement of specific cost reductions. In addition, some companies offer particularly high post-employment pay arrangements tied to the sale of the company, whereas

others offer no such arrangements. These factors are not necessarily negative features, but investors may want to understand whether the milestones driving the incentive plan align with the company's objectives.

- **Plans based on incentives from an earlier period in the company's life cycle.** A frequent example of such a plan relates to a company that has matured beyond its fast growth phase. The company's business may have matured, and competition may have limited the opportunity for market share gains. Investors may believe the company should become more focused on both returns and disciplined capital allocation. Even after the company communicates to the investor community a more returns-oriented strategy, the financial incentives in the remuneration plan may still be based purely on revenue growth. Investors may want to understand such potential misalignment of interests.

9.4 Investors in the Company

Examining the composition of investors in a company can be another source of useful insight for analysts. The behavior of these investors can result in both limitations and catalysts with regard to changes in the corporation. For example, cross-shareholdings are still prevalent in a number of markets. This situation occurs when a company, particularly a publicly listed one, holds a large, passive, minority stake in another company. Such holdings generally help to protect management from shareholder pressures because implicit in a cross-shareholding arrangement is the guarantee that the owner of the shares will support management on all voting issues. In effect, these shareholdings act as takeover defenses.

Similarly, the presence of a sizable affiliated stockholder (such as an individual, family trust, endowment, or private equity fund) can shield a company from the effects of voting by outside shareholders. As an example, a US consumer goods company has an affiliated charitable foundation that owns more than 20% of the outstanding shares. The company also has a provision in its corporate charter requiring that any changes to the charter must be approved by two-thirds of outstanding shares. As a result, it is virtually impossible for any measure to pass without the support of the foundation. The interests of the foundation thus conflict with the interests of the overall shareholder base. In effect, this single minority shareholder most likely holds the power to block the votes of the majority.

Analysts should note that market context is important in assessing the potential effects of affiliated stockholders. In certain countries, the presence of such shareholders is common, viewed by local market participants as a means of enhancing stability, strengthening the relationship between companies and their business partners, and fostering a long-term perspective by protecting the company against hostile takeover bids.

A final aspect of investor composition that has become increasingly relevant relates to activist shareholders, which we discussed earlier. The presence of activist shareholders can meaningfully and rapidly change the investment thesis for a company. Experienced activists, together with short-term-oriented investors who follow their activities, can create substantial turnover in a company's shareholder composition in a short amount of time. This is because an activist often serves as a catalyst for new strategic alternatives at a company and can attract new investors and/or arbitrageurs.

9.5 Strength of Shareholders' Rights

Within a framework of regional regulations and corporate governance codes, analysts may want to understand whether the shareholder rights of a particular company are strong, weak, or average compared with other companies. For example, if an analyst's viewpoint includes the possibility that a company will enter into a merger transaction in the future, he or she may want to understand whether there are significant structural obstacles to transactions that are embedded in the company's charter or bylaws. Similarly, if the thesis involves an outside catalyst, such as an activist shareholder who will introduce change at the company and improve performance, the analyst must take a position on whether shareholders are sufficiently empowered to advance such a change. If it is impossible for shareholders to remove directors from the board or to convene special stockholder meetings, it will be difficult for external initiatives to be successful.

In a number of developed markets, including the United Kingdom, the Netherlands, and Japan, regulatory agencies or stock exchanges have adopted governance codes, which are standards of governance reflecting local investors' expectations with regard to disclosure, board composition, shareholder rights, and other related issues. Often, these governance codes are implemented on a "comply or explain" basis, which indicates that standards are voluntary in nature. However, any deviation from the code must be explained by the company in a public disclosure. If a company has elected to deviate from the locally accepted governance practice, the analyst may want to understand the reasoning behind the decision.

9.6 Managing Long-Term Risks

Analysts may uncover useful insights regarding how a company manages various issues, such as long-term environmental risks, management of human capital, transparency, and treatment of investors and other stakeholders. Of particular note, the academic evidence linking these and other management quality issues to share price performance remains mixed, in part because indicators of management quality are often correlated with each other and, therefore, difficult to isolate. However, poor stakeholder relations and inadequate management of long-term risks have indeed had an enormous negative effect on share value in certain instances. Therefore, analysts may consider the assessment of such issues to be a useful component of their overall risk assessment of the company.

One way to evaluate management quality issues is to assess whether the company demonstrates a persistent pattern of wrongdoings, fines, regulatory penalties, investigations, and the like. A notable example is Toshiba Corp., beginning in 2015 when an investigation revealed that company managers had manufactured nearly $2 billion in profits since 2008. Poor internal controls allowed the fraudulent accounting to remain undetected for seven years. Over the next several years, Toshiba experienced the massive write-down of its Westinghouse Electric Co. unit, clashes with its auditors, negative equity, and legal disputes. Toshiba's dividend was withdrawn, Westinghouse filed for bankruptcy protection, and the company was forced to sell its profitable semiconductor unit along with other business units. Ultimately, the company's shares were moved to the second section of the Tokyo Stock Exchange, triggering forced sales from funds that track the first section (TOPIX) or the Nikkei 225 indexes. Even in the absence of circumstances as extreme as these, poor management of stakeholder interests can have a significant effect on company operations, reputation, and valuation.

9.7 Summary of Analyst Considerations

The analysis of corporate governance, stakeholder management, and other non-financial (often termed "extra-financial") considerations is inherently a subjective exercise. Governance practices that may raise the risk profile of one company may be perfectly acceptable in a different context, depending on geographic norms, mitigating circumstances, or the investor's risk tolerance and investment thesis. In this section, we have provided a basic framework for investors interested in uncovering incremental insights about a company by analyzing its governance standards, practices, and risks.

EXAMPLE 10

Analyst Considerations

1 Which of the following *best* describes dual-class share structures?
   - **A** Dual-class share structures can be easily changed over time.
   - **B** Company insiders can maintain significant power over the organization.
   - **C** Conflicts of interest between management and stakeholder groups are less likely than with single-share structures.

2 An investment analyst would likely be *most* concerned with an executive remuneration plan that:
   - **A** varies each year.
   - **B** is consistent with a company's competitors.
   - **C** is cash-based only, without an equity component.

3 Which of the following *best* describes activist shareholders? Activist shareholders:
   - **A** help stabilize a company's strategic direction.
   - **B** have little effect on the company's long-term investors.
   - **C** can alter the composition of a company's shareholder base.

Solution to 1:

B is correct. Under dual-class share systems, company founders or insiders may control board elections, strategic decisions, and other significant voting matters. A is incorrect because dual-share systems are virtually impossible to dismantle once adopted. C is incorrect because conflicts of interest between management and stakeholders are more likely than with single-share structures because of the potential control element under dual systems.

Solution to 2:

C is correct. If an executive remuneration plan offers cash only, the incentives between management and investors and other stakeholders may be misaligned. A is incorrect because a plan that varies over time would typically be of less concern to an analyst compared with one that did not change. B is incorrect because an analyst would likely be concerned if a company's executives were excessively compensated relative to competitors.

Solution to 3:

C is correct. The presence of activist shareholders can create substantial turnover in the company's shareholder composition. A is incorrect because the presence of activist shareholders can materially change a company's strategic direction. B is incorrect because long-term investors in a company need to consider how activist shareholders affect the company.

10

ESG CONSIDERATIONS FOR INVESTORS AND ANALYSTS

j describe environmental and social considerations in investment analysis

k describe how environmental, social, and governance factors may be used in investment analysis

10.1 Introduction to Environmental, Social, and Governance issues

The inclusion of governance factors in investment analysis has evolved considerably. Management and accountability structures are relatively transparent, and information regarding them is widely available. Also, the risks of poor corporate governance have long been understood by analysts and shareholders. In contrast, the practice of systematically considering environmental and social factors, which collectively with governance form the commonly used acronym "ESG," has evolved more slowly. Issues driving the inclusion of environmental and social information in the investment process include scarcity of natural resources, physical impacts of climate change, global economic and demographic trends, diversity and inclusion, treatment of workers, and the rise of social media. A non- exhaustive list of ESG issues are shown in Exhibit 1. Typically, a smaller set of ESG factors are material for each company, influenced by the business segments it operates in as well as its geography of operation.

Exhibit 1 Example of ESG Factors

Environmental Issues

- Climate change and carbon emissions
- Air and water pollution
- Biodiversity
- Deforestation
- Energy efficiency
- Waste management
- Water scarcity

Social Issues

- Customer satisfaction & product responsibility
- Data security and privacy
- Gender and diversity
- Occupational health & safety
- Treatment of workers
- Community relations & charitable activities
- Human rights
- Labor standards

Governance Issues

- Board composition (independence & diversity)
- Audit committee structure
- Bribery and corruption
- Executive compensation
- Shareholder rights
- Lobbying & political contributions
- Whistleblower schemes

ESG factors were once regarded as intangible or qualitative information. Increased corporate disclosures and refinements in the identification and analysis of ESG factors have resulted in increasingly quantifiable and decision-useful information. Social issues, such as human capital management, are less quantifiable; however, they are not necessarily more difficult to integrate as qualitative factors into investment analysis. The process of reflecting quantitative ESG-related information and data—both quantitative and qualitative—in financial valuation continues to evolve.

10.2 ESG Investment Strategies

There is a lack of consensus on several ESG strategy-related terms used in the investment community. For the purposes of this reading, we define ESG investment strategies as the following:

Responsible investing is the broadest (umbrella) term used to describe investment strategies that incorporate environmental, social, and governance (ESG) factors into their approaches. Overall, responsible investing includes the following:

- ESG integration
- Socially Responsible Investing (SRI)
- Thematic investing
- Impact investing

Sustainable investing is a term used in a similar context to responsible investing, but its key focus is on factoring in sustainability issues while investing.

ESG integration generally refers to the careful consideration of *material* ESG factors in the investment analysis and portfolio construction processes. **ESG investing** is often used interchangeably with ESG integration. A material factor is one whose omission or misstatement could influence an investment decision and ultimately the product's investment objectives. The materiality of ESG factors in investment analysis, particularly environmental and social factors, often differs meaningfully among sectors. An ESG factor is considered material when it has a potential to impact a company's ability to generate sustainable returns in the long term. For example, environmental factors, such as carbon emissions and water usage, will likely be material for utilities or mining companies yet are relatively inconsequential for financial institutions.

Socially responsible investing (SRI) is a related term that tends to have multiple meanings, creating confusion among investors. Socially responsible investing has traditionally referred to the practice of excluding investments in companies or industries, such as controversial weapons or tobacco, that deviate from an investor's beliefs, either moral or faith-based. The term has evolved to include investment objectives that promote positive environmental or social attributes, often by investing in companies with favorable environmental or social profiles.

Thematic investing refers to investment in themes or assets specifically related to ESG factors. This approach is often based on needs arising from economic or social trends.

Impact investing refers to investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.

GREEN FINANCE

Green finance is a responsible investing approach that utilizes financial instruments that support a green economy. According to the Organisation for Economic Co-Operation and Development (OECD), green finance relates to "achieving economic growth while reducing pollution and greenhouse gas emissions, minimising waste and improving efficiency in the use of natural resources."[5] As with other previously mentioned ESG terms, there are several definitions for green finance in practice. The primary investment vehicles used in green finance are **green bonds**, in which issuers earmark the proceeds toward environmental-related projects. Green bond issuance has grown significantly. According to the Climate Bonds Initiative, total worldwide green bond issuance for 2019 totaled USD255 billion, compared to USD171.1 billion for 2018. The United States continues to be the largest green bond market in the world, with China being the second largest, accounting for about 12% of global green bond issuance in 2019. China's first green bond was issued in July 2015 as the country prioritized environmental projects to address air and water pollution issues resulting from its rapid growth. In addition to green bonds, **sustainability linked loans** (including **green loans**) are also gradually being employed across the globe. Sustainability linked loans are any types of loan instruments and/or contingent facilities (such as bonding lines, guarantee lines, or letters of credit) that incentivize the borrower's achievement of ambitious, predetermined sustainability performance objectives.

---

[5] www.oecd-ilibrary.org/environment/green-finance-and-investment_24090344

10.3 ESG Investment Approaches

ESG investment approaches range from *value*-based to *values*-based. The objective of a *value*-based approach is to mitigate risks and identify opportunities by analyzing and incorporating material ESG considerations in addition to traditional financial metrics. Conversely, the objective of a *values*-based approach is to express the moral or faith-based beliefs of an investor. Between the value-based and the values-based approaches lie a continuum of approaches that strives for value creation through values investing.

There is a lack of consensus in the investment community on terminology to classify specific ESG investment approaches. For the purposes of this reading, we define six generic ESG investment approaches:

- Negative screening
- Positive screening
- ESG integration
- Thematic investing
- Engagement/active ownership
- Impact investing

Negative screening is the exclusion of certain sectors, companies, or practices from a fund or portfolio based on specific ESG criteria. Examples of a negative screen include excluding the fossil fuel extraction/production sector or excluding companies that deviate from globally accepted standards, in areas such as human rights or environmental management, from a portfolio. Many of these negative screens use a specific set of standards, such as the UN Global Compact's Ten Principles on human rights, labor, the environment, and corruption.

In contrast to negative screening, the **positive screening** approach focuses on investments that manage their material ESG risks well relative to industry peers. Positive screening, typically implemented through an ESG ranking or scoring approach, is the inclusion of certain sectors, companies, or practices in a fund or portfolio based on specific ESG criteria. For example, positive screening may include seeking companies that promote employee rights, exhibit diversity in the workplace and in board rooms, or perform well in customer safety. While seeking to invest in sectors, companies, or projects that demonstrate best-in-class ESG performance relative to industry peers, this approach does not exclude any sectors. Instead, it focuses on finding the companies within each sector that perform best on the chosen criteria. This approach typically maintains sector weightings comparable to a relative benchmark index to avoid excessively overweighting or underweighting risk exposures.

ESG integration entails a systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions for the purpose of achieving the product's stated investment objectives. There is an explicit inclusion of ESG factors into traditional financial analysis of individual stocks or bonds, such as inputs into cash flow forecasts, credit/default risk forecasts, and/or cost-of-capital estimates. The focus of ESG integration is to identify risks and opportunities arising from material ESG factors and to determine whether a company is properly managing these or not.

Thematic investing is investment in themes or assets specifically related to ESG factors, such as clean energy, green technology, sustainable agriculture, gender diversity, or affordable housing. This approach is often based on needs arising from economic or social trends. Two common investment themes focus on increased demand for energy and water, as well as the availability of alternative sources of each. Global economic development has raised the demand for energy at the same time as increased greenhouse gas emissions are widely believed to negatively affect the earth's climate. Similarly, rising global living standards and industrial needs have created a greater demand for water along with the need to prevent drought or increase access to clean drinking water in certain regions of the world. While these themes are based on trends related to environmental issues, social issues—such as access to affordable health care and nutrition, especially in the poorest countries in the world—are also of great interest to thematic investors.

Engagement/active ownership is the use of shareholder or bondholder power to influence corporate behavior through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is directed by ESG guidelines. Engagement/active ownership seeks to achieve targeted social or environmental objectives along with measurable financial returns. Engagement/active ownership can be executed through various asset classes and investment vehicles and often through direct transactions, such as venture capital investing. Collaborative engagement initiatives entail multiple investors collectively engaging with company management to influence positive action in managing their material ESG risks. Climate Action 100+, backed by more than 450 investors with over USD40 trillion in assets collectively under management (as of July 2020), is one such widely supported initiative that aims to influence the world's largest corporate greenhouse gas emitters to take necessary action on climate change. Other key ESG issues on which investors frequently engage are air pollution, plastic waste management, human and labor rights in the supply chain, and executive remuneration.

Impact investing is a relatively smaller segment of the broader sustainable and responsible investing market. Impact investments are investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.[6] An example would include investing in products or services that help achieve one (or more) of the 17 Sustainable Development Goals (SDGs) launched by the United Nations in 2015, such as "SDG 6: Clean Water and Sanitation—Ensure availability and sustainable management of water and sanitation for all" and "SDG 11: Sustainable Cities and Communities—Make cities and human settlements inclusive, safe, resilient and sustainable."

A summary of the six generic ESG Investment approaches and mapping to other classifications is shown in Exhibit 2.

---

[6] https://thegiin.org/impact-investing/

Exhibit 2 ESG Investment Approaches

		Mapping to Other Classifications	
ESG Investment Approach	**Description**	***Financial Analysts Journal***	***Global Sustainable Investment Review***
Negative screening	Excluding companies or sectors based on business activities or environmental or social concerns	Negative screening	Negative/exclusionary screening Norms-based screening
Positive screening	Including sectors or companies based on specific ESG criteria, typically ESG performance relative to industry peers	Positive screening Relative/best-in-class screening	Positive/Best-in-class screening
ESG integration	Systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions for the purpose of achieving or exceeding the product's stated investment objectives	Full integration Overlay/Portfolio tilt Risk factor/Risk premium	ESG integration
Thematic investing	Investing in themes or assets related to ESG factors	Thematic investing	Sustainability-themed investing
Engagement/active ownership	Using shareholder power to influence corporate behavior to achieve targeted ESG objectives along with financial returns	Engagement/Active ownership	Corporate engagement/Shareholder action
Impact investing	Investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return	N/A	Impact/Community investing

Notes: For information on the *Financial Analysts Journal* column under "Mapping to Other Classifications," see www.tandfonline.com/doi/full/10.2469/faj.v74.n3.2. For the *Global Sustainable Investment Review* column, see http://www.gsi-alliance.org/wp-content/uploads/2019/06/GSIR_Review2018F.pdf.

10.4 Catalysts for Growth in ESG Investing

ESG considerations have become increasingly relevant for two key reasons. First, ESG issues are having more material financial impacts on a company's fair value. Many investors have suffered substantial losses due to mismanagement of ESG issues by corporations, which resulted in environmental disasters, social controversies, or governance deficiencies. Second, a greater number of younger investors are increasingly demanding that their inherited wealth or their pension contributions be managed using investment strategies that systematically consider material ESG risks as well as negative environmental and societal impacts of their portfolio investments.

Historically, environmental and social issues, such as climate change, air pollution, and societal impacts of a company's products and services, have been treated as negative externalities—ones whose costs are not borne by the concerned company.

However, increased stakeholder awareness and strengthening regulations are internalizing environmental and societal costs onto the company's income statement either explicitly or implicitly by responsible investors.

At a *macro level*, such environmental risks as physical impacts of climate change—floods, droughts, wildfires, air pollution—are impacting our day-to-day lives both more frequently and on a larger scale than ever imagined. Wildfires in California, the Amazon forest, and Australia have caused loss of human life and billions of dollars in damages. Major cities, like Cape Town and Chennai, have had severe water shortages. The 2020 global slowdown due to Covid-19 demonstrated that such social factors as inequality, access to health care, and vulnerable labor can also have significant impacts on growth. While macro-level environmental and social risks have received widespread attention only since 2010, investors have been evaluating macro-level governance risks, such as quality of institutions and corruption, into their analysis since the 1990s.

At a *micro level*, there have been numerous examples of stakeholders penalizing companies that are unable to manage material ESG risks well. Facebook's Cambridge Analytica scandal, in which the personal data of over 80 million users were allegedly shared without consent and used in influencing voters, not only led to one of the largest US government fines in the technology sector (USD5 billion), but it also made a significant dent on user trust. Other examples of ESG risks materializing are shared later in this reading. As prudent fiduciaries, investors are paying more attention to these factors while constructing portfolios and engaging with company management.

Another investor development that has supported ESG growth is the adoption of more sophisticated views about sustainable growth and its effect on investment performance. To this end, some large institutional asset owners embrace the concept of being "universal owners." **Universal owners** are long-term investors, such as pension funds, that have significant assets invested in globally diversified portfolios. Given their size and scope, the investment portfolios of universal owners are linked to economic growth and unavoidably exposed to costs resulting from external factors or externalities, such as environmental damages. Some universal owners strive to positively influence the way companies conduct business to minimize exposure to ESG-related costs because both their funds' long-term performance and the interests of their beneficiaries are at stake. Institutions that adhere to the universal owner philosophy believe that sustainable global economic growth is essential to successful investment performance.

10.5 ESG Market Overview

Reflecting the growth of ESG-related information in the marketplace, the amount of global assets under management (AUM) dedicated to the consideration of ESG factors in portfolio selection and management has increased substantially. According to the Global Sustainable Investment Alliance (GSIA), a collaboration of organizations dedicated to advancing sustainable investing in the financial markets, nearly USD31 trillion of AUM were dedicated toward sustainable investment mandates at the start of 2018, a 34% increase in two years. Europe (49%) and the United States (39%) accounted for the vast majority of these AUM. Determining the true size of the ESG investment universe is difficult, however, because managers and investors define and implement sustainable and ESG mandates in many different ways. There are often differences regarding which ESG factors should be considered—as well as how they are considered—in the investment analysis and portfolio construction processes.

The increased interest in sustainable investing has led to increased corporate disclosures of ESG issues, as well as a growing number of companies that collect and analyze ESG data. In addition to the GSIA, several organizations have been formed to monitor and advance the mission of sustainable investing. For example, the Global Reporting Initiative (GRI), a non-profit organization formed in 1997, produces a sustainability reporting framework that measures and reports sustainability-related issues and performance. In 2006, the United Nations and a consortium of institutional investors launched the Principles for Responsible Investment (PRI) Initiative to support investors committed to including ESG issues in their investment decision-making and ownership practices. In 2011, the non-profit Sustainability Accounting Standards Board (SASB) was formed to develop sustainability accounting standards for companies when disclosing material ESG information. To help educate investors, CFA Institute and PRI published "Guidance and Case Studies for ESG Integration: Equities and Fixed Income" in 2018. In addition to these key organizations, several others exist that promote the advancement of sustainable investing.

10.6 ESG Factors in Investment Analysis

Environmental factors that are generally considered material in investment analysis include natural resource management, pollution prevention, water conservation, energy efficiency and reduced emissions, the existence of carbon assets, and adherence to environmental safety and regulatory standards. A specific concern among investors of energy companies is the existence of "stranded assets," which are carbon-intensive assets at risk of no longer being economically viable because of changes in regulation or investor sentiment. Analysts may find it difficult to assess potentially significant financial risks of energy companies because of limited information on the existence of these companies' carbon assets, as well as the difficulty in determining political and regulatory risks. Material environmental effects can arise from strategic or operational decisions based on inadequate governance processes or errors in judgment. For example, oil spills, industrial waste contamination events, and local resource depletion can result from poor environmental standards, breaches in safety standards, or unsustainable business models. Such events can be costly in terms of regulatory fines, litigation, clean-up costs, reputational risk, and resource management.

TOXIC EMISSIONS AND WASTE AS AN ENVIRONMENTAL RISK

Historically, environmental issues, such as toxic emissions and waste, have been treated as externalities and thus not fully provisioned for in a company's financial reporting. However, with growing awareness among all stakeholders, including the regulators, companies may now face financial liabilities associated with pollution, contamination, and the emission of toxic or carcinogenic substances and hence are required to manage these risks well. Gross mismanagement of these risks could not only lead to a permanent loss of a company's license to operate but also to other severe financial penalties.

In one notable example, the 2019 collapse of Dam I of the Córrego do Feijão Mine in Brumadinho, Brazil, resulted in the spillage of millions of tons of non-toxic mud. The plant and surrounding communities experienced the loss of 270 lives as well as the contamination of the nearby Paraopeba River. The mine was owned and operated by Vale, which has since been accused of hiding information about the dam's instability for many years to avoid hurting the company's reputation. Several employees from the company, including its ex-CEO, and its auditor, TÜV SÜD, were charged with murder and environmental crimes. Vale was also fined millions of dollars in addition to bearing clean-up costs.

Social factors considered in ESG implementation generally pertain to the management of the human capital of a business. These include human rights and welfare concerns in the workplace; data privacy and security; access to affordable health care products; and, in some cases, community impact. Staff turnover, worker training and safety, employee morale, ethics policies, employee diversity, and supply chain management can all affect a company's ability to sustain its competitive advantage. In addition, minimizing social risks can lower a company's costs (e.g., through higher employee productivity, lower employee turnover, and reduced litigation potential) and reduce its reputational risk.

DATA PRIVACY AND SECURITY AS A SOCIAL RISK

Data privacy and security focuses on how companies gather, use, and secure personally identifiable information and other meta-data collected from individuals. In some industries, such as internet software and services, this includes managing the risks associated with government requests that may result in violations of civil and political rights.

With the proliferation of the internet across the globe, more and more services are offered online. Consumers of such services are leaving a large digital footprint behind, often unknowingly. Some of this information may be personally identifiable in nature, leaving users vulnerable in case of theft or misuse. As per the "2019 Cost of a Data Breach Report" released by IBM and Ponemon Institute, the average cost of data breach is USD3.9 million, with cost per record lost being USD150.

Given the large amounts of sensitive data being managed by some of the largest internet and financial services companies today, mismanagement of data privacy and security risk can have materially damaging consequences both for a company's business model as well as financial performance. As an example, lax cybersecurity measures at Equifax, Inc., led to a data breach and the theft of identity and financial data belonging to more than 140 million US citizens in 2017. Equifax has since incurred hundreds of millions of dollars in expenses resulting from the breach and has faced numerous lawsuits and investigations.

As mentioned earlier, governance factors have long been recognized in investment analysis. Many performance indicators can help evaluate risks arising from governance issues, such as ownership structure, board independence and composition, and compensation. Although several governance factors may apply across industries and geographic regions, other factors are unique, such as systemic risk management for financial services companies. Wider adoption of responsible investing strategies enhances mainstream governance analysis via the addition of such factors as management oversight of environmental and social issues and tax transparency.

CORPORATE GOVERNANCE: AN ALL-ENCOMPASSING RISK

An example of investor loss related to corporate governance occurred in 2015 at German automaker Volkswagen. Volkswagen's "Dieselgate" scandal involved use of software to manipulate diesel vehicle emissions of more than 11 million cars globally. Specifically, many investors believed that inadequate governance oversight at Volkswagen permitted its diesel cars to pass emissions tests yet emit unlawful amounts of nitrogen oxide. Many investors attributed their losses to internal audit and compliance shortcomings as well as a lack of independence in Volkswagen's board of directors. The company has already incurred several billion dollars of civil and criminal penalties, and the scandal continues to pose significant legal and financial risks for the company.

One area of debate has been whether the consideration of ESG factors is consistent with fiduciary duty—particularly when overseeing and managing pension fund assets. Pension fund regulation regarding ESG considerations varies globally, however. PRI and the United Nations Environment Programme Finance Initiative (UNEP FI) promote the belief that ESG integration is a key part of investment analysis: "Investors that fail to incorporate ESG issues are failing their fiduciary duties and are increasingly likely to be subject to legal challenge."

EXAMPLE 11

ESG Investment Approach

The ESG investment approach that is *most* associated with excluding certain sectors or companies is:

A thematic investing.

B negative screening.

C positive screening.

Solution:

B is correct. Negative screening entails excluding certain companies or sectors, such as fossil fuel extraction, from a portfolio. A is incorrect because thematic investing typically focuses on investing in companies within a specific sector or following a specific theme, such as energy efficiency or climate change, as opposed to merely excluding a set of companies or industries from a portfolio. Likewise, C is incorrect because positive screening focuses on including companies that rank (or score) most favorably compared to their peers with regard to ESG factors.

SUMMARY

The investment community has increasingly recognized the importance of corporate governance as well as environmental and social considerations. Although practices concerning corporate governance (and ESG overall) will undoubtedly continue to evolve, investment analysts who have a good understanding of these concepts can better appreciate the implications of ESG considerations in investment decision making. The core concepts covered in this reading are as follows:

- Corporate governance can be defined as a system of controls and procedures by which individual companies are managed.
- There are many systems of corporate governance, most reflecting the influences of either shareholder theory or stakeholder theory, or both. Current trends, however, point to increasing convergence.
- A corporation's governance system is influenced by several stakeholder groups, and the interests of the groups often diverge or conflict.
- The primary stakeholder groups of a corporation consist of shareholders, creditors, managers and employees, the board of directors, customers, suppliers, and government/regulators.
- A principal–agent relationship (or agency relationship) entails a principal hiring an agent to perform a particular task or service. In a corporate structure, such relationships often lead to conflicts among various stakeholders.
- Stakeholder management involves identifying, prioritizing, and understanding the interests of stakeholder groups and on that basis managing the company's relationships with stakeholders. The framework of corporate governance and stakeholder management reflects a legal, contractual, organizational, and governmental infrastructure.

- Mechanisms of stakeholder management may include general meetings, a board of directors, the audit function, company reporting and transparency, related-party transactions, remuneration policies (including say on pay), and other mechanisms to manage the company's relationship with its creditors, employees, customers, suppliers, and regulators.
- A board of directors is the central pillar of the governance structure, serves as the link between shareholders and managers, and acts as the shareholders' internal monitoring tool within the company.
- The structure and composition of a board of directors vary across countries and companies. The number of directors may vary, and the board typically includes a mix of expertise levels, backgrounds, and competencies.
- Executive (internal) directors are employed by the company and are typically members of senior management. Non-executive (external) directors have limited involvement in daily operations but serve an important oversight role.
- Two primary duties of a board of directors are duty of care and duty of loyalty.
- A company's board of directors typically has several committees that are responsible for specific functions and report to the board. Although the types of committees may vary across organization, the most common are the audit committee, governance committee, remuneration (compensation) committee, nomination committee, risk committee, and investment committee.
- Stakeholder relationships and corporate governance are continually shaped and influenced by a variety of market and non-market factors.
- Shareholder engagement by a company can provide benefits that include building support against short-term activist investors, countering negative recommendations from proxy advisory firms, and receiving greater support for management's position.
- Shareholder activism encompasses a range of strategies that may be used by shareholders when seeking to compel a company to act in a desired manner.
- From a corporation's perspective, risks of poor governance include weak control systems; ineffective decision making; and legal, regulatory, reputational, and default risk. Benefits include better operational efficiency, control, and operating and financial performance, as well as lower default risk (or cost of debt).
- Key analyst considerations in corporate governance and stakeholder management include economic ownership and voting control, board of directors representation, remuneration and company performance, investor composition, strength of shareholders' rights, and the management of long-term risks.
- ESG investment approaches range from *value*-based to *values*-based. There are six broad ESG investment approaches: Negative screening, Positive screening, ESG integration, Thematic investing, Engagement/active ownership, and Impact investing.
- Historically, environmental and social issues, such as climate change, air pollution, and societal impacts of a company's products and services, have been treated as negative externalities. However, increased stakeholder awareness and strengthening regulations are internalizing environmental and societal costs onto the company's income statement by responsible investors.

PRACTICE PROBLEMS

1 Corporate governance:

- **A** complies with a set of global standards.
- **B** is independent of both shareholder theory and stakeholder theory.
- **C** seeks to minimize and manage conflicting interests between insiders and external shareholders.

2 Which group of company stakeholders would be *least* affected if the firm's financial position weakens?

- **A** Suppliers
- **B** Customers
- **C** Managers and employees

3 Which of the following represents a principal–agent conflict between shareholders and management?

- **A** Risk tolerance
- **B** Multiple share classes
- **C** Accounting and reporting practices

4 Which of the following issues discussed at a shareholders' general meeting would *most likely* require only a simple majority vote for approval?

- **A** Voting on a merger
- **B** Election of directors
- **C** Amendments to bylaws

5 Which of the following statements regarding stakeholder management is *most* accurate?

- **A** Company management ensures compliance with all applicable laws and regulations.
- **B** Directors are excluded from voting on transactions in which they hold material interest.
- **C** The use of variable incentive plans in executive remuneration is decreasing.

6 Which of the following represents a responsibility of a company's board of directors?

- **A** Implementation of strategy
- **B** Enterprise risk management
- **C** Considering the interests of shareholders only

7 Which of the following statements about non-market factors in corporate governance is *most* accurate?

- **A** Stakeholders can spread information quickly and shape public opinion.
- **B** A civil law system offers better protection of shareholder interests than does a common law system.
- **C** Vendors providing corporate governance services have limited influence on corporate governance practices.

8 Which of the following statements regarding corporate shareholders is *most* accurate?

- **A** Cross-shareholdings help promote corporate mergers.

© 2017 CFA Institute. All rights reserved.

- **B** Dual-class structures are used to align economic ownership with control.
- **C** Affiliated shareholders can protect a company against hostile takeover bids.

9 Which of the following statements about environmental, social, and governance (ESG) in investment analysis is correct?

- **A** ESG factors are strictly intangible in nature.
- **B** ESG terminology is easily distinguishable among investors.
- **C** Environmental and social factors have been adopted in investment analysis more slowly than governance factors.

10 Which of the following statements regarding ESG investment approaches is *most accurate*?

- **A** Negative screening is the most commonly applied method.
- **B** Thematic investing considers multiple factors.
- **C** Positive screening excludes industries with unfavorable ESG aspects.

SOLUTIONS

1 C is correct. Corporate governance is the arrangement of checks, balances, and incentives a company needs to minimize and manage the conflicting interests between insiders and external shareholders.

2 B is correct. Compared with other stakeholder groups, customers tend to be less concerned with, and affected by, a company's financial performance.

3 A is correct. Shareholder and manager interests can diverge with respect to risk tolerance. In some cases, shareholders with diversified investment portfolios can have a fairly high risk tolerance because specific company risk can be diversified away. Managers are typically more risk averse in their corporate decision making to better protect their employment status.

4 B is correct. The election of directors is considered an ordinary resolution and, therefore, requires only a simple majority of votes to be passed.

5 B is correct. Often, policies on related-party transactions require that such transactions or matters be voted on by the board (or shareholders), excluding the director holding the interest.

6 B is correct. The board typically ensures that the company has an appropriate enterprise risk management system in place.

7 A is correct. Social media has become a powerful tool for stakeholders to instantly broadcast information with little cost or effort and to compete with company management in influencing public sentiment.

8 C is correct. The presence of a sizable affiliated stockholder (such as an individual, family trust, endowment, or private equity fund) can shield a company from the effects of voting by outside shareholders.

9 C is correct. The risks of poor corporate governance have long been understood by analysts and shareholders. In contrast, the practice of considering environmental and social factors has been slower to take hold.

10 A is correct. Negative screening, which refers to the practice of excluding certain sectors, companies, or practices that violate accepted standards in such areas as human rights or environmental concerns, is the most common ESG investment approach.

READING 28

Uses of Capital

by John D. Stowe, PhD, CFA, and Jacques R. Gagné, FSA, CFA, CIPM

John D. Stowe, PhD, CFA, is at Ohio University (USA). Jacques R. Gagné, FSA, CFA, CIPM, is at ENAP (Canada).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the capital allocation process and basic principles of capital allocation;
☐	**b.** demonstrate the use of net present value (NPV) and internal rate of return (IRR) in allocating capital and describe the advantages and disadvantages of each method;
☐	**c.** describe expected relations among a company's investments, company value, and share price;
☐	**d.** describe types of real options relevant to capital investment;
☐	**e.** describe common capital allocation pitfalls.

INTRODUCTION

1

To achieve profitability and reach sustainability, a company must use its resources effectively and make long-term investments that increase revenues and profits. Investment opportunities returning long-term benefits and future cash flows greater than their associated cost to fund generate value for companies and corresponding wealth for their owners. Decisions on where and when to make long-term capital investments are, therefore, key management decisions central to a company's operating success and longevity. Understanding how companies allocate their capital among competing priorities and their resulting portfolios of investment activity is a fundamental area of knowledge for financial analysts for many reasons.

The allocation and investment of capital are important corporate activities. Capital investments (also referred to here as capital projects) are investments with a life of one year or longer, and they make up the long-term asset portion of the balance sheet. They can be so large that sound capital allocation decisions ultimately decide the future success of many corporations. Capital investments also describe a company's future prospects better than its working capital or capital structure, which

© 2021 CFA Institute. All rights reserved.

are intangible and often similar for companies. Analysts may attempt to estimate the process, within reason, for companies that are not too complex and, in doing so, better evaluate corporate decisions that extend to financing and other related activity. Analysts may also be able to appraise the quality of the company's capital allocation process—for example, on the basis of whether the company has an accounting focus or an economic focus. In doing so, analysts derive insights into how the company is creating value for investors.

This reading is organized as follows: Section 2 presents the steps taken by companies in a typical capital allocation process and the basic principles of capital allocation. In Section 3, we introduce basic investment decision criteria, and in Section 4, we discuss capital allocation options—known as real options—that allow their holders, in this case companies, to make decisions in the future that alter the value of their capital investment decisions made today. Section 5 covers a discussion of the common capital allocation pitfalls often made by companies, and a summary concludes the reading.

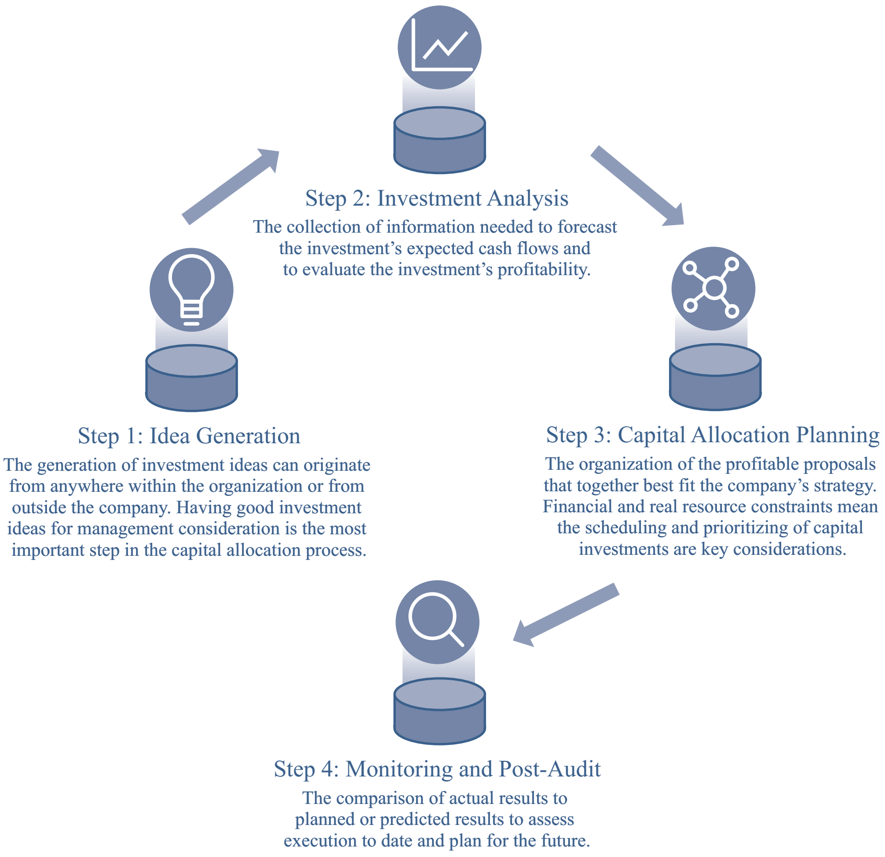
2

THE CAPITAL ALLOCATION PROCESS

a. describe the capital allocation process and basic principles of capital allocation

The **capital allocation process**—the process used by a company to make capital investment decisions—depends on factors specific to the company, such as management tenure and experience, the size and complexity of the capital investment being evaluated, and the size of the organization. The typical steps taken by a company in the capital allocation process are shown in Exhibit 1.

Exhibit 1 Steps in the Capital Allocation Process



Step 1: Idea Generation

The generation of investment ideas can originate from anywhere within the organization or from outside the company. Having good investment ideas for management consideration is the most important step in the capital allocation process.

Step 2: Investment Analysis

The collection of information needed to forecast the investment's expected cash flows and to evaluate the investment's profitability.

Step 3: Capital Allocation Planning

The organization of the profitable proposals that together best fit the company's strategy. Financial and real resource constraints mean the scheduling and prioritizing of capital investments are key considerations.

Step 4: Monitoring and Post-Audit

The comparison of actual results to planned or predicted results to assess execution to date and plan for the future.

Following the generation of investment ideas, an important part of the investment analysis is to accurately forecast the timing, duration, and volatility of each investment's expected cash flows in addition to their probability of occurrence. Capital allocation planning then involves the selection and prioritization of profitable investment opportunities that, when considered together, are most value enhancing for the firm's strategy. Some projects that look attractive when considered in isolation may be undesirable strategically. Opportunities that are not profitable or fail to generate returns sufficient to cover their associated cost of funding should not be pursued. In these cases, the firm is better off returning cash to shareholders as dividends or share repurchases.

Post-auditing capital projects and comparing an investment's realized revenues, expenses, and cash flows with the predictions are important for several reasons. First, they help monitor the forecasts and analysis that underlie the capital allocation process. Systematic errors, such as overly optimistic forecasts, become apparent. Second, they help improve business operations. If sales or costs are out of line, this process will focus attention on bringing performance closer to expectations if at all possible. Finally, monitoring and post-auditing recent capital investments will produce concrete ideas for future investments. Managers can decide to invest more heavily in profitable areas and scale down or cancel investments in areas that are disappointing.

Planning for capital investments can be very complex, often involving many persons inside and outside the company. Information across multiple disciplines such as marketing, technology, engineering, regulation, taxation, finance, production, and behavioral issues must be systematically gathered and evaluated by the company. Management's authority to make capital decisions depends on the size and complexity of the investment in question. Lower-level managers may have discretion to make decisions that involve less than a given amount of money or that do not exceed a given capital allocation. Larger and more complex decisions are reserved for top management, and some are so significant that the company's board of directors ultimately has the decision-making authority.

Like other business activities, the capital allocation activity is a cost–benefit exercise for the company. At the margin, the benefits from the improved decision making should exceed the costs of the capital allocation efforts.

Companies often put capital allocation projects into some basic categories for analysis. These classifications are summarized in Exhibit 2.

Exhibit 2 Types of Capital Projects

01.

Replacement Projects

Investment projects that maintain the existing size of the business.

02.

Expansion Projects

Investment projects that expand business size and often involve greater uncertainty and management consideration than replacement projects.

03.

New Products and Services

Investments in new products or services involving more stakeholders and higher degrees of risk (uncertainty).

04.

Regulatory, Safety, and Environmental Projects

Investment projects required by a third party, such as the government or an insurance company, to meet specified standards. These may not generate revenues and might not be otherwise undertaken by a company maximizing its own interests.

05.

Other

Investment projects, such as management pet projects and high-risk investments, which escape, or fall outside, normal project analyses.

Replacement projects are among the easier capital allocation decisions for management. If a piece of equipment breaks down or wears out, whether to replace it may not require in-depth analysis. If the expenditure is modest and if not investing has significant implications for production, operations, or sales, it would be a waste of corporate resources to overanalyze the decision. The investment should just be made.

For regulatory, safety, and environmental projects, often the company will accept the required investment and continue to operate. Occasionally, however, the cost of such projects is sufficiently high that the company would be better off to cease operating altogether or to shut down any part of the business that is related to the project.

Capital allocation relies on a few basic principles and usually involves the assumptions shown in Exhibit 3.

Exhibit 3 Capital Allocation Assumptions

1. Decisions are based on cash flows.

Decisions are not based on accounting concepts (net income or operating income), which are done on an accrual basis and subtract non-cash expenses, such as depreciation.

2. Cash flows are not accounting net income or operating income.

Accounting income reflects non-cash charges, such as depreciation, while economic income (cash flows plus change in company's market value) does not and is based on changes in the company's market value, not book value.

3. Cash flows are based on opportunity costs.

What incremental cash flows occur with the investment compared with what they would have been without the investment?

4. Cash flows are analyzed on an after-tax basis.

Decisions must reflect the impact of taxes.

5. Timing of cash flows is crucial

Detailed analysis is needed to accurately forecast the timing, duration, and volatility of expected cash flows and their probability of occurrence.

6. Financing costs are ignored.

Financing costs are reflected in the required rate of return and thus ignored in cash flows to avoid double-counting.

In ignoring financing costs, we referred to the rate used in discounting the cash flows as the "required rate of return." The required rate of return is the discount rate that investors, or suppliers of capital, require given the riskiness of the project. This discount rate is frequently called the "opportunity cost of funds" or the "cost of capital." If the company can invest elsewhere and earn a return of r or if the company can repay its sources of capital and save a cost of r, then r is the company's opportunity cost of funds. If the company cannot earn more than its opportunity cost of funds on an investment, it should not undertake that investment. Unless an investment earns more than the cost of funds from its suppliers of capital, the investment should not be undertaken. The cost of capital concept is discussed more extensively elsewhere. Regardless of what it is called, an economically sound discount rate is essential for making capital allocation decisions.

Although the principles of capital allocation are simple, they are easily confused in practice, leading to unfortunate decisions made by companies. The following are some important capital allocation concepts useful to financial analysts.

- A **sunk cost** is one that has already been incurred. One cannot change a sunk cost. Decisions made today, however, should be based on current and future cash flows and should not be affected by prior, or sunk, costs.
- An **opportunity cost** is what a resource is worth in its next-best use. For example, if a company uses some idle property, what should it record as the investment outlay: the purchase price several years ago, the current market value, or nothing? If the company invests $10 million, what is the opportunity cost? The answers to these questions are, respectively, the current market value and $10 million (which the company could invest elsewhere).

- An **incremental cash flow** is the cash flow that is realized because of an investment decision: the cash flow *with* a decision minus the cash flow *without* that decision. If opportunity costs are correctly assessed by the company, the incremental cash flows provide a sound basis for capital allocation.
- An **externality** is the effect of an investment on things other than the investment itself. Frequently, an investment affects the cash flows of other parts of the company, and these externalities can be positive or negative. A positive externality occurring within the company would be expected synergies with existing projects or business activities that result from making the investment. The production and sale of a new complimentary product or service may increase demand or lower costs for an existing product or service the company offers. **Cannibalization** is an example of a negative externality occurring within the company. Cannibalization occurs when an investment takes customers and sales away from another part of the company. If possible, companies should consider externalities in the investment decision. Often the adjustments are subjective because the company cannot get a precise estimate. Sometimes externalities occur outside the company. An investment might benefit (or harm) other companies or society at large, and yet the company is not compensated for these benefits (or charged for the costs).
- **Conventional cash flows** versus **nonconventional cash flows**: A conventional cash flow pattern is one with an initial outflow followed by a series of inflows. In a nonconventional cash flow pattern, the initial outflow is not followed by inflows only, but the cash flows can flip from being positive (inflows) to negative (outflows) again or possibly change signs several times. An investment that involved outlays (negative cash flows) for the first couple of years that were followed by positive cash flows would be considered to have a conventional pattern. If cash flows change signs once, the pattern is conventional. If cash flows change signs two or more times, the pattern is nonconventional.

  For example, a crude oil refining facility might require an initial investment outlay that produces positive cash flows for several years followed by additional investments to refurbish the facilities needed to produce additional positive cash flows. Or an investment in a nuclear power facility that generates electricity for a number of years might be followed by another cash outlay to decommission the power plant.

In assessing potential investment opportunities, several types of project interactions make the incremental cash flow analysis challenging for companies. The following are some of these interactions:

- **Independent projects** versus **mutually exclusive projects**: Independent projects are capital investments whose cash flows are independent of each other. Mutually exclusive projects compete directly with each other. For example, if Projects A and B are mutually exclusive, the company can choose to invest in A or B but cannot choose both. Sometimes management may be presented with several mutually exclusive projects and can choose only one from the group.
- **Project sequencing**: Many capital projects are sequenced over time, so that investing in a project creates the option to invest in future projects. For example, the company might invest in a project today and one year later invest in a second project if the financial results of the first project or new economic conditions are favorable. If the results of the first project or new economic conditions are not favorable, the company would not invest in the second project.
- **Unlimited funds** versus **capital rationing**: An unlimited funds environment assumes that the company can raise the funds it wants for all profitable projects simply by paying the required rate of return. Capital rationing exists when

the company has a fixed amount of funds to invest. If the company has more profitable projects than it has funds for, it must allocate the funds to achieve the maximum shareholder value subject to the funding constraints.

INVESTMENT DECISION CRITERIA

3

b demonstrate the use of net present value (NPV) and internal rate of return (IRR) in allocating capital and describe the advantages and disadvantages of each method

Several criteria can be used by companies to evaluate capital investment opportunities. The two most comprehensive measures to assess whether an investment is profitable or unprofitable are the **net present value** (NPV) and **internal rate of return** (IRR). An analyst should understand the economic logic behind each of these investment decision criteria as well as their strengths and limitations in practice.

3.1 Net Present Value

For a capital investment with one investment outlay, made initially, the net present value is the present value of the future after-tax cash flows minus the investment outlay, or

$$\text{NPV} = \sum_{t=1}^{n} \frac{\text{CF}_t}{(1+r)^t} - \text{Outlay}, \tag{1}$$

where

CF_t = After-tax cash flow at time t

r = Required rate of return for the investment

Outlay = Investment cash flow at time zero

To illustrate the net present value criterion, we will consider a simple example. Assume that Gerhardt Corporation is considering a capital investment of €50 million that will return after-tax cash flows of €16 million per year for the next four years plus another €20 million in Year 5. The company's required rate of return is 10%.

For the Gerhardt example, the NPV would be

$$\text{NPV} = \frac{16}{1.10^1} + \frac{16}{1.10^2} + \frac{16}{1.10^3} + \frac{16}{1.10^4} + \frac{20}{1.10^5} - 50.$$

$$\text{NPV} = 14.545 + 13.223 + 12.021 + 10.928 + 12.418 - 50.$$

$$\text{NPV} = 63.136 - 50 = €13.136 \text{ million}.$$

The investment has a total value, or present value of future cash flows, of €63.136 million. Since this investment can be acquired at a cost of €50 million, the company is giving up €50 million of its wealth in exchange for an investment worth €63.136 million. The investment increases the wealth of the company by a net amount of €13.136 million.

Because the NPV is the amount by which the company's wealth increases as a result of the investment, the decision rule for the NPV is as follows:

Invest if	$\text{NPV} > 0$.
Do not invest if	$\text{NPV} < 0$.

Positive-NPV investments are wealth increasing for the company and its shareholders, whereas negative-NPV investments are wealth decreasing for the company and its shareholders. In the rare case that NPV turns out to be zero, the project could be accepted because it meets the required rate of return. Keep in mind, however, that NPV analysis relies on estimated future cash flows. A zero-NPV project leaves no room for error.

Many investments have cash flow patterns in which outflows may occur not only at time zero but also at future dates. It is useful to consider the NPV to be the present value of all cash flows:

$$\text{NPV} = \text{CF}_0 + \frac{\text{CF}_1}{(1+r)^1} + \frac{\text{CF}_2}{(1+r)^2} + \cdots + \frac{\text{CF}_n}{(1+r)^n}, \text{ or}$$

$$\text{NPV} = \sum_{t=0}^{n} \frac{\text{CF}_t}{(1+r)^t}. \tag{2}$$

In Equation 2, the investment outlay, CF_0, is simply a negative cash flow. Future cash flows can also be negative.

3.2 Internal Rate of Return

The internal rate of return is one of the most frequently used concepts in capital allocation and in security analysis. For a capital investment with one investment outlay, made initially, the IRR is the discount rate that makes the present value of the future after-tax cash flows equal that investment outlay. Expressed in equation form, the IRR solves the following equation:

$$\sum_{t=1}^{n} \frac{\text{CF}_t}{(1+\text{IRR})^t} = \text{Outlay},$$

where IRR is the internal rate of return. The left-hand side of this equation is the present value of the capital investment's future cash flows, which, discounted at the IRR, equals the investment outlay. This equation will also be seen rearranged as

$$\sum_{t=1}^{n} \frac{\text{CF}_t}{(1+\text{IRR})^t} - \text{Outlay} = 0. \tag{3}$$

In this form, Equation 3 looks like the NPV equation, Equation 1, except that the discount rate is the IRR instead of r (the required rate of return). Discounted at the IRR, the NPV is equal to zero.

In the Gerhardt Corporation example, we want to find a discount rate that makes the total present value of all cash flows, the NPV, equal zero. In equation form, the IRR is the discount rate that solves the following equation:

$$-50 + \frac{16}{(1+\text{IRR})^1} + \frac{16}{(1+\text{IRR})^2} + \frac{16}{(1+\text{IRR})^3} + \frac{16}{(1+\text{IRR})^4} + \frac{20}{(1+\text{IRR})^5} = 0.$$

Algebraically, this equation would be difficult to solve. Without the use of a financial calculator, we would have to resort to trial and error, systematically choosing various discount rates until we find one, the IRR, that satisfies the equation. This process can be long and tedious. Fortunately, financial calculators can derive the IRR quickly. Even before using a calculator, however, we have some sense of what the IRR is. We previously discounted these cash flows at 10% and found the NPV to be €13.136 million. Since the NPV is positive, the IRR must be greater than 10%. If we use 20% as the discount rate, the NPV is –€0.543 million, so 20% is a little high. One might try several other discount rates until the NPV is equal to zero; this approach is illustrated in Exhibit 4.

Exhibit 4. Trial and Error Process for Finding IRR

Discount Rate (%)	NPV
10	13.136
20	−0.543
19	0.598
19.5	0.022
19.51	0.011
19.52	0.000

The IRR is 19.52%. To repeat, financial calculators and spreadsheet software can quickly calculate the IRR for us so that we do not have to go through this trial and error procedure ourselves.

The decision rule for the IRR is to invest if the IRR exceeds the required rate of return for a capital investment:

Invest if	IRR > r.
Do not invest if	IRR < r.

The required rate of return is often called the **hurdle rate**, the rate that a project's IRR must exceed for the project to be accepted by the company. In the unlikely event that the IRR is equal to r, the project is theoretically acceptable because it meets the required return. In fact, NPV equals zero when IRR equals r. In the Gerhardt example, since the IRR of 19.52% exceeds the project's required rate of return of 10%, Gerhardt should invest.

It is common to define the IRR as the discount rate that makes the present values of all cash flows sum to zero:

$$\sum_{t=0}^{n} \frac{\mathrm{CF}_t}{(1+\mathrm{IRR})^t} = 0. \tag{4}$$

Equation 4 is a more general version of Equation 3.

NPV and IRR criteria will usually indicate the same investment decision for a given capital investment. They will usually both recommend acceptance or rejection of the investment. In the case of mutually exclusive investment projects, a company could, on occasion, face the following situation: Project A might have a larger NPV than Project B, but Project B has a higher IRR than Project A. Since the company can invest in only one project, which should it be—Project A or Project B?

The correct choice is Project A, the one with the higher NPV. To understand why, consider this simple example. Suppose you could choose just one of two investments. The first allows you to double your initial outlay of $100 in just one year. The second requires an investment of $100,000, which will grow by 20% in a year. Which should you choose? The first investment gives you a profit of $100. The second gives you a $20,000 profit. Assuming you can access the required funds, the second choice is preferred. Even though it offers a smaller percentage return, it increases your wealth by much more.

When the choice is between two mutually exclusive projects and the NPV and IRR rank the two projects differently, the NPV criterion is strongly preferred. There are good reasons for this preference. The NPV shows the amount of gain, or wealth increase in company value, as a currency amount. The NPV assumes reinvestment of cash flows at the required rate of return, while the IRR assumes reinvestment at the IRR. In using the opportunity cost of funds, the reinvestment assumption of the NPV is the more economically realistic measure. Mathematically, whenever you discount a cash flow at a particular discount rate, you are implicitly assuming that you can reinvest a cash flow at that same discount rate. It is more realistic to assume reinvestment at a lower rate. The IRR does give a rate of return, but the IRR could be for a small investment size or for only a short period of time. As a practical matter, once a corporation has the data to calculate the NPV, it is fairly trivial to then calculate the IRR and other capital allocation criteria. However, the most appropriate and theoretically sound criterion is the NPV.

Another issue is that when the cash flows are nonconventional (i.e., they change sign more than once), there are multiple IRRs. This problem with the IRR criterion is known as the multiple IRR problem. We can illustrate this problem with the following nonconventional cash flow pattern:

Time	0	1	2
Cash flow	−1,000	5,000	−6,000

The IRR for these cash flows satisfies this equation:

$$-1{,}000 + \frac{5{,}000}{(1 + \text{IRR})^1} + \frac{-6{,}000}{(1 + \text{IRR})^2} = 0.$$

It turns out that there are two values of IRR that satisfy the equation: IRR = 1 = 100% and IRR = 2 = 200%. As a result, the IRR is not useful for nonconventional cash flow projects.

This brings up the following question: If NPV is always preferred over IRR for selecting projects, why do companies even bother with IRR? The answer is that many people find it easy to understand a rate of return. If they know that the required return is 10%, they can easily understand that a project returning more than 10% is desirable. If they are simply told the NPV amount, they may not find it as meaningful. In practice, most companies use both metrics. They typically use NPV to make the investment decisions, but they also report the IRR to help their audience understand.

3.3 Corporate Usage of Capital Allocation Methods

c describe expected relations among a company's investments, company value, and share price

Analysts should understand the basic logic of the various capital allocation criteria as well as the practicalities involved in their use at corporations. The usefulness of any analytical tool always depends on the specific application.

If the company can invest elsewhere and earn a return of *r* or if the company can repay its sources of capital and save a cost of *r*, then *r* is the company's opportunity cost of funds. If the company can make an investment that earns more than its opportunity cost of funds, then the investment is creating value for the company and corresponding wealth for shareholders and should be undertaken. Similarly, if the investment earns less than the company's opportunity cost of funds, the investment decreases value in the company, leaving shareholders less wealthy, and should not be undertaken.

The **return on invested capital** (ROIC) is a measure of the profitability of a company relative to the amount of capital invested by the equity- and debtholders. ROIC reflects how effectively a company's management is able to convert capital into profits. The ratio is calculated by dividing the after-tax net profit by the average book value of invested capital (common, preferred, and debt).

The ROIC measure is often compared with the associated **cost of capital** (COC), the required return used in the NPV calculation and the company's associated cost of funds. If the ROIC measure is higher than the COC, the company is generating a higher return for investors compared with the required return, thereby increasing the firm's value for shareholders. The inverse is true if the COC is higher than the ROIC.

Capital allocation criterion and analysis are thus essential tools for corporate managers. Because a corporation's investing decisions ultimately determine the value of its financial obligations, the corporation's investing processes and activity are vital. The NPV criterion is the criterion most directly related to stock prices. If a corporation invests in positive-NPV projects, they should add to the wealth of the company and the corresponding wealth of its shareholders. Example 1 illustrates this scenario for a company making a new investment.

EXAMPLE 1

NPVs and Stock Prices

Freitag Corporation is investing €600 million in distribution facilities. The present value of the future after-tax cash flows is estimated to be €850 million. Freitag has 200 million outstanding shares with a current market price of €32.00 per share. This investment is new information, and it is independent of other expectations about the company. What should be the investment's effect on the value of the company and the stock price?

Solution:

The NPV of the investment is €850 million − €600 million = €250 million. The total market value of the company prior to the investment is €32.00 × 200 million shares = €6,400 million. The value of the company should increase by €250 million, to €6,650 million. The price per share should increase by the NPV per share, or €250 million/200 million shares = €1.25 per share. The share price should increase from €32.00 to €33.25.

The effect of a capital investment's positive or negative NPV on share price is more complicated than in Example 1, in which the value of the stock increased by the investment's NPV.

The value of a company is the value of its existing investments plus the net present values of all its future investments, accounting for any externalities. If an analyst learns of an investment, the impact of that investment on the company's stock price will depend on whether the investment's profitability is more or less than expected. For example, an analyst could learn of a positive-NPV project, but if the project's profitability is less than expected, the company's stock price might drop because of this news. Alternatively, news of a particular capital project might be considered a signal about other capital investments underway or in the future. An investment that by itself might add, say, €0.25 to the value of the stock might signal the existence of other profitable projects. News of this investment might increase the stock price by far more than €0.25.

Inflation affects a company's capital allocation analysis in several ways. The first is whether the investment analysis is done in "nominal" terms or in "real" terms. Nominal cash flows include the effects of inflation, whereas real cash flows are adjusted downward to remove the effects of inflation. Companies may choose to do the analysis in either nominal or real terms, and sound decisions can be made either way.

The cash flows and discount rate used by the company should both be nominal or both be real. In other words, nominal cash flows should be discounted at a nominal discount rate, and real cash flows should be discounted at a real rate. The real rate, just like real cash flows, has had the effect of inflation removed.

Inflation reduces the value of depreciation tax savings to the company (unless the tax system adjusts depreciation for inflation), effectively increasing its real taxes. The effect of expected inflation is captured in the discounted cash flow analysis. If inflation is higher than expected, the profitability of the investment is correspondingly lower than expected. Inflation essentially shifts wealth from the taxpayer (i.e., company) to the government. Conversely, lower-than-expected inflation reduces real taxes for the company (the depreciation tax shelters are more valuable than expected) and results in higher-than-expected profitability of the investment and a corresponding wealth increase for the company.

Inflation does not affect all revenues and costs uniformly. The company's after-tax cash flows will be better or worse than expected depending on how particular sales outputs or cost inputs are affected. Furthermore, contracting with customers, suppliers, employees, and sources of capital can be complicated as inflation rises. Although the capital allocation model accommodates the effects of inflation, inflation complicates the capital allocation process (and the operations of a business, in general).

Finally, a company's capital allocation processes can demonstrate two things about the quality of management: the degree to which management embraces the goal of shareholder wealth maximization and management's effectiveness in pursuing that goal. Both are key for shareholders and potential investors in assessing management's success in value creation for the company.

4

REAL OPTIONS

d describe types of real options relevant to capital investment

Real options are options that allow companies to make decisions in the future that alter the value of capital investment decisions made today. Instead of making all capital investment decisions now, at time zero, a company can wait and make additional decisions at future dates when these future decisions are contingent on future economic events or information. Rather than one-time decisions, it is more reasonable to assume that a company is making decisions sequentially, some now and some in the future. A combination of optimal current and future decisions is what will maximize company value. Real options, by providing future decision-making flexibility to companies, can be an important piece of the value in many capital investments.

Real options are like financial options—except that they deal with real assets instead of financial assets.

EXAMPLE 2

Financial Option Similarity

A simple financial option could be a call option on a share of stock. Suppose the stock is selling for $50, you own a call option with an exercise (strike) price of $50, and the option expires in one year.

A What action should you take if the stock price increases to $60?

B What action should you take if the stock price falls to $40?

Solution:

A Since the stock price ($60) is greater than the option exercise price ($50), you should exercise the option, which allows you to buy the stock, worth $60, for a price of $50. In effect, you have a $10 gain less what you paid for the option.

B Since the stock price ($40) is less than the option exercise price ($50), you should not exercise the option. After all, why would you pay $50 for a stock that is now worth only $40? If you really want to own the stock, it would be cheaper to buy it in the market for $40.

Real options, like financial options, entail the right to make a decision but not the obligation to do so. The company should exercise a real option only if it is value enhancing.

Just as financial options are contingent on an underlying asset, real options are contingent on future events. The flexibility that real options give to companies can greatly enhance the NPV of the company's capital investments. The following are four types of real options.

4.1 Timing Options

Instead of investing now, the company can delay investing. Delaying an investment and basing the decision on hopefully improved information that you might have in, say, a year could help improve the NPV of the projects selected. Project sequencing options allow the company to defer the decision to invest in a future investment until the outcome of some or all of a current investment is known. Investments are sequenced over time, so that investing in a project creates the option to make future investments.

4.2 Sizing Options

If after investing the company can abandon the investment if the financial results are disappointing, it has an **abandonment option**. At some future date, if the cash flow from abandoning an investment exceeds the present value of the cash flows from continuing the investment, the company should exercise the abandonment option. Conversely, if the company can make additional investments when future financial results are strong, the company has a **growth option** or an expansion option. When estimating the cash flows from an expansion, the analyst must also be wary of cannibalization.

4.3 Flexibility Options

Once an investment is made, operational flexibilities besides abandonment or expansion may be available. For example, suppose demand exceeds capacity. Management may be able to exercise a **price-setting option**. By increasing prices, the company could benefit from the excess demand, which it cannot do by increasing production. There are also **production-flexibility options**, which offer the operational flexibility to alter production when demand varies from what is forecast. Even though it is expensive, the company can profit from working overtime or from adding additional shifts. The company can also work with customers and suppliers for their mutual benefit whenever a demand–supply mismatch occurs. This type of option also includes the possibility of using different inputs or producing different outputs.

4.4 Fundamental Options

In such cases as the aforementioned, there are options embedded in a project that can raise its value. In other cases, the whole investment is essentially an option. The payoffs from the investment are contingent on an underlying asset, just like most financial options. For example, the value of an oil well or refinery investment is contingent on the price of oil. The value of a gold mine is contingent on the price of gold. If oil prices were low, you likely would not choose to drill a well. If oil prices were high, you would go ahead and drill. Many R&D (research and development) projects also look like options.

There are several approaches a company's management may use in evaluating capital investments with real options. Analysis of real options can be very complicated for companies. Management can take the following common sense approaches to real option analysis:

1. Use DCF (discounted cash flow) analysis without considering options. If the NPV is positive without considering real options and the project has real options that would simply add more value, it is unnecessary to evaluate the options. Management should simply undertake the investment.
2. Consider the Project NPV = NPV (based on DCF alone) – Cost of options + Value of options. Calculate the NPV based on expected cash flows. Then simply add the value associated with real options less their incremental cost.
3. Use decision trees. Although they are not as conceptually sound as option pricing models, decision trees can capture the essence of many sequential decision-making problems for companies.
4. Use option pricing models.

In carrying out capital allocation, management must consider (1) a variety of real options that investments may possess and (2) a decision about how to reasonably value these options. Example 3 deals with production flexibility; in this case, an additional investment outlay gives the company an option to use alternative fuel sources.

EXAMPLE 3

Production-Flexibility Option

Auvergne AquaFarms has estimated the NPV of the expected cash flows from a new processing plant to be –€0.40 million. Auvergne is evaluating an incremental investment of €0.30 million that would give management the flexibility to switch among coal, natural gas, and oil as energy sources. The original plant relied only on coal. The option to switch to cheaper sources of energy when they are available has an estimated value of €1.20 million. What is the value of the new processing plant including this real option to use alternative energy sources?

Solution:

The NPV, including the real option, should be

$$\begin{aligned}\text{Project NPV} &= \text{NPV (based on DCF alone)} - \text{Cost of options} + \\ &\quad \text{Value of options.}\\ \text{Project NPV} &= -0.40 \text{ million} - 0.30 \text{ million} + 1.20 \text{ million}\\ &= €0.50 \text{ million.}\end{aligned}$$

Without the flexibility offered by the real option, the plant is unprofitable. The real option to adapt to cheaper energy sources adds enough to the value of this investment to give it a positive NPV. The company should undertake the investment, which would add to its value.

COMMON CAPITAL ALLOCATION PITFALLS

5

e describe common capital allocation pitfalls

Although the principles of capital allocation may be easy to understand, applying the principles to real-world investment opportunities can be challenging for companies. Some of the common capital allocation mistakes that companies make are listed here.

- **Not incorporating economic responses into the investment analysis:** Economic responses to an investment often affect its profitability, and these responses have to be correctly anticipated by companies. For example, in response to a successful investment, competitors can enter and reduce the investment's profitability. Companies that make highly profitable investments often find that a competitive marketplace eventually causes profitability to revert to normal levels.
- **Misusing capital allocation templates:** Because hundreds or even thousands of investments need to be analyzed over time, individuals assessing these opportunities on behalf of the company sometimes use standardized capital allocation templates in evaluating investments. This situation creates risks if the template model does not match the investment or if inappropriate information is used.
- **Pushing pet projects:** Often **pet projects** are selected at companies without undergoing normal capital allocation analysis. Or the pet project receives the analysis, but overly optimistic projections are used to inflate the investment's profitability.
- **Basing investment decisions on EPS, net income, or ROE:** Companies sometimes have incentives to boost earnings per share, net income, or return on equity. Many investments, even those with strong NPVs, do not increase these accounting numbers in the short run and may even reduce them. Paying too much attention to short-run accounting numbers can result in a company choosing investments that are not in the long-run economic interests of its shareholders.
- **Using IRR to make investment decisions:** The NPV criterion is economically sound. The IRR criterion is also sound for companies undertaking independent investments with conventional cash flow patterns. For investment opportunities that are mutually exclusive, or competitive with each other, using IRR criterion and not NPV can result in suboptimal investment decisions. Furthermore, the IRR methodology should never be used with nonconventional cash flow projects because it results in more than one solution.
- **Incorrectly accounting for cash flows:** In analyzing a complicated investment, individuals assessing the investment for the company may erroneously omit relevant cash flows, double-count cash flows, and mishandle taxes.

- **Over- or underestimating overhead costs:** In large companies, the cost of an investment may include the overhead it generates for such things as management time, information technology support, financial systems, and other support. Although these items are hard to estimate, over- or underestimating these overhead costs can lead to incorrect analysis by companies.
- **Not using the appropriate risk-adjusted discount rate:** The required rate of return for an investment should be based on its risk. If the company is financing an investment with debt (or with equity), the investment's required rate of return—not the company's **cost of debt** (or **cost of equity**)—should still be used. Similarly, a high-risk investment being considered should be discounted not at the company's overall cost of capital but at the investment's required rate of return. Discount rate errors have a huge impact on the computed NPVs of long-lived investments.
- **Overspending and underspending the capital allocation:** Spending all the investment allocation just because it is available is another common mistake made by companies. In a well-run company, managers will return excess funds whenever their profitable investments cost less than their allocation, and managers will make a sound case for extra funds if their allocation is too small.
- **Failing to consider investment alternatives or alternative states:** Generating good investment ideas is the most basic step in the capital allocation process, and many good alternatives are never even considered at some companies. In addition, many companies fail to consider differing states of the world, which can and should be incorporated through breakeven, scenario, and simulation analyses.
- **Incorrectly handling sunk costs and opportunity costs:** Ignoring sunk costs is difficult for companies to do. Further, not identifying the economic alternatives (real and financial) that are the opportunity costs is probably the biggest failure by companies in their analyses. Only costs that change with the decision are relevant.

SUMMARY

Capital allocation is the process that companies use for decision making on capital investments—those investments with a life of a year or longer. This reading developed the principles behind the basic capital allocation model, the cash flows that go into the model, and several extensions of the basic model.

- Capital allocation supports the most critical investments for many corporations—their investments in long-term assets. The principles of capital allocation have been applied to other corporate investing and financing decisions and to security analysis and portfolio management.
- The typical steps in the capital allocation process are (1) generating ideas, (2) analyzing investment opportunities, (3) planning the capital allocation, and (4) monitoring and post-auditing.
- Types of investments appropriate for the capital allocation process can be categorized as (1) replacement, (2) expansion, (3) new products and services, and (4) regulatory, safety, and environmental.

- Capital allocation decisions are based on incremental after-tax cash flows discounted at the opportunity cost of funds. Financing costs are ignored because both the cost of debt and the cost of other capital are captured in the discount rate.
- The net present value is the present value of all after-tax cash flows, or

  $$\mathrm{NPV} = \sum_{t=0}^{n} \frac{\mathrm{CF}_t}{(1+r)^t},$$

  where the investment outlays are negative cash flows included in the CF_t's and where r is the required rate of return for the investment.
- The IRR is the discount rate that makes the present value of all future cash flows sum to zero. This equation can be solved for the IRR:

  $$\sum_{t=0}^{n} \frac{\mathrm{CF}_t}{(1+\mathrm{IRR})^t} = 0.$$

- The capital allocation decision rules are to invest if the NPV > 0 or if the IRR > r.
- For mutually exclusive investments that are ranked differently by the NPV and IRR, the NPV criterion is more economically sound.
- The fact that projects with positive NPVs theoretically increase the value of the company and the value of its stock could explain the popularity of NPV as an evaluation method.
- Real options can be classified as (1) timing options; (2) sizing options, which can be abandonment options or growth (expansion) options; (3) flexibility options, which can be price-setting options or production-flexibility options; and (4) fundamental options.

PRACTICE PROBLEMS

1 The net present value (NPV) of an investment is equal to the sum of the expected cash flows discounted at the:

A internal rate of return.

B risk-free rate.

C opportunity cost of capital.

2 A $2.2 million investment will result in the following cash flows:

Year	Year-End Cash Flow (millions)
1	$1.3
2	$1.6
3	$1.9
4	$0.8

Using an 8% opportunity cost of capital, the investment's net present value is *closest* to:

A $2.47 million.

B $3.40 million.

C $4.67 million.

3 The internal rate of return (IRR) is *best* described as the:

A opportunity cost of capital.

B time-weighted rate of return.

C discount rate that makes the net present value equal to zero.

4 A three-year investment requires an initial outlay of £1,000. It is expected to provide three year-end cash flows of £200 plus a net salvage value of £700 at the end of three years. Its internal rate of return is *closest* to:

A 10%.

B 11%.

C 20%.

5 Given the following cash flows for a capital investment, calculate the NPV and IRR. The required rate of return is 8%.

Year	0	1	2	3	4	5
Cash flow	–50,000	15,000	15,000	20,000	10,000	5,000

	NPV	IRR
A	$1,905	10.9%
B	$1,905	26.0%
C	$3,379	10.9%

© 2021 CFA Institute. All rights reserved.

6 An investment of \$100 generates after-tax cash flows of \$40 in Year 1, \$80 in Year 2, and \$120 in Year 3. The required rate of return is 20%. The net present value is *closest* to:

- **A** \$42.22.
- **B** \$58.33.
- **C** \$68.52.

7 An investment of \$150,000 is expected to generate an after-tax cash flow of \$100,000 in one year and another \$120,000 in two years. The cost of capital is 10%. What is the internal rate of return?

- **A** 28.39%.
- **B** 28.59%.
- **C** 28.79%.

8 Kim Corporation is considering an investment of 750 million won with expected after-tax cash inflows of 175 million won per year for seven years. The required rate of return is 10%. What is the investment's:

	NPV?	IRR?
A	102 million won	14.0%
B	157 million won	23.3%
C	193 million won	10.0%

9 Erin Chou is reviewing a profitable investment that has a conventional cash flow pattern. If the cash flows for the investment, initial outlay, and future after-tax cash flows all double, Chou would predict that the IRR would:

- **A** increase and the NPV would increase.
- **B** stay the same and the NPV would increase.
- **C** stay the same and the NPV would stay the same.

10 Catherine Ndereba is an energy analyst tasked with evaluating a crude oil exploration and production company. The company previously announced that it plans to embark on a new project to drill for oil offshore. As a result of this announcement, the stock price ran up by 10%. After conducting her analysis, Ms. Ndereba concludes that the project does indeed have a positive NPV. Which statement is true?

- **A** The stock price should remain where it is because Ms. Ndereba's analysis confirms that the recent run-up was justified.
- **B** The stock price should go even higher now that an independent source has confirmed that the NPV is positive.
- **C** The stock price could remain steady, move higher, or move lower.

11 The Bearing Corp. invests only in positive-NPV projects. Which of the following statements is true?

- **A** Bearing's return on invested capital (ROIC) is greater than its cost of capital (COC).
- **B** Bearing's COC is greater than its ROIC.
- **C** We can't reach any conclusions about the relationship between the company's ROIC and COC.

12 Investments 1 and 2 have similar outlays, although the patterns of future cash flows are different. The cash flows as well as the NPV and IRR for the two investments are shown below. For both investments, the required rate of return is 10%.

	Cash Flows						
Year	**0**	**1**	**2**	**3**	**4**	**NPV**	**IRR (%)**
Investment 1	–50	20	20	20	20	13.40	21.86
Investment 2	–50	0	0	0	100	18.30	18.92

The two projects are mutually exclusive. What is the appropriate investment decision?

A Invest in both investments.

B Invest in Investment 1 because it has the higher IRR.

C Invest in Investment 2 because it has the higher NPV.

13 Consider the two investments below. The cash flows as well as the NPV and IRR for the two investments are given. For both investments, the required rate of return is 10%.

	Cash Flows						
Year	**0**	**1**	**2**	**3**	**4**	**NPV**	**IRR (%)**
Investment 1	–100	36	36	36	36	14.12	16.37
Investment 2	–100	0	0	0	175	19.53	15.02

What discount rate would result in the same NPV for both investments?

A A rate between 0.00% and 10.00%

B A rate between 10.00% and 15.02%

C A rate between 15.02% and 16.37%

14 Wilson Flannery is concerned that the following investment has multiple IRRs.

Year	**0**	**1**	**2**	**3**
Cash flows	–50	100	0	–50

How many discount rates produce a zero NPV for this investment?

A One, a discount rate of 0%

B Two, discount rates of 0% and 32%

C Two, discount rates of 0% and 62%

15 With regard to capital allocation, an appropriate estimate of the incremental cash flows from an investment is *least likely* to include:

A externalities.

B interest costs.

C opportunity costs.

The following information relates to Questions 16–18

Bouchard Industries is a Canadian company that manufactures gutters for residential houses. Its management believes it has developed a new process that produces a superior product. The company must make an initial investment of C$190 million to begin production. If demand is high, cash flows are expected to be C$40 million per year. If demand is low, cash flows will be only C$20 million per year. Management believes there is an equal chance that demand will be high or low. The investment also gives the company a production-flexibility option allowing the company to add shifts at the end of the first year if demand turns out to be high. If the company exercises this option, net cash flows would increase by an additional C$5 million in Years 2–10. Bouchard's opportunity cost of funds is 10%.

The internal auditor for Bouchard Industries has made two suggestions for improving capital allocation processes at the company. The internal auditor's suggestions are as follows:

Suggestion 1	"In order to treat all capital allocation proposals in a fair manner, the investments should all use the risk-free rate for the required rate of return."
Suggestion 2	"When rationing capital, it is better to choose the portfolio of investments that maximizes the company NPV than the portfolio that maximizes the company IRR."

16 What is the NPV (C$ millions) of the original project for Bouchard Industries without considering the production-flexibility option?

A –C$6.11 million

B –C$5.66 million

C C$2.33 million

17 What is the NPV (C$ millions) of the optimal set of investment decisions for Bouchard Industries including the production-flexibility option?

A –C$6.34 million

B C$7.43 million

C C$31.03 million

18 Should the capital allocation committee accept the internal auditor's suggestions?

A No for Suggestions 1 and 2

B No for Suggestion 1 and yes for Suggestion 2

C Yes for Suggestion 1 and no for Suggestion 2

SOLUTIONS

1 C is correct. The NPV sums the investment's expected cash flows (CF) discounted at the opportunity cost of capital. The NPV calculation is

$$\text{NPV} = \sum_{t=0}^{N} \frac{\text{CF}_t}{(1+r)^t},$$

where

CF_t = the expected net cash flow at time t
N = the investment's projected life
r = the discount rate or opportunity cost of capital

2 A is correct.

$$\text{NPV} = -\$2.2 + \frac{\$1.3}{(1.08)} + \frac{\$1.6}{(1.08)^2} + \frac{\$1.9}{(1.08)^3} + \frac{\$0.8}{(1.08)^4} = \$2.47 \text{ million.}$$

3 C is correct. The internal rate of return is computed by identifying all cash flows and solving for the rate that makes the net present value of those cash flows equal to zero.

4 B is correct. IRR is determined by setting the net present value equal to zero for the cash flows shown in the following table.

Year	Cash Flow (£)
0	–1,000
1	200
2	200
3	900

5 C is correct.

$$\text{NPV} = -50{,}000 + \frac{15{,}000}{1.08} + \frac{15{,}000}{1.08^2} + \frac{20{,}000}{1.08^3} + \frac{10{,}000}{1.08^4} + \frac{5{,}000}{1.08^5}.$$

$$\text{NPV} = -50{,}000 + 13{,}888.89 + 12{,}860.08 + 15{,}876.64 + 7{,}350.30 + 3{,}402.92.$$

$$\text{NPV} = -50{,}000 + 53{,}378.83 = 3{,}378.83.$$

The IRR, found with a financial calculator, is 10.88%.

6 B is correct.

$$\text{NPV} = \sum_{t=0}^{3} \frac{\text{CF}_t}{(1+r)^t} = -100 + \frac{40}{1.20} + \frac{80}{1.20^2} + \frac{120}{1.20^3} = \$58.33.$$

7 C is correct. The IRR can be found using a financial calculator or by trial and error. Using either method, the IRR is 28.79%. The cost of capital, which is stated as 10%, is not used to solve the problem.

Year	Cash Flow	Present Value			
		28.19%	28.39%	28.59%	28.79%
0	−150,000	−150,000	−150,000	−150,000	−150,000
1	100,000	78,009	77,888	77,767	77,646
2	120,000	73,025	72,798	72,572	72,346
Total		1,034	686	338	−8

A more precise IRR of 28.7854% has a total PV closer to zero.

8 A is correct.

$$\text{NPV} = -750 + \sum_{t=1}^{7} \frac{175}{1.10^t} = -750 + 851.97 = 101.97 \text{ million won.}$$

The IRR, found with a financial calculator, is 14.02%. (PV is −750, $N = 7$, and PMT = 175.)

9 B is correct. The IRR would stay the same because both the initial outlay and the after-tax cash flows double, so the return on each dollar invested would remain the same. All the cash flows and their present values double. The difference between the total present value of the future cash flows and the initial outlay (the NPV) also doubles.

10 C is correct. There are many factors that can affect the stock price, including whether Ms. Ndereba's analysis indicates that the project is more or less profitable than investors expected.

11 A is correct. Since all of Bearing's projects have a positive NPV, they are all providing a return that is greater than the opportunity cost of capital. Therefore, the ROIC must be greater than the COC.

12 C is correct. When valuing mutually exclusive investments, the decision should be made with the NPV method because this method uses the most realistic discount rate—namely, the opportunity cost of funds. In this example, the reinvestment rate for the NPV method (here, 10%) is more realistic than the reinvestment rate for the IRR method (here, 21.86% or 18.92%).

13 B is correct. For these investments, a discount rate of 13.16% would yield the same NPV for both (an NPV of 6.73).

14 C is correct. Discount rates of 0% and approximately 61.8% both give an NPV of zero.

Rate	0%	20%	40%	60%	61.8%	80%	100%
NPV	0.00	4.40	3.21	0.29	0.00	−3.02	−6.25

15 B is correct. Costs to finance the investment are taken into account when the cash flows are discounted at the appropriate cost of capital; including interest costs in the cash flows would result in double-counting the cost of debt.

16 B is correct.

If demand is "high," the NPV is as follows:

$$\text{NPV} = -190 + \sum_{t=1}^{10} \frac{40}{1.10^t} = \text{C\$55.78 million.}$$

If demand is "low," the NPV is

$$\text{NPV} = -190 + \sum_{t=1}^{10} \frac{20}{1.10^t} = -\text{C\$67.11 million.}$$

The expected NPV is 0.50(55.78) + 0.50(–67.11) = –C$5.66 million.

17 B is correct. The additional NPV of adding shifts if demand is "high" is

$$\text{NPV} = \sum\nolimits_{t=2}^{10} \frac{5}{1.10^t} = \text{C\$26.18 million.}$$

If demand is low, the production-flexibility option will not be exercised. The optimal decision is to add shifts only if demand is high.

Because the production-flexibility option is exercised only when demand is high, which happens 50% of the time, the expected present value of adding shifts is

NPV = 0.50(26.18) = C$13.09 million.

The total NPV of the initial project and the production-flexibility option is

NPV = –C$5.66 million + C$13.09 million = C$7.43 million.

The option to add shifts, handled optimally, adds sufficient value to make this a positive-NPV project.

18 B is correct. In valuing investments, expected cash flows should be discounted at required rates of return that reflect their risk, not at a risk-free rate that ignores risk. NPV is superior to IRR. Choosing projects based on IRR might cause the company to concentrate on short-term investments that do not maximize the company's NPV.

READING

29

Sources of Capital

by Edgar A. Norton, Jr., PhD, CFA, Kenneth L. Parkinson, MBA, CCM, Pamela Peterson Drake, PhD, CFA, and John D. Stowe, PhD, CFA

Edgar A. Norton, Jr., PhD, CFA, is at Illinois State University (USA). Kenneth L. Parkinson, MBA, CCM, is at Treasury Information Services, LLC (USA). Pamela Peterson Drake, PhD, CFA, is at James Madison University (USA). John D. Stowe, PhD, CFA, is at Ohio University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe types of financing methods and considerations in their selection;
☐	**b.** describe primary and secondary sources of liquidity and factors that influence a company's liquidity position;
☐	**c.** compare a company's liquidity position with that of peer companies;
☐	**d.** evaluate choices of short-term funding.

INTRODUCTION

1

Raising capital is a fundamental business activity, and companies have multiple short-term and long-term financing choices. Short-term funds without explicit interest rates, such as accounts payable, are part of **working capital management**, which is the management of short-term assets and liabilities. Other debt and equity obligations used to finance the business longer term are considered part of the firm's **capital structure**. The goal of effective working capital management is to ensure that a company has adequate, ready access to the funds necessary for day-to-day operations, while at the same time making sure that the company's assets are invested in the most productive way. The goal of capital structure management is to balance the risks and costs of the firm's long-term finances. In this reading, we examine a variety of debt and equity claims that companies rely on for their sources of capital. This reading also considers sources of liquidity and how to judge the liquidity positions of firms.

© 2021 CFA Institute. All rights reserved.

2

CORPORATE FINANCING OPTIONS

a Describe types of financing methods and considerations in their selection

Businesses have a complex set of assets and liabilities, as illustrated by the financial statements of any publicly traded company. The specific instruments chosen depend on their costs and risks to the borrower (company) and the returns and risks to the investor (equity owner or debtholder), in addition to market conditions, regulatory requirements, and the services of agents, brokers, dealers, and financial intermediaries working in this market. In addition, many other financing options that were also considered are likely not shown on the balance sheet. Here we overview the types of short- and long-term financing sources available to companies. When a company issues stocks and bonds, it is using external sources for funding. Funding can also be generated internally by reinvesting cash flow from operations, delaying payment of liabilities such as accounts payable, or selling or borrowing against assets such as accounts receivable.

Although companies might have an identified target capital structure, in practice, they rarely issue securities in amounts that precisely maintain that target capital structure. For example, if target weights are 20% debt, 20% preferred stock, and 60% common stock, the company would not issue securities in exactly these proportions every time it needs financing. It might issue debt now and common stock or preferred stock the next time it raises capital. However, it tries to keep the overall capital structure near the target over time.

It is important to recognize that the risks of each source of capital are different for the company and for the investor. Debt, for example, is a safer investment than common stock for the investor because it has a higher priority in claims in the event of financial distress. As a result, investors are usually willing to settle for a lower required return from debt securities than from common equity. This is also why the cost of debt to the issuing company is lower than the cost of equity, even on a before-tax basis. From the company's point of view, however, while debt financing is cheaper than equity financing, it is also riskier in that having more debt increases the probability of financial distress.

Exhibit 1 shows some of the major internal and external sources of capital. For large, profitable companies, internal financing is frequently the primary source of funds for financing growth. Short-term sources include financing from operating activities such as accounts payable, short-term loans, and short-term instruments sold, or issued, by the company in the capital markets. To fund longer-term investment activity or for larger amounts of capital, companies have several external financing options. Long-term sources include bonds and leases, common shareholders' equity, and, to lesser extents, hybrids such as preferred shares and convertible debt.

Each of the common sources of finance in Exhibit 1 provide funding for a business and investment opportunities for outside investors.

Exhibit 1 Internal and External Funding Sources

Internal	External		
	Financial intermediaries	Capital markets	Other
After-tax operating cash flows Accounts payable Accounts receivable Inventory & marketable securities	Uncommitted lines of credit Committed lines of credit Revolving credit Secured loans Factoring	Commercial paper Public & private debt Hybrid securities Preferred equity Convertibles Common equity	Leasing

2.1 Internal Financing

Companies can generate internal financing and liquidity from shorter-term operating activities in several ways. These include

- generating more after-tax operating cash flow;
- increasing working capital efficiency, such as extending a company's payables period, reducing its receivables period, or shortening its asset conversion cycle; and
- converting liquid assets such as receivables, inventories, and marketable securities to cash.

Operating Cash Flows

These are the company's after-tax operating cash flows less interest and dividend payments (adjusted for taxes) that can be used to invest in assets and are equal to net income plus depreciation charges minus dividend payments. A company with higher, more predictable after-tax operating cash flows has greater ability to internally finance itself.

Accounts Payable

Accounts payable are amounts due suppliers of goods and services that have not been paid. They arise from **trade credit**, which is a spontaneous form of credit in which a purchaser of the goods or service is, effectively, financing its purchase by delaying the date on which payment is made. Trade credit might involve a delay of payment, with a discount for early payment. The terms of the latter form of credit are generally stated in the discount form: A discount from the purchase price is allowed if payment is received within a specified number of days; otherwise, the full amount is due by a specified date. For example, the terms "2/10, net 30" indicate that a 2% discount is available if the account is paid within 10 days; otherwise, the full amount is due by the 30th day. The terms will differ among industries, influenced by tradition within the industry, terms of competitors, and current interest rates.

Accounts Receivable

Accounts receivable can be thought of as being the opposite of accounts payable. Instead of representing amounts due to suppliers, accounts receivable are amounts owed by customers. In general, businesses prefer to delay paying what they owe, but prefer to receive what is owed to them as quickly as possible. One company's accounts payable represent another company's accounts receivable. The sooner a company can collect what it is owed, the lesser its need to finance its operations in some other way.

Inventory

Like accounts receivable, inventory is a current asset on the balance sheet. These are goods waiting to be sold. Investing in and holding inventory costs money. Companies would prefer not to put a lot of money into inventory when that money could be used for more productive means. The longer the inventory remains unsold, the longer that money is tied up and not usable for other purposes. As a result, an efficient company holds as little inventory as is necessary and sells or turns over the inventory as quickly as possible. Many companies use a just-in-time inventory system, which means the inventory arrives from suppliers when it is needed. That way, it does not sit in storage waiting to be used or sold. This approach makes the supply chain much more efficient, but it also increases the risk of a critical component or item being out of stock when needed. Once inventory is sold, the purchase amount moves into accounts receivable, and the money becomes available once the customer makes payment.

Marketable Securities

Marketable securities are financial instruments, such as stocks and bonds, that can be quickly sold and converted to cash. They are listed as a current asset because the company intends to liquidate them within a year. They could consist of short-term debt that matures within a year, long-term debt, and even common stocks (which have no maturity date) that will be sold within a year. Companies often invest in marketable securities to earn a rate of return that is greater than what they would earn by holding cash. Companies can sell marketable securities if they need funds for any reason, such as when making a capital expenditure.

2.2 Financial Intermediaries

Financial intermediaries can also be a means of financing. These can be bank or non-bank lenders. The main types of short-term bank financing include uncommitted bank lines of credit, committed bank lines of credit and revolving credit agreements, or revolvers. The latter two types can be unsecured or secured, depending on the company's financial strength and the general credit situation, which can vary from country to country. Uncommitted lines and revolvers are more common in the United States, whereas regular committed lines are more common in other parts of the world.

Uncommitted lines of credit are, as the name suggests, the least reliable form of bank borrowing. A bank can offer an uncommitted line of credit for an extended period of time but reserves the right to refuse to honor any request for use of the line. In other words, an uncommitted line is very unstable and is only as good as the bank's desire to offer it. Therefore, companies should not rely very much on uncommitted lines. In fact, banks will not "officially" acknowledge that an uncommitted line is usable, which means that uncommitted lines cannot be shown as a financial reserve in a footnote to the company's financial statements. The primary attraction of uncommitted lines is that they do not require any compensation other than interest.

Committed (regular) lines of credit are the form of bank line of credit that most companies refer to as regular lines of credit. They are more reliable than uncommitted lines because of the bank's formal commitment, which can be verified through an acknowledgment letter as part of the annual financial audit and can be footnoted in the company's annual report. These lines of credit are in effect for 364 days, less than a full year. This term length benefits companies by minimizing amounts needed to meet bank capital requirements. For commitments of a year or longer, banks require more capital. This also effectively ensures that they are short-term liabilities, usually classified as notes payable or the equivalent, on the financial statements.

Regular lines are unsecured and are pre-payable without any penalties. The borrowing rate paid by the company is a negotiated item. The most common interest rates negotiated are borrowing at the bank's prime rate or at a money market rate plus a spread. The most common money market rate is a benchmark reference rate plus a spread. The spread varies depending on the borrower's **creditworthiness**, which is the perceived ability of the borrower to pay its debt obligations in a timely manner and represents the ability of a company to withstand adverse impacts on its cash flows. Regular lines, unlike uncommitted lines, require compensation, usually in the form of a commitment fee to the lender. The fee is typically a fractional percent (e.g., 0.50%) of the full amount or the unused amount of the line, depending on bank–company negotiations.

Revolving credit agreements, also referred to as "revolvers," are the most reliable form of short-term bank borrowing. They involve formal legal agreements that define the aspects of the agreement. These agreements are similar to those for regular lines with respect to borrowing rates, compensation, and being unsecured. Revolvers differ in that they are in effect for multiple years (e.g., three to five years) and can have optional medium-term loan features. In addition, they are often used for much larger amounts than a regular line, and these larger amounts are spread out among more than one bank. With revolvers, borrowers draw down and pay back amounts periodically.

Secured ("asset-based") loans are loans in which the lender requires the company to provide collateral in the form of an asset, such as a fixed asset that the company owns or high-quality receivables and inventory. These assets are pledged against the loan, and the lender files a lien against them. This lien becomes part of the borrower's financial record and is shown on its credit report. Companies that lack sufficient credit quality to qualify for unsecured loans might arrange for secured loans. For example, a company can use its accounts receivable to generate cash flow through the **assignment of accounts receivable**, which is the use of these receivables as collateral for a loan.

A company can also sell its accounts receivable to a lender (called a factor), typically at a substantial discount. In an assignment arrangement, the company remains responsible for the collection of the accounts, whereas in a **factoring arrangement**, the company shifts the credit granting and collection process to the lender or factor. The cost of this credit (i.e., the amount of the discount) depends on the credit quality of the accounts and the costs of collection. Similarly, inventory can be used in different ways as collateral for a loan.

Web-based lenders and **non-bank lenders** are recent innovations not typically used by larger companies. Web-based lenders operate primarily on the internet, offering loans in relatively small amounts, typically to small businesses in need of cash. Non-bank lenders also lend to businesses, but unlike typical banks, which make loans and take deposits, these lenders only make loans.

Many short-term loans have historically been tied to Libor (London Inter-bank Offered Rate), such as Libor plus 0.50% (50 basis points). However, in 2021, Libor and other similar global benchmark rates were replaced with alternative reference rates in the United States, Europe, Switzerland, Great Britain, and Japan. Both bank and non-bank lenders often charge additional commissions and fees beyond the quoted interest rate.

2.3 Capital Markets

Commercial Paper

Commercial paper is a short-term, unsecured instrument typically issued by large and well-rated companies. Commercial paper can be sold directly to investors or through dealers. To avoid registration costs with national regulatory agencies, maturities for commercial paper typically range from a few days up to 270 days. Although a significant amount of asset-backed commercial paper exists, most commercial paper is unsecured, with no specific collateral. It is a feasible source of funds for large, highly creditworthy companies. Issuers of commercial paper are often required to have a backup line of credit. The short-term nature of commercial paper, along with the creditworthiness of the borrower and the backup line of credit, generally make commercial paper a low-risk investment for investors.

Next we look at debt and common equity to contrast these two very different forms of financing, before turning to hybrid securities, which have characteristics of both debt and equity.

Debt vs. Equity Financing

Businesses are generally financed longer term using a combination of debt and equity securities. There are key differences between debt and equity obligations. Exhibit 2 summarizes these differences.

Exhibit 2 Debt vs. Equity Summary

	Debt	**Equity**
Legal Agreement	Company has a contractual obligation to debtholders	Shareholders are residual owners of the company
Claim Priority	Debtholder interest and principal payments have priority	Residual claimants to distributions and corporate assets
Distributions	Periodic, contractual interest payments and repayment of principal at maturity	Discretionary dividend payments declared by the Board of Directors
Taxation (some variation across jurisdictions)	Interest payments are tax-deductible expenses	Dividend payments and share repurchases are not tax-deductible expenses
Term	Stated term to maturity	No finite term
Voting Rights	Generally no voting rights	Voting rights
Cost to Company	Generally lower cost to company	Generally higher cost to company
Investor Risk	Generally lower risk to investors	Generally higher risk to investors

One key difference is that the company has a binding contract with debtholders that requires their claims be fully paid before the company can make distributions to equity owners. In other words, debtholders have a prior legal claim on the company's cash flows and assets over the claims of equity owners. Another difference is that interest payments to debtholders are generally treated as a tax-deductible expense for the company, while dividend payments to shareholders are not. Finally, equity holders represent a more permanent source of capital and have voting rights to elect the board of directors, which oversees the management of the company. In contrast, debt represents a generally cheaper financing source for companies and a lower risk investment than equity.

Long-Term Debt

Long-term debt has a maturity of at least one year. Conventionally, money market instruments have a maturity of less than a year, notes have maturities from one to ten years, and bonds have maturities of at least ten years. Because of their long maturities, bonds are riskier from both an interest rate and credit risk perspective than notes or money market instruments. Hence, bond lenders (investors) and borrowers (companies) agree to bond covenants that are detailed contracts specifying the rights of the lender and restrictions on the borrower (company). These covenants regulate the company's use and disposition of assets, restrict its ability to pay dividends, and restrict its ability to issue additional debt that might dilute the value of a bond.

Public debt is negotiable and approved for sale on open markets. In this context, a negotiable instrument is a written document describing the promise to pay that is transferable and can be sold to another party. Private debt can also be negotiable; however, private debt does not trade on a market. As a result, private debt is more difficult for the holder to sell. While some private debt instruments are not negotiable, such as savings bonds and certificates of deposit, private debt issued by businesses can usually be sold by one party to another.

Common Equity

Common equity, or common shares, represents ownership in a company and is considered a more permanent source of capital. Shareholders have claims on the company's profits after its obligations under other contractual claims are satisfied. Shareholders receive dividends distributed by the company and are entitled to the residual value of the assets if the company goes out of business. Shareholders elect the company's board of directors and thereby, conceptually, at least, have control over how the company is managed and operated.

In addition to publicly traded equity, private equity is an important corporate financing alternative in many countries. Private equity is equity that is not registered with the national security regulatory agency and cannot be traded in the public equity markets. Much private equity consists of family and personal investments in small private businesses. Also, large organizations that are not publicly traded are called private equity firms and provide private equity financing to companies. Some of these private equity firms invest in venture capital firms, and others own a portfolio of publicly and non-publicly traded companies.

Preferred Equity

Preferred shares, or preferred equity, are hybrid securities that are issued by companies and have characteristics of both bonds and common equity. Like interest obligations on debt, dividends on preferred shares are often fixed but can be variable. For example, "participating" preferred shares pay dividends based on a company's profitability. However, preferred dividends paid are not a tax-deductible expense for the company, and unlike with debt instruments, a company can choose to defer or decline to pay dividends on preferred equity. In some cases, preferred dividends are cumulative, meaning the company must pay any skipped preferred dividend payments before it can make any dividend payments on common equity. In case of business failure, bondholders have seniority (have a prior claim) over preferred shareholders on assets or cash flows, and preferred shareholders have similar seniority over common shareholders. Historically, the number of common stock issues by companies and their related proceeds has been substantially greater than financing from preferred issues.

Other Hybrid Securities

Convertible debt and convertible preferreds are hybrid securities issued by companies and are convertible into a fixed number of the companies' common shares. These securities are considered hybrids because they are neither pure equity nor pure debt in nature. They are similar to regular preferred equity or debt but provide the owner with an option to convert to common stock. If the company's share price remains low, there may be no financial benefit to converting, and convertible owners will retain ownership of their convertible security while continuing to receive its dividend or interest payments. If the share price rises sufficiently, however, convertible owners may choose to convert their security to common stock in the company, given the appreciation in value of their investment. The option to convert is a valuable right. As a result, companies can offer smaller dividends or interest payments on convertible securities.

2.4 Other Financing

Leasing Obligations

Some companies might also be able to finance by leasing. In a leasing arrangement, the purchase of an asset and its financing are bundled instead of being separate transactions. In some cases, the leasing of real property or equipment might offer lower joint costs for the company and for suppliers of capital than buying and financing separately. The lease is a debt instrument where the asset owner (the lessor) gives another party (the lessee) the right to use the asset. The lessee agrees to make a set of contractually fixed payments. The leasing contract specifies the length of time the lessee can use the asset, whether the lessee is responsible for maintenance of the asset, and whether the lessee can buy the asset at the end of the leasing period and if so, at what price.

2.5 Considerations Affecting Financing Choices

A company's financing choices depend on the nature of its needs and business, the general economic environment, and market conditions. In different parts of the world, the dominant sources of external financing can be 1) the capital markets, 2) large commercial banks, or 3) the sovereign state. Many countries are assigned to one of these three cases, but in reality, most countries are varying blends of these. Because companies in a particular industry have many common traits, the business models in an industry frequently exhibit similar patterns of external financing.

In addition to geographic location and industry, firm-specific and general economic considerations affect financing choices. These are summarized in Exhibit 3.

Exhibit 3 Financing Considerations

Firm Specific	Macroeconomic
Company size	Taxation
Riskiness of assets	Inflation
Assets for collateral	Government policy
Public versus private equity	Monetary policy
Asset liability management	
Debt maturity structure	
Currency risks	
Agency costs	

Exhibit 3 (Continued)

Firm Specific	Macroeconomic
Bankruptcy costs	
Flotation costs	

2.5.1 *Firm-Specific Financing Considerations*

Company Size The size of a company affects its financing options. Small companies generally do not have access to funding sources such as commercial paper and publicly issued debt and equity. Smaller companies might not have the positive and predictable cash flows needed to meet fixed, periodic debt payments. Large companies with strong operating cash flows can rely more heavily on internal financing, while smaller companies, especially those that are younger and faster growing, or without cash flow, cannot satisfy their capital needs internally and must rely more on external financing—in particular, private equity. Because smaller companies typically lack the positive, predictable cash flows often required by lenders, and because debt financing may not be available or is too expensive, equity financing is usually the predominant source of financing.

Riskiness of Assets Companies with higher volatility of operating cash flows rely primarily on equity financing and tend to use little if any debt financing. High levels of uncertainty or volatility in operating cash flows can make servicing debt obligations challenging. Companies having a high degree of business risk may seek to minimize financial risks and leverage. Business risks are non-financial risks for the company and include technological change, political risks, marketing and production uncertainties, managerial risks, and many others. For example, Moderna, an American biotechnology company best known for using a unique technology to develop a vaccine for COVID-19, has some lease obligations but does not use other kinds of debt financing. NIO Inc., a Chinese electric car manufacturer, used very little debt financing until unit sales and revenues began to accelerate.

Assets for Collateral Real property and equipment are good collateral for mortgages and asset-backed bonds. In the case of bankruptcy, mortgage bondholders have first claim on the collateral and are still general creditors for any claims left unsatisfied by repossession of the collateral. Assets that are unique, highly specialized, and intangible might not be valuable as collateral. Sound collateral can increase a company's access to debt financing as well as reduce its costs. Macy's, for example, operates department stores across the United States and owns a significant amount of real estate, which it can easily borrow against. In contrast, Workhorse Group, a US manufacturer of medium-duty electric trucks, has little revenues and assets that cannot be easily converted to other uses. As a result, the company has more difficulty issuing debt at favorable terms.

Public vs. Private Equity For companies that have issued public equity, shares are available and easily traded through public stock markets. For companies that finance through private equity, shares are available directly from the company or from non-public transactions. The extraordinary size of many publicly traded companies could not be attained without public trading of their securities. Many private companies are simply too small to sell their debt and equity securities publicly. An advantage large, publicly traded companies have is that their securities are liquid and trade frequently, while private company claims are not liquid and do not trade in secondary markets. Some large private equity firms are organized as limited partnerships and own a portfolio of large companies, some of which were once publicly traded, were bought by the private equity firm, and are no longer traded. One perceived advantage private equity firms have is that they can reduce the agency costs associated with manager and shareholder conflicts of interest.

Asset Liability Management Businesses tend to match the maturity structures of their assets and liabilities. A mismatch of asset and liability maturity structures can be problematic. For example, if a company finances its long-term assets with short-term obligations, its profitability might be threatened if the cost of its short-term financing increases. Commercial banks carefully manage the maturity structure of their asset portfolios with the maturity structure of their financing. Non-financial businesses such as utilities, manufacturing firms, real estate firms, and others that have long-term assets can safely use more long-term financing than firms with short-term assets.

Debt Maturity Structure If interest rates on short-term debt are generally less than those on long-term debt, a less expensive option for the company would be to finance with short-term debt and continually refinance with new short-term debt whenever current debt matures. This risk, sometimes called rollover risk, materializes when interest rates go up or when company-specific or general economic conditions cause the company to be unable to refinance or issue new debt.

Currency Risks A company's business revenues and its financing can be in different currencies. For example, suppose the company's business operations are in the domestic currency and its debt financing is in a foreign currency. If the domestic currency declines in value, the company might not be able to repay its debts. Short-term debt contracts can be hedged in derivatives markets with currency swaps, options, and forwards. However, because the market for long-term derivatives is relatively small, the company's best option to hedge currency risks is to execute its debt and equity financing in the same currency as its business operations. Large, well-rated multinational companies will finance themselves in different currencies to match the currency exposure of their local operations.

Agency Costs Debtholders can suffer losses if the borrowing company increases its riskiness beyond its original expectations. This can occur in several ways. A major risk is asset substitution, wherein the company replaces assets that were fairly safe with much riskier assets. The risk of financial distress or default rises for unsecured lenders if the company distributes too much cash to owners or managers, or pledges significant assets to secured creditors. The company can also become riskier if it uses more financial leverage, replacing equity financing with increased levels of new debt financing. Debtholders try to protect their interests by having bond covenants that restrict the company from taking certain actions that would be detrimental to bondholders. Debt covenants are often used in bank revolvers and some lines of credit.

Bankruptcy Costs The costs and risks of financing a business are normally shared among debtholders and shareholders. Bankruptcy introduces additional costs, with some company resources being consumed by third parties, the legal system, and other administrative costs of bankruptcy, which is a net loss to the company's suppliers of capital. The possibility of bankruptcy and the costs of financial distress can also affect customers, suppliers, managers, employees, and the community at large. Because of these costs, suppliers of capital might be averse to financial structures that risk bankruptcy.

Flotation Costs A publicly traded company incurs flotation costs when it issues new debt or equity securities. These costs can include various expenses that are company specific, such as legal fees, registration fees, audit fees, and underwriting fees. Flotation costs as a percentage of the capital raised are generally lower for debt offerings than for equity offerings. Flotation costs increase the cost of financing and can affect a company's financing decisions.

2.5.2 *General Economic Considerations*

In addition to company-specific considerations, economic factors and governmental policies external to the company affect its financing choices.

Taxation The cost of debt financing is typically lower than the cost of equity financing. If interest payments on debt financing are tax-deductible and distributions on equity financing are not, the after-tax cost of debt financing is even more attractive, relative to equity financing, for profitable companies. By allowing interest payments to be tax-deductible, governments provide companies an incentive to use more financial leverage. However, greater use of debt by companies increases the risk of financial distress, particularly during economic downturns.

Inflation Normally, the nominal rate of return is a sum of the real rate and the expected inflation rate. However, uncertainty over inflation rates can make longer-term, fixed-rate contracts unattractive to the company (borrower) or investor (lender). If inflation is expected to rise, companies would prefer to borrow at a fixed rate, while investors would prefer to lend at a variable rate. If inflation is expected to decline, the opposite would be true.

Government Policy Governmental fiscal policies can be used to stimulate the economy or to subsidize specific industries. A government agency, for example, can directly make loans at lower interest rates than would be required by capital markets. The government can provide loan guarantees where the principal and interest payments are guaranteed, which would increase capital amounts and lower the interest rates on the guaranteed loans.

Monetary Policy Central banks around the world have at times driven interest rates to historically low and even negative levels. The intent was to stimulate economies by making access to capital cheaper for companies and to give investors an incentive to take more risk with their cash. Many companies used the opportunity to increase leverage at low rates to refinance existing debt and repurchase common stock with the proceeds rather than make capital investments.

EXAMPLE 1

Estimating External Financing Needs

Fllama Company is planning its sources of financing for a substantial investment of £30 million next year. The assumptions for Fllama's plan are the following:

- Investments of £5 million in receivables and £5 million in inventory. Fixed capital investments of £20 million, including £5 million to replace depreciated equipment and £15 million of net new investments. Total investments will be £30 million.
- Net income of £10 million and depreciation charges of £5 million. Dividend payments will be £4 million.
- Short-term financing from accounts payable of £3 million is expected. The company will use receivables as collateral for another £3 million loan. The company will also issue a £4 million short-term note to a commercial bank.
- Any additional external financing needed can be raised from an increase in long-term bonds. If additional financing is not needed, any excess funds will be used to repurchase common shares.

Part A: Describe how Fllama would determine its financing needs.

Part B: How much, if any, does Fllama need to issue in long-term bonds?

A Fllama does not need to issue any bonds.

B Fllama will need to issue £4 million of bonds.

C Fllama will need to issue £9 million of bonds.

Solution:

Part A: Fllama should begin by considering its various sources of net cash to determine whether its net cash is sufficient to meet its investment needs. The facts state that the company will generate net income and will also have depreciation charges and make dividend payments. Because depreciation is a non-cash expense, it should be added to net income as a source of cash. Dividends, however, are paid from net income, are a use of cash, and reduce the net cash available for funding. Other sources of cash include the amounts generated from accounts payable, accounts receivable, and the short-term note from the bank.

Part B: C is correct. Fllama must issue £9 million of bonds.

Source	Amount (£)
Internal Financing	
Operating cash flow (Net income + depreciation – dividends)	11,000,000
Accounts payable	3,000,000
Short-Term Intermediary Financing	
Bank loan against receivables	3,000,000
Short-term note	4,000,000
Total sources	21,000,000

Because the company requires £30 million of financing and the planned sources total £21 million, Fllama will need to issue £9 million of new bonds.

3 MANAGING AND MEASURING LIQUIDITY

b describe primary and secondary sources of liquidity and factors that influence a company's liquidity position

Liquidity is the extent to which a company is able to meet its short-term obligations using cash flows and those assets that can be readily transformed into cash. When we evaluate the liquidity of an asset, we focus on two dimensions: the type of asset and the speed at which the asset can be converted to cash, either by sale or financing. Unlike many aspects of corporate finance, corporate liquidity management does not involve a great deal of theory or generally accepted principles. For companies that have the luxury of large excesses of cash, liquidity is typically taken for granted, and the focus is on putting the excess liquidity to its most productive use. In the event no productive use within the company can be identified, shareholders will pressure companies to return this cash to owners in the form of share buybacks or dividends. On the other hand, when a company faces tighter financial situations, having effective liquidity management is important to ensure solvency. Unfortunately, this recognition comes too late for some companies, with bankruptcy and possible liquidation representing the companies' final choice.

3.1 Defining Liquidity Management

Liquidity management refers to an organization's ability to generate cash when and where it is needed. Liquidity refers to the cash balances, borrowing capacity, and ability to convert other assets or extend other liabilities into cash for use in keeping the entity solvent (i.e., being able to pay bills and continue in operation). For the most part, we associate liquidity with short-term assets and liabilities, yet longer-term assets such as marketable securities can be converted into cash to provide liquidity. In addition, longer-term liabilities can also be renegotiated to reduce the drain on cash, thereby providing liquidity by preserving the limited supply of cash. Of course, these last two methods might come at a price because they tend to reduce the company's overall financial strength.

The challenges of managing liquidity include developing, implementing, and maintaining a liquidity policy. To do this effectively, a company must manage all its key sources of liquidity efficiently. These key sources can vary from company to company, but they generally include primary sources of liquidity, such as cash balances, and secondary sources of liquidity, such as selling assets.

3.1.1 *Primary Sources of Liquidity*

Primary sources of liquidity represent the most readily accessible resources available. They can be held as cash or as near-cash securities. Primary sources include the following:

- Free cash flow, which is the firm's after-tax operating cash flow less planned short- and long-term investments. For a profitable firm, free cash flow provides substantial liquidity. A rapidly growing firm has less free cash flow because of the investments required to facilitate the firm's growth.
- Ready cash balances, which is cash available in bank accounts, resulting from payment collections, investment income, liquidation of near-cash securities (i.e., those with maturities of fewer than 90 days), and other cash flows.
- Short-term funds, which can include items such as trade credit, bank lines of credit, and short-term investment portfolios.
- Cash flow management, which is the company's effectiveness in its cash management system and practices, and the degree of decentralization of the collections or payments processes. The more decentralized the system of collections, for example, the more likely the company will be to have cash tied up in the system and not available for use.

These sources represent liquidity that is typical for most companies. They represent funds that are readily accessible at relatively low cost.

3.1.2 *Secondary Sources of Liquidity*

The main difference between primary and secondary sources of liquidity is that using a primary source is not likely to affect the normal operations of the company, whereas using a secondary source might result in a change in the company's financial and operating positions. Secondary sources include

- negotiating debt contracts, relieving pressures from high interest payments or principal repayments, and negotiating contracts with customers and suppliers;

- liquidating assets, which depends on the degree to which short-term and/or long-term assets can be liquidated and converted into cash without substantial loss in value; and
- filing for bankruptcy protection and reorganization.

The use of secondary sources might signal a company's deteriorating financial health and provide liquidity at a high price—the cost of giving up a company asset to produce emergency cash. This last source, reorganization through bankruptcy, can also be considered a liquidity tool because a company under bankruptcy protection that generates operating cash will be liquid and generally able to continue business operations until a restructuring has been devised and approved. However, this option is likely to be at a significant cost, or disadvantage, to existing equityholders.

Example 2 shows the net proceeds from the primary and secondary sources of liquidity for a company in a financial crisis. The example also shows the liquidation costs incurred by the company when using these sources to raise funds. These costs can include the fees and commissions involved with the asset sale as well as any discount in asset value due to illiquidity issues.

EXAMPLE 2

Estimating Costs of Liquidity for OM Distributors

One of the companies in your portfolio, OM Distributors (OM), is having a liquidity crisis. You have identified four potential actions that OM could take to raise funds. You have estimated, in local currency, the fair value of the assets and the liquidation costs OM might incur.

Source of Funds	Fair Value (Local, millions)	Liquidation Cost (%)
Sell short-term marketable securities	10	0
Sell select inventories and receivables	20	10
Sell excess real estate property	50	15
Sell a subsidiary of the firm	30	20

The liquidation costs include the fees and commissions of selling an asset as well as any reduction in the value of the asset, because it is an illiquid asset being sold quickly. In this case, liquidation costs for marketable securities are rounded to 0%.

Net of liquidation costs how much liquidity can OM raise if all four sources of funds are used, and what are the total liquidation costs incurred by OM? In local currency, these amounts are:

A 110 million, 9.5 million

B 94.5 million, 15.5 million

C 125.5 million, 15.5 million

Solution:

The costs and net funds raised are summarized in this table:

Source of Funds	Fair Value (local, millions)	Liquidation Costs %	Liquidation Costs (local, millions)	Net Proceeds (local, millions)
Marketable securities	10	0	0	10
Inventories & receivables	20	10	2	18
Real estate property	50	15	7.5	42.5
Subsidiary of the firm	30	20	6	24
Total			15.5	94.5

B is correct. Shown in the table above in local currency terms, the total net proceeds are 94.5 million, and the total liquidation costs incurred are 15.5 million.

3.1.3 *Drags and Pulls on Liquidity*

Cash flow transactions—that is, cash receipts and disbursements—have significant effects on a company's liquidity position. We refer to these effects as drags and pulls on liquidity. A **drag on liquidity** is when receipts lag, creating pressure from the decreased available funds; a **pull on liquidity** is when disbursements are paid too quickly or trade credit availability is limited, requiring companies to expend funds before they receive funds from sales that could cover the liability.

Major drags on receipts involve pressures from credit management and deterioration in other assets and include the following:

- *Uncollected receivables.* The longer these are outstanding, the greater the risk that they will not be collected at all. They are indicated by the large number of days of receivables and high levels of bad debt expenses. Just as the drags on receipts might cause increased pressure on working capital, pulls on outgoing payments could have similar effects.
- *Obsolete inventory.* If inventory stands unused for long periods, it might be an indication that it is no longer usable. Slow inventory turnover ratios can also indicate obsolete inventory. Once identified, obsolete inventory should be attended to as soon as possible to minimize storage and other costs.
- *Tight credit.* When economic conditions make capital scarcer, short-term debt becomes more expensive to arrange and use. Attempting to smooth out peak borrowings can help blunt the impact of tight credit, as can improving the company's collections.

In many cases, drags can be alleviated by stricter enforcement of credit and collection practices.

However, managing the cash outflows might be as important as managing the inflows. If suppliers and other vendors who offer credit terms perceive a weakened financial position or are unfamiliar with a company, they might restrict payment terms so much that the company's liquidity reserves are stretched thin. Major pulls on payments include the following:

- *Making payments early.* By paying vendors, employees, or others before the due dates, companies forgo the use of funds. Effective payment management means not making early payments. Payables managers typically hold payments until they can be made by the due date.

- *Reduced credit limits.* If a company has a history of making late payments, suppliers might cut the amount of credit they will allow to be outstanding at any time, which can squeeze the company's liquidity. Some companies try to extend payment periods as long as possible, disregarding the possible impact of reduced credit limits.
- *Limits on short-term lines of credit.* If a company's bank reduces the line of credit it offers the company, a liquidity squeeze might result. Credit line restrictions can be government mandated, market related, or simply company specific. Many companies try to avert this situation by establishing credit lines far in excess of what they are likely to need. This "overbanking" approach is often commonplace in emerging economies or even in more-developed countries when the banking system is not sound and the economy is shaky.
- *Low liquidity positions.* Many companies face chronic liquidity shortages, often because of their particular industry or from their weaker financial position. The major remedy for this situation is, of course, to improve the company's financial position, perhaps by issuing additional equity or hybrid securities. If not, the company could be heavily affected by interest rates and credit availability, in which case it might have to turn to secured borrowing to obtain working capital funds. Therefore, these companies must identify ahead of time the assets that can be used to help support their short-term borrowing activities.

Identifying these drags and pulls as soon as possible is critical, including before they happen or when they have just arisen. In Example 3, an analyst is trying to identify changes in the firm that are affecting the its liquidity position.

EXAMPLE 3

Drags and Pulls on Liquidity

Arya Rajavade's firm is experiencing liquidity challenges. Several things might be contributing to this. Rajavade is reviewing three notable changes that have been suggested to her as drags or pulls on liquidity:

- The increasing days in receivables are a drag on liquidity.
- Lower inventory turnover is a drag on liquidity
- Increased credit limits by lenders is a pull on liquidity.

Which of these does not contribute to the firm's tight liquidity situation?

A The change in days in receivables

B The change in inventory turnover

C The change in credit limits

Solution:

C is correct. The increase in credit limits is not a pull on liquidity but is in fact the opposite; it provides liquidity.

3.2 Measuring Liquidity

c compare a company's liquidity position with that of peer companies

Liquidity contributes to a company's creditworthiness. Creditworthiness allows the company to obtain lower borrowing costs and better terms for trade credit and contributes to the company's investment flexibility, enabling it to exploit profitable opportunities.

The less liquid the company, the greater the risk it will suffer financial distress or, in the extreme case, insolvency or bankruptcy. Because debt obligations are paid with cash, the company's cash flows ultimately determine solvency. Immediate sources of funds for paying bills are cash on hand, proceeds from the sale of marketable securities, and the collection of accounts receivable. Additional liquidity comes from inventory that can be sold and thus converted into cash, either directly through cash sales or indirectly through credit sales (i.e., accounts receivable).

At some point, however, a company might have too much invested in low- and non-earning assets. Cash, marketable securities, accounts receivable, and inventory represent a company's liquidity, but these investments are low earning relative to the long-term, capital investment opportunities that companies might have available.

Various financial ratios can be used to assess a company's liquidity as well as its management of assets over time. Here we will look at some of these ratios in a little more detail. Introduced in the coverage of Financial Reporting and Analysis, several liquidity and activity ratios are summarized in Exhibit 4.

Exhibit 4 Ratios Used for Assessing Company Liquidity

Liquidity Ratios
$\text{Current ratio} = \dfrac{\text{Current assets}}{\text{Current liabilities}}$
$\text{Quick ratio} = \dfrac{\text{Cash} + \text{Short-term marketable instruments} + \text{Receivables}}{\text{Current liabilities}}$
$\text{Cash ratio} = \dfrac{\text{Cash} + \text{Short-term marketable instruments}}{\text{Current liabilities}}$
Activity Ratios
$\text{Accounts receivable turnover} = \dfrac{\text{Credit sales}}{\text{Average receivables}}$
$\text{Inventory turnover} = \dfrac{\text{Cost of goods sold}}{\text{Average inventory}}$
$\text{Number of days of receivables} = \dfrac{\text{Average accounts receivable}}{\text{Average day's sales on credit}} = \dfrac{\text{Average accounts receivable}}{\text{Sales on credit/365}}$
$\text{Number of days of inventory} = \dfrac{\text{Average inventory}}{\text{Average day's cost of goods sold}} = \dfrac{\text{Average inventory}}{\text{Cost of goods sold/365}}$

(continued)

Exhibit 4 (Continued)

$$\text{Number of days of payables} = \frac{\text{Average accounts payable}}{\text{Average day's purchases}}$$

$$= \frac{\text{Average accounts payable}}{\text{Purchases/365}}$$

Cash conversion cycle = Days of inventory + Days of receivables − Days of payables

We calculate **liquidity ratios** to measure a company's ability to meet short-term obligations to creditors as they mature or come due. This form of liquidity analysis focuses on the relationship between current assets and current liabilities and the rapidity with which receivables and inventory can be converted into cash during normal business operations. The levels of these ratios, trends or changes in the ratios over time, and comparisons with competitors or the industry are used to judge a firm's liquidity position.

In addition to looking at the relationships among these balance sheet accounts, we can also estimate **activity ratios**, which measure how well key current assets are managed over time. These ratios use information from the income statement and the balance sheet to help tell the story of how well a company is managing its liquid assets.

Some of the major applications of this type of analysis include performance evaluation, monitoring, creditworthiness assessment, and financial projections. But ratios are useful only when they can be compared. The comparison should be done in two ways: comparisons over time for the same company and over time for the company compared with its peer group. Peer groups can include competitors from the same industry as the company as well as other companies of comparable size with comparable financial situations.

Consider Daimler AG, a producer of cars, trucks, and vans. We can see the change in the company's current, quick, and cash ratios over a decade, the fiscal years 2010 through 2019, in Exhibit 5, Panel A. Here, we see that the current ratio and the quick ratio increased over the time period. The cash ratio, over the decade, did not show the upward trend that the quick and current ratios exhibited. We can see what is driving these trends in the calculation of the cash conversion cycle in Panel B of the exhibit. Over the 10 years, the cash conversion cycle increased by more than 40 days. The slight increase in days of payables outstanding would have decreased the cash conversion cycle slightly. The days of inventory increased by approximately 9 days, and the days of receivables increased by approximately 37 days. Although the increase in inventory certainly puts a demand on liquidity, for Daimler, the impact of the increase in receivables was more dramatic. Over the decade, however, based on the liquidity and activity ratios, this company seemed to have good control of its liquidity position.

Exhibit 5 Liquidity Analysis of Daimler AG, 10 Years Ending December 2019

	2019	**2018**	**2017**	**2016**	**2015**	**2014**	**2013**	**2012**	**2011**	**2010**
Panel A: Current, Quick, and Cash Ratios, December 2010–2019										
Current Ratio	1.21	1.24	1.22	1.21	1.19	1.15	1.19	1.15	1.11	1.07
Quick Ratio	0.93	0.94	0.93	0.91	0.88	0.84	0.90	0.85	0.80	0.80
Cash Ratio	0.25	0.25	0.26	0.25	0.23	0.23	0.30	0.26	0.20	0.25

Exhibit 5 (Continued)

	2019	2018	2017	2016	2015	2014	2013	2012	2011	2010
Panel B: Days of Inventory, Receivables, Payables, and Cash Conversion Cycle, December 2010–2019										
Days of Inventory on Hand	75.2	74.9	71.6	73.9	69.1	68.5	69.2	71.5	71.2	66.6
+ Days of Receivables	136.8	127.9	118.8	118.4	104.6	102.6	105.5	104.6	104.2	100.0
– Days of Payables Outstanding	34.1	35.2	33.6	32.8	31.3	33.4	35.5	37.4	37.5	31.6
= Cash Conversion Cycle	177.9	167.6	156.8	159.5	142.4	137.7	139.2	138.7	137.9	135.0

Now consider Walmart, Target Corporation, Kohl's Corporation, and Costco Wholesale Corporation. Selected ratios for these large US discount retailers are shown in Exhibit 6. The data are from the fiscal year ending January 2020, except for Costco, whose fiscal year ended in August 2019. We see some differences among these four competitors. These differences can be explained, in part, by the retailers' different product mixes (e.g., Walmart and Costco have more sales from grocery lines than the other two), as well as their different inventory management systems and different inventory suppliers. None of the four firms invests heavily in accounts receivable, and customers generally pay with credit cards. The different need for liquidity can also be explained, in part, by the companies' different operating cycles. The most striking difference is for Kohl's. Walmart, Target, and Costco have investments in inventory that are largely matched and paid for by accounts payable. Kohl's, notably, has the highest investment in inventory, which is not financed by payables. Because it includes more of such items as clothing, Kohl's inventory is also more seasonal than that of the other companies. Hence, Kohl's has the highest current ratio because much of its inventory must be financed by non-current liabilities.

Exhibit 6 Liquidity Ratios among Discount Retailers

	Company			
Ratio for January 2020 Fiscal Year	**Walmart**	**Target**	**Kohl's**	**Costco**
Current ratio	0.79	0.89	1.68	1.01
Quick ratio	0.22	0.27	0.40	0.52
Days of inventory on hand	41.0	59.1	98.0	30.8
Days of receivables	4.4	4.5	0.4	3.8
Days of payables outstanding	43.5	56.2	33.3	31.4
Cash conversion cycle	1.9	7.4	65.1	3.2

EXAMPLE 4

Measuring Liquidity

Given the following ratios, how well has the company been managing its liquidity?

	Current Fiscal Year		Average for the Previous Five Fiscal Years	
Ratio	**Company**	**Industry**	**Company**	**Industry**
Current ratio	1.9	2.5	1.1	2.3
Quick ratio	0.7	1.0	0.4	0.9
Days of receivables	39.0	34.0	44.0	32.5
Days of inventory on hand	41.0	30.3	45.0	27.4
Days of payables outstanding	34.3	36.0	29.4	35.5

Solution:

The ratios should be compared in two ways: over time (the trend over time would typically be examined) and with the trend in the industry. In all ratios shown here, the current year shows improvement over the previous years in terms of increased liquidity. In each case, however, the company remains behind the industry average in terms of liquidity. A brief snapshot such as this example could be the starting point in assessing management's ability to deliver future improvement and reach or beat the industry standards.

4

EVALUATING SHORT-TERM FINANCING CHOICES

d evaluate choices of short-term funding

Because so many short-term financing alternatives are available, the company must know how various financing instruments and markets can affect its liquidity position and risks. The costs of the financing alternatives can also be crucial.

When companies do not explore their options sufficiently, their liquidity might be inadequate, and they might not take advantage of cost savings that some forms of borrowing offer. This situation can arise if a company's treasurer is not familiar with the common forms of short-term borrowing or has not formulated an effective borrowing strategy. Given the various forms of short-term borrowing, having a planned strategy is essential for a borrower to avoid getting stuck in an uneconomical situation. If a company spends too little time establishing a sound strategy for its short-term borrowing, in a crisis, it might not be able to borrow at all, from any source.

The major objectives of a short-term borrowing strategy include the following:

- Ensuring that sufficient capacity exists to handle peak cash needs
- Maintaining sufficient sources of credit to be able to fund ongoing cash needs
- Ensuring that rates obtained are cost-effective and do not substantially exceed market averages

In addition, several factors will influence a company's short-term borrowing strategies, such as the following:

- *Size and creditworthiness.* A borrower's size can dictate the options available. Larger companies can take advantage of economies of scale to access commercial paper, banker's acceptances, and so on. The lender's size is also an important criterion, because larger banks have higher house or legal lending limits. The borrower's creditworthiness will determine the rate, compensation, or even whether the loan will be made at all.
- *Legal and regulatory considerations.* Some countries impose constraints on how much a company can borrow and under what terms it can borrow. Structure is usually greater for companies operating in developed countries with well-defined legal systems than for those operating in countries with emerging economies. Also, in developed countries, some industries are strongly regulated. Regulated companies, such as utilities and banks, might be restricted in how much they can borrow and the kind of borrowing they can engage in. Utilities typically have steady and predictable income streams and are often allowed to borrow larger amounts. Banks in the United States have come under greater scrutiny since the financial crisis and recession of 2007–2009. They are now required to have a greater amount of equity capital.
- *Sufficient access.* Borrowers should diversify to have adequate alternatives and not be too reliant on one lender or form of lending if the amount of their borrowing is very large. Even so, using one alternative primarily, but often with more than one provider, is typical for borrowers. Borrowers should be ready to go to other sources and know how to do so. Borrowers should not stay too long with just one source or with the lowest rates. Many borrowers are usually prepared (somewhat) to trade rates for certainty. Market power can help a company dictate more favorable terms for accounts payable, allowing it to delay payments without having to forgo discounts. Having marketable securities on hand that can be quickly sold to meet short-term needs is also helpful.
- *Flexibility of borrowing options.* Flexibility means the ability to manage maturities efficiently; that is, there should not be any "big" days, when significant amounts of loans mature. To effectively manage loan maturities, borrowers need active maturity management, awareness of the market conditions (e.g., knowing when the market or certain maturities should be avoided), and the ability to pre-pay loans when unexpected cash receipts occur.

Borrowing strategies, like investment strategies, can be either passive or active. Passive strategies usually involve minimal activity, with one source or type of borrowing and with little (if any) planning. A passive strategy is often reactive in responding to immediate needs for liquidity. Passive strategies are characterized by steady, often routine rollovers of borrowings for the same amount of funds each time, without much comparison shopping. Passive strategies might also arise when borrowing is restricted, such as when borrowers are limited to one or two lenders by agreement (e.g., in a secured loan arrangement).

Active strategies are usually more flexible and reflect planning, reliable forecasting, and comparison pricing. With active strategies, borrowers are more in control and do not fall into the rollover "trap" that is possible with passive strategies.

Many active strategies are matching strategies. Matching borrowing strategies function in a manner similar to matching investment strategies: loans are scheduled to mature when large cash receipts are expected. These receipts can pay back the loan, so the company does not have to invest the funds at potentially lower rates than the borrowing cost, thereby creating unnecessary costs.

EXAMPLE 5

Evaluating Short-Term Choices

When contemplating the possibilities for short-term financing needs, a company should consider which of the following items?

A Cost of the funds borrowed

B The flexibility offered by the source

C The ease with which the funds can be accessed

D Any legal or regulatory constraints that might favor one source over another

E All of the above

Solution

The correct answer is E. Cost of funds is the most obvious item to consider, but borrowing at a slightly higher cost sometimes makes sense when taking all the other items into consideration.

EXAMPLE 6

Meeting Short-Term Financing Need

The Keown Corp. has accounts payable of €2 million with terms of 2/10, net 30. Accounts receivable also stands at €2 million. In addition, the company has €5 million in marketable securities. Keown has a short-term need of €200,000 to meet payroll. Which of the following options makes the most sense for raising the €200,000?

A The company should issue long-term debt.

B The company should issue common stock.

C The company should delay paying accounts payable and forgo the 2% discount.

D The company should sell some of its accounts receivable to a factor at a 10% discount.

E The company should sell some of its marketable securities at a 0.5% brokerage cost.

Solution:

A and B would not be appropriate for raising €200,000 for a short-term need. These options take time to arrange, and they are more appropriate for long-term capital needs and for much larger financing amounts.

C, D, and E are all appropriate options for meeting short-term financing needs. However, C and D are costly.

The options for raising €200,000 are summarized in this table:

Source of Funds	Action	Liquidation Costs %	 €
C. Accounts payable (2/10, net 30)	Delay €200,000 in payment, and forgo 2% discount	2.0	4,000
D. Accounts receivable	Sell €222,222 in value at 10% discount to raise €200,000	10.0	22,222
E. Marketable securities	Sell €200,000 in value	0.5	1,000

Choosing C means forgoing a 2% discount, which on €200,000 amounts to a cost of €4,000. To net €200,000 using option D, the company would have to sell €222,222 of accounts receivable to a factor, representing a cost of €22,222. E appears to be the best choice. Marketable securities are liquid and can be easily sold for market value, less the relatively minor brokerage cost of €1,000.

SUMMARY

In this reading, we considered key aspects of capital alternatives and short-term financial management: the financing choices available to a company and effective liquidity management. Both are critical in ensuring a company's solvency and ability to remain in business. If done improperly, the results can be disastrous for the company.

This reading covered the following:

- Describing internal and external sources of capital and the considerations that lead to their selection
- Describing primary and secondary sources of liquidity and factors that can enhance a company's liquidity position
- Understanding how to evaluate a company's liquidity position and comparing it to peer companies
- Evaluating the short-term financing choices available to a company based on their characteristics and their effective costs.

PRACTICE PROBLEMS

1 Two analysts are discussing the costs of external financing sources. The first states that the company's bonds have a known interest rate but that the interest rate on accounts payable and the interest rate on equity financing are not specified. They are implicitly zero. Upon hearing this, the second analyst advocates financing the firm with greater amounts of accounts payable and common shareholders equity. Is the second analyst correct in his analysis?

A He is correct in his analysis of accounts payable only.

B He is correct in his analysis of common equity financing only.

C He is not correct in his analysis of either accounts payable or equity financing.

2 A company has arranged a $20 million line of credit with a bank, allowing the company the flexibility to borrow and repay any amount of funds as long as the balance does not exceed the line of credit. These arrangements are called:

A convertibles.

B factoring.

C revolvers.

3 SOA Company needs to raise 75 million, in local currency, for substantial new investments next year. Specific details, all in local currency, are as follows:

Investments of 10 million in receivables and 15 million in inventory. Fixed capital investments of 50 million, including 10 million to replace depreciated equipment and 40 million of net new investments.

Net income is expected to be 30 million, and dividend payments will be 12 million. Depreciation charges will be 10 million.

Short-term financing from accounts payable of 6 million is expected. The firm will use receivables as collateral for an 8 million loan. The firm will also issue a 14 million short-term note to a commercial bank.

Any additional external financing needed can be raised from an increase in long-term bonds. If additional financing is not needed, any excess funds will be used to repurchase common shares.

What additional financing does SOA require?

A SOA will need to issue 19 million of bonds.

B SOA will need to issue 26 million of bonds.

C SOA can repurchase 2 million of common shares.

4 Kwam Solutions must raise €120 million. Kwam has two primary sources of liquidity: €60 million of marketable securities (which can be sold with minimal liquidation/brokerage costs) and €30 million of bonds (which can be sold with 3% liquidation costs). Kwam can sell some or all of either of these portfolios. Kwam has a secondary source of liquidity, which would be to sell a large piece of real estate valued at €70 million (which would incur 10% liquidation costs). If Kwam sells the real estate, it must be sold entirely (a fractional sale is not possible). What is the lowest cost strategy for raising the needed €120 million?

A Sell €60 million of the marketable securities, €30 million of the bonds, and €34.3 million of the real estate property.

B Sell the real estate property and €50 million of the marketable securities.

© 2021 CFA Institute. All rights reserved.

C Sell the real estate property and €57 million of the marketable securities.

5 A company increasing its credit terms for customers from 1/10, net 30, to 1/10, net 60, will *most likely* experience:

A an increase in cash on hand.

B a lower level of uncollectible accounts.

C an increase in the average collection period.

6 Paloma Villarreal has received three suggestions from her staff about how to address her firm's liquidity problems.

Suggestion 1. Reduce the firm's inventory turnover rate.

Suggestion 2. Reduce the average collection period on accounts receivable.

Suggestion 3. Accelerate the payments on accounts payable by paying invoices before their due dates.

Which suggestion should Villarreal employ to improve the firm's liquidity position?

A Suggestion 1

B Suggestion 2

C Suggestion 3

7 Selected liquidity ratios for three firms in the leisure products industry are given in the table below. The most recent fiscal year ratio is shown, along with the average of the previous five years.

	Company H		Company J		Company S	
	Most recent	Five-year average	Most recent	Five-year average	Most recent	Five-year average
Current ratio	5.37	2.51	3.67	3.04	3.05	2.53
Quick ratio	5.01	2.19	2.60	2.01	1.78	1.44
Cash ratio	3.66	0.97	1.96	1.28	0.96	0.67

Relative to their peers and relative to their own prior performance, which company is in the most liquid position?

A Company H

B Company J

C Company S

8 An analyst is examining the cash conversion cycles and their components for three companies that she covers in the leisure products industry. She believes that changes in the investments in these working capital accounts can reveal liquidity stresses on a company.

	2021	2020	2019	2018	2017	2016
Company H						
Days of Inventory on Hand	68.4	70.5	60	57.8	59.8	59.8
+ Days of Receivables	101.8	103.4	95.6	92.4	94.7	93.3
– Days of Payables Outstanding	52.1	54.6	48	41.9	36.8	35.9
= Cash Conversion Cycle	118.1	119.3	107.6	108.3	117.7	117.2
Company J						

(continued)

	2021	2020	2019	2018	2017	2016
Days of Inventory on Hand	105.6	101.4	96.3	105.2	103.2	101.4
+ Days of Receivables	27.7	29.4	32.9	36.3	37.8	38
– Days of Payables Outstanding	36.6	38.5	35.3	39.3	37.8	40.2
= Cash Conversion Cycle	96.7	92.3	93.9	102.2	103.2	99.2
Company S						
Days of Inventory on Hand	135.8	131	118.9	69.2	63.4	81.7
+ Days of Receivables	49.1	42.5	54.2	36.2	29.1	38.3
– Days of Payables Outstanding	30.9	27.9	34.6	29.8	31.8	35.9
= Cash Conversion Cycle	154.0	145.6	138.5	75.6	60.7	84.1

Which company's operating cycle appears to have caused the most liquidity stress?

A Company H's

B Company J's

C Company S's

SOLUTIONS

1 C is correct. Although accounts payable do not charge an explicit interest rate, the cost of accounts payable is reflected in the costs of the services or products purchased and in the costs of any discounts not taken. Accounts payable can have a very high implicit cost. Similarly, equity financing is not free. A required return is expected on shareholder financing just as on any other form of financing.

2 C is correct. A revolver is a short-term borrowing facility in which a bank allows the firm to borrow and repay loans during the life of the line of credit.

3 A is correct. SOA must issue 19 million of bonds.

Source	Amount (local, millions)
Accounts payable	6
Bank loan against receivables	8
Short-term note	14
Net income + depreciation – dividends	28
Total sources	56

The firm requires 75 million of financing in local currency terms. Given the planned sources (before bond financing or repurchases) total 56 million, SOA will need to issue 19 million of new bonds.

4 C is correct. Kwam must sell the entire real estate property because the two primary sources (marketable securities and bonds) will not raise the needed €120 million. A is incorrect because it assumes a fractional real estate sale. The real estate sale will raise a net of €63 million (€70 million minus 10% liquidation expenses). To raise the rest of the funds needed (€120 million - €63 million = €57 million), Kwam can sell €57 million of marketable securities, which have minimal liquidation/brokerage costs.

5 C is correct. A longer average collection period will certainly occur. Higher cash balances and a lower level of uncollectible accounts will not occur.

6 B is correct. Reducing the average collection period would speed up receipts and improve the firm's liquidity position. The other two suggestions would worsen the firm's liquidity position.

7 A is correct. Relative to peers, Company H has the highest set of ratios. Relative to historical average ratios, Company H's recent ratios show the greatest increases. The cash ratio is the most relevant for judging liquidity, and Company H's cash ratio is quite high.

8 C is correct. Company S's cash conversion cycle nearly doubled over recent years, while the cash conversion cycles for Companies H and J are nearly unchanged. The days of inventory on hand and days of receivables both increased substantially for Company S, and its days of payables outstanding decreased very slightly. The net effect was the large increase in the cash conversion cycle. Although changes occurred in the components of the cash conversion cycles for Companies H and J, the net effect on their cash conversion cycles was small.

Glossary

A priori probability A probability based on logical analysis rather than on observation or personal judgment.

Abandonment option The ability to terminate an investment at some future time if the financial results are disappointing.

Abnormal return The amount by which a security's actual return differs from its expected return, given the security's risk and the market's return.

Absolute advantage A country's ability to produce a good or service at a lower absolute cost than its trading partner.

Absolute dispersion The amount of variability present without comparison to any reference point or benchmark.

Absolute frequency The actual number of observations counted for each unique value of the variable (also called raw frequency).

Accelerated book build An offering of securities by an investment bank acting as principal that is accomplished in only one or two days.

Accelerated methods Depreciation methods that allocate a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Accounting costs Monetary value of economic resources used in performing an activity. These can be explicit, out-of-pocket, current payments, or an allocation of historical payments (depreciation) for resources. They do not include implicit opportunity costs.

Accounting profit Income as reported on the income statement, in accordance with prevailing accounting standards, before the provisions for income tax expense. Also called *income before taxes* or *pretax income*.

Accounts payable Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Accrued expenses Liabilities related to expenses that have been incurred but not yet paid as of the end of an accounting period—an example of an accrued expense is rent that has been incurred but not yet paid, resulting in a liability "rent payable." Also called *accrued liabilities*.

Accrued interest Interest earned but not yet paid.

Acquisition method A method of accounting for a business combination where the acquirer is required to measure each identifiable asset and liability at fair value. This method was the result of a joint project of the IASB and FASB aiming at convergence in standards for the accounting of business combinations.

Action lag Delay from policy decisions to implementation.

Active investment An approach to investing in which the investor seeks to outperform a given benchmark.

Active return The return on a portfolio minus the return on the portfolio's benchmark.

Activity ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory. Also called *asset utilization ratios* or *operating efficiency ratios*.

Add-on rates Bank certificates of deposit, repos, and indexes such as Libor and Euribor are quoted on an add-on rate basis (bond equivalent yield basis).

Addition rule for probabilities A principle stating that the probability that *A* or *B* occurs (both occur) equals the probability that *A* occurs, plus the probability that *B* occurs, minus the probability that both *A* and *B* occur.

Agency bond See *quasi-government bond*.

Agency costs Costs associated with the conflict of interest present between principals and agents when a company is managed by non-owners. Agency costs result from the inherent conflicts of interest between managers, bondholders, and equity owners.

Agency costs of debt Costs arising from conflicts of interest between managers and debtholders.

Agency costs of equity The smaller the stake managers have in the company, the less their share in bearing the cost of excessive perquisite consumption—consequently, the less their desire to give their best efforts in running the company.

Agency RMBS In the United States, securities backed by residential mortgage loans and guaranteed by a federal agency or guaranteed by either of the two GSEs (Fannie Mae and Freddie Mac).

Aggregate demand The quantity of goods and services that households, businesses, government, and non-domestic customers want to buy at any given level of prices.

Aggregate demand curve Inverse relationship between the price level and real output.

Aggregate income The value of all the payments earned by the suppliers of factors used in the production of goods and services.

Aggregate output The value of all the goods and services produced during a specified period.

Aggregate supply The quantity of goods and services producers are willing to supply at any given level of price.

Aggregate supply curve The level of domestic output that companies will produce at each price level.

All-or-nothing (AON) orders An order that includes the instruction to trade only if the trade fills the entire quantity (size) specified.

Allocationally efficient A characteristic of a market, a financial system, or an economy that promotes the allocation of resources to their highest value uses.

Alternative data Non-traditional data types generated by the use of electronic devices, social media, satellite and sensor networks, and company exhaust.

Alternative hypothesis The hypothesis that is accepted if the null hypothesis is rejected.

Alternative investment markets Market for investments other than traditional securities investments (i.e., traditional common and preferred shares and traditional fixed income instruments). The term usually encompasses direct and indirect investment in real estate (including timberland and farmland) and commodities (including precious metals); hedge funds, private equity, and other investments requiring specialized due diligence.

Alternative trading systems Trading venues that function like exchanges but that do not exercise regulatory authority over their subscribers except with respect to the conduct

of the subscribers' trading in their trading systems. Also called *electronic communications networks* or *multilateral trading facilities*.

American depository receipt A US dollar-denominated security that trades like a common share on US exchanges.

American depository share The underlying shares on which American depository receipts are based. They trade in the issuing company's domestic market.

American-style Type of option contract that can be exercised at any time up to the option's expiration date.

Amortisation The process of allocating the cost of intangible long-term assets having a finite useful life to accounting periods; the allocation of the amount of a bond premium or discount to the periods remaining until bond maturity.

Amortised cost The historical cost (initially recognised cost) of an asset, adjusted for amortisation and impairment.

Amortizing bond Bond with a payment schedule that calls for periodic payments of interest and repayments of principal.

Amortizing loans Loans with a payment schedule that calls for periodic payments of interest and repayments of principal.

Analysis of variance (ANOVA) The analysis that breaks the total variability of a dataset (such as observations on the dependent variable in a regression) into components representing different sources of variation. With reference to regression, ANOVA provides the inputs for an *F*-test of the significance of the regression as a whole, as well as the inputs for the coefficient of determination and the standard error of the estimate.

Analytical duration The use of mathematical formulas to estimate the impact of benchmark yield-to-maturity changes on bond prices.

Anchoring and adjustment bias An information-processing bias in which the use of a psychological heuristic influences the way people estimate probabilities.

Annual percentage rate The cost of borrowing expressed as a yearly rate.

Annuity A finite set of level sequential cash flows.

Annuity due An annuity having a first cash flow that is paid immediately.

Anomalies Apparent deviations from market efficiency.

Antidilutive With reference to a transaction or a security, one that would increase earnings per share (EPS) or result in EPS higher than the company's basic EPS—antidilutive securities are not included in the calculation of diluted EPS.

Arbitrage 1) The simultaneous purchase of an undervalued asset or portfolio and sale of an overvalued but equivalent asset or portfolio, in order to obtain a riskless profit on the price differential. Taking advantage of a market inefficiency in a risk-free manner. 2) The condition in a financial market in which equivalent assets or combinations of assets sell for two different prices, creating an opportunity to profit at no risk with no commitment of money. In a well-functioning financial market, few arbitrage opportunities are possible. 3) A risk-free operation that earns an expected positive net profit but requires no net investment of money.

Arbitrage-free pricing The overall process of pricing derivatives by arbitrage and risk neutrality. Also called the *principle of no arbitrage*.

Arbitrageurs Traders who engage in arbitrage. See *arbitrage*.

Arithmetic mean The sum of the observations divided by the number of observations.

Artificial intelligence Computer systems that exhibit cognitive and decision-making ability comparable (or superior) to that of humans.

Ask The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size). See *offer*.

Ask size The maximum quantity of an asset that pertains to a specific ask price from a trader. For example, if the ask for a share issue is $30 for a size of 1,000 shares, the trader is offering to sell at $30 up to 1,000 shares.

Asset allocation The process of determining how investment funds should be distributed among asset classes.

Asset-backed securities A type of bond issued by a legal entity called a *special purpose entity* (SPE) on a collection of assets that the SPE owns. Also, securities backed by receivables and loans other than mortgages.

Asset-based valuation models Valuation based on estimates of the market value of a company's assets.

Asset beta The unlevered beta; which reflects the business risk of the assets; the assets' systematic risk.

Asset class A group of assets that have similar characteristics, attributes, and risk–return relationships.

Asset swap Converts the periodic fixed coupon of a specific bond to an MRR plus or minus a spread.

Asset utilization ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Assets Resources controlled by an enterprise as a result of past events and from which future economic benefits to the enterprise are expected to flow.

Assignment of accounts receivable The use of accounts receivable as collateral for a loan.

Asymmetric information The differential of information between corporate insiders and outsiders regarding the company's performance and prospects. Managers typically have more information about the company's performance and prospects than owners and creditors.

At the money An option in which the underlying's price equals the exercise price.

Auction A type of bond issuing mechanism often used for sovereign bonds that involves bidding.

Autarkic price The price of a good or service in an autarkic economy.

Autarky A state in which a country does not trade with other countries.

Automatic stabilizer A countercyclical factor that automatically comes into play as an economy slows and unemployment rises.

Availability bias An information-processing bias in which people take a heuristic approach to estimating the probability of an outcome based on how easily the outcome comes to mind.

Available-for-sale Under US GAAP, debt securities not classified as either held-to-maturity or held-for-trading securities. The investor is willing to sell but not actively planning to sell. In general, available-for-sale debt securities are reported at fair value on the balance sheet, with unrealized gains included as a component of other comprehensive income.

Average fixed cost Total fixed cost divided by quantity produced.

Average life See *weighted average life*.

Average product Measures the productivity of inputs on average and is calculated by dividing total product by the total number of units for a given input that is used to generate that output.

Average revenue Total revenue divided by quantity sold.

Average total cost Total cost divided by quantity produced.

Average variable cost Total variable cost divided by quantity produced.

Back-testing With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backfill bias A bias introduced if a fund backfills its past performance data to the database when it first starts reporting actual performance.

Backtesting With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backup line of credit A type of credit enhancement provided by a bank to an issuer of commercial paper to ensure that the issuer will have access to sufficient liquidity to repay maturing commercial paper if issuing new paper is not a viable option.

Backwardation A condition in the futures markets in which the spot price exceeds the futures price, the forward curve is downward sloping, and the convenience yield is high.

Balance of payments A double-entry bookkeeping system that summarizes a country's economic transactions with the rest of the world for a particular period of time, typically a calendar quarter or year.

Balance of trade deficit When the domestic economy is spending more on non-domestic goods and services than non-domestic economies are spending on domestic goods and services.

Balance sheet The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet). Also called *statement of financial position* or *statement of financial condition.*

Balance sheet ratios Financial ratios involving balance sheet items only.

Balanced With respect to a government budget, one in which spending and revenues (taxes) are equal.

Balloon payment Large payment required at maturity to retire a bond's outstanding principal amount.

Bar chart A chart for plotting the frequency distribution of categorical data, where each bar represents a distinct category and each bar's height is proportional to the frequency of the corresponding category. In technical analysis, a bar chart that plots four bits of data for each time interval—the high, low, opening, and closing prices. A vertical line connects the high and low prices. A cross-hatch left indicates the opening price and a cross-hatch right indicates the closing price.

Barter economy An economy where economic agents as house-holds, corporations, and governments "pay" for goods and services with another good or service.

Base-rate neglect A type of representativeness bias in which the base rate or probability of the categorization is not adequately considered.

Base rates The reference rate on which a bank bases lending rates to all other customers.

Basic EPS Net earnings available to common shareholders (i.e., net income minus preferred dividends) divided by the weighted average number of common shares outstanding.

Basis point Used in stating yield spreads, one basis point equals one-hundredth of a percentage point, or 0.01%.

Basket of listed depository receipts An exchange-traded fund (ETF) that represents a portfolio of depository receipts.

Bayes' formula The rule for updating the probability of an event of interest—given a set of prior probabilities for the event, information, and information given the event—if you receive new information.

Bearer bonds Bonds for which ownership is not recorded; only the clearing system knows who the bond owner is.

Bearish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from above; this movement is considered bearish. A **death cross** is a bearish crossover based on 50-day and 200-day moving averages.

Behavioral finance A field of finance that examines the psychological variables that affect and often distort the investment decision making of investors, analysts, and portfolio managers.

Behind the market Said of prices specified in orders that are worse than the best current price; e.g., for a limit buy order, a limit price below the best bid.

Benchmark A comparison portfolio; a point of reference or comparison.

Benchmark issue The latest sovereign bond issue for a given maturity. It serves as a benchmark against which to compare bonds that have the same features but that are issued by another type of issuer.

Benchmark rate Typically the yield-to-maturity on a government bond having the same or close to the same time-to-maturity.

Benchmark spread The yield spread over a specific benchmark, usually measured in basis points.

Bernoulli random variable A random variable having the outcomes 0 and 1.

Bernoulli trial An experiment that can produce one of two outcomes.

Best bid The highest bid in the market.

Best effort offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-efforts offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-in-class An ESG implementation approach that seeks to identify the most favorable companies in an industry based on ESG considerations.

Best offer The lowest offer (ask price) in the market.

Beta A measure of the sensitivity of a given investment or portfolio to movements in the overall market.

Bid The price at which a dealer or trader is willing to buy an asset, typically qualified by a maximum quantity.

Bid–ask spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid–offer spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid size The maximum quantity of an asset that pertains to a specific bid price from a trader.

Big Data The vast amount of data being generated by industry, governments, individuals, and electronic devices that arises from both traditional and non-traditional data sources.

Bilateral loan A loan from a single lender to a single borrower.

Bimodal A distribution that has two most frequently occurring values.

Binomial model A model for pricing options in which the underlying price can move to only one of two possible new prices.

Binomial random variable The number of successes in n Bernoulli trials for which the probability of success is constant for all trials and the trials are independent.

Binomial tree The graphical representation of a model of asset price dynamics in which, at each period, the asset moves up with probability p or down with probability $(1 - p)$.

Bitcoin A cryptocurrency using blockchain technology that was created in 2009.

Bivariate correlation See *Pearson correlation*.

Block brokers A broker (agent) that provides brokerage services for large-size trades.

Blockchain A type of digital ledger in which information is recorded sequentially and then linked together and secured using cryptographic methods.

Blue chip Widely held large market capitalization companies that are considered financially sound and are leaders in their respective industry or local stock market.

Bollinger Bands A price-based technical analysis indicator consisting of a line representing the moving average, a higher line representing the moving average plus a set number of standard deviations from the average (for the same number of periods as was used to calculate the moving average), and a lower line representing the moving average minus the same number of standard deviations.

Bond Contractual agreement between the issuer and the bondholders.

Bond equivalent yield A calculation of yield that is annualized using the ratio of 365 to the number of days to maturity. Bond equivalent yield allows for the restatement and comparison of securities with different compounding periods.

Bond indenture The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *trust deed*.

Bond market vigilantes Bond market participants who might reduce their demand for long-term bonds, thus pushing up their yields.

Bond yield plus risk premium approach An estimate of the cost of common equity that is produced by summing the before-tax cost of debt and a risk premium that captures the additional yield on a company's stock relative to its bonds. The additional yield is often estimated using historical spreads between bond yields and stock yields.

Bonus issue of shares A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Book building Investment bankers' process of compiling a "book" or list of indications of interest to buy part of an offering.

Book value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. Also called *carrying value*.

Boom An expansionary phase characterized by economic growth "testing the limits" of the economy.

Bootstrap A resampling method that repeatedly draws samples with replacement of the selected elements from the original observed sample. Bootstrap is usually conducted by using computer simulation and is often used to find standard error or construct confidence intervals of population parameters.

Bottom-up analysis An investment selection approach that focuses on company-specific circumstances rather than emphasizing economic cycles or industry analysis.

Box and whisker plot A graphic for visualizing the dispersion of data across quartiles. It consists of a "box" with "whiskers" connected to the box.

Breakdown A breakdown occurs when the price of an asset moves below a support level.

Breakeven point The number of units produced and sold at which the company's net income is zero (Revenues = Total cost); in the case of perfect competition, the quantity at which price, average revenue, and marginal revenue equal average total cost.

Breakout A breakout occurs when the price of an asset moves above a resistance level.

Bridge financing Interim financing that provides funds until permanent financing can be arranged.

Broad money Encompasses narrow money plus the entire range of liquid assets that can be used to make purchases.

Broker 1) An agent who executes orders to buy or sell securities on behalf of a client in exchange for a commission. 2) See *futures commission merchants*.

Broker–dealer A financial intermediary (often a company) that may function as a principal (dealer) or as an agent (broker) depending on the type of trade.

Brokered market A market in which brokers arrange trades among their clients.

Brownfield investment Investing in *existing* infrastructure assets.

Bubble line chart A line chart that uses varying-sized bubbles to represent a third dimension of the data. The bubbles are sometimes color-coded to present additional information.

Budget surplus/deficit The difference between government revenue and expenditure for a stated fixed period of time.

Bullet bond Bond in which the principal repayment is made entirely at maturity.

Bullish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from below; this movement is considered bullish. A **golden cross** is a bullish crossover based on 50-day and 200-day moving averages.

Business cycles Are recurrent expansions and contractions in economic activity affecting broad segments of the economy.

Business risk The risk associated with operating earnings. Operating earnings are uncertain because total revenues and many of the expenditures contributed to produce those revenues are uncertain.

Buy-side firm An investment management company or other investor that uses the services of brokers or dealers (i.e., the client of the sell side firms).

Buyback A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Call An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call market A market in which trades occur only at a particular time and place (i.e., when the market is called).

Call money rate The interest rate that buyers pay for their margin loan.

Call option An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call protection The time during which the issuer of the bond is not allowed to exercise the call option.

Callable bond A bond containing an embedded call option that gives the issuer the right to buy the bond back from the investor at specified prices on pre-determined dates.

Calmar ratio A comparison of the average annual compounded return to the maximum drawdown risk.

Candlestick chart A price chart with four bits of data for each time interval. A candle indicates the opening and closing price for the interval. The body of the candle is shaded if the opening price was higher than the closing price, and the body is white (or clear) if the opening price was lower than the closing price. Vertical lines known as wicks or shadows extend from the top and bottom of the candle to indicate, respectively, the high and low prices for the interval.

Cannibalization Cannibalization occurs when an investment takes customers and sales away from another part of the company.

Cap rate A metric by which real estate managers are often judged; the annual rent actually earned (net of any vacancies) divided by the price originally paid for the property.

Capacity The ability of the borrower to make its debt payments on time.

Capital account A component of the balance of payments account that measures transfers of capital.

Capital allocation line (CAL) A graph line that describes the combinations of expected return and standard deviation of return available to an investor from combining the optimal portfolio of risky assets with the risk-free asset.

Capital allocation process The process that companies use for decision making on capital investments—those projects with a life of one year or longer.

Capital asset pricing model (CAPM) An equation describing the expected return on any asset (or portfolio) as a linear function of its beta relative to the market portfolio.

Capital consumption allowance A measure of the wear and tear (depreciation) of the capital stock that occurs in the production of goods and services.

Capital deepening investment Increases the stock of capital relative to labor.

Capital expenditure Expenditure on physical capital (fixed assets).

Capital light Capital light businesses require little incremental investment in fixed assets or working capital to enable revenue growth.

Capital loss ratio The percentage of capital in deals that have been realized below cost, net of any recovered proceeds, divided by total invested capital.

Capital market expectations An investor's expectations concerning the risk and return prospects of asset classes.

Capital market line (CML) The line with an intercept point equal to the risk-free rate that is tangent to the efficient frontier of risky assets; represents the efficient frontier when a risk-free asset is available for investment.

Capital market securities Securities with maturities at issuance longer than one year.

Capital markets Financial markets that trade securities of longer duration, such as bonds and equities.

Capital rationing A capital rationing environment assumes that the company has a fixed amount of funds to invest.

Capital restrictions Controls placed on foreigners' ability to own domestic assets and/or domestic residents' ability to own foreign assets.

Capital stock The accumulated amount of buildings, machinery, and equipment used to produce goods and services.

Capital structure The mix of debt and equity that a company uses to finance its business; a company's specific mixture of long-term financing.

Carry The net of the costs and benefits of holding, storing, or "carrying" an asset.

Carrying amount The amount at which an asset or liability is valued according to accounting principles.

Carrying value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. For a bond, the purchase price plus (or minus) the amortized amount of the discount (or premium).

Cartel Participants in collusive agreements that are made openly and formally.

Cash collateral account Form of external credit enhancement whereby the issuer immediately borrows the credit-enhancement amount and then invests that amount, usually in highly rated short-term commercial paper.

Cash conversion cycle A financial metric that measures the length of time required for a company to convert cash invested in its operations to cash received as a result of its operations; equal to days of inventory on hand + days of sales outstanding – number of days of payables. Also called *net operating cycle*.

Cash flow additivity principle The principle that dollar amounts indexed at the same point in time are additive.

Cash flow from operating activities The net amount of cash provided from operating activities.

Cash flow from operations The net amount of cash provided from operating activities.

Cash flow yield The internal rate of return on a series of cash flows.

Cash market securities Money market securities settled on a "same day" or "cash settlement" basis.

Cash markets See *spot markets*.

Cash prices See *spot prices*.

Cash-settled forwards See *non-deliverable forwards*.

Catch-up clause A clause in an agreement that favors the GP. For a GP who earns a 20% performance fee, a catch-up clause allows the GP to receive 100% of the distributions above the hurdle rate *until* she receives 20% of the profits generated, and then every excess dollar is split 80/20 between the LPs and GP.

Categorical data Values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize (also called **qualitative data**).

CBOE Volatility Index (VIX) A measure of near-term market volatility as conveyed by S&P 500 stock index option prices.

Central bank funds market The market in which deposit-taking banks that have an excess reserve with their national central bank can lend money to banks that need funds for maturities ranging from overnight to one year. Called the Federal or Fed funds market in the United States.

Central bank funds rate Interest rate at which central bank funds are bought (borrowed) and sold (lent) for maturities ranging from overnight to one year. Called Federal or Fed funds rate in the United States.

Central banks The dominant bank in a country, usually with official or semi-official governmental status.

Central limit theorem The theorem that states the sum (and the mean) of a set of independent, identically distributed random variables with finite variances is normally distributed, whatever distribution the random variables follow.

Certificate of deposit An instrument that represents a specified amount of funds on deposit with a bank for a specified maturity and interest rate. CDs are issued in various denominations and can be negotiable or non-negotiable.

Change in polarity principle A tenet of technical analysis that states that once a support level is breached, it becomes a resistance level. The same holds true for resistance levels: Once breached, they become support levels.

Change of control put A covenant giving bondholders the right to require the issuer to buy back their debt, often at par or at some small premium to par value, in the event that the borrower is acquired.

Character The quality of a debt issuer's management.

Chartist An individual who uses charts or graphs of a security's historical prices or levels to forecast its future trends.

Chi-square test of independence A statistical test for detecting a potential association between categorical variables.

Classical cycle Refers to fluctuations in the level of economic activity when measured by GDP in volume terms.

Classified balance sheet A balance sheet organized so as to group together the various assets and liabilities into subcategories (e.g., current and noncurrent).

Clawback A requirement that the general partner return any funds distributed as incentive fees until the limited partners have received back their initial investment and a percentage of the total profit.

Clearing The process by which the exchange verifies the execution of a transaction and records the participants' identities.

Clearing instructions Instructions that indicate how to arrange the final settlement ("clearing") of a trade.

Clearinghouse An entity associated with a futures market that acts as middleman between the contracting parties and guarantees to each party the performance of the other.

Closed economy An economy that does not trade with other countries; an *autarkic economy*.

Closed-end fund A mutual fund in which no new investment money is accepted. New investors invest by buying existing shares, and investors in the fund liquidate by selling their shares to other investors.

Cluster sampling A procedure that divides a population into subpopulation groups (clusters) representative of the population and then randomly draws certain clusters to form a sample.

Clustered bar chart See *grouped bar chart*.

Co-investing In co-investing, the investor invests in assets *indirectly* through the fund but also possesses rights (known as co-investment rights) to invest *directly* in the same assets. Through co-investing, an investor is able to make an investment *alongside* a fund when the fund identifies deals.

Code of ethics An established guide that communicates an organization's values and overall expectations regarding member behavior. A code of ethics serves as a general guide for how community members should act.

Coefficient of determination The percentage of the variation of the dependent variable that is explained by the independent variable. Also referred to as the "R-squared" or "R^2."

Coefficient of variation The ratio of a set of observations' standard deviation to the observations' mean value.

Cognitive cost The effort involved in processing new information and updating beliefs.

Cognitive dissonance The mental discomfort that occurs when new information conflicts with previously held beliefs or cognitions.

Cognitive errors Behavioral biases resulting from faulty reasoning; cognitive errors stem from basic statistical, information-processing, or memory errors.

Coincident economic indicators Turning points that are usually close to those of the overall economy; they are believed to have value for identifying the economy's present state.

Collateral Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Collateral manager Buys and sells debt obligations for and from the CDO's portfolio of assets (i.e., the collateral) to generate sufficient cash flows to meet the obligations to the CDO bondholders.

Collateral trust bonds Bonds secured by securities, such as common shares, other bonds, or other financial assets.

Collateralized bond obligations A structured asset-backed security that is collateralized by a pool of bonds.

Collateralized debt obligation Generic term used to describe a security backed by a diversified pool of one or more debt obligations.

Collateralized loan obligations A structured asset-backed security that is collateralized by a pool of loans.

Collateralized mortgage obligation A security created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collateralized mortgage obligations Securities created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collaterals Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Combination A listing in which the order of the listed items does not matter.

Combination formula (binomial formula) The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is ${}_nC_r = \binom{n}{r} = \dfrac{n!}{(n-r)!r!}$.

Commercial paper A short-term, negotiable, unsecured promissory note that represents a debt obligation of the issuer.

Committed (Regular) lines of credit A bank commitment to extend credit up to a pre-specified amount; the commitment is considered a short-term liability and is usually in effect for 364 days (one day short of a full year).

Committed capital The amount that the limited partners have agreed to provide to the private equity fund.

Commodity swap A swap in which the underlying is a commodity such as oil, gold, or an agricultural product.

Common market Level of economic integration that incorporates all aspects of the customs union and extends it by allowing free movement of factors of production among members.

Common shares A type of security that represent an ownership interest in a company.

Common-size analysis The restatement of financial statement items using a common denominator or reference item that allows one to identify trends and major differences; an example is an income statement in which all items are expressed as a percent of revenue.

Common stock See *common shares.*

Company analysis Analysis of an individual company.

Comparable company A company that has similar business risk, usually in the same industry and preferably with a single line of business.

Comparative advantage A country's ability to produce a good or service at a lower relative cost, or opportunity cost, than its trading partner.

Competitive strategy A company's plans for responding to the threats and opportunities presented by the external environment.

Complement The event not-S, written S^C, given the event S. Note that $P(S) + P(S^C) = 1$.

Complements Goods that tend to be used together; technically, two goods whose cross-price elasticity of demand is negative.

Complete markets Informally, markets in which the variety of distinct securities traded is so broad that any desired payoff in a future state-of-the-world is achievable.

Component cost of capital The rate of return required by suppliers of capital for an individual source of a company's funding, such as debt or equity.

Compounding The process of accumulating interest on interest.

Comprehensive income The change in equity of a business enterprise during a period from nonowner sources; includes all changes in equity during a period except those resulting from investments by owners and distributions to owners; comprehensive income equals net income plus other comprehensive income.

Conditional expected value The expected value of a stated event given that another event has occurred.

Conditional pass-through covered bonds Covered bonds that convert to pass-through securities after the original maturity date if all bond payments have not yet been made and the sponsor is in default.

Conditional probability The probability of an event given (conditioned on) another event.

Conditional variances The variance of one variable, given the outcome of another.

Confidence level The complement of the level of significance.

Confirmation bias A belief perseverance bias in which people tend to look for and notice what confirms their beliefs, to ignore or undervalue what contradicts their beliefs, and to misinterpret information as support for their beliefs.

Confusion matrix A type of contingency table used for evaluating the performance of a classification model.

Conservatism bias A belief perseverance bias in which people maintain their prior views or forecasts by inadequately incorporating new information.

Consolidation The movement of a stock's price within a well-defined range of trading levels for a period of time. The price consolidates between a support level and a resistance level.

Constant-yield price trajectory A graph that illustrates the change in the price of a fixed-income bond over time assuming no change in yield-to-maturity. The trajectory shows the "pull to par" effect on the price of a bond trading at a premium or a discount to par value.

Constituent securities With respect to an index, the individual securities within an index.

Consumer surplus The difference between the value that a consumer places on units purchased and the amount of money that was required to pay for them.

Contango A condition in the futures markets in which the spot price is lower than the futures price, the forward curve is upward sloping, and there is little or no convenience yield.

Contingency provision Clause in a legal document that allows for some action if a specific event or circumstance occurs.

Contingency table A tabular format that displays the frequency distributions of two or more categorical variables simultaneously and is used for finding patterns between the variables. A contingency table for two categorical variables is also known as a two-way table.

Contingent claim A security whose value is based on some other underlying security.

Contingent claims Derivatives in which the payoffs occur if a specific event occurs; generally referred to as options.

Contingent convertible bonds Bonds that automatically convert into equity if a specific event or circumstance occurs, such as the issuer's equity capital falling below the minimum requirement set by the regulators. Also called *CoCos.*

Continuation pattern A type of pattern used in technical analysis to predict the resumption of a market trend that was in place prior to the formation of a pattern.

Continuous data Data that can be measured and can take on any numerical value in a specified range of values.

Continuous random variable A random variable for which the range of possible outcomes is the real line (all real numbers between $-\infty$ and $+\infty$) or some subset of the real line.

Continuous trading market A market in which trades can be arranged and executed any time the market is open.

Continuously compounded return The natural logarithm of 1 plus the holding period return, or equivalently, the natural logarithm of the ending price over the beginning price.

Contra account An account that offsets another account.

Contract rate See *mortgage rate.*

Contraction The period of a business cycle after the peak and before the trough; often called a *recession* or, if exceptionally severe, called a *depression*.

Contraction risk The risk that when interest rates decline, the security will have a shorter maturity than was anticipated at the time of purchase because borrowers refinance at the new, lower interest rates.

Contractionary Tending to cause the real economy to contract.

Contractionary fiscal policy A fiscal policy that has the objective to make the real economy contract.

Contracts for differences See *non-deliverable forwards.*

Contribution margin The amount available for fixed costs and profit after paying variable costs; revenue minus variable costs.

Controlling shareholders A particular shareholder or block of shareholders holding a percentage of shares that gives them significant voting power.

Convenience sampling A procedure of selecting an element from a population on the basis of whether or not it is accessible to a researcher or how easy it is for a researcher to access the element.

Convenience yield A non-monetary advantage of holding an asset.

Conventional bond See *plain vanilla bond*.

Conventional cash flows A conventional cash flow pattern is one with an initial outflow followed by a series of inflows.

Convergence The tendency for differences in output per capita across countries to diminish over time. In technical analysis, the term describes the case when an indicator moves in the same manner as the security being analyzed.

Conversion premium The difference between the convertible bond's price and its conversion value.

Conversion price For a convertible bond, the price per share at which the bond can be converted into shares.

Conversion ratio For a convertible bond, the number of common shares that each bond can be converted into.

Conversion value For a convertible bond, the current share price multiplied by the conversion ratio.

Convertible bond Bond that gives the bondholder the right to exchange the bond for a specified number of common shares in the issuing company.

Convertible preference shares A type of equity security that entitles shareholders to convert their shares into a specified number of common shares.

Convexity adjustment For a bond, one half of the annual or approximate convexity statistic multiplied by the change in the yield-to-maturity squared.

Core inflation Refers to the inflation rate calculated based on a price index of goods and services except food and energy.

Corporate governance The system of internal controls and procedures by which individual companies are managed.

Correlation A measure of the linear relationship between two random variables.

Correlation coefficient A number between −1 and +1 that measures the consistency or tendency for two investments to act in a similar way. It is used to determine the effect on portfolio risk when two assets are combined.

Cost averaging The periodic investment of a fixed amount of money.

Cost of capital The cost of financing to a company; the rate of return that suppliers of capital require as compensation for their contribution of capital.

Cost of carry See *carry*.

Cost of debt The cost of debt financing to a company, such as when it issues a bond or takes out a bank loan; the rate of return required by a company's bondholders.

Cost of equity The rate of return required by a company's common stockholders.

Cost of preferred stock The cost to a company of issuing preferred stock; the dividend yield that a company must commit to pay preferred stockholders.

Cost structure The mix of a company's variable costs and fixed costs.

Counterparty risk The risk that the other party to a contract will fail to honor the terms of the contract.

Coupon rate The interest rate promised in a contract; this is the rate used to calculate the periodic interest payments.

Cournot assumption Assumption in which each firm determines its profit-maximizing production level assuming that the other firms' output will not change.

Covariance A measure of the co-movement (linear association) between two random variables.

Covariance matrix A matrix or square array whose entries are covariances; also known as a variance–covariance matrix.

Covenants The terms and conditions of lending agreements that the issuer must comply with; they specify the actions that an issuer is obligated to perform (affirmative covenant) or prohibited from performing (negative covenant).

Covered bond Debt obligation secured by a segregated pool of assets called the cover pool. The issuer must maintain the value of the cover pool. In the event of default, bondholders have recourse against both the issuer and the cover pool.

Covered bonds Senior debt obligations issued by a financial institution and backed by a segregated pool of assets. Covered bond investors have recourse to both the underlying assets and the financial institution.

Credit analysis The evaluation of credit risk; the evaluation of the creditworthiness of a borrower or counterparty.

Credit default swap (CDS) A type of credit derivative in which one party, the credit protection buyer who is seeking credit protection against a third party, makes a series of regularly scheduled payments to the other party, the credit protection seller. The seller makes no payments until a credit event occurs.

Credit derivatives A contract in which one party has the right to claim a payment from another party in the event that a specific credit event occurs over the life of the contract.

Credit enhancements Provisions that may be used to reduce the credit risk of a bond issue.

Credit-linked coupon bond Bond for which the coupon changes when the bond's credit rating changes.

Credit-linked note (CLN) Fixed-income security in which the holder of the security has the right to withhold payment of the full amount due at maturity if a credit event occurs.

Credit migration risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *downgrade risk*.

Credit risk The risk of loss caused by a counterparty's or debtor's failure to make a promised payment. Also called *default risk*.

Credit spread option An option on the yield spread on a bond.

Credit tranching A structure used to redistribute the credit risk associated with the collateral; a set of bond classes created to allow investors a choice in the amount of credit risk that they prefer to bear.

Creditworthiness The perceived ability of the borrower to pay its debt obligations in a timely manner; it represents the ability of a company to withstand adverse impacts on its cash flows.

Critical values Values of the test statistic at which the decision changes from fail to reject the null hypothesis to reject the null hypothesis.

Cross-default Covenant or contract clause that specifies a borrower is considered in default if they default on another debt obligation.

Cross-default provisions Provisions whereby events of default, such as non-payment of interest on one bond, trigger default on all outstanding debt; implies the same default probability for all issues.

Cross-price elasticity of demand The percentage change in quantity demanded for a given percentage change in the price of another good; the responsiveness of the demand for Product A that is associated with the change in price of Product B.

Cross-sectional analysis Analysis that involves comparisons across individuals in a group over a given time period or at a given point in time.

Cross-sectional data A list of the observations of a specific variable from multiple observational units at a given point in time. The observational units can be individuals, groups, companies, trading markets, regions, etc.

Crossing networks Trading systems that match buyers and sellers who are willing to trade at prices obtained from other markets.

Crowding out The thesis that government borrowing may divert private sector investment from taking place.

Cryptocurrency An electronic medium of exchange that lacks physical form.

Cryptography An algorithmic process to encrypt data, making the data unusable if received by unauthorized parties.

Cumulative absolute frequency Cumulates (i.e., adds up) in a frequency distribution the absolute frequencies as one moves from the first bin to the last bin.

Cumulative distribution function A function giving the probability that a random variable is less than or equal to a specified value.

Cumulative frequency distribution chart A chart that plots either the cumulative absolute frequency or the cumulative relative frequency on the y-axis against the upper limit of the interval and allows one to see the number or the percentage of the observations that lie below a certain value.

Cumulative preference shares Preference shares for which any dividends that are not paid accrue and must be paid in full before dividends on common shares can be paid.

Cumulative relative frequency A sequence of partial sums of the relative frequencies in a frequency distribution.

Cumulative voting A voting process whereby shareholders can accumulate and vote all their shares for a single candidate in an election, as opposed to having to allocate their voting rights evenly among all candidates.

Currencies Monies issued by national monetary authorities.

Currency option bonds Bonds that give bondholders the right to choose the currency in which they want to receive interest payments and principal repayments.

Currency swap A swap in which each party makes interest payments to the other in different currencies.

Current account A component of the balance of payments account that measures the flow of goods and services.

Current assets Assets that are expected to be consumed or converted into cash in the near future, typically one year or less. *Also called liquid assets.*

Current cost With reference to assets, the amount of cash or cash equivalents that would have to be paid to buy the same or an equivalent asset today; with reference to liabilities, the undiscounted amount of cash or cash equivalents that would be required to settle the obligation today.

Current government spending With respect to government expenditures, spending on goods and services that are provided on a regular, recurring basis including health, education, and defense.

Current liabilities Short-term obligations, such as accounts payable, wages payable, or accrued liabilities, that are expected to be settled in the near future, typically one year or less.

Current yield The sum of the coupon payments received over the year divided by the flat price; also called the *income* or *interest yield* or *running yield*.

Curve duration The sensitivity of the bond price (or the market value of a financial asset or liability) with respect to a benchmark yield curve.

Customs union Extends the free trade area (FTA) by not only allowing free movement of goods and services among members, but also creating a common trade policy against nonmembers.

CVaR Conditional VaR, a tail loss measure. The weighted average of all loss outcomes in the statistical distribution that exceed the VaR loss.

Cyclical A cyclical company is one whose profits are strongly correlated with the strength of the overall economy.

Cyclical companies Companies with sales and profits that regularly expand and contract with the business cycle or state of economy.

Daily settlement See *mark to market* and *marking to market.*

Dark pools Alternative trading systems that do not display the orders that their clients send to them.

Data A collection of numbers, characters, words, and text—as well as images, audio, and video—in a raw or organized format to represent facts or information.

Data mining The practice of determining a model by extensive searching through a dataset for statistically significant patterns. Also called *data snooping.*

Data science An interdisciplinary field that brings computer science, statistics, and other disciplines together to analyze and produce insights from Big Data.

Data snooping The practice of determining a model by extensive searching through a dataset for statistically significant patterns.

Data table see **two-dimensional rectangular array**.

Day order An order that is good for the day on which it is submitted. If it has not been filled by the close of business, the order expires unfilled.

Days of inventory on hand An activity ratio equal to the number of days in the period divided by inventory turnover over the period.

Dealers A financial intermediary that acts as a principal in trades.

Dealing securities Securities held by banks or other financial intermediaries for trading purposes.

Death cross See **Bearish crossover**.

Debentures Type of bond that can be secured or unsecured.

Debt rating An independent measure of the quality and safety of a company's debt based upon an analysis of the company's ability to pay the promised cash flows in full and on time. It includes an analysis of any indentures.

Debt-rating approach A method for estimating a company's before-tax cost of debt based on the yield on comparably rated bonds for maturities that closely match that of the company's existing debt.

Debt tax shield The tax benefit from interest paid on debt being tax deductible from income, equal to the marginal tax rate multiplied by the value of the debt.

Debt-to-assets ratio A solvency ratio calculated as total debt divided by total assets.

Debt-to-capital ratio A solvency ratio calculated as total debt divided by total debt plus total shareholders' equity.

Debt-to-equity ratio A solvency ratio calculated as total debt divided by total shareholders' equity.

Deciles Quantiles that divide a distribution into 10 equal parts.

Declaration date The day that the corporation issues a statement declaring a specific dividend.

Decreasing returns to scale When a production process leads to increases in output that are proportionately smaller than the increase in inputs.

Deductible temporary differences Temporary differences that result in a reduction of or deduction from taxable income in a future period when the balance sheet item is recovered or settled.

Deep learning Machine learning using neural networks with many hidden layers.

Deep learning nets Machine learning using neural networks with many hidden layers.

Default probability The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default risk*.

Default risk The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default probability*.

Default risk premium An extra return that compensates investors for the possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount.

Defensive companies Companies with sales and profits that have little sensitivity to the business cycle or state of the economy.

Defensive interval ratio A liquidity ratio that estimates the number of days that an entity could meet cash needs from liquid assets; calculated as (cash + short-term marketable investments + receivables) divided by daily cash expenditures.

Deferred coupon bond Bond that pays no coupons for its first few years but then pays a higher coupon than it otherwise normally would for the remainder of its life. Also called *split coupon bond*.

Deferred income A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred tax assets A balance sheet asset that arises when an excess amount is paid for income taxes relative to accounting profit. The taxable income is higher than accounting profit and income tax payable exceeds tax expense. The company expects to recover the difference during the course of future operations when tax expense exceeds income tax payable.

Deferred tax liabilities A balance sheet liability that arises when a deficit amount is paid for income taxes relative to accounting profit. The taxable income is less than the accounting profit and income tax payable is less than tax expense. The company expects to eliminate the liability over the course of future operations when income tax payable exceeds tax expense.

Defined benefit pension plans Plans in which the company promises to pay a certain annual amount (defined benefit) to the employee after retirement. The company bears the investment risk of the plan assets.

Defined contribution pension plans Individual accounts to which an employee and typically the employer makes contributions during their working years and expect to draw on the accumulated funds at retirement. The employee bears the investment and inflation risk of the plan assets.

Deflation Negative inflation.

Degree of confidence The probability that a confidence interval includes the unknown population parameter.

Degree of financial leverage (DFL) The ratio of the percentage change in net income to the percentage change in operating income; the sensitivity of the cash flows available to owners when operating income changes.

Degree of operating leverage (DOL) The ratio of the percentage change in operating income to the percentage change in units sold; the sensitivity of operating income to changes in units sold.

Degree of total leverage The ratio of the percentage change in net income to the percentage change in units sold; the sensitivity of the cash flows to owners to changes in the number of units produced and sold.

Degrees of freedom The number of independent variables used in defining sample statistics, such as variance, and the probability distributions they measure.

Delta The sensitivity of the derivative price to a small change in the value of the underlying asset.

Demand curve Graph of the inverse demand function. A graph showing the demand relation, either the highest quantity willingly purchased at each price or the highest price willingly paid for each quantity.

Demand function A relationship that expresses the quantity demanded of a good or service as a function of own-price and possibly other variables.

Demand shock A typically unexpected disturbance to demand, such as an unexpected interruption in trade or transportation.

Dependent With reference to events, the property that the probability of one event occurring depends on (is related to) the occurrence of another event.

Dependent variable The variable whose variation about its mean is to be explained by the regression; the left-side variable in a regression equation. Also referred to as the *explained variable*.

Depository bank A bank that raises funds from depositors and other investors and lends it to borrowers.

Depository institutions Commercial banks, savings and loan banks, credit unions, and similar institutions that raise funds from depositors and other investors and lend it to borrowers.

Depository receipt A security that trades like an ordinary share on a local exchange and represents an economic interest in a foreign company.

Depreciation The process of systematically allocating the cost of long-lived (tangible) assets to the periods during which the assets are expected to provide economic benefits.

Derivative pricing rule A pricing rule used by crossing networks in which a price is taken (derived) from the price that is current in the asset's primary market.

Derivatives A financial instrument whose value depends on the value of some underlying asset or factor (e.g., a stock price, an interest rate, or exchange rate).

Descriptive statistics Measures that summarize central tendency and spread variation in the data's distribution.

Diffuse prior The assumption of equal prior probabilities.

Diffusion index Reflects the proportion of the index's components that are moving in a pattern consistent with the overall index.

Diluted EPS The EPS that would result if all dilutive securities were converted into common shares.

Diluted shares The number of shares that would be outstanding if all potentially dilutive claims on common shares (e.g., convertible debt, convertible preferred stock, and employee stock options) were exercised.

Diminishing balance method An accelerated depreciation method, i.e., one that allocates a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Diminishing marginal productivity Describes a state in which each additional unit of input produces less output than previously.

Direct format With reference to the cash flow statement, a format for the presentation of the statement in which cash flow from operating activities is shown as operating cash receipts less operating cash disbursements. Also called *direct method*.

Direct investing Occurs when an investor makes a direct investment in an asset without the use of an intermediary.

Direct method See *direct format.*

Direct taxes Taxes levied directly on income, wealth, and corporate profits.

Direct write-off method An approach to recognizing credit losses on customer receivables in which the company waits until such time as a customer has defaulted and only then recognizes the loss.

Discount To reduce the value of a future payment in allowance for how far away it is in time; to calculate the present value of some future amount. Also, the amount by which an instrument is priced below its face (par) value.

Discount margin See *required margin.*

Discount rates In general, the interest rates used to calculate present values. In the money market, however, a discount rate is a specific type of quoted rate.

Discounted cash flow models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security.

Discouraged worker A person who has stopped looking for a job or has given up seeking employment.

Discrete data Numerical values that result from a counting process; therefore, practically speaking, the data are limited to a finite number of values.

Discrete random variable A random variable that can take on at most a countable number of possible values.

Discriminatory pricing rule A pricing rule used in continuous markets in which the limit price of the order or quote that first arrived determines the trade price.

Diseconomies of scale Increase in cost per unit resulting from increased production.

Dispersion The variability of a population or sample of observations around the central tendency.

Display size The size of an order displayed to public view.

Disposition effect As a result of loss aversion, an emotional bias whereby investors are reluctant to dispose of losers. This reluctance results in an inefficient and gradual adjustment to deterioration in fundamental value.

Distributed ledger A type of database that may be shared among entities in a network.

Distributed ledger technology Technology based on a distributed ledger.

Divergence In technical analysis, a term that describes the case when an indicator moves differently from the security being analyzed.

Diversification ratio The ratio of the standard deviation of an equally weighted portfolio to the standard deviation of a randomly selected security.

Dividend A distribution paid to shareholders based on the number of shares owned.

Dividend discount model (DDM) A present value model that estimates the intrinsic value of an equity share based on the present value of its expected future dividends.

Dividend payout ratio The ratio of cash dividends paid to earnings for a period.

Divisor A number (denominator) used to determine the value of a price return index. It is initially chosen at the inception of an index and subsequently adjusted by the index provider, as necessary, to avoid changes in the index value that are unrelated to changes in the prices of its constituent securities.

Doji In the Japanese terminology used in candlestick charting, the doji signifies that after a full day of trading, the positive price influence of buyers and the negative price influence of sellers exactly counteracted each other—with opening and closing prices that are virtually equal—which suggests that the market under analysis is in balance.

Domestic content provisions Stipulate that some percentage of the value added or components used in production should be of domestic origin.

Double bottom In technical analysis, a reversal pattern that is formed when the price reaches a low, rebounds, and then declines back to the first low level. A double bottom is used to predict a change from a downtrend to an uptrend.

Double coincidence of wants A prerequisite to barter trades, in particular that both economic agents in the transaction want what the other is selling.

Double declining balance depreciation An accelerated depreciation method that involves depreciating the asset at double the straight-line rate. This rate is multiplied by the book value of the asset at the beginning of the period (a declining balance) to calculate depreciation expense.

Double top In technical analysis, a reversal pattern that is formed when an uptrend reverses twice at roughly the same high price level. A double top is used to predict a change from an uptrend to a downtrend.

Down transition probability The probability that an asset's value moves down in a model of asset price dynamics.

Downgrade risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *credit migration risk.*

Downside risk Risk of incurring returns below a specified value.

Downtrend A pattern that occurs when the price of an asset moves lower over a period of time.

Drag on liquidity When receipts lag, creating pressure from the decreased available funds.

Drawdown A percentage peak-to-trough reduction in net asset value.

Dual-currency bonds Bonds that make coupon payments in one currency and pay the par value at maturity in another currency.

DuPont analysis An approach to decomposing return on investment, e.g., return on equity, as the product of other financial ratios.

Duration A measure of the approximate sensitivity of a security to a change in interest rates (i.e., a measure of interest rate risk).

Duration gap A bond's Macaulay duration minus the investment horizon.

Dutch Book Theorem A result in probability theory stating that inconsistent probabilities create profit opportunities.

Early repayment option See *prepayment option*.

Earnings per share The amount of income earned during a period per share of common stock.

Earnings surprise The portion of a company's earnings that is unanticipated by investors and, according to the efficient market hypothesis, merits a price adjustment.

Economic costs All the remuneration needed to keep a productive resource in its current employment or to acquire the resource for productive use; the sum of total accounting costs and implicit opportunity costs.

Economic indicator A variable that provides information on the state of the overall economy.

Economic loss The amount by which accounting profit is less than normal profit.

Economic profit Equal to accounting profit less the implicit opportunity costs not included in total accounting costs; the difference between total revenue (TR) and total cost (TC). Also called *abnormal profit* or *supernormal profit*.

Economic stabilization Reduction of the magnitude of economic fluctuations.

Economic union Incorporates all aspects of a common market and in addition requires common economic institutions and coordination of economic policies among members.

Economies of scale Reduction in cost per unit resulting from increased production.

Effective annual rate The amount by which a unit of currency will grow in a year with interest on interest included.

Effective convexity A *curve convexity* statistic that measures the secondary effect of a change in a benchmark yield curve on a bond's price.

Effective duration The sensitivity of a bond's price to a change in a benchmark yield curve.

Effective interest rate The borrowing rate or market rate that a company incurs at the time of issuance of a bond.

Efficient market A market in which asset prices reflect new information quickly and rationally.

Elastic Said of a good or service when the magnitude of elasticity is greater than one.

Elasticity The percentage change in one variable for a percentage change in another variable; a general measure of how sensitive one variable is to a change in the value of another variable.

Elasticity of demand A measure of the sensitivity of quantity demanded to a change in a product's own price: $\%\Delta Q^D/\%\Delta P$.

Elasticity of supply A measure of the sensitivity of quantity supplied to a change in price: $\%\Delta Q^S/\%\Delta P$.

Electronic communications networks See *alternative trading systems*.

Embedded option Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Embedded options Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Emotional biases Behavioral biases resulting from reasoning influenced by feelings; emotional biases stem from impulse or intuition.

Empirical duration The use of statistical methods and historical bond prices to estimate the price–yield relationship for a specific bond or portfolio of bonds.

Empirical probability The probability of an event estimated as a relative frequency of occurrence.

Employed The number of people with a job.

Endowment bias An emotional bias in which people value an asset more when they hold rights to it than when they do not.

Engagement/active ownership An ESG investment approach that uses shareholder power to influence corporate behavior through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is directed by ESG guidelines.

Enterprise risk management An overall assessment of a company's risk position. A centralized approach to risk management sometimes called firmwide risk management.

Enterprise value A measure of a company's total market value from which the value of cash and short-term investments have been subtracted.

Equal weighting An index weighting method in which an equal weight is assigned to each constituent security at inception.

Equipment trust certificates Bonds secured by specific types of equipment or physical assets.

Equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Equity risk premium The expected return on equities minus the risk-free rate; the premium that investors demand for investing in equities.

Equity swap A swap transaction in which at least one cash flow is tied to the return to an equity portfolio position, often an equity index.

Error term The difference between an observation and its expected value, where the expected value is based on the true underlying population relation between the dependent and independent variables. Also known simply as the *error*.

ESG An acronym that encompasses environmental, social, and governance.

ESG integration An ESG investment approach that focuses on systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions for the purpose of achieving the product's stated investment objectives. Used interchangeably with **ESG investing**.

ESG investing The consideration of environmental, social, and governance factors in the investment process. Used interchangeably with **ESG integration**.

Estimate The particular value calculated from sample observations using an estimator.

Estimated parameters With reference to a regression analysis, the estimated values of the population intercept and population slope coefficients in a regression.

Estimator An estimation formula; the formula used to compute the sample mean and other sample statistics are examples of estimators.

Ethical principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Ethics The study of moral principles or of making good choices. Ethics encompasses a set of moral principles and rules of conduct that provide guidance for our behavior.

Eurobonds Type of bond issued internationally, outside the jurisdiction of the country in whose currency the bond is denominated.

European-style Said of an option contract that can only be exercised on the option's expiration date.

Event Any outcome or specified set of outcomes of a random variable.

Ex-dividend date The first date that a share trades without (i.e., "ex") the dividend.

Excess kurtosis Degree of kurtosis (fatness of tails) relative to the kurtosis of the normal distribution.

Exchanges Places where traders can meet to arrange their trades.

Execution instructions Instructions that indicate how to fill an order.

Exercise The process of using an option to buy or sell the underlying.

Exercise price The fixed price at which an option holder can buy or sell the underlying. Also called *strike price*, *striking price*, or *strike*.

Exercise value The value obtained if an option is exercised based on current conditions. Also known as *intrinsic value*.

Exhaustive Covering or containing all possible outcomes.

Expansion The period of a business cycle after its lowest point and before its highest point.

Expansionary Tending to cause the real economy to grow.

Expansionary fiscal policy Fiscal policy aimed at achieving real economic growth.

Expected inflation The level of inflation that economic agents expect in the future.

Expected loss Default probability times loss severity given default.

Expected return on the portfolio ($E(R_p)$) The weighted average of the expected returns (R_1 to R_n) on the component securities using their respective weights (w_1 to w_n).

Expected value The probability-weighted average of the possible outcomes of a random variable.

Expenses Outflows of economic resources or increases in liabilities that result in decreases in equity (other than decreases because of distributions to owners); reductions in net assets associated with the creation of revenues.

Export subsidy Paid by the government to the firm when it exports a unit of a good that is being subsidized.

Exports Goods and services that an economy sells to other countries.

Extension risk The risk that when interest rates rise, fewer prepayments will occur because homeowners are reluctant to give up the benefits of a contractual interest rate that now looks low. As a result, the security becomes longer in maturity than anticipated at the time of purchase.

Externalities Spillover effects of production and consumption activities onto others who are not directly involved in a particular transaction, activity, or decision.

Externality An effect of a market transaction that is borne by parties other than those who transacted.

Extra dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Extreme value theory A branch of statistics that focuses primarily on extreme outcomes.

Face value The amount of cash payable by a company to the bondholders when the bonds mature; the promised payment at maturity separate from any coupon payment.

Factoring arrangement When a company sells its accounts receivable to a lender (known as a factor) who assumes responsibility for the credit granting and collection process.

Fair value The amount at which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's-length transaction; the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants.

False discovery approach An adjustment in the p-values for tests performed multiple times.

False discovery rate The rate of Type I errors in testing a null hypothesis multiple times for a given level of significance.

Fat-Tailed Describes a distribution that has fatter tails than a normal distribution (also called leptokurtic).

Fed funds rate The US interbank lending rate on overnight borrowings of reserves.

Federal funds rate The US interbank lending rate on overnight borrowings of reserves.

Fiat money Money that is not convertible into any other commodity.

Fiduciary An entity designated to represent the rights and responsibilities of a beneficiary whose assets they are managing, such as a bond trustee acting on behalf of fixed-income investors.

Fiduciary call A combination of a European call and a risk-free bond that matures on the option expiration day and has a face value equal to the exercise price of the call.

FIFO method The first in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the order in which they were placed in inventory.

Fill or kill See *immediate or cancel order*.

Finance lease From the lessee perspective, under US GAAP, a type of lease which is more akin to the purchase of an asset by the lessee. From the lessor perspective, under IFRS, a lease which "transfers substantially all the risks and rewards incidental to ownership of an underlying asset."

Financial account A component of the balance of payments account that records investment flows.

Financial distress Heightened uncertainty regarding a company's ability to meet its various obligations because of diminished earnings power or actual current losses.

Financial flexibility The ability to react and adapt to financial adversity and opportunities.

Financial leverage The extent to which a company can effect, through the use of debt, a proportional change in the return on common equity that is greater than a given proportional change in operating income; also, short for the financial leverage ratio.

Financial leverage ratio A measure of financial leverage calculated as average total assets divided by average total equity.

Financial risk The risk that environmental, social, or governance risk factors will result in significant costs or other losses to a company and its shareholders; the risk arising from a company's obligation to meet required payments under its financing agreements.

Financing activities Activities related to obtaining or repaying capital to be used in the business (e.g., equity and long-term debt).

Fintech Technological innovation in the design and delivery of financial services and products in the financial industry.

Firm commitment offering See *underwritten offering*.

First-degree price discrimination Where a monopolist is able to charge each customer the highest price the customer is willing to pay.

First lien debt Debt secured by a pledge of certain assets that could include buildings, but it may also include property and equipment, licenses, patents, brands, etc.

First mortgage debt Debt secured by a pledge of a specific property.

Fiscal multiplier The ratio of a change in national income to a change in government spending.

Fiscal policy The use of taxes and government spending to affect the level of aggregate expenditures.

Fisher effect The thesis that the real rate of interest in an economy is stable over time so that changes in nominal interest rates are the result of changes in expected inflation.

Fixed charge coverage A solvency ratio measuring the number of times interest and lease payments are covered by operating income, calculated as (EBIT + lease payments) divided by (interest payments + lease payments).

Fixed costs Costs that remain at the same level regardless of a company's level of production and sales.

Fixed-for-floating interest rate swap An interest rate swap in which one party pays a fixed rate and the other pays a floating rate, with both sets of payments in the same currency. Also called *plain vanilla swap* or *vanilla swap*.

Fixed-rate perpetual preferred stock Nonconvertible, noncallable preferred stock that has a fixed dividend rate and no maturity date.

Flag A technical analysis continuation pattern formed by parallel trendlines, typically over a short period.

Flat price The full price of a bond minus the accrued interest; also called the *quoted* or *clean* price.

Float-adjusted market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by adjusting its market capitalization for its market float.

Floaters See *floating-rate notes*.

Floating-rate notes Notes on which interest payments are not fixed but instead vary from period to period depending on the current level of a reference interest rate.

Flotation cost Fees charged to companies by investment bankers and other costs associated with raising new capital.

Foreclosure Allows the lender to take possession of a mortgaged property if the borrower defaults and then sell it to recover funds.

Foreign currency reserves Holding by the central bank of non-domestic currency deposits and non-domestic bonds.

Foreign direct investment Direct investment by a firm in one country (the source country) in productive assets in a foreign country (the host country).

Foreign exchange gains (or losses) Gains (or losses) that occur when the exchange rate changes between the investor's currency and the currency that foreign securities are denominated in.

Foreign portfolio investment Shorter-term investment by individuals, firms, and institutional investors (e.g., pension funds) in foreign financial instruments such as foreign stocks and foreign government bonds.

Forward commitments Class of derivatives that provides the ability to lock in a price to transact in the future at a previously agreed-upon price.

Forward contract An agreement between two parties in which one party, the buyer, agrees to buy from the other party, the seller, an underlying asset at a later date for a price established at the start of the contract.

Forward curve A series of forward rates, each having the same time frame.

Forward market For future delivery, beyond the usual settlement time period in the cash market.

Forward price The fixed price or rate at which the transaction scheduled to occur at the expiration of a forward contract will take place. This price is agreed on at the initiation date of the contract.

Forward rate The interest rate on a bond or money market instrument traded in a forward market. A forward rate can be interpreted as an incremental, or marginal, return for extending the time-to-maturity for an additional time period.

Forward rate agreements A forward contract calling for one party to make a fixed interest payment and the other to make an interest payment at a rate to be determined at the contract expiration.

Fractile A value at or below which a stated fraction of the data lies. Also called quantile.

Fractional reserve banking Banking in which reserves constitute a fraction of deposits.

Framing bias An information-processing bias in which a person answers a question differently based on the way in which it is asked (framed).

Free cash flow The actual cash that would be available to the company's investors after making all investments necessary to maintain the company as an ongoing enterprise (also referred to as free cash flow to the firm); the internally generated funds that can be distributed to the company's investors (e.g., shareholders and bondholders) without impairing the value of the company.

Free cash flow hypothesis The hypothesis that higher debt levels discipline managers by forcing them to make fixed debt service payments and by reducing the company's free cash flow.

Free-cash-flow-to-equity models Valuation models based on discounting expected future free cash flow to equity.

Free float The number of shares that are readily and freely tradable in the secondary market.

Free trade When there are no government restrictions on a country's ability to trade.

Free trade areas One of the most prevalent forms of regional integration, in which all barriers to the flow of goods and services among members have been eliminated.

Frequency distribution A tabular display of data constructed either by counting the observations of a variable by distinct values or groups or by tallying the values of a numerical variable into a set of numerically ordered bins (also called a one-way table).

Frequency polygon A graph of a frequency distribution obtained by drawing straight lines joining successive points representing the class frequencies.

Frequency table A representation of the frequency of occurrence of two discrete variables.

Full price The price of a security with accrued interest; also called the *invoice* or *dirty* price.

Fund investing In fund investing, the investor invests in assets indirectly by contributing capital to a fund as part of a group of investors. Fund investing is available for all major alternative investment types.

Fundamental analysis The examination of publicly available information and the formulation of forecasts to estimate the intrinsic value of assets.

Fundamental value The underlying or true value of an asset based on an analysis of its qualitative and quantitative characteristics. Also called *intrinsic value*.

Fundamental weighting An index weighting method in which the weight assigned to each constituent security is based on its underlying company's size. It attempts to address the disadvantages of market-capitalization weighting by using measures that are independent of the constituent security's price.

Funds of hedge funds Funds that hold a portfolio of hedge funds, more commonly shortened to *funds of funds*.

Fungible Freely exchangeable, interchangeable, or substitutable with other things of the same type. Money and commodities are the most common examples.

Future value (FV) The amount to which a payment or series of payments will grow by a stated future date.

Futures contract A variation of a forward contract that has essentially the same basic definition but with some additional features, such as a clearinghouse guarantee against credit losses, a daily settlement of gains and losses, and an organized electronic or floor trading facility.

Futures price The agreed-upon price of a futures contract.

FX swap The combination of a spot and a forward FX transaction.

G-spread The yield spread in basis points over an actual or interpolated government bond.

Gains Asset inflows not directly related to the ordinary activities of the business.

Game theory The set of tools decision makers use to incorporate responses by rival decision makers into their strategies.

Gamma A numerical measure of how sensitive an option's delta (the sensitivity of the derivative's price) is to a change in the value of the underlying.

Gap opening A gap is an area of a chart where a security's price either rises or falls from the previous day's close with no trading occurring in between. A gap opening is the start of a new trading session with a gap.

GDP deflator A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

General partner (GP) The partner that runs the business and theoretically bears unlimited liability for the business's debts and obligations.

Geometric mean A measure of central tendency computed by taking the nth root of the product of n non-negative values.

Giffen goods Goods that are consumed more as the price of the good rises because it is a very inferior good whose income effect overwhelms its substitution effect when price changes.

Gilts Bonds issued by the UK government.

Global depository receipt A depository receipt that is issued outside of the company's home country and outside of the United States.

Global minimum-variance portfolio The portfolio on the minimum-variance frontier with the smallest variance of return.

Global registered share A common share that is traded on different stock exchanges around the world in different currencies.

Gold standard With respect to a currency, if a currency is on the gold standard a given amount can be converted into a prespecified amount of gold.

Golden Cross See **Bullish crossover**.

Good-on-close An execution instruction specifying that an order can only be filled at the close of trading. Also called *market on close*.

Good-on-open An execution instruction specifying that an order can only be filled at the opening of trading.

Good-till-cancelled order An order specifying that it is valid until the entity placing the order has cancelled it (or, commonly, until some specified amount of time such as 60 days has elapsed, whichever comes sooner).

Goodwill An intangible asset that represents the excess of the purchase price of an acquired company over the value of the net assets acquired.

Government equivalent yield A yield that restates a yield-to-maturity based on a 30/360 day count to one based on actual/actual.

Green bonds Bonds used in green finance whereby the proceeds are earmarked toward environmental-related products.

Green finance A type of finance that addresses environmental concerns while achieving economic growth.

Green loans Any loan instruments made available exclusively to finance or re-finance, in whole or in part, new and/or existing eligible green projects. Green loans are commonly aligned in the market with the Green Loan Principles.

Greenfield investment Investing in infrastructure assets *that are to be constructed*.

Grey market The forward market for bonds about to be issued. Also called "when issued" market.

Gross domestic product The market value of all final goods and services produced within the economy during a given period (output definition) or, equivalently, the aggregate income earned by all households, all companies, and the government within the economy during a given period (income definition).

Gross margin Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit margin The ratio of gross profit to revenues.

Grouped bar chart A bar chart for showing joint frequencies for two categorical variables (also known as a **clustered bar chart**).

Grouping by function With reference to the presentation of expenses in an income statement, the grouping together of expenses serving the same function, e.g. all items that are costs of goods sold.

Grouping by nature With reference to the presentation of expenses in an income statement, the grouping together of expenses by similar nature, e.g., all depreciation expenses.

Growth cycle Refers to fluctuations in economic activity around the long-term potential trend growth level, focusing on how much actual economic activity is below or above trend growth in economic activity.

Growth cyclical A term sometimes used to describe companies that are growing rapidly on a long-term basis but that still experience above-average fluctuation in their revenues and profits over the course of a business cycle.

Growth investors With reference to equity investors, investors who seek to invest in high-earnings-growth companies.

Growth option The ability to make additional investments in a project at some future time if the financial results are strong. Also called *expansion option*.

Growth rate cycle Refers to fluctuations in the growth rate of economic activity.

Guarantee certificate A type of structured financial instrument that provides investors with capital protection. It combines a zero-coupon bond and a call option on some underlying asset.

Haircut See *repo margin*.

Halo effect An emotional bias that extends a favorable evaluation of some characteristics to other characteristics.

Hard-bullet covered bonds Covered bonds for which a bond default is triggered and bond payments are accelerated in the event of sponsor default if payments do not occur according to the original maturity schedule.

Harmonic mean A type of weighted mean computed as the reciprocal of the arithmetic average of the reciprocals.

Head and shoulders pattern In technical analysis, a reversal pattern that is formed in three parts: a left shoulder, a head, and a right shoulder. A head and shoulders pattern is used to predict a change from an uptrend to a downtrend.

Headline inflation Refers to the inflation rate calculated based on the price index that includes all goods and services in an economy.

Heat map A type of graphic that organizes and summarizes data in a tabular format and represents it using a color spectrum.

Hedge funds Private investment vehicles that typically use leverage, derivatives, and long and short investment strategies.

Hedge portfolio A hypothetical combination of the derivative and its underlying that eliminates risk.

Held-to-maturity Debt (fixed-income) securities that a company intends to hold to maturity; these are presented at their original cost, updated for any amortisation of discounts or premiums.

Herding Clustered trading that may or may not be based on information.

Heteroskedasticity A violation of the simple linear regression assumption of homoskedasticity; with heteroskedasticity, the variability of the residuals is not equal across all observations.

Hidden order An order that is exposed not to the public but only to the brokers or exchanges that receive it.

High-frequency trading A form of algorithmic trading that makes use of vast quantities of data to execute trades on ultra-high-speed networks in fractions of a second.

High-water mark The highest value, net of fees, which a fund has reached in history. It reflects the highest cumulative return used to calculate an incentive fee.

Hindsight bias A bias with selective perception and retention aspects in which people may see past events as having been predictable and reasonable to expect.

Histogram A chart that presents the distribution of numerical data by using the height of a bar or column to represent the absolute frequency of each bin or interval in the distribution.

Historical cost In reference to assets, the amount paid to purchase an asset, including any costs of acquisition and/or preparation; with reference to liabilities, the amount of proceeds received in exchange in issuing the liability.

Historical equity risk premium approach An estimate of a country's equity risk premium that is based on the historical averages of the risk-free rate and the rate of return on the market portfolio.

Holder-of-record date The date that a shareholder listed on the corporation's books will be deemed to have ownership of the shares for purposes of receiving an upcoming dividend.

Holding period return The return that an investor earns during a specified holding period; a synonym for total return.

Home bias A preference for securities listed on the exchanges of one's home country.

Homogeneity of expectations The assumption that all investors have the same economic expectations and thus have the same expectations of prices, cash flows, and other investment characteristics.

Homoskedasticity An assumption of simple linear regression such that the variability of the residuals is equal across all observations.

Horizon yield The internal rate of return between the total return (the sum of reinvested coupon payments and the sale price or redemption amount) and the purchase price of the bond.

Horizontal analysis Common-size analysis that involves comparing a specific financial statement with that statement in prior or future time periods; also, cross-sectional analysis of one company with another.

Horizontal demand schedule Implies that at a given price, the response in the quantity demanded is infinite.

Hostile takeover An attempt by one entity to acquire a company without the consent of the company's management.

Household A person or a group of people living in the same residence, taken as a basic unit in economic analysis.

Human capital The accumulated knowledge and skill that workers acquire from education, training, or life experience and the corresponding present value of future earnings to be generated by said skilled individual.

Hurdle rate The minimum rate of return on investment that a fund must reach before a GP receives carried interest.

Hypothesis A proposed explanation or theory that can be tested.

Hypothesis testing The process of testing of hypotheses about one or more populations using statistical inference.

I-spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Iceberg order An order in which the display size is less than the order's full size.

If-converted method A method for accounting for the effect of convertible securities on earnings per share (EPS) that specifies what EPS would have been if the convertible securities had been converted at the beginning of the period, taking account of the effects of conversion on net income and the weighted average number of shares outstanding.

Illusion of control bias A bias in which people tend to believe that they can control or influence outcomes when, in fact, they cannot.

Immediate or cancel order An order that is valid only upon receipt by the broker or exchange. If such an order cannot be filled in part or in whole upon receipt, it cancels immediately. Also called *fill or kill*.

Impact investing Investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.

Impact lag The lag associated with the result of actions affecting the economy with delay.

Implicit price deflator for GDP A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

Implicit selection bias One type of selection bias introduced through the presence of a threshold that filters out some unqualified members.

Implied forward rates Calculated from spot rates, an implied forward rate is a breakeven reinvestment rate that links the return on an investment in a shorter-term zero-coupon bond to the return on an investment in a longer-term zero-coupon bond.

Implied volatility The volatility that option traders use to price an option, implied by the price of the option and a particular option-pricing model.

Import license Specifies the quantity of a good that can be imported into a country.

Imports Goods and services that a domestic economy (i.e., house-holds, firms, and government) purchases from other countries.

In the money Options that, if exercised, would result in the value received being worth more than the payment required to exercise.

Income Increases in economic benefits in the form of inflows or enhancements of assets, or decreases of liabilities that result in an increase in equity (other than increases resulting from contributions by owners).

Income elasticity of demand A measure of the responsiveness of demand to changes in income, defined as the percentage change in quantity demanded divided by the percentage change in income.

Income tax paid The actual amount paid for income taxes in the period; not a provision, but the actual cash outflow.

Income tax payable The income tax owed by the company on the basis of taxable income.

Increasing marginal returns When the marginal product of a resource increases as additional units of that input are employed.

Increasing returns to scale When a production process leads to increases in output that are proportionately larger than the increase in inputs.

Incremental cash flow The cash flow that is realized because of a decision; the changes or increments to cash flows resulting from a decision or action.

Indenture Legal contract that describes the form of a bond, the obligations of the issuer, and the rights of the bondholders. Also called the *trust deed*.

Independent With reference to events, the property that the occurrence of one event does not affect the probability of another event occurring. With reference to two random variables X and Y, they are independent if and only if $P(X,Y) = P(X)P(Y)$.

Independent projects Independent projects are capital investments whose cash flows are independent of each other.

Independent variable A variable used to explain the dependent variable in a regression; a right-side variable in a regression equation. Also referred to as the *explanatory variable*.

Independently and identically distributed With respect to random variables, the property of random variables that are independent of each other but follow the identical probability distribution.

Index-linked bond Bond for which coupon payments and/or principal repayment are linked to a specified index.

Indexing An investment strategy in which an investor constructs a portfolio to mirror the performance of a specified index.

Indicator variable A variable that takes on only one of two values, 0 or 1, based on a condition. In simple linear regression, the slope is the difference in the dependent variable for the two conditions. Also referred to as a *dummy variable*.

Indifference curve A curve representing all the combinations of two goods or attributes such that the consumer is entirely indifferent among them.

Indirect format With reference to cash flow statements, a format for the presentation of the statement which, in the operating cash flow section, begins with net income then shows additions and subtractions to arrive at operating cash flow. Also called *indirect method*.

Indirect method See *indirect format*.

Indirect taxes Taxes such as taxes on spending, as opposed to direct taxes.

Industry A group of companies offering similar products and/or services.

Industry analysis The analysis of a specific branch of manufacturing, service, or trade.

Inelastic Said of a good or service when the magnitude of elasticity is less than one. Insensitive to price changes.

Inferior goods A good whose consumption decreases as income increases.

Inflation The percentage increase in the general price level from one period to the next; a sustained rise in the overall level of prices in an economy.

Inflation-linked bond Type of index-linked bond that offers investors protection against inflation by linking the bond's coupon payments and/or the principal repayment to an index of consumer prices. Also called *linkers*.

Inflation premium An extra return that compensates investors for expected inflation.

Inflation reports A type of economic publication put out by many central banks.

Inflation uncertainty The degree to which economic agents view future rates of inflation as difficult to forecast.

Information cascade The transmission of information from those participants who act first and whose decisions influence the decisions of others.

Information-motivated traders Traders that trade to profit from information that they believe allows them to predict future prices.

Informationally efficient market A market in which asset prices reflect new information quickly and rationally.

Initial coin offering An unregulated process whereby companies raise capital by selling crypto tokens to investors in exchange for fiat money or another agreed-upon cryptocurrency.

Initial margin The amount that must be deposited in a clearinghouse account when entering into a futures contract.

Initial margin requirement The margin requirement on the first day of a transaction as well as on any day in which additional margin funds must be deposited.

Initial public offering (IPO) The first issuance of common shares to the public by a formerly private corporation.

Input productivity The amount of output produced by workers in a given period of time—for example, output per hour worked; measures the efficiency of labor.

Intangible assets Assets lacking physical substance, such as patents and trademarks.

Interbank market The market of loans and deposits between banks for maturities ranging from overnight to one year.

Intercept The expected value of the dependent variable when the independent variable in a simple linear regression is equal to zero.

Interest Payment for lending funds.

Interest coverage A solvency ratio calculated as EBIT divided by interest payments.

Interest-only mortgage A loan in which no scheduled principal repayment is specified for a certain number of years.

Interest rate A rate of return that reflects the relationship between differently dated cash flows; a discount rate.

Interest rate effect The effect through which price level changes, through demand for money, impact interest rate, which in turn impacts investment and consumption.

Interest rate swap A swap in which the underlying is an interest rate. Can be viewed as a currency swap in which both currencies are the same and can be created as a combination of currency swaps.

Intermarket analysis A field within technical analysis that combines analysis of the major categories of securities—namely, equities, bonds, currencies, and commodities—to identify market trends and possible inflections in trends.

Internal rate of return The discount rate that makes net present value equal 0; the discount rate that makes the present value of an investment's costs (outflows) equal to the present value of the investment's benefits (inflows).

Internet of Things A network arrangement of structures and devices whereby the objects on the network are able to interact and share information.

Interpolated spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Interquartile range The difference between the third and first quartiles of a dataset.

Interval With reference to grouped data, a set of values within which an observation falls.

Intrinsic value See *exercise value*.

Inventory investment Net change in business inventory.

Inventory turnover An activity ratio calculated as cost of goods sold divided by average inventory.

Inverse demand function A restatement of the demand function in which price is stated as a function of quantity.

Inverse floater A type of leveraged structured financial instrument. The cash flows are adjusted periodically and move in the opposite direction of changes in the reference rate.

Inverse transformation method A method using randomly generated numbers from the continuous uniform distribution to generate random observations from any distribution.

Investing activities Activities associated with the acquisition and disposal of property, plant, and equipment; intangible assets; other long-term assets; and both long-term and short-term investments in the equity and debt (bonds and loans) issued by other companies.

Investment banks Financial intermediaries that provide advice to their mostly corporate clients and help them arrange transactions such as initial and seasoned securities offerings.

Investment policy statement (IPS) A written planning document that describes a client's investment objectives and risk tolerance over a relevant time horizon, along with constraints that apply to the client's portfolio.

Investment property Property used to earn rental income or capital appreciation (or both).

Jackknife A resampling method that repeatedly draws samples by taking the original observed data sample and leaving out one observation at a time (without replacement) from the set.

January effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January. Also called *turn-of-the-year effect*.

Joint frequencies The entry in the cells of the contingency table that represent the joining of one variable from a row and the other variable from a column to count observations.

Joint probability The probability of the joint occurrence of stated events.

Joint probability function A function giving the probability of joint occurrences of values of stated random variables.

Judgmental Sampling A procedure of selectively handpicking elements from the population based on a researcher's knowledge and professional judgment.

Key rate duration A method of measuring the interest rate sensitivities of a fixed-income instrument or portfolio to shifts in key points along the yield curve.

Keynesians Economists who believe that fiscal policy can have powerful effects on aggregate demand, output, and employment when there is substantial spare capacity in an economy.

Kurtosis The statistical measure that indicates the combined weight of the tails of a distribution relative to the rest of the distribution.

Labor force The portion of the working age population (over the age of 16) that is employed or is available for work but not working (unemployed).

Labor productivity The quantity of goods and services (real GDP) that a worker can produce in one hour of work.

Lagging economic indicators Turning points that take place later than those of the overall economy; they are believed to have value in identifying the economy's past condition.

Law of demand The principle that as the price of a good rises, buyers will choose to buy less of it, and as its price falls, they will buy more.

Law of diminishing marginal returns The observation that a variable factor's marginal product must eventually fall as more of it is added to a fixed amount of the other factors.

Law of diminishing returns The smallest output that a firm can produce such that its long run average costs are minimized.

Law of one price The condition in a financial market in which two equivalent financial instruments or combinations of financial instruments can sell for only one price. Equivalent to the principle that no arbitrage opportunities are possible.

Lead underwriter The lead investment bank in a syndicate of investment banks and broker–dealers involved in a securities underwriting.

Leading economic indicators Turning points that usually precede those of the overall economy; they are believed to have value for predicting the economy's future state, usually near-term.

Lease A contract that conveys the right to use an asset for a period of time in exchange for consideration.

Legal tender Something that must be accepted when offered in exchange for goods and services.

Lender of last resort An entity willing to lend money when no other entity is ready to do so.

Leptokurtic Describes a distribution that has fatter tails than a normal distribution (also called fat-tailed).

Lessee The party obtaining the use of an asset through a lease.

Lessor The owner of an asset that grants the right to use the asset to another party.

Letter of credit Form of external credit enhancement whereby a financial institution provides the issuer with a credit line to reimburse any cash flow shortfalls from the assets backing the issue.

Level of significance The probability of a Type I error in testing a hypothesis.

Leverage In the context of corporate finance, leverage refers to the use of fixed costs within a company's cost structure. Fixed costs that are operating costs (such as depreciation or rent) create operating leverage. Fixed costs that are financial costs (such as interest expense) create financial leverage.

Leveraged buyout A transaction whereby the target company's management team converts the target to a privately held company by using heavy borrowing to finance the purchase of the target company's outstanding shares.

Liabilities Present obligations of an enterprise arising from past events, the settlement of which is expected to result in an outflow of resources embodying economic benefits; creditors' claims on the resources of a company.

LIFO layer liquidation With respect to the application of the LIFO inventory method, the liquidation of old, relatively low-priced inventory; happens when the volume of sales rises above the volume of recent purchases so that some sales are made from relatively old, low-priced inventory. Also called *LIFO liquidation*.

LIFO method The last in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the reverse order the items were placed in inventory (i.e., inventory produced or acquired last are assumed to be sold first).

LIFO reserve The difference between the reported LIFO inventory carrying amount and the inventory amount that would have been reported if the FIFO method had been used (in other words, the FIFO inventory value less the LIFO inventory value).

Likelihood The probability of an observation, given a particular set of conditions.

Limit down A limit move in the futures market in which the price at which a transaction would be made is at or below the lower limit.

Limit order Instructions to a broker or exchange to obtain the best price immediately available when filling an order, but in no event accept a price higher than a specified (limit) price when buying or accept a price lower than a specified (limit) price when selling.

Limit order book The book or list of limit orders to buy and sell that pertains to a security.

Limit up A limit move in the futures market in which the price at which a transaction would be made is at or above the upper limit.

Limitations on liens Meant to put limits on how much secured debt an issuer can have.

Limited partners Partners with limited liability. Limited partnerships in hedge and private equity funds are typically restricted to investors who are expected to understand and to be able to assume the risks associated with the investments.

Limited partnership agreement (LPA) A legal document that outlines the rules of the partnership and establishes the framework that ultimately guides the fund's operations throughout its life.

Lin-log model A regression model in which the independent variable is in logarithmic form.

Line chart A type of graph used to visualize ordered observations. In technical analysis, a plot of price data, typically closing prices, with a line connecting the points.

Linear interpolation The estimation of an unknown value on the basis of two known values that bracket it, using a straight line between the two known values.

Linear regression Regression that models the straight-line relationship between the dependent and independent variables. Also known as *least squares regression* and *ordinary least squares regression*.

Linear scale A scale in which equal distances correspond to equal absolute amounts. Also called an *arithmetic scale*.

Linker See *inflation-linked bond*.

Liquid market Said of a market in which traders can buy or sell with low total transaction costs when they want to trade.

Liquidation To sell the assets of a company, division, or subsidiary piecemeal, typically because of bankruptcy; the form of bankruptcy that allows for the orderly satisfaction of creditors' claims after which the company ceases to exist.

Liquidity The extent to which a company is able to meet its short-term obligations using cash flows and those assets that can be readily transformed into cash.

Liquidity premium An extra return that compensates investors for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

Liquidity ratios Financial ratios measuring the company's ability to meet its short-term obligations to creditors.

Liquidity risk The risk that a financial instrument cannot be purchased or sold without a significant concession in price due to the size of the market.

Liquidity trap A condition in which the demand for money becomes infinitely elastic (horizontal demand curve) so that injections of money into the economy will not lower interest rates or affect real activity.

Load fund A mutual fund in which, in addition to the annual fee, a percentage fee is charged to invest in the fund and/or for redemptions from the fund.

Loan-to-value ratio The amount of the property's mortgage divided by the purchase price

Locked limit A condition in the futures markets in which a transaction cannot take place because the price would be beyond the limits.

Lockup period The minimum holding period before investors are allowed to make withdrawals or redeem shares from a fund.

Log-lin model A regression model in which the dependent variable is in logarithmic form.

Log-log model A regression model in which both the dependent and independent variables are in logarithmic form. Also known as the *double-log model*.

Logarithmic scale A scale in which equal distances represent equal proportional changes in the underlying quantity.

London interbank offered rate (Libor) Collective name for multiple rates at which a select set of banks believes they could borrow unsecured funds from other banks in the London interbank market for different currencies and different borrowing periods ranging from overnight to one year.

Long The buyer of a derivative contract. Also refers to the position of owning a derivative.

Long-lived assets Assets that are expected to provide economic benefits over a future period of time, typically greater than one year. Also called *long-term assets*.

Long position A position in an asset or contract in which one owns the asset or has an exercisable right under the contract.

Long-run average total cost The curve describing average total cost when no costs are considered fixed.

Look-ahead bias A bias caused by using information that was unavailable on the test date.

Loss aversion The tendency of people to dislike losses more than they like comparable gains.

Loss-aversion bias A bias in which people tend to strongly prefer avoiding losses as opposed to achieving gains.

Loss severity Portion of a bond's value (including unpaid interest) an investor loses in the event of default.

Losses Asset outflows not directly related to the ordinary activities of the business.

Lower bound The lowest possible value of an option.

M^2 A measure of what a portfolio would have returned if it had taken on the same total risk as the market index.

M^2 alpha Difference between the risk-adjusted performance of the portfolio and the performance of the benchmark.

Macaulay duration The approximate amount of time a bond would have to be held for the market discount rate at purchase to be realized if there is a single change in interest rate. It indicates the point in time when the coupon reinvestment and price effects of a change in yield-to-maturity offset each other.

Machine learning Computer based techniques that seek to extract knowledge from large amounts of data by "learning" from known examples and then generating structure or predictions. ML algorithms aim to "find the pattern, apply the pattern."

Macroeconomics The branch of economics that deals with aggregate economic quantities, such as national output and national income.

Maintenance covenants Covenants in bank loan agreements that require the borrower to satisfy certain financial ratio tests while the loan is outstanding.

Maintenance margin The minimum amount that is required by a futures clearinghouse to maintain a margin account and to protect against default. Participants whose margin balances drop below the required maintenance margin must replenish their accounts.

Maintenance margin requirement The margin requirement on any day other than the first day of a transaction.

Management buy-in Leveraged buyout in which the current management team is being replaced and the acquiring team will be involved in managing the company.

Management buyout A leveraged buyout event in which a group of investors consisting primarily of the company's existing management purchases at least controlling interest in its outstanding shares. At the extreme, they may purchase all shares and take the company private.

Management fee A fee based on assets under management or committed capital, as applicable—also called a *base fee*.

MAR ratio A variation of the Calmar ratio that uses a full investment history and the average drawdown.

Margin The amount of money that a trader deposits in a margin account. The term is derived from the stock market practice in which an investor borrows a portion of the money required to purchase a certain amount of stock. In futures markets, there is no borrowing so the margin is more of a down payment or performance bond.

Margin bond A cash deposit required by the clearinghouse from the participants to a contract to provide a credit guarantee. Also called a *performance bond*.

Margin call A request for the short to deposit additional funds to bring their balance up to the initial margin.

Margin loan Money borrowed from a broker to purchase securities.

Marginal cost The cost of producing an additional unit of a good.

Marginal frequencies The sums determined by adding joint frequencies across rows or across columns in a contingency table.

Marginal product Measures the productivity of each unit of input and is calculated by taking the difference in total product from adding another unit of input (assuming other resource quantities are held constant).

Marginal propensity to consume The proportion of an additional unit of disposable income that is consumed or spent; the change in consumption for a small change in income.

Marginal propensity to save The proportion of an additional unit of disposable income that is saved (not spent).

Marginal revenue The change in total revenue divided by the change in quantity sold; simply, the additional revenue from selling one more unit.

Marginal value curve A curve describing the highest price consumers are willing to pay for each additional unit of a good.

Mark to market The revaluation of a financial asset or liability to its current market value or fair value.

Mark-to-market Refers to the current expected fair market value for which a given security would likely be available for purchase or sale if traded in current market conditions.

Market anomaly Change in the price or return of a security that cannot directly be linked to current relevant information known in the market or to the release of new information into the market.

Market bid–ask spread The difference between the best bid and the best offer.

Market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its market capitalization by the total market capitalization (sum of the market capitalization) of all securities in the index. Also called *value weighting*.

Market discount rate The rate of return required by investors given the risk of the investment in a bond; also called the *required yield* or the *required rate of return*.

Market float The number of shares that are available to the investing public.

Market liquidity risk The risk that the price at which investors can actually transact—buying or selling—may differ from the price indicated in the market.

Market model A regression equation that specifies a linear relationship between the return on a security (or portfolio) and the return on a broad market index.

Market multiple models Valuation models based on share price multiples or enterprise value multiples.

Market-on-close An execution instruction specifying that an order can only be filled at the close of trading.

Market order Instructions to a broker or exchange to obtain the best price immediately available when filling an order.

Market-oriented investors With reference to equity investors, investors whose investment disciplines cannot be clearly categorized as value or growth.

Market rate of interest The rate demanded by purchasers of bonds, given the risks associated with future cash payment obligations of the particular bond issue.

Market reference rate (MRR) Collective name for a set of rates covering different currencies for different maturities, ranging from overnight to one year.

Market risk The risk that arises from movements in interest rates, stock prices, exchange rates, and commodity prices.

Market value The price at which an asset or security can currently be bought or sold in an open market.

Marketable limit order A buy limit order in which the limit price is placed above the best offer, or a sell limit order in which the limit price is placed below the best bid. Such orders generally will partially or completely fill right away.

Markowitz efficient frontier The graph of the set of portfolios offering the maximum expected return for their level of risk (standard deviation of return).

Matching principle The accounting principle that expenses should be recognized in the same period in which the associated revenue is recognized.

Matrix pricing Process of estimating the market discount rate and price of a bond based on the quoted or flat prices of more frequently traded comparable bonds.

Maturity premium An extra return that compensates investors for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended.

Maturity structure A factor explaining the differences in yields on similar bonds; also called *term structure*.

Mean absolute deviation With reference to a sample, the mean of the absolute values of deviations from the sample mean.

Mean square error (MSE) The sum of squares error divided by the degrees of freedom, $n - k - 1$; in a simple linear regression, $n - k - 1 = n - 2$.

Mean square regression (MSR) The sum of squares regression divided by the number of independent variables k; in a simple linear regression, $k = 1$.

Mean–variance analysis An approach to portfolio analysis using expected means, variances, and covariances of asset returns.

Measure of central tendency A quantitative measure that specifies where data are centered.

Measure of value A standard for measuring value; a function of money.

Measures of location Quantitative measures that describe the location or distribution of data. They include not only measures of central tendency but also other measures, such as percentiles.

Median The value of the middle item of a set of items that has been sorted into ascending or descending order (i.e., the 50th percentile).

Medium of exchange Any asset that can be used to purchase goods and services or to repay debts; a function of money.

Medium-term note A corporate bond offered continuously to investors by an agent of the issuer, designed to fill the funding gap between commercial paper and long-term bonds.

Mental accounting bias An information-processing bias in which people treat one sum of money differently from another equal-sized sum based on which mental account the money is assigned to.

Menu costs A cost of inflation in which businesses constantly have to incur the costs of changing the advertised prices of their goods and services.

Mesokurtic Describes a distribution with kurtosis equal to that of the normal distribution, namely, kurtosis equal to three.

Microeconomics The branch of economics that deals with markets and decision making of individual economic units, including consumers and businesses.

Minimum efficient scale The smallest output that a firm can produce such that its long-run average total cost is minimized.

Minimum-variance portfolio The portfolio with the minimum variance for each given level of expected return.

Minority shareholders Particular shareholders or a block of shareholders holding a small proportion of a company's outstanding shares, resulting in a limited ability to exercise control in voting activities.

Minsky moment Named for Hyman Minksy. A point in a business cycle when, after individuals become overextended in borrowing to finance speculative investments, people start realizing that something is likely to go wrong and a panic ensues, leading to asset sell-offs.

Modal interval With reference to grouped data, the interval containing the greatest number of observations (i.e., highest frequency).

Mode The most frequently occurring value in a distribution.

Modern portfolio theory (MPT) The analysis of rational portfolio choices based on the efficient use of risk.

Modified duration A measure of the percentage price change of a bond given a change in its yield-to-maturity.

Momentum oscillator A graphical representation of market sentiment that is constructed from price data and calculated so that it oscillates either between a low and a high or around some number.

Monetarists Economists who believe that the rate of growth of the money supply is the primary determinant of the rate of inflation.

Monetary policy Actions taken by a nation's central bank to affect aggregate output and prices through changes in bank reserves, reserve requirements, or its target interest rate.

Monetary transmission mechanism The process whereby a central bank's interest rate gets transmitted through the economy and ultimately affects the rate of increase of prices.

Monetary union An economic union in which the members adopt a common currency.

Money A generally accepted medium of exchange and unit of account.

Money convexity For a bond, the annual or approximate convexity multiplied by the full price.

Money creation The process by which changes in bank reserves translate into changes in the money supply.

Money duration A measure of the price change in units of the currency in which the bond is denominated given a change in its yield-to-maturity.

Money market The market for short-term debt instruments (one-year maturity or less).

Money market securities Fixed-income securities with maturities at issuance of one year or less.

Money multiplier Describes how a change in reserves is expected to affect the money supply; in its simplest form, 1 divided by the reserve requirement.

Money neutrality The thesis that an increase in the money supply leads in the long-run to an increase in the price level, while leaving real variables like output and employment unaffected.

Money-weighted return The internal rate of return on a portfolio, taking account of all cash flows.

Moneyness The relationship between the price of the underlying and an option's exercise price.

Monopolistic competition Highly competitive form of imperfect competition; the competitive characteristic is a notably large number of firms, while the monopoly aspect is the result of product differentiation.

Monopoly In pure monopoly markets, there are no substitutes for the given product or service. There is a single seller, which exercises considerable power over pricing and output decisions.

Monte Carlo simulation A technique involving the use of computer software to represent the operation of a complex financial system. A characteristic feature is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Moral principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Mortgage-backed securities Debt obligations that represent claims to the cash flows from pools of mortgage loans, most commonly on residential property.

Mortgage loan A loan secured by the collateral of some specified real estate property that obliges the borrower to make a predetermined series of payments to the lender.

Mortgage pass-through security A security created when one or more holders of mortgages form a pool of mortgages and sell shares or participation certificates in the pool.

Mortgage rate The interest rate on a mortgage loan; also called *contract rate* or *note rate*.

Moving average The average of the closing price of a security over a specified number of periods. With each new period, the average is recalculated.

Moving-average convergence/divergence oscillator A momentum oscillator that is based on the difference between short-term and long-term moving averages of a security's price.

Multi-factor model A model that explains a variable in terms of the values of a set of factors.

Multi-market indexes Comprised of indexes from different countries, designed to represent multiple security markets.

Multi-step format With respect to the format of the income statement, a format that presents a subtotal for gross profit (revenue minus cost of goods sold).

Multilateral trading facilities See *alternative trading systems*.

Multinational corporation A company operating in more than one country or having subsidiary firms in more than one country.

Multinomial formula (general formula for labeling problems) The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by $\dfrac{n!}{n_1!n_2!\ldots n_k!}$.

Multiple of invested capital (MOIC) A simplified calculation that measures the total value of all distributions and residual asset values relative to an initial total investment; also known as a *money multiple*.

Multiple testing problem The risk of getting statistically significant test results when performing a test multiple times.

Multiplication rule for counting If one task can be done in n_1 ways, and a second task, given the first, can be done in n_2 ways, and a third task, given the first two tasks, can be done in n_3 ways, and so on for k tasks, then the number of ways the k tasks can be done is $(n_1)(n_2)(n_3) \ldots (n_k)$.

Multiplication rule for expected value of the product of uncorrelated random variables This rule is $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated.

Multiplication rule for independent events The rule that when two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B.

Multiplication rule for probability The rule that the joint probability of events A and B equals the probability of A given B times the probability of B.

Multiplier models Valuation models based on share price multiples or enterprise value multiples.

Multivariate distribution A probability distribution that specifies the probabilities for a group of related random variables.

Multivariate normal distribution A probability distribution for a group of random variables that is completely defined by the means and variances of the variables plus all the correlations between pairs of the variables.

Muni See *municipal bond*.

Municipal bond A type of non-sovereign bond issued by a state or local government in the United States. It very often (but not always) offers income tax exemptions.

Mutual fund A comingled investment pool in which investors in the fund each have a pro-rata claim on the income and value of the fund.

Mutually exclusive Indicates that only one event can occur at a time.

Mutually exclusive projects Mutually exclusive projects compete directly with each other. For example, if Projects A and B are mutually exclusive, you can choose A or B but you cannot choose both.

***n* Factorial** For a positive integer n, the product of the first n positive integers; 0 factorial equals 1 by definition. n factorial is written as $n!$.

Narrow money The notes and coins in circulation in an economy, plus other very highly liquid deposits.

Nash equilibrium When two or more participants in a non-coop-erative game have no incentive to deviate from their respective equilibrium strategies given their opponent's strategies.

National income The income received by all factors of production used in the generation of final output. National income equals gross domestic product (or, in some countries, gross national product) minus the capital consumption allowance and a statistical discrepancy.

Natural language processing Computer programs developed to analyze and interpret human language.

Natural rate of unemployment Effective unemployment rate, below which pressure emerges in labor markets.

Negative screening An ESG investment style that focuses on the exclusion of certain sectors, companies, or practices in a fund or portfolio on the basis of specific ESG criteria.

Net book value The remaining (undepreciated) balance of an asset's purchase cost. For liabilities, the face value of a bond minus any unamortized discount, or plus any unamortized premium.

Net exports The difference between the value of a country's exports and the value of its imports (i.e., value of exports minus imports).

Net income The difference between revenue and expenses; what remains after subtracting all expenses (including depreciation, interest, and taxes) from revenue.

Net present value The present value of an investment's cash inflows (benefits) minus the present value of its cash outflows (costs).

Net profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also called *profit margin* or *return on sales*.

Net realisable value Estimated selling price in the ordinary course of business less the estimated costs necessary to make the sale.

Net revenue Revenue after adjustments (e.g., for estimated returns or for amounts unlikely to be collected).

Net tax rate The tax rate net of transfer payments.

Neural networks Computer programs based on how our own brains learn and process information.

Neutral rate of interest The rate of interest that neither spurs on nor slows down the underlying economy.

No-load fund A mutual fund in which there is no fee for investing in the fund or for redeeming fund shares, although there is an annual fee based on a percentage of the fund's net asset value.

Node Each value on a binomial tree from which successive moves or outcomes branch.

Nominal data Categorical values that are not amenable to being organized in a logical order. An example of nominal data is the classification of publicly listed stocks into sectors.

Nominal GDP The value of goods and services measured at current prices.

Nominal risk-free interest rate The sum of the real risk-free interest rate and the inflation premium.

Non-accelerating inflation rate of unemployment Effective unemployment rate below which pressure emerges in labor markets.

Non-agency RMBS In the United States, securities issued by private entities that are not guaranteed by a federal agency or a GSE.

Non-Bank lender Unlike typical banks, which make loans and take deposits, these lenders only make loans. Although they lend money to companies, they usually specialize in issuing mortgages to first-time homebuyers and parties looking to refinance an existing mortgage.

Non-cumulative preference shares Preference shares for which dividends that are not paid in the current or subsequent periods are forfeited permanently (instead of being accrued and paid at a later date).

Non-current assets Assets that are expected to benefit the company over an extended period of time (usually more than one year).

Non-current liabilities Obligations that broadly represent a probable sacrifice of economic benefits in periods generally greater than one year in the future.

Non-cyclical A company whose performance is largely independent of the business cycle. Also known as defensive.

Non-deliverable forwards Cash-settled forward contracts, used predominately with respect to foreign exchange forwards. Also called *contracts for differences*.

Non-financial risks Risks that arise from sources other than changes in the external financial markets, such as changes in accounting rules, legal environment, or tax rates.

Non-participating preference shares Preference shares that do not entitle shareholders to share in the profits of the company. Instead, shareholders are only entitled to receive a fixed dividend payment and the par value of the shares in the event of liquidation.

Non-probability sampling A sampling plan dependent on factors other than probability considerations, such as a sampler's judgment or the convenience to access data.

Non-recourse loan A loan in which the lender does not have a shortfall claim against the borrower, so the lender can look only to the property to recover the outstanding mortgage balance.

Non-renewable resources Finite resources that are depleted once they are consumed, such as oil and coal.

Non-sovereign bond A bond issued by a government below the national level, such as a province, region, state, or city.

Nonconventional cash flows In a nonconventional cash flow pattern, the initial outflow is not followed by inflows only, but the cash flows can flip from positive (inflows) to negative (outflows) again or even change signs several times.

Nonparametric test A test that is not concerned with a parameter or that makes minimal assumptions about the population from which a sample comes.

Nonsystematic risk Unique risk that is local or limited to a particular asset or industry that need not affect assets outside of that asset class.

Normal distribution A continuous, symmetric probability distribution that is completely described by its mean and its variance.

Normal goods Goods that are consumed in greater quantities as income increases.

Normal profit The level of accounting profit needed to just cover the implicit opportunity costs ignored in accounting costs.

Notching Ratings adjustment methodology where specific issues from the same borrower may be assigned different credit ratings.

Note rate See *mortgage rate*.

Notice period The length of time (typically 30 to 90 days) in advance that investors may be required to notify a fund of their intent to redeem some or all of their investment.

Notional principal An imputed principal amount.

Null hypothesis The hypothesis that is tested.

Numerical data Values that represent measured or counted quantities as a number. Also called **quantitative data**.

Objective probabilities Probabilities that generally do not vary from person to person; includes a priori and empirical probabilities.

Observation The value of a specific variable collected at a point in time or over a specified period of time.

Odds against *E* The reciprocal of odds for *E*.

Odds for *E* The probability of *E* divided by 1 minus the probability of *E*.

Off-the-run Seasoned government bonds are off-the-run securities; they are not the most recently issued or the most actively traded.

Offer The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size).

Official interest rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks. Also called *official policy rate* or *policy rate*.

Official policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Oligopoly Market structure with a relatively small number of firms supplying the market.

On-the-run The most recently issued and most actively traded sovereign securities.

One-dimensional array The simplest format for representing a collection of data of the same data type.

One-sided hypothesis test A test in which the null hypothesis is rejected only if the evidence indicates that the population parameter is greater than or less than the hypothesized parameter; occurs when the alternative hypothesis is stated either as greater than or less than the hypothesized population parameter.

Open economy An economy that trades with other countries.

Open-end fund A mutual fund that accepts new investment money and issues additional shares at a value equal to the net asset value of the fund at the time of investment.

Open interest The number of outstanding contracts in a clearinghouse at any given time. The open interest figure changes daily as some parties open up new positions, while other parties offset their old positions.

Open market operations The purchase or sale of bonds by the national central bank to implement monetary policy. The bonds traded are usually sovereign bonds issued by the national government.

Operating activities Activities that are part of the day-to-day business functioning of an entity, such as selling inventory and providing services.

Operating breakeven The number of units produced and sold at which the company's operating profit is zero (revenues = operating costs).

Operating cash flow The net amount of cash provided from operating activities.

Operating efficiency ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Operating lease A type of lease which is more akin to the rental of the underlying asset.

Operating leverage The use of fixed costs in operations.

Operating profit A company's profits on its usual business activities before deducting taxes. Also called *operating income*.

Operating profit margin A profitability ratio calculated as operating income (i.e., income before interest and taxes) divided by revenue. Also called *operating margin*.

Operating risk The risk attributed to the operating cost structure, in particular the use of fixed costs in operations; the risk arising from the mix of fixed and variable costs; the risk that a company's operations may be severely affected by environmental, social, and governance risk factors.

Operational independence A bank's ability to execute monetary policy and set interest rates in the way it thought would best meet the inflation target.

Operational risk The risk that arises from inadequate or failed people, systems, and internal policies, procedures, and processes, as well as from external events that are beyond the control of the organization but that affect its operations.

Operationally efficient Said of a market, a financial system, or an economy that has relatively low transaction costs.

Opportunity cost The value that investors forgo by choosing a particular course of action; the value of something in its best alternative use.

Optimal capital structure The capital structure at which the value of the company is maximized.

Option A financial instrument that gives one party the right, but not the obligation, to buy or sell an underlying asset from or to another party at a fixed price over a specific period of time. Also referred to as *contingent claim* or *option contract.*

Option-adjusted price The value of the embedded option plus the flat price of the bond.

Option-adjusted spread OAS = Z-spread – Option value (in basis points per year).

Option-adjusted yield The required market discount rate whereby the price is adjusted for the value of the embedded option.

Option contract See *option*.

Option premium The amount of money a buyer pays and seller receives to engage in an option transaction.

Order A specification of what instrument to trade, how much to trade, and whether to buy or sell.

Order-driven markets A market (generally an auction market) that uses rules to arrange trades based on the orders that traders submit; in their pure form, such markets do not make use of dealers.

Order precedence hierarchy With respect to the execution of orders to trade, a set of rules that determines which orders execute before other orders.

Ordinal data Categorical values that can be logically ordered or ranked.

Ordinary annuity An annuity with a first cash flow that is paid one period from the present.

Ordinary shares Equity shares that are subordinate to all other types of equity (e.g., preferred equity). Also called *common stock* or *common shares.*

Organized exchange A securities marketplace where buyers and seller can meet to arrange their trades.

Other comprehensive income Items of comprehensive income that are not reported on the income statement; comprehensive income minus net income.

Out-of-sample test A test of a strategy or model using a sample outside the period on which the strategy or model was developed.

Out of the money Options that, if exercised, would require the payment of more money than the value received and therefore would not be currently exercised.

Outcome A possible value of a random variable.

Over-the-counter (OTC) market A decentralized market where buy and sell orders initiated from various locations are matched through a communications network.

Overbought A market condition in which market sentiment is thought to be unsustainably bullish.

Overcollateralization Form of internal credit enhancement that refers to the process of posting more collateral than needed to obtain or secure financing.

Overconfidence bias A bias in which people demonstrate unwarranted faith in their own intuitive reasoning, judgments, and/or cognitive abilities.

Overfitting An undesirable result from fitting a model so closely to a dataset that it does not perform well on new data.

Oversold A market condition in which market sentiment is thought to be unsustainably bearish.

Own price The price of a good or service itself (as opposed to the price of something else).

Own-price elasticity of demand The percentage change in quantity demanded for a percentage change in good's own price, holding all other things constant.

Owners' equity The excess of assets over liabilities; the residual interest of shareholders in the assets of an entity after deducting the entity's liabilities. Also called *shareholders' equity* or *shareholders' funds*.

***p*-Value** The smallest level of significance at which the null is rejected.

Paired comparisons test See *test of the mean of the differences.*

Panel data A mix of time-series and cross-sectional data that contains observations through time on characteristics of across multiple observational units.

Par curve A sequence of yields-to-maturity such that each bond is priced at par value. The bonds are assumed to have the same currency, credit risk, liquidity, tax status, and annual yields stated for the same periodicity.

Par value The amount of principal on a bond.

Parallel shift A parallel yield curve shift implies that all rates change by the same amount in the same direction.

Parameter A descriptive measure computed from or used to describe a population of data, conventionally represented by Greek letters.

Parametric test Any test (or procedure) concerned with parameters or whose validity depends on assumptions concerning the population generating the sample.

Pari passu Covenant or contract clause that ensures a debt obligation is treated the same as the borrower's other senior debt instruments and is not subordinated to similar obligations.

Partial duration See *key rate duration.*

Participating preference shares Preference shares that entitle shareholders to receive the standard preferred dividend plus the opportunity to receive an additional dividend if the company's profits exceed a pre-specified level.

Pass-through rate The coupon rate of a mortgage pass-through security.

Passive investment A buy and hold approach in which an investor does not make portfolio changes based on short-term expectations of changing market or security performance.

Payable date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payment date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payments system The system for the transfer of money.

Pearson correlation A parametric measure of the relationship between two variables.

Pecking order theory The theory that managers consider how their actions might be interpreted by outsiders and thus order their preferences for various forms of corporate financing. Forms of financing that are least visible to outsiders (e.g., internally generated funds) are most preferable to managers, and those that are most visible (e.g., equity) are least preferable.

Peer company See *comparable company*.

Peer group A group of companies engaged in similar business activities whose economics and valuation are influenced by closely related factors.

Pennant A technical analysis continuation pattern formed by trendlines that converge to form a triangle, typically over a short period.

Per capita real GDP Real GDP divided by the size of the population, often used as a measure of a country's average standard of living.

Per unit contribution margin The amount that each unit sold contributes to covering fixed costs—that is, the difference between the price per unit and the variable cost per unit.

Percentiles Quantiles that divide a distribution into 100 equal parts that sum to 100.

Perfect capital markets Markets in which, by assumption, there are no taxes, transaction costs, or bankruptcy costs and in which all investors have equal ("symmetric") information.

Perfect competition A market structure in which the individual firm has virtually no impact on market price, because it is assumed to be a very small seller among a very large number of firms selling essentially identical products.

Perfectly elastic When the quantity demanded or supplied of a given good is infinitely sensitive to a change in the value of a specified variable (e.g., price).

Perfectly inelastic When the quantity demanded or supplied of a given good is completely insensitive to a change in the value of a specified variable (e.g., price).

Performance bond See *margin bond.*

Performance evaluation The measurement and assessment of the outcomes of investment management decisions.

Performance fee Fees paid to the general partner from the limited partner(s) based on realized net profits.

Period costs Costs (e.g., executives' salaries) that cannot be directly matched with the timing of revenues and which are thus expensed immediately.

Periodicity The assumed number of periods in the year; typically matches the frequency of coupon payments.

Permanent differences Differences between tax and financial reporting of revenue (expenses) that will not be reversed at some future date. These result in a difference between the company's effective tax rate and statutory tax rate and do not result in a deferred tax item.

Permissioned networks Networks that are fully open only to select participants on a DLT network.

Permissionless networks Networks that are fully open to any user on a DLT network.

Permutation An ordered listing.

Permutation formula The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is ${}_nP_r = \dfrac{n!}{(n-r)!}$.

Perpetual bonds Bonds with no stated maturity date.

Perpetuity A perpetual annuity, or a set of never-ending level sequential cash flows, with the first cash flow occurring one period from now. A bond that does not mature.

Personal income A broad measure of household income that includes all income received by households, whether earned or unearned; measures the ability of consumers to make purchases.

Pet projects Investments in which influential managers want the corporation to invest. Often, unfortunately, pet projects are selected without undergoing normal capital allocation analysis.

Plain vanilla bond Bond that makes periodic, fixed coupon payments during the bond's life and a lump-sum payment of principal at maturity. Also called *conventional bond*.

Platykurtic Describes a distribution that has relatively less weight in the tails than the normal distribution (also called thin-tailed).

Point estimate A single numerical estimate of an unknown quantity, such as a population parameter.

Policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Population All members of a specified group.

Portfolio company In private equity, the company in which the private equity fund is investing.

Portfolio demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments.

Portfolio planning The process of creating a plan for building a portfolio that is expected to satisfy a client's investment objectives.

Position The quantity of an asset that an entity owns or owes.

Positive screening The process of including sectors or companies based on specific ESG criteria, typically ESG performance relative to industry peers

Posterior probability An updated probability that reflects or comes after new information.

Potential GDP The level of real GDP that can be produced at full employment; measures the productive capacity of the economy.

Power of a test The probability of correctly rejecting the null—that is, rejecting the null hypothesis when it is false.

Precautionary money balances Money held to provide a buffer against unforeseen events that might require money.

Preference shares A type of equity interest which ranks above common shares with respect to the payment of dividends and the distribution of the company's net assets upon liquidation. They have characteristics of both debt and equity securities. Also called *preferred stock*.

Preferred stock See *preference shares*.

Premium In the case of bonds, premium refers to the amount by which a bond is priced above its face (par) value. In the case of an option, the amount paid for the option contract.

Prepaid expense A normal operating expense that has been paid in advance of when it is due.

Prepayment option Contractual provision that entitles the borrower to prepay all or part of the outstanding mortgage principal prior to the scheduled due date when the principal must be repaid. Also called *early repayment option*.

Prepayment penalty mortgages Mortgages that stipulate a monetary penalty if a borrower prepays within a certain time period after the mortgage is originated.

Prepayment risk The uncertainty that the timing of the actual cash flows will be different from the scheduled cash flows as set forth in the loan agreement due to the borrowers' ability to alter payments, usually to take advantage of interest rate movements.

Present value (PV) The present discounted value of future cash flows: For assets, the present discounted value of the future net cash inflows that the asset is expected to generate; for liabilities, the present discounted value of the future net cash outflows that are expected to be required to settle the liabilities.

Present value models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security. Also called *discounted cash flow models*.

Pretax margin A profitability ratio calculated as earnings before taxes divided by revenue.

Price elasticity of demand Measures the percentage change in the quantity demanded, given a percentage change in the price of a given product.

Price index Represents the average prices of a basket of goods and services.

Price limits Limits imposed by a futures exchange on the price change that can occur from one day to the next.

Price multiple A ratio that compares the share price with some sort of monetary flow or value to allow evaluation of the relative worth of a company's stock.

Price priority The principle that the highest priced buy orders and the lowest priced sell orders execute first.

Price relative A ratio of an ending price over a beginning price; it is equal to 1 plus the holding period return on the asset.

Price return Measures *only* the price appreciation or percentage change in price of the securities in an index or portfolio.

Price return index An index that reflects *only* the price appreciation or percentage change in price of the constituent securities. Also called *price index*.

Price-setting option The option to adjust prices when demand varies from what is forecast.

Price stability In economics, refers to an inflation rate that is low on average and not subject to wide fluctuation.

Price takers Producers that must accept whatever price the market dictates.

Price to book value A valuation ratio calculated as price per share divided by book value per share.

Price to cash flow A valuation ratio calculated as price per share divided by cash flow per share.

Price to earnings ratio (P/E ratio or P/E) The ratio of share price to earnings per share.

Price to sales A valuation ratio calculated as price per share divided by sales per share.

Price value of a basis point A version of money duration, it is an estimate of the change in the full price of a bond given a 1 basis point change in the yield-to-maturity.

Price weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its price by the sum of all the prices of the constituent securities.

Priced risk Risk for which investors demand compensation for bearing (e.g., equity risk, company-specific factors, macroeconomic factors).

Primary bond market A market in which issuers first sell bonds to investors to raise capital.

Primary capital markets (primary markets) The market where securities are first sold and the issuers receive the proceeds.

Primary dealer Financial institution that is authorized to deal in new issues of sovereign bonds and that serves primarily as a trading counterparty of the office responsible for issuing sovereign bonds.

Primary market The market where securities are first sold and the issuers receive the proceeds.

Prime brokers Brokers that provide services that commonly include custody, administration, lending, short borrowing, and trading.

Principal The amount of funds originally invested in a project or instrument; the face value to be paid at maturity.

Principal–agent relationship A relationship in which a principal hires an agent to perform a particular task or service; also known as an *agency relationship*.

Principal amount Amount that an issuer agrees to repay the debtholders on the maturity date.

Principal business activity The business activity from which a company derives a majority of its revenues and/or earnings.

Principal value Amount that an issuer agrees to repay the debtholders on the maturity date.

Principle of no arbitrage See *arbitrage-free pricing*.

Prior probabilities Probabilities reflecting beliefs prior to the arrival of new information.

Priority of claims Priority of payment, with the most senior or highest ranking debt having the first claim on the cash flows and assets of the issuer.

Private equity fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *venture capital fund*.

Private equity funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *venture capital funds*.

Private equity securities Securities that are not listed on public exchanges and have no active secondary market. They are issued primarily to institutional investors via non-public offerings, such as private placements.

Private investment in public equity (PIPE) An investment in the equity of a publicly traded firm that is made at a discount to the market value of the firm's shares.

Private placement Typically, a non-underwritten, unregistered offering of securities that are sold only to an investor or a small group of investors. It can be accomplished directly between the issuer and the investor(s) or through an investment bank.

Probability A number between 0 and 1 describing the chance that a stated event will occur.

Probability density function A function with non-negative values such that probability can be described by areas under the curve graphing the function.

Probability distribution A distribution that specifies the probabilities of a random variable's possible outcomes.

Probability function A function that specifies the probability that the random variable takes on a specific value.

Probability sampling A sampling plan that allows every member of the population to have an equal chance of being selected.

Probability tree diagram A diagram with branches emanating from nodes representing either mutually exclusive chance events or mutually exclusive decisions.

Production-flexibility options The options to alter production when demand varies from what is forecast.

Production function Provides the quantitative link between the levels of output that the economy can produce and the inputs used in the production process.

Productivity The amount of output produced by workers during a given period—for example, output per hour worked measures the efficiency of labor.

Profession An occupational group that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition.

Profit The return that owners of a company receive for the use of their capital and the assumption of financial risk when making their investments.

Profit and loss (P&L) statement A financial statement that provides information about a company's profitability over a stated period of time. Also called the *income statement*.

Profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses.

Profitability ratios Ratios that measure a company's ability to generate profitable sales from its resources (assets).

Project sequencing To defer the decision to invest in a future project until the outcome of some or all of a current investment is known. Investments are sequenced over time, so that making an investment creates the option to invest in future projects.

Promissory note A written promise to pay a certain amount of money on demand.

Property, plant, and equipment Tangible assets that are expected to be used for more than one period in either the production or supply of goods or services, or for administrative purposes.

Prospectus The document that describes the terms of a new bond issue and helps investors perform their analysis on the issue.

Protective put An option strategy in which a long position in an asset is combined with a long position in a put.

Proxy contest Corporate takeover mechanism in which shareholders are persuaded to vote for a group seeking a controlling position on a company's board of directors.

Proxy voting A process that enables shareholders who are unable to attend a meeting to authorize another individual to vote on their behalf.

Public offer See *public offering*.

Public offering An offering of securities in which any member of the public may buy the securities. Also called *public offer*.

Public–private partnership (PPP) An agreement between the public sector and the private sector to finance, build, and operate public infrastructure, such as hospitals and toll roads.

Pull on liquidity When disbursements are paid too quickly or trade credit availability is limited, requiring companies to expend funds before they receive funds from sales that could cover the liability.

Pure discount bonds See *zero-coupon bond*.

Put An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Put–call–forward parity The relationship among puts, calls, and forward contracts.

Put–call parity An equation expressing the equivalence (parity) of a portfolio of a call and a bond with a portfolio of a put and the underlying, which leads to the relationship between put and call prices.

Put/call ratio A technical analysis indicator that evaluates market sentiment based on the volume of put options traded divided by the volume of call options traded for a particular financial instrument.

Put option An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Putable bonds Bonds that give the bondholder the right to sell the bond back to the issuer at a predetermined price on specified dates.

Qualitative data see **categorical data**.

Quantile A value at or below which a stated fraction of the data lies. Also referred to as a fractile.

Quantitative data see **numerical data**.

Quantitative easing An expansionary monetary policy based on aggressive open market purchase operations.

Quantity equation of exchange An expression that over a given period, the amount of money used to purchase all goods and services in an economy, $M \times V$, is equal to monetary value of this output, $P \times Y$.

Quantity theory of money Asserts that total spending (in money terms) is proportional to the quantity of money.

Quartiles Quantiles that divide a distribution into four equal parts.

Quasi-fixed cost A cost that stays the same over a range of production but can change to another constant level when production moves outside of that range.

Quasi-government bond A bond issued by an entity that is either owned or sponsored by a national government. Also called *agency bond*.

Quintiles Quantiles that divide a distribution into five equal parts.

Quota rents Profits that foreign producers can earn by raising the price of their goods higher than they would without a quota.

Quotas Government policies that restrict the quantity of a good that can be imported into a country, generally for a specified period of time.

Quote-driven market A market in which dealers acting as principals facilitate trading.

Quoted interest rate A quoted interest rate that does not account for compounding within the year. Also called *stated annual interest rate*.

Quoted margin The specified yield spread over the reference rate, used to compensate an investor for the difference in the credit risk of the issuer and that implied by the reference rate.

Random number An observation drawn from a uniform distribution.

Random number generator An algorithm that produces uniformly distributed random numbers between 0 and 1.

Random variable A quantity whose future outcomes are uncertain.

Range The difference between the maximum and minimum values in a dataset.

Raw data Data available in their original form as collected.

Real exchange rate effect The effect through which changing price level impacts real exchange rate which in turn impacts net exports and aggregate demand.

Real GDP The value of goods and services produced, measured at base year prices.

Real income Income adjusted for the effect of inflation on the purchasing power of money. Also known as the *purchasing power of income*. If income remains constant and a good's price falls, real income is said to rise, even though the number of monetary units (e.g., dollars) remains unchanged.

Real interest rate Nominal interest rate minus the expected rate of inflation.

Real risk-free interest rate The single-period interest rate for a completely risk-free security if no inflation were expected.

Realizable (settlement) value With reference to assets, the amount of cash or cash equivalents that could currently be obtained by selling the asset in an orderly disposal; with reference to liabilities, the undiscounted amount of cash or cash equivalents expected to be paid to satisfy the liabilities in the normal course of business.

Rebalancing Adjusting the weights of the constituent securities in an index.

Rebalancing policy The set of rules that guide the process of restoring a portfolio's asset class weights to those specified in the strategic asset allocation.

Recession A period during which real GDP decreases (i.e., negative growth) for at least two successive quarters, or a period of significant decline in total output, income, employment, and sales usually lasting from six months to a year.

Recognition lag The lag in government response to an economic problem resulting from the delay in confirming a change in the state of the economy.

Recourse loan A loan in which the lender has a claim against the borrower for any shortfall between the outstanding mortgage balance and the proceeds received from the sale of the property.

Redemption yield See *yield-to-maturity*.

Redemptions Withdrawals of funds by investors, as allowed by the notice period and other terms in the partnership agreement.

Refinancing rate A type of central bank policy rate.

Registered bonds Bonds for which ownership is recorded by either name or serial number.

Regression analysis A tool for examining whether a variable is useful for explaining another variable.

Regression coefficients The intercept and slope coefficient of a regression.

Regret The feeling that an opportunity has been missed; typically, an expression of *hindsight bias*.

Regret-aversion bias An emotional bias in which people tend to avoid making decisions that will result in action out of fear that the decision will turn out poorly.

Relative dispersion The amount of dispersion relative to a reference value or benchmark.

Relative frequency The absolute frequency of each unique value of the variable divided by the total number of observations of the variable.

Relative price The price of a specific good or service in comparison with those of other goods and services.

Relative strength analysis A comparison of the performance of one asset with the performance of another asset or a benchmark, based on changes in the ratio of the two assets' prices over time.

Relative strength index (RSI) A technical analysis momentum oscillator that compares a security's gains with its losses over a set period.

Renewable resources Resources that can be replenished, such as a forest.

Rent Payment for the use of property.

Reorganization Agreements made by a company in bankruptcy under which a company's capital structure is altered and/or alternative arrangements are made for debt repayment; US Chapter 11 bankruptcy. The company emerges from bankruptcy as a going concern.

Replication The creation of an asset or portfolio from another asset, portfolio, and/or derivative.

Repo A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy back the same security from the purchaser at an agreed-on price and future date. The party who sells the

security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repo margin The difference between the market value of the security used as collateral and the value of the loan. Also called *haircut*.

Repo rate The interest rate on a repurchase agreement.

Representativeness bias A belief perseverance bias in which people tend to classify new information based on past experiences and classifications.

Repurchase agreement A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy the same security back from the purchaser at an agreed-on price and future date. The party who sells the security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repurchase date The date when the party who sold the security at the inception of a repurchase agreement buys back the security from the cash lending counterparty.

Repurchase price The price at which the party who sold the security at the inception of the repurchase agreement buys back the security from the cash lending counterparty.

Required margin The yield spread over or under the reference rate such that an FRN is priced at par value on a rate reset date.

Required rate of return See *market discount rate*.

Required yield See *market discount rate*.

Required yield spread The difference between the yield-to-maturity on a new bond and the benchmark rate; additional compensation required by investors for the difference in risk and tax status of a bond relative to a government bond. Sometimes called the *spread over the benchmark*.

Resampling A statistical method that repeatedly draws samples from the original observed data sample for the statistical inference of population parameters.

Reserve accounts Form of internal credit enhancement that relies on creating accounts and depositing in these accounts cash that can be used to absorb losses. Also called *reserve funds*.

Reserve funds See *reserve accounts*.

Reserve requirement The requirement for banks to hold reserves in proportion to the size of deposits.

Residual The difference between an observation and its predicted value, where the predicted value is based on the estimated linear relation between the dependent and independent variables using sample data.

Resistance In technical analysis, a price range in which selling activity is sufficient to stop the rise in the price of a security.

Responsible investing An overall (umbrella) term for several investment strategies—ESG investing, SRI investing, thematic investing, and impact investing—that incorporate ESG analysis in their investment processes.

Restricted payments A bond covenant meant to protect creditors by limiting how much cash can be paid out to shareholders over time.

Retracement In technical analysis, a reversal in the movement of a security's price such that it is counter to the prevailing longer-term price trend.

Return-generating model A model that can provide an estimate of the expected return of a security given certain parameters and estimates of the values of the independent variables in the model.

Return on assets (ROA) A profitability ratio calculated as net income divided by average total assets; indicates a company's net profit generated per dollar invested in total assets.

Return on equity (ROE) A profitability ratio calculated as net income divided by average shareholders' equity.

Return on invested capital A measure of the profitability of a company relative to the amount of capital invested by the equity- and debtholders.

Return on sales An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also referred to as *net profit margin*.

Return on total capital A profitability ratio calculated as EBIT divided by the sum of short- and long-term debt and equity.

Revaluation model Under IFRS, the process of valuing long-lived assets at fair value, rather than at cost less accumulated depreciation. Any resulting profit or loss is either reported on the income statement and/or through equity under revaluation surplus.

Revenue The amount charged for the delivery of goods or services in the ordinary activities of a business over a stated period; the inflows of economic resources to a company over a stated period.

Reversal pattern A type of pattern used in technical analysis to predict the end of a trend and a change in the direction of a security's price.

Reverse repo A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse repurchase agreement A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse stock split A reduction in the number of shares outstanding with a corresponding increase in share price, but no change to the company's underlying fundamentals.

Revolving credit agreements (also known as revolvers) The most reliable form of short-term bank borrowing facilities; they are in effect for multiple years (e.g., three to five years) and can have optional medium-term loan features.

Rho The sensitivity of the option price to the risk-free rate.

Ricardian equivalence An economic theory that implies that it makes no difference whether a government finances a deficit by increasing taxes or issuing debt.

Risk Exposure to uncertainty. The chance of a loss or adverse outcome as a result of an action, inaction, or external event.

Risk averse The assumption that an investor will choose the least risky alternative.

Risk aversion The degree of an investor's inability and unwillingness to take risk.

Risk budgeting The establishment of objectives for individuals, groups, or divisions of an organization that takes into account the allocation of an acceptable level of risk.

Risk exposure The state of being exposed or vulnerable to a risk. The extent to which an organization is sensitive to underlying risks.

Risk governance The top-down process and guidance that directs risk management activities to align with and support the overall enterprise.

Risk management The process of identifying the level of risk an organization wants, measuring the level of risk the organization currently has, taking actions that bring the actual

level of risk to the desired level of risk, and monitoring the new actual level of risk so that it continues to be aligned with the desired level of risk.

Risk management framework The infrastructure, process, and analytics needed to support effective risk management in an organization.

Risk-neutral pricing Sometimes said of derivatives pricing, uses the fact that arbitrage opportunities guarantee that a risk-free portfolio consisting of the underlying and the derivative must earn the risk-free rate.

Risk-neutral probabilities Weights that are used to compute a binomial option price. They are the probabilities that would apply if a risk-neutral investor valued an option.

Risk premium An extra return expected by investors for bearing some specified risk.

Risk shifting Actions to change the distribution of risk outcomes.

Risk tolerance The amount of risk an investor is willing and able to bear to achieve an investment goal.

Risk transfer Actions to pass on a risk to another party, often, but not always, in the form of an insurance policy.

Robo-adviser A machine-based analytical tool or service that provides technology-driven investment solutions through online platforms.

Rule of 72 The principle that the approximate number of years necessary for an investment to double is 72 divided by the stated interest rate.

Running yield See *current yield*.

Safety-first rules Rules for portfolio selection that focus on the risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Sales Generally, a synonym for revenue; “sales” is generally understood to refer to the sale of goods, whereas “revenue” is understood to include the sale of goods or services.

Sales risk Uncertainty with respect to the quantity of goods and services that a company is able to sell and the price it is able to achieve; the risk related to the uncertainty of revenues.

Sample A subset of a population.

Sample correlation coefficient A standardized measure of how two variables in a sample move together. It is the ratio of the sample covariance to the product of the two variables’ standard deviations.

Sample covariance A measure of how two variables in a sample move together.

Sample excess kurtosis A sample measure of the degree of a distribution’s kurtosis in excess of the normal distribution’s kurtosis.

Sample mean The sum of the sample observations divided by the sample size.

Sample selection bias Bias introduced by systematically excluding some members of the population according to a particular attribute—for example, the bias introduced when data availability leads to certain observations being excluded from the analysis.

Sample-size neglect A type of representativeness bias in which financial market participants incorrectly assume that small sample sizes are representative of populations (or “real” data).

Sample skewness A sample measure of the degree of asymmetry of a distribution.

Sample standard deviation The positive square root of the sample variance.

Sample statistic A quantity computed from or used to describe a sample.

Sample variance The sum of squared deviations around the mean divided by the degrees of freedom.

Sampling The process of obtaining a sample.

Sampling distribution The distribution of all distinct possible values that a statistic can assume when computed from samples of the same size randomly drawn from the same population.

Sampling error The difference between the observed value of a statistic and the estimate resulting from using subsets of the population.

Sampling plan The set of rules used to select a sample.

Say on pay A process whereby shareholders may vote on executive remuneration (compensation) matters.

Scatter plot A chart in which two variables are plotted along the axis and points on the chart represent pairs of the two variables. In regression, the dependent variable is plotted on the vertical axis and the independent variable is plotted along the horizontal axis. Also known as a scattergram and a *scatter diagram.*

Scatter plot matrix A tool for organizing scatter plots between pairs of variables, making it easy to inspect all pairwise relationships in one combined visual.

Scenario analysis Analysis that shows the changes in key financial quantities that result from given (economic) events, such as the loss of customers, the loss of a supply source, or a catastrophic event; a risk management technique involving examination of the performance of a portfolio under specified situations. Closely related to stress testing.

Screening The application of a set of criteria to reduce a set of potential investments to a smaller set having certain desired characteristics.

Seasoned offering An offering in which an issuer sells additional units of a previously issued security.

Second-degree price discrimination When the monopolist charges different per-unit prices using the quantity purchased as an indicator of how highly the customer values the product.

Second lien A secured interest in the pledged assets that ranks below first lien debt in both collateral protection and priority of payment.

Secondary bond markets Markets in which existing bonds are traded among investors.

Secondary market The market where securities are traded among investors.

Secondary precedence rules Rules that determine how to rank orders placed at the same time.

Sector A group of related industries (GICS definition).

Sector indexes Indexes that represent and track different economic sectors—such as consumer goods, energy, finance, health care, and technology—on either a national, regional, or global basis.

Secured (“asset-based”) loan A loan that is backed by specific, secured company assets.

Secured bonds Bonds secured by assets or financial guarantees pledged to ensure debt repayment in case of default.

Secured debt Debt in which the debtholder has a direct claim—a pledge from the issuer—on certain assets and their associated cash flows.

Securitization A process that involves moving assets into a special legal entity, which then uses the assets as guarantees to secure a bond issue.

Securitized assets Assets that are typically used to create asset-backed bonds; for example, when a bank securitizes a pool of loans, the loans are said to be securitized.

Security characteristic line A plot of the excess return of a security on the excess return of the market.

Security market index A portfolio of securities representing a given security market, market segment, or asset class.

Security market line (SML) The graph of the capital asset pricing model.

Security selection The process of selecting individual securities; typically, security selection has the objective of generating superior risk-adjusted returns relative to a portfolio's benchmark.

Self-attribution bias A bias in which people take too much credit for successes (*self-enhancing*) and assign responsibility to others for failures (*self-protecting*).

Self-control bias A bias in which people fail to act in pursuit of their long-term, overarching goals because of a lack of self-discipline.

Self-investment limits With respect to investment limitations applying to pension plans, restrictions on the percentage of assets that can be invested in securities issued by the pension plan sponsor.

Sell-side firm A broker/dealer that sells securities and provides independent investment research and recommendations to their clients (i.e., buy-side firms).

Semi-strong-form efficient market A market in which security prices reflect all publicly known and available information.

Semiannual bond basis yield An annual rate having a periodicity of two; also known as a *semiannual bond equivalent yield*.

Semiannual bond equivalent yield See *semiannual bond basis yield*.

Seniority ranking Priority of payment of various debt obligations.

Sensitivity analysis Analysis that shows the range of possible outcomes as specific assumptions are changed.

Separately managed account (SMA) An investment portfolio managed exclusively for the benefit of an individual or institution.

Serial maturity structure Structure for a bond issue in which the maturity dates are spread out during the bond's life; a stated number of bonds mature and are paid off each year before final maturity.

Settlement The process that occurs after a trade is completed, the securities are passed to the buyer, and payment is received by the seller.

Settlement date Date when the buyer makes cash payment and the seller delivers the security.

Settlement price The official price, designated by the clearinghouse, from which daily gains and losses will be determined and marked to market.

Share repurchase A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Shareholder activism Strategies used by shareholders to attempt to compel a company to act in a desired manner.

Shareholder engagement The process whereby companies engage with their shareholders.

Shareholders' equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Sharpe ratio The average return in excess of the risk-free rate divided by the standard deviation of return; a measure of the average excess return earned per unit of standard deviation of return.

Shelf registration A type of public offering that allows the issuer to file a single, all-encompassing offering circular that covers a series of bond issues.

Short The seller of an asset or derivative contract. Also refers to the position of being short an asset or derivative contract.

Short position A position in an asset or contract in which one has sold an asset one does not own, or in which a right under a contract can be exercised against oneself.

Short-run average total cost The curve describing average total cost when some costs are considered fixed.

Short selling A transaction in which borrowed securities are sold with the intention to repurchase them at a lower price at a later date and return them to the lender.

Shortfall risk The risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Shutdown point The point at which average revenue is equal to the firm's average variable cost.

Side letters Side agreements created between the GP and specific LPs. These agreements exist *outside* the LPA. These agreements provide additional terms and conditions related to the investment agreement.

Simple interest The interest earned each period on the original investment; interest calculated on the principal only.

Simple linear regression (SLR) A regression that summarizes the relation between the dependent variable and a single independent variable.

Simple random sample A subset of a larger population created in such a way that each element of the population has an equal probability of being selected to the subset.

Simple random sampling The procedure of drawing a sample to satisfy the definition of a simple random sample.

Simple yield The sum of the coupon payments plus the straight-line amortized share of the gain or loss, divided by the flat price.

Simulation Computer-generated sensitivity or scenario analysis that is based on probability models for the factors that drive outcomes.

Simulation trial A complete pass through the steps of a simulation.

Single-step format With respect to the format of the income statement, a format that does not subtotal for gross profit (revenue minus cost of goods sold).

Sinking fund arrangement Provision that reduces the credit risk of a bond issue by requiring the issuer to retire a portion of the bond's principal outstanding each year.

Situational influences External factors, such as environmental or cultural elements, that shape our behavior.

Skewed Not symmetrical.

Skewness A quantitative measure of skew (lack of symmetry); a synonym of skew. It is computed as the average cubed deviation from the mean standardized by dividing by the standard deviation cubed.

Slope coefficient The coefficient of an independent variable that represents the average change in the dependent variable for a one-unit change in the independent variable.

Small country A country that is a price taker in the world market for a product and cannot influence the world market price.

Smart beta Involves the use of simple, transparent, rules-based strategies as a basis for investment decisions.

Smart contract A computer program that is designed to self-execute on the basis of pre-specified terms and conditions agreed to by parties to a contract.

Socially responsible investing An investment approach that excludes investments in companies or industries that deviate from an organization's beliefs and sometimes includes investments with favorable environmental or social profiles.

Soft-bullet covered bonds Covered bonds for which bond default and payment acceleration of bond cash flows may be delayed upon sponsor default until a new final maturity date is reached.

Solvency With respect to financial statement analysis, the ability of a company to fulfill its long-term obligations.

Solvency ratios Ratios that measure a company's ability to meet its long-term obligations.

Solvency risk The risk that an organization does not survive or succeed because it runs out of cash, even though it might otherwise be solvent.

Sovereign bond A bond issued by a national government. Also called "Sovereign."

Spearman rank correlation coefficient A measure of correlation applied to ranked data.

Special dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Special purpose entity A non-operating entity created to carry out a specified purpose, such as leasing assets or securitizing receivables; can be a corporation, partnership, trust, or limited liability partnership formed to facilitate a specific type of business activity. Also called *special purpose vehicle*, *special purpose company*, or *variable interest entity*.

Special purpose vehicle See *special purpose entity*.

Specific identification method An inventory accounting method that identifies which specific inventory items were sold and which remained in inventory to be carried over to later periods.

Speculative demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments. Also called *portfolio demand for money*.

Speculative money balances Monies held in anticipation that other assets will decline in value.

Split coupon bond See *deferred coupon bond*.

Sponsored A type of depository receipt in which the foreign company whose shares are held by the depository has a direct involvement in the issuance of the receipts.

Spot curve A sequence of yields-to-maturity on zero-coupon bonds. Sometimes called *zero* or *strip curve* (because coupon payments are "stripped" off the bonds).

Spot markets Markets in which assets are traded for immediate delivery.

Spot prices The price of an asset for immediately delivery.

Spot rates A sequence of market discount rates that correspond to the cash flow dates; yields-to-maturity on zero-coupon bonds maturing at the date of each cash flow.

Spread In general, the difference in yield between different fixed-income securities. Often used to refer to the difference between the yield-to-maturity and the benchmark.

Spread over the benchmark See *required yield spread*.

Spread risk Bond price risk arising from changes in the yield spread on credit-risky bonds; reflects changes in the market's assessment and/or pricing of credit migration (or downgrade) risk and market liquidity risk.

Spurious correlation Refers to: 1) correlation between two variables that reflects chance relationships in a particular dataset; 2) correlation induced by a calculation that mixes each of two variables with a third variable; and 3) correlation between two variables arising not from a direct relation between them but from their relation to a third variable.

Stacked bar chart An alternative form for presenting the frequency distribution of two categorical variables, where bars representing the sub-groups are placed on top of each other to form a single bar. Each sub-section is shown in a different color to represent the contribution of each sub-group, and the overall height of the stacked bar represents the marginal frequency for the category.

Stackelberg model A prominent model of strategic decision making in which firms are assumed to make their decisions sequentially.

Stagflation The combination of a high inflation rate with a high level of unemployment and a slowdown of the economy.

Staggered boards Election process whereby directors are typically divided into multiple classes that are elected separately in consecutive years—that is, one class every year.

Stakeholder management The identification, prioritization, and understanding of the interests of stakeholder groups and managing the company's relationships with these groups.

Stakeholders Individuals or groups of individuals who may be affected either directly or indirectly by a decision and thus have an interest, or stake, in the decision.

Standard deviation The positive square root of the variance; a measure of dispersion in the same units as the original data.

Standard error of the estimate A measure of the fit of a regression line, calculated as the square root of the mean square error. Also known as the *standard error of the regression* and the *root mean square error*.

Standard error of the forecast A measure of the uncertainty associated with a forecasted value of the dependent variable that depends on the standard error of the estimate, the variability of the independent variable, the deviation of the forecasted independent variable from the mean in the regression, and the number of observations.

Standard error of the slope coefficient The standard error of the slope, which in a simple linear regression is the ratio of the model's standard error of the estimate (s_e) to the square root of the variation of the independent variable.

Standard normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Standardizing A transformation that involves subtracting the mean and dividing the result by the standard deviation.

Standards of conduct Behaviors required by a group; established benchmarks that clarify or enhance a group's code of ethics.

Standing limit orders A limit order at a price below market and which therefore is waiting to trade.

Stated annual interest rate A quoted interest rate that does not account for compounding within the year. Also called *quoted interest rate*.

Statement of changes in equity (statement of owners' equity) A financial statement that reconciles the beginning-of-period and end-of-period balance sheet values of

shareholders' equity; provides information about all factors affecting shareholders' equity. Also called *statement of owners' equity*.

Statement of financial condition The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of financial position The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of operations A financial statement that provides information about a company's profitability over a stated period of time.

Static trade-off theory of capital structure A theory pertaining to a company's optimal capital structure. The optimal level of debt is found at the point where additional debt would cause the costs of financial distress to increase by a greater amount than the benefit of the additional tax shield.

Statistic A summary measure of a sample of observations.

Statistically significant A result indicating that the null hypothesis can be rejected; with reference to an estimated regression coefficient, frequently understood to mean a result indicating that the corresponding population regression coefficient is different from zero.

Status quo bias An emotional bias in which people do nothing (i.e., maintain the status quo) instead of making a change.

Statutory voting A common method of voting where each share represents one vote.

Step-up coupon bond Bond for which the coupon, fixed or floating, increases by specified margins at specified dates.

Stochastic oscillator A momentum indicator that compares a particular closing price of a security to a range of the security's prices over a certain period of time.

Stock dividend A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Stock split An increase in the number of shares outstanding with a consequent decrease in share price, but no change to the company's underlying fundamentals.

Stop-loss order See *stop order*.

Stop order An order in which a trader has specified a stop price condition. Also called *stop-loss order*.

Store of value The quality of tending to preserve value.

Store of wealth Goods that depend on the fact that they do not perish physically over time, and on the belief that others would always value the good.

Straight-line method A depreciation method that allocates evenly the cost of a long-lived asset less its estimated residual value over the estimated useful life of the asset.

Straight voting A shareholder voting process in which shareholders receive one vote for each share owned.

Strategic analysis Analysis of the competitive environment with an emphasis on the implications of the environment for corporate strategy.

Strategic asset allocation The set of exposures to IPS-permissible asset classes that is expected to achieve the client's long-term objectives given the client's investment constraints.

Stratified random sampling A procedure that first divides a population into subpopulations (strata) based on classification criteria and then randomly draws samples from each stratum in sizes proportional to that of each stratum in the population.

Street convention A yield measure that neglects weekends and holidays; the internal rate of return on cash flows assuming payments are made on the scheduled dates, even when the scheduled date falls on a weekend or holiday.

Stress testing A specific type of scenario analysis that estimates losses in rare and extremely unfavorable combinations of events or scenarios.

Strong-form efficient market A market in which security prices reflect all public and private information.

Structural (or cyclically adjusted) budget deficit The deficit that would exist if the economy was at full employment (or full potential output).

Structural subordination Arises in a holding company structure when the debt of operating subsidiaries is serviced by the cash flow and assets of the subsidiaries before funds can be passed to the holding company to service debt at the parent level.

Structured data Data that are highly organized in a predefined manner, usually with repeating patterns.

Structured financial instrument A financial instrument that shares the common attribute of repackaging risks. Structured financial instruments include asset-backed securities, collateralized debt obligations, and other structured financial instruments such as capital protected, yield enhancement, participation, and leveraged instruments.

Subjective probability A probability drawing on personal or subjective judgment.

Subordinated debt A class of unsecured debt that ranks below a firm's senior unsecured obligations.

Subordination Form of internal credit enhancement that relies on creating more than one bond tranche and ordering the claim priorities for ownership or interest in an asset between the tranches. The ordering of the claim priorities is called a senior/subordinated structure, where the tranches of highest seniority are called senior followed by subordinated or junior tranches. Also called *credit tranching*.

Substitutes Said of two goods or services such that if the price of one increases the demand for the other tends to increase, holding all other things equal (e.g., butter and margarine).

Sum of squares error (SSE) The sum of the squared deviations of (1) the value of the dependent variable and (2) the value of the dependent variable based on the estimated regression line. Also referred to as the *residual sum of squares*.

Sum of squares regression (SSR) The sum of the squared deviations of (1) the value of the dependent variable based on the estimated regression line and (2) the mean of the dependent variable.

Sum of squares total (SST) The sum of the squared deviations of the dependent variable from its mean; the variation of the dependent variable. Also referred to as the *total sum of squares*.

Sunk cost A cost that has already been incurred.

Supervised learning A machine learning approach that makes use of labeled training data.

Supply shock A typically unexpected disturbance to supply.

Support In technical analysis, a price range in which buying activity is sufficient to stop the decline in the price of a security.

Support tranches Classes or tranches in CMOs that protect PAC tranches from prepayment risk.

Supranational bond A bond issued by a supranational agency such as the World Bank.

Surety bond Form of external credit enhancement whereby a rated and regulated insurance company guarantees to reimburse bondholders for any losses incurred up to a maximum amount if the issuer defaults.

Survey approach An estimate of the equity risk premium that is based on estimates provided by a panel of finance experts.

Survivorship bias Relates to the inclusion of only current investment funds in a database. As such, the returns of funds that are no longer available in the marketplace (have been liquidated) are excluded from the database. In addition, *backfill bias* is another problem, whereby certain surviving hedge funds may be added to databases and various hedge fund indexes only after they are initially successful and start to report their returns.

Sustainability linked loans These are any types of loan instruments and/or contingent facilities (such as bonding lines, guarantee lines, or letters of credit) that incentivize the borrower's achievement of ambitious, pre-determined sustainability performance objectives.

Sustainable growth rate The rate of dividend (and earnings) growth that can be sustained over time for a given level of return on equity, keeping the capital structure constant and without issuing additional common stock.

Sustainable investing A term used in a similar context to responsible investing, but its key focus is on factoring in sustainability issues while investing.

Sustainable rate of economic growth The rate of increase in the economy's productive capacity or potential GDP.

Swap contract An agreement between two parties to exchange a series of future cash flows.

Syndicated loan A loan from a group of lenders to a single borrower.

Syndicated offering A bond issue underwritten by a group of investment banks.

Systematic risk Risk that affects the entire market or economy; it cannot be avoided and is inherent in the overall market. Systematic risk is also known as non-diversifiable or market risk.

Systematic sampling A procedure of selecting every *k*th member until reaching a sample of the desired size. The sample that results from this procedure should be approximately random.

Tactical asset allocation The decision to deliberately deviate from the strategic asset allocation in an attempt to add value based on forecasts of the near-term relative performance of asset classes.

Tactical asset allocation (TAA) A portfolio strategy that shifts the percentages of assets held in various asset classes (or categories) to take advantage of market opportunities. Allocation shifts can occur within an asset class or across asset classes.

Tag cloud see **word cloud**.

Target capital structure A company's chosen proportions of debt and equity.

Target independent A bank's ability to determine the definition of inflation that they target, the rate of inflation that they target, and the horizon over which the target is to be achieved.

Target semideviation A measure of downside risk, calculated as the square root of the average of the squared deviations of observations below the target (also called target downside deviation).

Tariffs Taxes that a government levies on imported goods.

Tax base The amount at which an asset or liability is valued for tax purposes.

Tax expense An aggregate of an entity's income tax payable (or recoverable in the case of a tax benefit) and any changes in deferred tax assets and liabilities. It is essentially the income tax payable or recoverable if these had been determined based on accounting profit rather than taxable income.

Tax loss carry forward A taxable loss in the current period that may be used to reduce future taxable income.

Taxable income The portion of an entity's income that is subject to income taxes under the tax laws of its jurisdiction.

Taxable temporary differences Temporary differences that result in a taxable amount in a future period when determining the taxable profit as the balance sheet item is recovered or settled.

Technical analysis A form of security analysis that uses price and volume data, often displayed graphically, in decision making.

Technology The process a company uses to transform inputs into outputs.

Tender offer Corporate takeover mechanism that involves shareholders selling their interests directly to the group seeking to gain control.

Tenor The time-to-maturity for a bond or derivative contract. Also called *term to maturity*.

Term maturity structure Structure for a bond issue in which the bond's notional principal is paid off in a lump sum at maturity.

Term structure See *maturity structure*.

Term structure of credit spreads The relationship between the spreads over the "risk-free" (or benchmark) rates and times-to-maturity.

Term structure of yield volatility The relationship between the volatility of bond yields-to-maturity and times-to-maturity.

Terminal stock value The expected value of a share at the end of the investment horizon—in effect, the expected selling price. Also called *terminal value*.

Terminal value The expected value of a share at the end of the investment horizon—in effect, the expected selling price.

Terms of trade The ratio of the price of exports to the price of imports, representing those prices by export and import price indexes, respectively.

Test of the mean of the differences A statistical test for differences based on paired observations drawn from samples that are dependent on each other.

Text analytics The use of computer programs to analyze and derive meaning from typically large, unstructured text- or voice-based datasets.

Thematic investing An ESG investment approach that focuses on investing in themes or assets specifically relating to ESG factors, such as clean energy, green technology, or sustainable agriculture.

Thin-Tailed Describes a distribution that has relatively less weight in the tails than the normal distribution (also called platykurtic).

Third-degree price discrimination When the monopolist segregates customers into groups based on demographic or other characteristics and offers different pricing to each group.

Time-period bias The possibility that when we use a time-series sample, our statistical conclusion may be sensitive to the starting and ending dates of the sample.

Time-series data A sequence of observations for a single observational unit of a specific variable collected over time and at discrete and typically equally spaced intervals of time (such as daily, weekly, monthly, annually, or quarterly).

Time tranching The creation of classes or tranches in an ABS/MBS that possess different (expected) maturities.

Time value The difference between the market price of the option and its intrinsic value.

Time value decay Said of an option when, at expiration, no time value remains and the option is worth only its exercise value.

Time value of money The principles governing equivalence relationships between cash flows with different dates.

Time-weighted rate of return The compound rate of growth of one unit of currency invested in a portfolio during a stated measurement period; a measure of investment performance that is not sensitive to the timing and amount of withdrawals or additions to the portfolio.

Tokenization The process of representing ownership rights to physical assets on a blockchain or distributed ledger.

Top-down analysis An investment selection approach that begins with consideration of macroeconomic conditions and then evaluates markets and industries based upon such conditions.

Total comprehensive income The change in equity during a period resulting from transaction and other events, other than those changes resulting from transactions with owners in their capacity as owners.

Total cost The summation of all costs, for which costs are classified as fixed or variable.

Total factor productivity A scale factor that reflects the portion of growth unaccounted for by explicit factor inputs (e.g., capital and labor).

Total fixed cost The summation of all expenses that do not change as the level of production varies.

Total invested capital The sum of market value of common equity, book value of preferred equity, and face value of debt.

Total probability rule A rule explaining the unconditional probability of an event in terms of probabilities of the event conditional on mutually exclusive and exhaustive scenarios.

Total probability rule for expected value A rule explaining the expected value of a random variable in terms of expected values of the random variable conditional on mutually exclusive and exhaustive scenarios.

Total return Measures the price appreciation, or percentage change in price of the securities in an index or portfolio, plus any income received over the period.

Total return index An index that reflects the price appreciation or percentage change in price of the constituent securities plus any income received since inception.

Total return swap A swap in which one party agrees to pay the total return on a security. Often used as a credit derivative, in which the underlying is a bond.

Total variable cost The summation of all variable expenses.

Tracking error The standard deviation of the differences between a portfolio's returns and its benchmark's returns; a synonym of active risk.

Tracking risk The standard deviation of the differences between a portfolio's returns and its benchmarks returns. Also called *tracking error*.

Trade creation When regional integration results in the replacement of higher cost domestic production by lower cost imports from other members.

Trade credit A spontaneous form of credit in which a purchaser of the goods or service is financing its purchase by delaying the date on which payment is made.

Trade diversion When regional integration results in lower-cost imports from non-member countries being replaced with higher-cost imports from members.

Trade payables Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Trade protection Government policies that impose restrictions on trade, such as tariffs and quotas.

Trade surplus (deficit) When the value of exports is greater (less) than the value of imports.

Trading securities Under US GAAP, a category of debt securities held by a company with the intent to trade them. Also called *held-for-trading securities*.

Traditional investment markets Markets for traditional investments, which include all publicly traded debts and equities and shares in pooled investment vehicles that hold publicly traded debts and/or equities.

Transactions money balances Money balances that are held to finance transactions.

Transfer payments Welfare payments made through the social security system that exist to provide a basic minimum level of income for low-income households.

Transparency Said of something (e.g., a market) in which information is fully disclosed to the public and/or regulators.

Treasury stock method A method for accounting for the effect of options (and warrants) on earnings per share (EPS) that specifies what EPS would have been if the options and warrants had been exercised and the company had used the proceeds to repurchase common stock.

Tree-Map Another graphical tool for displaying categorical data. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group.

Trend A long-term pattern of movement in a particular direction.

Treynor ratio A measure of risk-adjusted performance that relates a portfolio's excess returns to the portfolio's beta.

Triangle pattern In technical analysis, a continuation chart pattern that forms as the range between high and low prices narrows, visually forming a triangle.

Trimmed mean A mean computed after excluding a stated small percentage of the lowest and highest observations.

Trimodal A distribution that has the three most frequently occurring values.

Triple bottom In technical analysis, a reversal pattern that results when the price forms three troughs at roughly the same price level. A triple bottom is used to predict a change from a downtrend to an uptrend.

Triple top In technical analysis, a reversal pattern that results when the price forms three peaks at roughly the same price level. A triple top is used to predict a change from an uptrend to a downtrend.

True yield The internal rate of return on cash flows using the actual calendar, including weekends and bank holidays.

Trust deed The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *bond indenture*.

Turn-of-the-year effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January.

Two-dimensional rectangular array A popular form for organizing data for processing by computers or for presenting data visually. It is comprised of columns and rows to hold multiple variables and multiple observations, respectively (also called a **data table**).

Two-fund separation theorem The theory that all investors regardless of taste, risk preferences, and initial wealth will hold a combination of two portfolios or funds: a risk-free asset and an optimal portfolio of risky assets.

Two-sided hypothesis test A test in which the null hypothesis is rejected in favor of the alternative hypothesis if the evidence indicates that the population parameter is either smaller or larger than a hypothesized value; occurs when the alternative hypothesis is stated as not equal to the hypothesized population parameters.

Two-way table See *contingency table*.

Two-week repo rate The interest rate on a two-week repurchase agreement; may be used as a policy rate by a central bank.

Type I error The error of rejecting a true null hypothesis; a false positive.

Type II error The error of not rejecting a false null hypothesis; false negative.

Unanticipated (unexpected) inflation The component of inflation that is a surprise.

Uncommitted lines of credit The least reliable form of bank borrowing in which a bank offers, without formal commitment, a line of credit for an extended period of time but reserves the right to refuse any request for its use.

Unconditional probability The probability of an event *not* conditioned on another event.

Underemployed A person who has a job but has the qualifications to work a significantly higher-paying job.

Underlying An asset that trades in a market in which buyers and sellers meet, decide on a price, and the seller then delivers the asset to the buyer and receives payment. The underlying is the asset or other derivative on which a particular derivative is based. The market for the underlying is also referred to as the *spot market*.

Underwriter A firm, usually an investment bank, that takes the risk of buying the newly issued securities from the issuer and then reselling them to investors or to dealers, thus guaranteeing the sale of the securities at the offering price negotiated with the issuer.

Underwritten offering A type of securities issue mechanism in which the investment bank guarantees the sale of the securities at an offering price that is negotiated with the issuer. Also known as *firm commitment offering*.

Unearned revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service. Also called *deferred revenue* or *deferred income*.

Unemployed People who are actively seeking employment but are currently without a job.

Unemployment rate The ratio of unemployed to the labor force.

Unexpected inflation The component of inflation that is a surprise.

Unimodal A distribution with a single value that is most frequently occurring.

Unit elastic An elasticity with a magnitude of negative one. Also called *unitary elastic*.

Unit labor cost The average labor cost to produce one unit of output.

Unit normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Unitranche debt Consists of a hybrid or blended loan structure that combines different tranches of secured and unsecured debt into a single loan with a single, blended interest rate.

Units-of-production method A depreciation method that allocates the cost of a long-lived asset based on actual usage during the period.

Univariate distribution A distribution that specifies the probabilities for a single random variable.

Universal owners Long-term investors, such as pension funds, that have significant assets invested in globally diversified portfolios.

Unlimited funds An unlimited funds environment assumes that the company can raise the funds it wants for all profitable investments simply by paying the required rate of return.

Unsecured debt Debt that gives the debtholder only a general claim on an issuer's assets and cash flow.

Unsponsored A type of depository receipt in which the foreign company whose shares are held by the depository has no involvement in the issuance of the receipts.

Unstructured data Data that do not follow any conventionally organized forms.

Unsupervised learning A machine learning approach that does not make use of labeled training data.

Up transition probability The probability that an asset's value moves up.

Uptrend A pattern that occurs when the price of an asset moves higher over a period of time.

Validity instructions Instructions which indicate when the order may be filled.

Valuation allowance A reserve created against deferred tax assets, based on the likelihood of realizing the deferred tax assets in future accounting periods.

Valuation ratios Ratios that measure the quantity of an asset or flow (e.g., earnings) in relation to the price associated with a specified claim (e.g., a share or ownership of the enterprise).

Value at risk A money measure of the minimum value of losses expected during a specified time period at a given level of probability.

Value investors With reference to equity investors, investors who are focused on paying a relatively low share price in relation to earnings or assets per share.

VaR See *value at risk*.

Variable A characteristic or quantity that can be measured, counted, or categorized and that is subject to change (also called a field, an attribute, or a feature).

Variable costs Costs that fluctuate with the level of production and sales.

Variance The expected value (the probability-weighted average) of squared deviations from a random variable's expected value.

Variation margin Additional margin that must be deposited in an amount sufficient to bring the balance up to the initial margin requirement.

Veblen goods Goods that increase in desirability with increasing price.

Vega A measure of the sensitivity of an option's price to changes in the underlying's volatility.

Venture capital Investments that provide "seed" or startup capital, early-stage financing, or later-stage financing (including mezzanine-stage financing) to companies that are in early development stages and require additional capital for expansion or preparation for an initial public offering.

Venture capital fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *private equity fund*.

Venture capital funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *private equity funds.*

Vertical analysis Common-size analysis using only one reporting period or one base financial statement; for example, an income statement in which all items are stated as percentages of sales.

Vertical demand schedule Implies that some fixed quantity is demanded, regardless of price.

Visual technique The most common and readily available method of initial data assessment. Experts in pattern recognition maintain that the visual (or "eyeball") technique is still the most effective way of searching for recognizable patterns.

Visualization The presentation of data in a pictorial or graphical format for the purpose of increasing understanding and for gaining insights into the data.

Volatility The standard deviation of the continuously compounded returns on the underlying asset.

Voluntarily unemployed A person voluntarily outside the labor force, such as a jobless worker refusing an available vacancy.

Voluntary export restraint A trade barrier under which the exporting country agrees to limit its exports of the good to its trading partners to a specific number of units.

Vote by proxy A mechanism that allows a designated party—such as another shareholder, a shareholder representative, or management—to vote on the shareholder's behalf.

Warrant Attached option that gives its holder the right to buy the underlying stock of the issuing company at a fixed exercise price until the expiration date.

Waterfall Represents the distribution method that defines the order in which allocations are made to LPs and GPs. There are two major types of waterfall: *deal by deal* (or *American*) and *whole of fund* (or *European*).

Weak-form efficient market hypothesis The belief that security prices fully reflect all past market data, which refers to all historical price and volume trading information.

Wealth effect An increase (decrease) in household wealth increases (decreases) consumer spending out of a given level of current income.

Web-based lender A lender that operates primarily on the internet, offering loans in relatively small amounts, typically to small businesses in need of cash.

Weighted average cost method An inventory accounting method that averages the total cost of available inventory items over the total units available for sale.

Weighted average cost of capital A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average cost of capital (WACC) A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average coupon rate Weighting the mortgage rate of each mortgage loan in the pool by the percentage of the mortgage outstanding relative to the outstanding amount of all the mortgages in the pool.

Weighted average life A measure that gives investors an indication of how long they can expect to hold the MBS before it is paid off; the convention-based average time to receipt of all principal repayments. Also called *average life*.

Weighted average maturity Weighting the remaining number of months to maturity for each mortgage loan in the pool by the amount of the outstanding mortgage balance.

Weighted mean An average in which each observation is weighted by an index of its relative importance.

Winsorized mean A mean computed after assigning a stated percentage of the lowest values equal to one specified low value and a stated percentage of the highest values equal to one specified high value.

Word cloud A visual device for representing textual data, which consists of words extracted from a source of textual data. The size of each distinct word is proportional to the frequency with which it appears in the given text (also known as **tag cloud**).

Working capital The difference between current assets and current liabilities.

Working capital management The management of a company's short-term assets (such as inventory) and short-term liabilities (such as money owed to suppliers).

World price The price prevailing in the world market.

Yield duration The sensitivity of the bond price with respect to the bond's own yield-to-maturity.

Yield to maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield to redemption* or *redemption yield*.

Yield-to-maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield-to-redemption* or *redemption yield*.

Yield-to-redemption See *yield-to-maturity*.

Yield-to-worst The lowest of the sequence of yields-to-call and the yield-to-maturity.

Zero-coupon bond A bond that does not pay interest during its life. It is issued at a discount to par value and redeemed at par. Also called *pure discount bond*.

Zero-volatility spread (Z-spread) Calculates a constant yield spread over a government (or interest rate swap) spot curve.

CFA Institute®
CFA Program

CORPORATE ISSUERS, EQUITY, AND FIXED INCOME

CFA® Program Curriculum
2022 • LEVEL I • VOLUME 4

WILEY

CONTENTS

How to Use the CFA Program Curriculum **xi**
Background on the CBOK **xi**
Organization of the Curriculum **xii**
Features of the Curriculum **xii**
Designing Your Personal Study Program **xiii**
CFA Institute Learning Ecosystem (LES) **xiv**
Prep Providers **xv**
Feedback **xvi**

Corporate Issuers

Study Session 10 **Corporate Issuers (2)** **3**

Reading 30 **Cost of Capital-Foundational Topics** **5**
Introduction **5**
Cost of Capital **6**
  Taxes and the Cost of Capital **8**
Costs of the Various Sources of Capital **9**
  Cost of Debt **9**
  Cost of Preferred Stock **12**
  Cost of Common Equity **13**
Estimating Beta **18**
  Estimating Beta for Public Companies **18**
  Estimating Beta for Thinly Traded and Nonpublic Companies **19**
Flotation Costs **22**
Methods in Use **25**
Summary **25**
Practice Problems **28**
Solutions **34**

Reading 31 **Capital Structure** **39**
Introduction **39**
Capital Structure and Company Life Cycle **40**
  Background **40**
  Start-Ups **42**
  Growth Businesses **42**
  Mature Businesses **43**
  Unique Situations **45**
Modigliani–Miller Propositions **47**
  MM Proposition I without Taxes: Capital Structure Irrelevance **48**
  MM Proposition II without Taxes: Higher Financial Leverage Raises the Cost of Equity **48**
  MM Propositions with Taxes: Taxes, Cost of Capital, and Value of the Company **51**
  Costs of Financial Distress **53**

Optimal and Target Capital Structure 54
Factors Affecting Capital Structure Decisions 57
    Capital Structure Policies and Target Capital Structures 58
    Financing Capital Investments 60
    Market Conditions 61
    Information Asymmetries and Signaling 62
Agency Costs 63
Stakeholder Interests 65
    Shareholder vs. Stakeholder Theory 65
    Debt vs. Equity Conflict 66
    Preferred Shareholders 70
    Private Equity Investors/Controlling Shareholders 71
    Bank and Private Lenders 72
    Other Stakeholders 72
Summary 77
Practice Problems 79
Solutions 83

Reading 32 **Measures of Leverage** 85
Introduction 85
Leverage 86
Business and Sales Risks 88
    Business Risk and Its Components 88
    Sales Risk 89
Operating Risk and the Degree of Operating Leverage 90
Financial Risk, the Degree of Financial Leverage and the Leveraging Role of Debt 97
Total Leverage and the Degree of Total Leverage 101
Breakeven Points and Operating Breakeven Points 104
The Risks of Creditors and Owners 107
Summary 108
Practice Problems 110
Solutions 114

Equity Investments

Study Session 11 **Equity Investments (1)** 119

Reading 33 **Market Organization and Structure** 121
Introduction 122
The Functions of the Financial System 122
    Helping People Achieve Their Purposes in Using the Financial System 123
    Determining Rates of Return 128
    Capital Allocation Efficiency 129
Assets and Contracts 130
    Classifications of Assets and Markets 130
Securities 133
    Fixed Income 133
    Equities 134

Pooled Investments 135
Currencies, Commodities, and Real Assets 136
Commodities 136
Real Assets 137
Contracts 139
Forward Contracts 140
Futures Contracts 141
Swap Contracts 143
Option Contracts 143
Other Contracts 144
Financial Intermediaries 145
Brokers, Exchanges, and Alternative Trading Systems 145
Dealers 146
Arbitrageurs 147
Securitizers, Depository Institutions and Insurance Companies 149
Depository Institutions and Other Financial Corporations 150
Insurance Companies 152
Settlement and Custodial Services and Summary 153
Summary 155
Positions and Short Positions 155
Short Positions 157
Leveraged Positions 158
Orders and Execution Instructions 161
Execution Instructions 162
Validity Instructions and Clearing Instructions 165
Stop Orders 165
Clearing Instructions 167
Primary Security Markets 167
Public Offerings 167
Private Placements and Other Primary Market Transactions 169
Importance of Secondary Markets to Primary Markets 170
Secondary Security Market and Contract Market Structures 171
Trading Sessions 171
Execution Mechanisms 172
Market Information Systems 175
Well-functioning Financial Systems 175
Market Regulation 178
Summary 181
Practice Problems 184
Solutions 191

Reading 34 **Security Market Indexes** 195
Introduction 195
Index Definition and Calculations of Value and Returns 196
Calculation of Single-Period Returns 197
Calculation of Index Values over Multiple Time Periods 199
Index Construction 200
Target Market and Security Selection 200
Index Weighting 201

Index Management: Rebalancing and Reconstitution 209
    Rebalancing 209
    Reconstitution 210
Uses of Market Indexes 211
    Gauges of Market Sentiment 212
    Proxies for Measuring and Modeling Returns, Systematic Risk, and Risk-Adjusted Performance 212
    Proxies for Asset Classes in Asset Allocation Models 212
    Benchmarks for Actively Managed Portfolios 212
    Model Portfolios for Investment Products 213
Equity indexes 213
    Broad Market Indexes 213
    Multi-Market Indexes 213
    Sector Indexes 215
    Style Indexes 215
Fixed-income indexes 216
    Construction 216
    Types of Fixed-Income Indexes 217
Indexes for Alternative Investments 219
    Commodity Indexes 219
    Real Estate Investment Trust Indexes 220
    Hedge Fund Indexes 221
Summary 223
Practice Problems 225
Solutions 231

Reading 35 **Market Efficiency** 235
Introduction 235
The Concept of Market Efficiency 237
    The Description of Efficient Markets 237
    Market Value versus Intrinsic Value 239
Factors Affecting Market Efficiency Including Trading Costs 240
    Market Participants 241
    Information Availability and Financial Disclosure 242
    Limits to Trading 243
    Transaction Costs and Information-Acquisition Costs 243
Forms of Market Efficiency 244
    Weak Form 245
    Semi-Strong Form 246
    Strong Form 248
Implications of the Efficient Market Hypothesis 249
    Fundamental Analysis 249
    Technical Analysis 250
    Portfolio Management 250
Market Pricing Anomalies - Time Series and Cross-Sectional 250
    Time-Series Anomalies 251
    Cross-Sectional Anomalies 254
Other Anomalies, Implications of Market Pricing Anomalies 254
    Closed-End Investment Fund Discounts 255

Earnings Surprise 255
Initial Public Offerings (IPOs) 256
Predictability of Returns Based on Prior Information 257
Implications for Investment Strategies 257
Behavioral Finance 257
Loss Aversion 258
Herding 258
Overconfidence 258
Information Cascades 259
Other Behavioral Biases 260
Behavioral Finance and Investors 260
Behavioral Finance and Efficient Markets 260
Summary 260
Practice Problems 263
Solutions 266

Study Session 12 **Equity Investments (2)** 269

Reading 36 **Overview of Equity Securities** 271
Importance of Equity Securities 271
Equity Securities in Global Financial Markets 272
Characteristics of Equity Securities 277
Common Shares 278
Preference Shares 280
Private Versus Public Equity Securities 282
Non-Domestic Equity Securities 285
Direct Investing 286
Depository Receipts 287
Risk and Return Characteristics 290
Return Characteristics of Equity Securities 290
Risk of Equity Securities 291
Equity and Company Value 292
Accounting Return on Equity 293
The Cost of Equity and Investors' Required Rates of Return 298
Summary 299
Practice Problems 301
Solutions 305

Reading 37 **Introduction to Industry and Company Analysis** 307
Introduction 308
Uses of Industry Analysis 308
Approaches to Identifying Similar Companies 309
Products and/or Services Supplied 310
Business-Cycle Sensitivities 310
Statistical Similarities 312
Industry Classification Systems 313
Commercial Industry Classification Systems 313
Constructing a Peer Group 318

Describing and Analyzing an Industry and Principles of Strategic Analysis 321
  Principles of Strategic Analysis 323
  Barriers to Entry 324
  Industry Concentration 326
  Industry Capacity 328
  Market Share Stability 330
  Price Competition 331
  Industry Life Cycle 332
External Influences on Industry 336
  Macroeconomic Influences 337
  Technological Influences 337
  Demographic Influences 338
  Governmental Influences 339
  Social Influences 340
  Environmental Influences 340
  Industry Comparison 343
Company Analysis 346
  Elements That Should Be Covered in a Company Analysis 347
  Spreadsheet Modeling 349
Summary 350
Practice Problems 354
Solutions 358

Reading 38 **Equity Valuation: Concepts and Basic Tools** 361
Introduction 362
Estimated Value and Market Price 363
Categories of Equity Valuation Models 364
Background for the Dividend Discount Model 366
  Dividends: Background for the Dividend Discount Model 366
Dividend Discount Model (DDM) and Free-Cash-Flow-to-Equity Model (FCFE) 369
Preferred Stock Valuation 373
The Gordon Growth Model 375
Multistage Dividend Discount Models 380
Multipler Models and Relationship Among Price Multiples, Present Value Models, and Fundamentals 385
  Relationships among Price Multiples, Present Value Models, and Fundamentals 386
Method of Comparables and Valuation Based on Price Multiples 389
  Illustration of a Valuation Based on Price Multiples 392
Enterprise Value 394
Asset-Based Valuation 397
Summary 401
Practice Problems 403
Solutions 409

Fixed Income

Study Session 13 **Fixed Income (1)** **415**

Reading 39 **Fixed-Income Securities: Defining Elements** **417**
Introduction and Overview of a Fixed-Income Security **417**
    Overview of a Fixed-Income Security **418**
Bond Indenture **424**
    Bond Indenture **424**
Legal, Regulatory, and Tax Considerations **432**
    Tax Considerations **435**
Principal Repayment Structures **437**
    Principal Repayment Structures **437**
Coupon Payment Structures **441**
    Floating-Rate Notes **442**
    Step-Up Coupon Bonds **442**
    Credit-Linked Coupon Bonds **443**
    Payment-in-Kind Coupon Bonds **443**
    Deferred Coupon Bonds **444**
    Index-Linked Bonds **444**
Callable and Putable Bonds **448**
    Callable Bonds **448**
    Putable Bonds **450**
Convertible Bonds **451**
Summary **454**
Practice Problems **457**
Solutions **461**

Reading 40 **Fixed-Income Markets: Issuance, Trading, and Funding** **465**
Introduction **465**
Classification of Fixed-Income Markets **466**
    Classification of Fixed-Income Markets **466**
    Fixed-Income Indexes **473**
    Investors in Fixed-Income Securities **474**
Primary Bond Markets **475**
    Primary Bond Markets **475**
Secondary Bond Markets **480**
Sovereign Bonds **483**
    Characteristics of Sovereign Bonds **483**
    Credit Quality of Sovereign Bonds **484**
    Types of Sovereign Bonds **485**
Non-Sovereign, Quasi-Government, and Supranational Bonds **487**
    Non-Sovereign Bonds **487**
    Quasi-Government Bonds **488**
    Supranational Bonds **488**
Corporate Debt: Bank Loans, Syndicated Loans, and Commercial Paper **489**
    Bank Loans and Syndicated Loans **489**
    Commercial Paper **490**

O indicates an optional segment

Corporate Debt: Notes and Bonds 493
  Maturities 493
  Coupon Payment Structures 493
  Principal Repayment Structures 494
  Asset or Collateral Backing 494
  Contingency Provisions 495
  Issuance, Trading, and Settlement 495
Structured Financial Instruments 497
  Capital Protected Instruments 497
  Yield Enhancement Instruments 498
  Participation Instruments 498
  Leveraged Instruments 499
Short-Term Bank Funding Alternatives 500
  Retail Deposits 500
  Short-Term Wholesale Funds 501
Repurchase and Reverse Repurchase Agreements 502
  Structure of Repurchase and Reverse Repurchase Agreements 503
  Credit Risk Associated with Repurchase Agreements 504
Summary 506
Practice Problems 509
Solutions 513

Reading 41 **Introduction to Fixed-Income Valuation** 517
Introduction 517
Bond Prices and the Time Value of Money 518
  Bond Pricing with a Market Discount Rate 518
  Yield-to-Maturity 522
  Relationships between the Bond Price and Bond Characteristics 523
  Pricing Bonds Using Spot Rates 527
Prices and Yields: Conventions For Quotes and Calculations 529
  Flat Price, Accrued Interest, and the Full Price 529
  Matrix Pricing 533
  Annual Yields for Varying Compounding Periods in the Year 536
  Yield Measures for Fixed-Rate Bonds 539
  Yield Measures for Floating-Rate Notes 542
  Yield Measures for Money Market Instruments 546
The Maturity Structure of Interest Rates 550
Yield Spreads 558
  Yield Spreads over Benchmark Rates 558
  Yield Spreads over the Benchmark Yield Curve 560
Summary 563
Practice Problems 566
Solutions 575

Reading 42 **Introduction to Asset-Backed Securities** 589
Introduction: Benefits of Securitization 589
  Benefits of Securitization for Economies and Financial Markets 590
How Securitization Works 591
  An Example of a Securitization 592

- Parties to a Securitization and Their Roles 593
Structure of a Securitization 595
- Key Role of the Special Purpose Entity 597
Residential Mortgage Loans 600
- Maturity 601
- Interest Rate Determination 601
- Amortization Schedule 602
- Prepayment Options and Prepayment Penalties 602
- Rights of the Lender in a Foreclosure 603
Mortgage Pass-Through Securities 604
- Mortgage Pass-Through Securities 605
Collateralized Mortgage Obligations and Non-Agency RMBS 612
- Sequential-Pay CMO Structures 612
- CMO Structures Including Planned Amortization Class and Support Tranches 614
- Other CMO Structures 617
- Non-Agency Residential Mortgage-Backed Securities 618
Commercial Mortgage-Backed Securities 619
- Credit Risk 619
- CMBS Structure 620
Non-Mortgage Asset-Backed Securities 623
- Auto Loan ABS 624
- Credit Card Receivable ABS 626
Collateralized Debt Obligations 628
- CDO Structure 628
- An Example of a CDO Transaction 629
Covered Bonds 631
Summary 632
Practice Problems 636
Solutions 642

Glossary **G-1**

How to Use the CFA Program Curriculum

Congratulations on your decision to enter the Chartered Financial Analyst (CFA®) Program. This exciting and rewarding program of study reflects your desire to become a serious investment professional. You are embarking on a program noted for its high ethical standards and the breadth of knowledge, skills, and abilities (competencies) it develops. Your commitment should be educationally and professionally rewarding.

The credential you seek is respected around the world as a mark of accomplishment and dedication. Each level of the program represents a distinct achievement in professional development. Successful completion of the program is rewarded with membership in a prestigious global community of investment professionals. CFA charterholders are dedicated to life-long learning and maintaining currency with the ever-changing dynamics of a challenging profession. CFA Program enrollment represents the first step toward a career-long commitment to professional education.

The CFA exam measures your mastery of the core knowledge, skills, and abilities required to succeed as an investment professional. These core competencies are the basis for the Candidate Body of Knowledge (CBOK™). The CBOK consists of four components:

- A broad outline that lists the major CFA Program topic areas (www.cfainstitute.org/programs/cfa/curriculum/cbok);
- Topic area weights that indicate the relative exam weightings of the top-level topic areas (www.cfainstitute.org/programs/cfa/curriculum);
- Learning outcome statements (LOS) that advise candidates about the specific knowledge, skills, and abilities they should acquire from readings covering a topic area (LOS are provided in candidate study sessions and at the beginning of each reading); and
- CFA Program curriculum that candidates receive upon exam registration.

Therefore, the key to your success on the CFA exams is studying and understanding the CBOK. The following sections provide background on the CBOK, the organization of the curriculum, features of the curriculum, and tips for designing an effective personal study program.

BACKGROUND ON THE CBOK

CFA Program is grounded in the practice of the investment profession. CFA Institute performs a continuous practice analysis with investment professionals around the world to determine the competencies that are relevant to the profession, beginning with the Global Body of Investment Knowledge (GBIK®). Regional expert panels and targeted surveys are conducted annually to verify and reinforce the continuous feedback about the GBIK. The practice analysis process ultimately defines the CBOK. The CBOK reflects the competencies that are generally accepted and applied by investment professionals. These competencies are used in practice in a generalist context and are expected to be demonstrated by a recently qualified CFA charterholder.

© 2021 CFA Institute. All rights reserved.

The CFA Institute staff—in conjunction with the Education Advisory Committee and Curriculum Level Advisors, who consist of practicing CFA charterholders—designs the CFA Program curriculum in order to deliver the CBOK to candidates. The exams, also written by CFA charterholders, are designed to allow you to demonstrate your mastery of the CBOK as set forth in the CFA Program curriculum. As you structure your personal study program, you should emphasize mastery of the CBOK and the practical application of that knowledge. For more information on the practice analysis, CBOK, and development of the CFA Program curriculum, please visit www.cfainstitute.org.

ORGANIZATION OF THE CURRICULUM

The Level I CFA Program curriculum is organized into 10 topic areas. Each topic area begins with a brief statement of the material and the depth of knowledge expected. It is then divided into one or more study sessions. These study sessions should form the basic structure of your reading and preparation. Each study session includes a statement of its structure and objective and is further divided into assigned readings. An outline illustrating the organization of these study sessions can be found at the front of each volume of the curriculum.

The readings are commissioned by CFA Institute and written by content experts, including investment professionals and university professors. Each reading includes LOS and the core material to be studied, often a combination of text, exhibits, and in-text examples and questions. End of Reading Questions (EORQs) followed by solutions help you understand and master the material. The LOS indicate what you should be able to accomplish after studying the material. The LOS, the core material, and the EORQs are dependent on each other, with the core material and EORQs providing context for understanding the scope of the LOS and enabling you to apply a principle or concept in a variety of scenarios.

The entire readings, including the EORQs, are the basis for all exam questions and are selected or developed specifically to teach the knowledge, skills, and abilities reflected in the CBOK.

You should use the LOS to guide and focus your study because each exam question is based on one or more LOS and the core material and practice problems associated with the LOS. As a candidate, you are responsible for the entirety of the required material in a study session.

We encourage you to review the information about the LOS on our website (www.cfainstitute.org/programs/cfa/curriculum/study-sessions), including the descriptions of LOS “command words” on the candidate resources page at www.cfainstitute.org.

FEATURES OF THE CURRICULUM

End of Reading Questions/Solutions *All End of Reading Questions (EORQs) as well as their solutions are part of the curriculum and are required material for the exam.* In addition to the in-text examples and questions, these EORQs help demonstrate practical applications and reinforce your understanding of the concepts presented. Some of these EORQs are adapted from past CFA exams and/or may serve as a basis for exam questions.

Glossary For your convenience, each volume includes a comprehensive Glossary. Throughout the curriculum, a **bolded** word in a reading denotes a term defined in the Glossary.

Note that the digital curriculum that is included in your exam registration fee is searchable for key words, including Glossary terms.

LOS Self-Check We have inserted checkboxes next to each LOS that you can use to track your progress in mastering the concepts in each reading.

Source Material The CFA Institute curriculum cites textbooks, journal articles, and other publications that provide additional context or information about topics covered in the readings. As a candidate, you are not responsible for familiarity with the original source materials cited in the curriculum.

Note that some readings may contain a web address or URL. The referenced sites were live at the time the reading was written or updated but may have been deactivated since then.

Some readings in the curriculum cite articles published in the *Financial Analysts Journal*®, which is the flagship publication of CFA Institute. Since its launch in 1945, the *Financial Analysts Journal* has established itself as the leading practitioner-oriented journal in the investment management community. Over the years, it has advanced the knowledge and understanding of the practice of investment management through the publication of peer-reviewed practitioner-relevant research from leading academics and practitioners. It has also featured thought-provoking opinion pieces that advance the common level of discourse within the investment management profession. Some of the most influential research in the area of investment management has appeared in the pages of the *Financial Analysts Journal*, and several Nobel laureates have contributed articles.

Candidates are not responsible for familiarity with *Financial Analysts Journal* articles that are cited in the curriculum. But, as your time and studies allow, we strongly encourage you to begin supplementing your understanding of key investment management issues by reading this, and other, CFA Institute practice-oriented publications through the Research & Analysis webpage (www.cfainstitute.org/en/research).

Errata The curriculum development process is rigorous and includes multiple rounds of reviews by content experts. Despite our efforts to produce a curriculum that is free of errors, there are times when we must make corrections. Curriculum errata are periodically updated and posted by exam level and test date online (www.cfainstitute.org/en/programs/submit-errata). If you believe you have found an error in the curriculum, you can submit your concerns through our curriculum errata reporting process found at the bottom of the Curriculum Errata webpage.

DESIGNING YOUR PERSONAL STUDY PROGRAM

Create a Schedule An orderly, systematic approach to exam preparation is critical. You should dedicate a consistent block of time every week to reading and studying. Complete all assigned readings and the associated problems and solutions in each study session. Review the LOS both before and after you study each reading to ensure that

you have mastered the applicable content and can demonstrate the knowledge, skills, and abilities described by the LOS and the assigned reading. Use the LOS self-check to track your progress and highlight areas of weakness for later review.

Successful candidates report an average of more than 300 hours preparing for each exam. Your preparation time will vary based on your prior education and experience, and you will probably spend more time on some study sessions than on others.

You should allow ample time for both in-depth study of all topic areas and additional concentration on those topic areas for which you feel the least prepared.

CFA INSTITUTE LEARNING ECOSYSTEM (LES)

As you prepare for your exam, we will email you important exam updates, testing policies, and study tips. Be sure to read these carefully.

Your exam registration fee includes access to the CFA Program Learning Ecosystem (LES). This digital learning platform provides access, even offline, to all of the readings and End of Reading Questions found in the print curriculum organized as a series of shorter online lessons with associated EORQs. This tool is your one-stop location for all study materials, including practice questions and mock exams.

The LES provides the following supplemental study tools:

Structured and Adaptive Study Plans The LES offers two ways to plan your study through the curriculum. The first is a structured plan that allows you to move through the material in the way that you feel best suits your learning. The second is an adaptive study plan based on the results of an assessment test that uses actual practice questions.

Regardless of your chosen study path, the LES tracks your level of proficiency in each topic area and presents you with a dashboard of where you stand in terms of proficiency so that you can allocate your study time efficiently.

Flashcards and Game Center The LES offers all the Glossary terms as Flashcards and tracks correct and incorrect answers. Flashcards can be filtered both by curriculum topic area and by action taken—for example, answered correctly, unanswered, and so on. These Flashcards provide a flexible way to study Glossary item definitions.

The Game Center provides several engaging ways to interact with the Flashcards in a game context. Each game tests your knowledge of the Glossary terms a in different way. Your results are scored and presented, along with a summary of candidates with high scores on the game, on your Dashboard.

Discussion Board The Discussion Board within the LES provides a way for you to interact with other candidates as you pursue your study plan. Discussions can happen at the level of individual lessons to raise questions about material in those lessons that you or other candidates can clarify or comment on. Discussions can also be posted at the level of topics or in the initial Welcome section to connect with other candidates in your area.

Practice Question Bank The LES offers access to a question bank of hundreds of practice questions that are in addition to the End of Reading Questions. These practice questions, only available on the LES, are intended to help you assess your mastery of individual topic areas as you progress through your studies. After each practice question, you will receive immediate feedback noting the correct response and indicating the relevant assigned reading so you can identify areas of weakness for further study.

Mock Exams The LES also includes access to three-hour Mock Exams that simulate the morning and afternoon sessions of the actual CFA exam. These Mock Exams are intended to be taken after you complete your study of the full curriculum and take practice questions so you can test your understanding of the curriculum and your readiness for the exam. If you take these Mock Exams within the LES, you will receive feedback afterward that notes the correct responses and indicates the relevant assigned readings so you can assess areas of weakness for further study. We recommend that you take Mock Exams during the final stages of your preparation for the actual CFA exam. For more information on the Mock Exams, please visit www.cfainstitute.org.

PREP PROVIDERS

You may choose to seek study support outside CFA Institute in the form of exam prep providers. After your CFA Program enrollment, you may receive numerous solicitations for exam prep courses and review materials. When considering a prep course, make sure the provider is committed to following the CFA Institute guidelines and high standards in its offerings.

Remember, however, that there are no shortcuts to success on the CFA exams; reading and studying the CFA Program curriculum *is* the key to success on the exam. The CFA Program exams reference only the CFA Institute assigned curriculum; no prep course or review course materials are consulted or referenced.

SUMMARY

Every question on the CFA exam is based on the content contained in the required readings and on one or more LOS. Frequently, an exam question is based on a specific example highlighted within a reading or on a specific practice problem and its solution. To make effective use of the CFA Program curriculum, please remember these key points:

1. All pages of the curriculum are required reading for the exam.
2. All questions, problems, and their solutions are part of the curriculum and are required study material for the exam. These questions are found at the end of the readings in the print versions of the curriculum. In the LES, these questions appear directly after the lesson with which they are associated. The LES provides immediate feedback on your answers and tracks your performance on these questions throughout your study.
3. We strongly encourage you to use the CFA Program Learning Ecosystem. In addition to providing access to all the curriculum material, including EORQs, in the form of shorter, focused lessons, the LES offers structured and adaptive study planning, a Discussion Board to communicate with other candidates, Flashcards, a Game Center for study activities, a test bank of practice questions, and online Mock Exams. Other supplemental study tools, such as eBook and PDF versions of the print curriculum, and additional candidate resources are available at www.cfainstitute.org.
4. Using the study planner, create a schedule and commit sufficient study time to cover the study sessions. You should also plan to review the materials, answer practice questions, and take Mock Exams.
5. Some of the concepts in the study sessions may be superseded by updated rulings and/or pronouncements issued after a reading was published. Candidates are expected to be familiar with the overall analytical framework contained in the assigned readings. Candidates are not responsible for changes that occur after the material was written.

FEEDBACK

At CFA Institute, we are committed to delivering a comprehensive and rigorous curriculum for the development of competent, ethically grounded investment professionals. We rely on candidate and investment professional comments and feedback as we work to improve the curriculum, supplemental study tools, and candidate resources.

Please send any comments or feedback to info@cfainstitute.org. You can be assured that we will review your suggestions carefully. Ongoing improvements in the curriculum will help you prepare for success on the upcoming exams and for a lifetime of learning as a serious investment professional.

Corporate Issuers

STUDY SESSION

Study Session 9	Corporate Issuers (1)
Study Session 10	Corporate Issuers (2)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to evaluate a company's corporate governance; to demonstrate methods used to make capital investment; to evaluate the management of working capital; estimate a company's cost of capital; and to evaluate a company's operating and financial leverage.

Some academic studies have shown that well governed companies may perform better in financial terms. Increasingly, investment approaches that consider environmental, social, and governance factors, known as ESG, are being adopted. In addition to good governance practices, management decisions regarding investment and financing also play a central role in corporate profitability and performance. To remain in business as a going concern and to increase shareholder value over time, a company's management must consistently identify and invest in profitable long-term capital projects relative to cost of capital (financing) and make optimal use of leverage and working capital in day to day operations.

© 2021 CFA Institute. All rights reserved.

CORPORATE ISSUERS

STUDY SESSION

10

Corporate Issuers (2)

This study session begins with practical techniques to estimate a company's, or project's, cost of capital, a key input used in both corporate decision-making and investor analysis. Methods to estimate the costs of the various sources of capital are covered. Next, capital structure considerations and the Modigliani-Miller propositions are discussed, including factors affecting the use of leverage by companies. Examples of potential stakeholder conflicts that arise with financing decisions are also examined. The session concludes with coverage of the various types of leverage (operating, financial, total), measures of leverage, and the impact that leverage may have on a company's earnings and financial ratios.

READING ASSIGNMENTS

Reading 30	Cost of Capital-Foundational Topics by Yves Courtois, CMT, MRICS, CFA, Gene C. Lai, PhD, and Pamela Peterson Drake, PhD, CFA
Reading 31	Capital Structure Raj Aggarwal, PhD, CFA, Glen D. Campbell, MBA, Pamela Peterson Drake, PhD, CFA, Adam Kobor, PhD, CFA, and Gregory Noronha, PhD, CFA
Reading 32	Measures of Leverage by Pamela Peterson Drake, PhD, CFA, Raj Aggarwal, PhD, CFA, Cynthia Harrington, CFA, and Adam Kobor, PhD, CFA

© 2021 CFA Institute. All rights reserved.

READING

30

Cost of Capital-Foundational Topics

by Yves Courtois, CMT, MRICS, CFA, Gene C. Lai, PhD, and
Pamela Peterson Drake, PhD, CFA

Yves Courtois, CMT, MRICS, CFA, is at KPMG (Luxembourg). Gene C. Lai, PhD, is at the University of North Carolina at Charlotte (USA). Pamela Peterson Drake, PhD, CFA, is at James Madison University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** calculate and interpret the weighted average cost of capital (WACC) of a company;
☐	**b.** describe how taxes affect the cost of capital from different capital sources;
☐	**c.** calculate and interpret the cost of debt capital using the yield-to-maturity approach and the debt-rating approach;
☐	**d.** calculate and interpret the cost of noncallable, nonconvertible preferred stock;
☐	**e.** calculate and interpret the cost of equity capital using the capital asset pricing model approach and the bond yield plus risk premium approach;
☐	**f.** explain and demonstrate beta estimation for public companies, thinly traded public companies, and nonpublic companies;
☐	**g.** explain and demonstrate the correct treatment of flotation costs.

INTRODUCTION

1

A company grows by making investments that are expected to increase revenues and profits. It acquires the capital or funds necessary to make such investments by borrowing (i.e., using debt financing) or by using funds from the owners (i.e., equity financing). By applying this capital to investments with long-term benefits, the company is producing value today. How much value? The answer depends not only on the investments' expected future cash flows but also on the cost of the funds. Borrowing is not costless, nor is using owners' funds.

© 2021 CFA Institute. All rights reserved.

The cost of this capital is an important ingredient in both investment decision making by the company's management and the valuation of the company by investors. If a company invests in projects that produce a return in excess of the cost of capital, the company has created value; in contrast, if the company invests in projects whose returns are less than the cost of capital, the company has destroyed value. Therefore, the estimation of the cost of capital is a central issue in corporate financial management and for an analyst seeking to evaluate a company's investment program and its competitive position.

Cost of capital estimation is a challenging task. As we have already implied, the cost of capital is not observable but, rather, must be estimated. Arriving at a cost of capital estimate requires a multitude of assumptions and estimates. Another challenge is that the cost of capital that is appropriately applied to a specific investment depends on the characteristics of that investment: The riskier the investment's cash flows, the greater its cost of capital. In reality, a company must estimate project-specific costs of capital. What is often done, however, is to estimate the cost of capital for the company as a whole and then adjust this overall corporate cost of capital upward or downward to reflect the risk of the contemplated project relative to the company's average project.

This reading is organized as follows: In Section 2, we introduce the cost of capital and its basic computation. Section 3 presents a selection of methods for estimating the costs of the various sources of capital: debt, preferred stock, and common equity. For the latter, two approaches for estimating the equity risk premium are mentioned. Section 4 discusses beta estimation, a key input in using the CAPM to calculate the cost of equity, and Section 5 examines the correct treatment of flotation, or capital issuance, costs. Section 6 highlights methods used by corporations, and a summary concludes the reading.

2

COST OF CAPITAL

a calculate and interpret the weighted average cost of capital (WACC) of a company

The **cost of capital** is the rate of return that the suppliers of capital—lenders and owners—require as compensation for their contribution of capital. Another way of looking at the cost of capital is that it is the opportunity cost of funds for the suppliers of capital: A potential supplier of capital will not voluntarily invest in a company unless its return meets or exceeds what the supplier could earn elsewhere in an investment of comparable risk. In other words, to raise new capital, the issuer must price the security to offer a level of expected return that is competitive with the expected returns being offered by similarly risky securities.

A company typically has several alternatives for raising capital, including issuing equity, debt, and hybrid instruments that share characteristics of both debt and equity, such as preferred stock and convertible debt. Each source selected becomes a component of the company's funding and has a cost (required rate of return) that may be called a **component cost of capital**. Because we are using the cost of capital in the evaluation of investment opportunities, we are dealing with a *marginal* cost—what it would cost to raise additional funds for the potential investment project. Therefore, the cost of capital that the investment analyst is concerned with is a marginal cost, and the required return on a security is the issuer's marginal cost for raising additional capital of the same type.

The cost of capital of a company is the required rate of return that investors demand for the average-risk investment of a company. A company with higher-than-average-risk investments must pay investors a higher rate of return, competitive with other securities of similar risk, which corresponds to a higher cost of capital. Similarly, a company with lower-than-average-risk investments will have lower rates of return demanded by investors, resulting in a lower associated cost of capital. The most common way to estimate this required rate of return is to calculate the marginal cost of each of the various sources of capital and then calculate a weighted average of these costs. You will notice that the debt and equity costs of capital and the tax rate are all understood to be "marginal" rates: the cost or tax rate for additional capital.

The weighted average is referred to as the **weighted average cost of capital** (WACC). The WACC is also referred to as the marginal cost of capital (MCC) because it is the cost that a company incurs for additional capital. Further, this is the current cost: what it would cost the company today.

The weights are the proportions of the various sources of capital that the company uses to support its investment program. It is important to note that the weights should represent the company's **target capital structure**, not the current capital structure. A company's target capital structure is its chosen (or targeted) proportions of debt and equity, whereas its current capital structure is the company's actual weighting of debt and equity. For example, suppose the current capital structure is one-third debt, one-third preferred stock, and one-third common stock. Now suppose the new investment will be financed by issuing more debt so that capital structure changes to one-half debt, one-fourth preferred stock, and one-fourth common stock. Those new weights (i.e., the target weights) should be used to calculate the WACC.

Taking the sources of capital to be common stock, preferred stock, and debt and allowing for the fact that in some jurisdictions, interest expense may be tax deductible, the expression for WACC is

$$\text{WACC} = w_d r_d(1 - t) + w_p r_p + w_e r_e, \tag{1}$$

where

w_d = the target proportion of debt in the capital structure when the company raises new funds

r_d = the before-tax marginal cost of debt

t = the company's marginal tax rate

w_p = the target proportion of preferred stock in the capital structure when the company raises new funds

r_p = the marginal cost of preferred stock

w_e = the target proportion of common stock in the capital structure when the company raises new funds

r_e = the marginal cost of common stock

Note that preferred stock is also referred to as preferred equity, and common stock is also referred to as common equity, or equity.

EXAMPLE 1

Computing the Weighted Average Cost of Capital

Assume that ABC Corporation has the following capital structure: 30% debt, 10% preferred stock, and 60% common stock, or equity. Also assume that interest expense is tax deductible. ABC Corporation wishes to maintain these proportions as it raises new funds. Its before-tax cost of debt is 8%, its cost of preferred stock is 10%, and its cost of equity is 15%. If the company's marginal tax rate is 40%, what is ABC's weighted average cost of capital?

Solution:

The weighted average cost of capital is

$$\text{WACC} = (0.3)(0.08)(1 - 0.40) + (0.1)(0.1) + (0.6)(0.15)$$

$$= 11.44\%.$$

The cost for ABC Corporation to raise new funds while keeping its current capital structure is 11.44%.

Self reflection:

Maintaining its current capital structure, what happens to ABC's weighted average cost of capital if component costs increase or decrease? What happens if the company's marginal tax rate increases or decreases?

There are important points concerning the calculation of the WACC as shown in Equation 1 that the analyst must be familiar with. The next section addresses the key issue of taxes.

2.1 Taxes and the Cost of Capital

b describe how taxes affect the cost of capital from different capital sources

The marginal cost of debt financing is the cost of debt after considering the allowable deduction for interest on debt based on the country's tax law. If interest cannot be deducted for tax purposes, the tax rate applied is zero, so the effective marginal cost of debt is equal to r_d in Equation 1. If interest can be deducted in full, the tax deductibility of debt reduces the effective marginal cost of debt to reflect the income shielded from taxation (often referred to as the tax shield) and the marginal cost of debt is $r_d(1 - t)$. For example, suppose a company pays €1 million in interest on its €10 million of debt. The cost of this debt is not €1 million, because this interest expense reduces taxable income by €1 million, resulting in a lower tax. If the company has a marginal tax rate of 40%, this €1 million of interest costs the company (€1 million)(1 – 0.4) = €0.6 million because the interest reduces the company's tax bill by €0.4 million. In this case, the before-tax cost of debt is 10%, whereas the after-tax cost of debt is (€0.6 million)/(€10 million) = 6%, which can also be calculated as 10%(1 – 0.4).

In jurisdictions in which a tax deduction for a business's interest expense is allowed, there may be reasons why additional interest expense is not tax deductible (e.g., not having sufficient income to offset with interest expense). If the above company with €10 million in debt were in that position, its effective marginal cost of debt would be 10% rather than 6% because any additional interest expense would not be deductible for tax purposes. In other words, if the limit on tax deductibility is reached, the marginal cost of debt is the cost of debt without any adjustment for a tax shield.

EXAMPLE 2

Incorporating the Effect of Taxes on the Costs of Capital

Jorge Ricard, a financial analyst, is estimating the costs of capital for the Zeale Corporation. In the process of this estimation, Ricard has estimated the before-tax costs of capital for Zeale's debt and equity as 4% and 6%, respectively. What are the after-tax costs of debt and equity if there is no limit to the tax deductibility of interest and Zeale's marginal tax rate is:

1 30%?

2 48%?

	Marginal Tax Rate	After-Tax Cost of Debt	After-Tax Cost of Equity
Solution to 1:	30%	0.04(1 – 0.30) = 2.80%.	6%
Solution to 2:	48%	0.04(1 – 0.48) = 2.08%.	6%

Note: There is no adjustment for taxes in the case of equity; the before-tax cost of equity is equal to the after-tax cost of equity.

COSTS OF THE VARIOUS SOURCES OF CAPITAL

3

Each source of capital has a different cost because of the differences among the sources, such as risk, seniority, contractual commitments, and potential value as a tax shield. We focus on the costs of three primary sources of capital: debt, preferred stock, and common equity.

3.1 Cost of Debt

c calculate and interpret the cost of debt capital using the yield-to-maturity approach and the debt-rating approach

The **cost of debt** is the cost of debt financing to a company when it issues a bond or takes out a bank loan. That cost is equal to the risk-free rate plus a premium for risk. A company that is perceived to be very risky would have a higher cost of debt than one that presents little investment risk. Factors that might affect the level of investment risk include profitability, stability of profits, and the degree of financial leverage. In general, the cost of debt would be higher for companies that are unprofitable, whose profits are not stable, or that are already using a lot of debt in their capital structure. We discuss two methods to estimate the before-tax cost of debt, r_d: the yield-to-maturity approach and debt-rating approach.

3.1.1 *Yield-to-Maturity Approach*

The before-tax required return on debt is typically estimated using the expected **yield to maturity** (YTM) of the company's debt based on current market values. YTM is the annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. In other words, it is the yield, r_d, that equates the present value of the bond's promised payments to its market price:

$$P_0 = \frac{PMT_1}{\left(1+\frac{r_d}{2}\right)} + \ldots + \frac{PMT_n}{\left(1+\frac{r_d}{2}\right)^n} + \frac{FV}{\left(1+\frac{r_d}{2}\right)^n} = \left[\sum_{t=1}^{n}\frac{PMT_t}{\left(1+\frac{r_d}{2}\right)^t}\right] + \frac{FV}{\left(1+\frac{r_d}{2}\right)^n}, \qquad (2)$$

where

P_0 = the current market price of the bond
PMT_t = the interest payment in period t
r_d = the yield to maturity
n = the number of periods remaining to maturity
FV = the maturity value of the bond

In this valuation equation, the constant 2 reflects the assumption that the bond pays interest semi-annually (which is the case in many but not all countries) and that any intermediate cash flows (i.e., the interest payments prior to maturity) are reinvested at the rate $r_d/2$ semi-annually.

Example 3a illustrates the calculation of the after-tax cost of debt.

EXAMPLE 3A

Calculating the After-Tax Cost of Debt

Valence Industries issues a bond to finance a new project. It offers a 10-year, $1,000 face value, 5% semi-annual coupon bond. Upon issue, the bond sells at $1,025. What is Valence's before-tax cost of debt? If Valence's marginal tax rate is 35%, what is Valence's after-tax cost of debt?

Solution:

The following are given:

$$PV = \$1{,}025.$$

$$FV = \$1{,}000.$$

$$PMT = 5\% \text{ of } 1{,}000 \div 2 = \$25.$$

$$n = 10 \times 2 = 20.$$

$$\$1{,}025 = \left[\sum_{t=1}^{20} \frac{\$25}{(1+i)^t}\right] + \frac{\$1{,}000}{(1+i)^{20}}.$$

Before proceeding to solve the problem, we already know that the before-tax cost of debt must be less than 5% because the present value of the bond is greater than the face value. We can use a financial calculator to solve for i, the six-month yield. Because $i = 2.342\%$, the before-tax cost of debt is $r_d = 2.342\% \times 2 = 4.684\%$, and Valence's after-tax cost of debt is $r_d(1 - t) = 0.04684(1 - 0.35) = 0.03045$, or 3.045%.

3.1.2 *Debt-Rating Approach*

When a reliable current market price for a company's debt is not available, the **debt-rating approach** can be used to estimate the before-tax cost of debt. Based on a company's debt rating, we estimate the before-tax cost of debt by using the yield on comparably rated bonds for maturities that closely match that of the company's existing debt.

Suppose a company's capital structure includes debt with an average maturity of 10 years and the company's marginal tax rate is 35%. If the company's rating is AAA and the yield on debt with the same debt rating and similar maturity is 4%, the company's after-tax cost of debt is

$$r_d(1 - t) = 0.04(1 - 0.35) = 2.6\%.$$

EXAMPLE 3B

Calculating the After-Tax Cost of Debt

Elttaz Company's capital structure includes debt with an average maturity of 15 years. The company's rating is A1, and it has a marginal tax rate of 18%. If the yield on comparably rated A1 bonds with similar maturity is 6.1%, what is Elttaz's after-tax cost of debt?

Solution:

Elttaz's after-tax cost of debt is

$$r_d(1 - t) = 0.061(1 - 0.18) = 5.0\%.$$

A consideration when using this approach is that debt ratings are ratings of the debt issue itself, with the issuer being only one of the considerations. Other factors, such as debt seniority and security, also affect ratings and yields, so care must be taken to consider the likely type of debt to be issued by the company in determining the comparable debt rating and yield. The debt-rating approach is a simple example of pricing on the basis of valuation-relevant characteristics, which in bond markets has been known as evaluated pricing or **matrix pricing**.

3.1.3 *Issues in Estimating the Cost of Debt*

There are other issues to consider when estimating the cost of debt. Among these are whether the debt is fixed rate or floating rate, whether it has option-like features, whether it is unrated, and whether the company uses leases instead of typical debt.

3.1.3.1 Fixed-Rate Debt vs. Floating-Rate Debt Up to now, we have assumed that the interest on debt is a fixed amount each period. We can observe market yields of the company's existing debt or market yields of debt of similar risk in estimating the before-tax cost of debt. However, the company may also issue floating-rate debt, in which the interest rate adjusts periodically according to a prescribed index, such as the prime rate, over the life of the instrument.

Estimating the cost of a floating-rate security is difficult because the cost of this form of capital over the long term depends not only on the current yields but also on the future yields. The analyst may use the current term structure of interest rates and term structure theory to assign an average cost to such instruments.

3.1.3.2 Debt with Optionlike Features How should an analyst determine the cost of debt when the company uses debt with option-like features, such as call, conversion, or put provisions? Clearly, options affect the value of debt. For example, a callable bond would have a yield greater than a similar noncallable bond of the same issuer because bondholders want to be compensated for the call risk associated with the bond. In a similar manner, the put feature of a bond, which provides the investor with an option to sell the bond back to the issuer at a predetermined price, has the effect of lowering the yield on a bond below that of a similar nonputable bond. Likewise, convertible bonds, which give investors the option of converting the bonds into common stock, lower the yield on the bonds below that of similar nonconvertible bonds.

If the company already has debt outstanding incorporating optionlike features that the analyst believes are representative of the future debt issuance of the company, the analyst may simply use the yield to maturity on such debt in estimating the cost of debt.

If the analyst believes that the company will add or remove option features in future debt issuance, the analyst can make market value adjustments to the current YTM to reflect the value of such additions or deletions. The technology for such adjustments is an advanced topic that is outside the scope of this coverage.

3.1.3.3 Nonrated Debt If a company does not have any debt outstanding or if the yields on the company's existing debt are not available, the analyst may not always be able to use the yield on similarly rated debt securities. It may be the case that the company does not have rated bonds. Although researchers offer approaches for estimating a company's "synthetic" debt rating based on financial ratios, these methods are imprecise because debt ratings incorporate not only financial ratios but also information about the particular bond issue and the issuer that are not captured in financial ratios. A further discussion of these methods is outside the scope of this reading.

3.1.3.4 Leases A lease is a contractual obligation that can substitute for other forms of borrowing. This is true whether the lease is an **operating lease** or a **finance lease** (also called a capital lease). If the company uses leasing as a source of capital, the cost of these leases should be included in the cost of capital. The cost of this form of borrowing is similar to that of the company's other long-term borrowing.

3.2 Cost of Preferred Stock

d calculate and interpret the cost of noncallable, nonconvertible preferred stock

The **cost of preferred stock** is the cost that a company has committed to pay preferred stockholders as a preferred dividend when it issues preferred stock. In the case of nonconvertible, noncallable preferred stock that has a fixed dividend rate and no maturity date (**fixed-rate perpetual preferred stock**), we can use the formula for the value of a preferred stock:

$$P_p = \frac{D_p}{r_p},$$

where

P_p = the current preferred stock price per share
D_p = the preferred stock dividend per share
r_p = the cost of preferred stock

We can rearrange this equation to solve for the cost of preferred stock:

$$r_p = \frac{D_p}{P_p}. \tag{3}$$

Therefore, the cost of preferred stock is the preferred stock's dividend per share divided by the current preferred stock's price per share. Unlike interest on debt, the dividend on preferred stock is not tax-deductible by the company; therefore, there is no adjustment to the cost for taxes.

A preferred stock may have a number of features that affect its yield and hence its cost. These features include a call option, cumulative dividends, participating dividends, adjustable-rate dividends, and convertibility into common stock. When estimating a yield based on current yields of the company's preferred stock, we must make appropriate adjustments for the effects of these features on the yield of an issue. For example, if the company has callable, convertible preferred stock outstanding yet it is expected that the company will issue only noncallable, nonconvertible preferred stock in the future, we would have to either use the current yields on comparable companies' noncallable, nonconvertible preferred stock or estimate the yield on preferred equity using methods outside the scope of this coverage.

EXAMPLE 4

Calculating the Cost of Preferred Stock

Consider a company that has one issue of preferred stock outstanding with a $3.75 cumulative dividend. If the price of this stock is $80, what is the estimate of its cost of preferred stock?

Solution:

Cost of preferred stock = $3.75/$80 = 4.6875%.

EXAMPLE 5

Choosing the Best Estimate of the Cost of Preferred Stock

Wim Vanistendael is finance director of De Gouden Tulip N.V., a leading Dutch flower producer and distributor. He has been asked by the CEO to calculate the cost of preferred stock and has recently obtained the following information:

- The issue price of preferred stock was €3.5 million, and the preferred dividend is 5%.
- If the company issued new preferred stock today, the preferred dividend yield would be 6.5%.
- The company's marginal tax rate is 30.5%.

What is the cost of preferred stock for De Gouden Tulip N.V.?

Solution:

If De Gouden Tulip were to issue new preferred stock today, the dividend yield would be close to 6.5%. The current terms thus prevail over the past terms when evaluating the actual cost of preferred stock. The cost of preferred stock for De Gouden Tulip is, therefore, 6.5%. Because preferred dividends offer no tax shield, there is no adjustment made on the basis of the marginal tax rate.

3.3 Cost of Common Equity

e calculate and interpret the cost of equity capital using the capital asset pricing model approach and the bond yield plus risk premium approach

The cost of common equity, r_e, usually referred to simply as the **cost of equity**, is the rate of return required by a company's common stockholders. A company may increase common equity through the reinvestment of earnings—that is, retained earnings—or through the issuance of new shares of stock.

The estimation of the cost of equity is challenging because of the uncertain nature of the future cash flows in terms of the amount and timing. Commonly used approaches for estimating the cost of equity include the capital asset pricing model (CAPM) method and the bond yield plus risk premium (BYPRP) method. In practice, analysts may use more than one approach to develop the cost of equity. A survey of analysts showed that the CAPM approach is used by 68% of respondents, whereas a build-up approach (bond yield plus a premium) is used by 43% of respondents (Pinto, Robinson, and Stowe 2019).

3.3.1 *Capital Asset Pricing Model Approach*

In the CAPM approach, we use the basic relationship from the capital asset pricing model theory that the expected return on a stock, $E(R_i)$, is the sum of the risk-free rate of interest, R_F, and a premium for bearing the stock's market risk, $\beta_i(R_M - R_F)$. Note that this premium incorporates the stock's return sensitivity to changes in the market return, or market-related risk, known as β_i, or beta:

$$E(R_i) = R_F + \beta_i\left[E(R_M) - R_F\right], \tag{4}$$

where

β_i = the return sensitivity of stock *i* to changes in the market return
$E(R_M)$ = the expected return on the market
$E(R_M) - R_F$ = the expected market risk premium

A risk-free asset is defined here as an asset that has no default risk. A common proxy for the risk-free rate is the yield on a default-free government debt instrument. In general, the selection of the appropriate risk-free rate should be guided by the duration of projected cash flows. For example, for the evaluation of a project with an estimated useful life of 10 years, the rate on the 10-year Treasury bond would be an appropriate proxy to use.

USING THE CAPM TO ESTIMATE THE COST OF EQUITY

1. Valence Industries wants to know its cost of equity. Its chief financial officer (CFO) believes the risk-free rate is 5%, the market risk premium is 7%, and Valence's equity beta is 1.5. What is Valence's cost of equity using the CAPM approach?

Solution:

The cost of equity for Valence is 5% + 1.5(7%) = 15.5%.

2. Exxon Mobil Corporation, BP p.l.c., and Total S.A. are three "super major" integrated oil and gas companies headquartered, respectively, in the United States, the United Kingdom, and France. An analyst estimates that the market risk premium in the United States, the United Kingdom, and the eurozone are, respectively, 4.4%, 5.5%, and 5.9%. Other information is summarized in Exhibit 1.

Exhibit 1. ExxonMobil, BP, and Total

Company	Beta	Estimated Market Risk Premium (%)	Risk-Free Rate (%)
Exxon Mobil Corporation	0.90	4.4	2.8
BP p.l.c.	0.78	5.5	2.0
Total S.A.	0.71	5.9	1.7

Source: Bloomberg; Fernandez, Pershn, and Acin (2018) survey.

Using the capital asset pricing model, calculate the cost of equity for:

1. Exxon Mobil Corporation.
2. BP p.l.c.
3. Total S.A.

Solution to 1:

The cost of equity for ExxonMobil is 2.8% + 0.90(4.4%) = 6.76%.

Solution to 2:

The cost of equity for BP is 2.0% + 0.78(5.5%) = 6.29%.

Solution to 3:

The cost of equity for Total is 1.7% + 0.71(5.9%) = 5.89%.

The expected market risk premium, or $E(R_M - R_F)$, is the premium that investors demand for investing in a market portfolio relative to the risk-free rate. When using the CAPM to estimate the cost of equity, in practice we typically estimate beta relative to an equity market index. In that case, the market premium estimate we are using is actually an estimate of the **equity risk premium** (ERP). Therefore, we are using the terms market risk premium and equity risk premium interchangeably.

An alternative to the CAPM to accommodate risks that may not be captured by the market portfolio alone is a multifactor model that incorporates factors that may be other sources of **priced risk** (risk for which investors demand compensation for bearing), including macroeconomic factors and company-specific factors. In general,

$$\begin{aligned} E(R_i) = R_F &+ \beta_{i1}(\text{Factor risk premium})_1 \\ &+ \beta_{i2}(\text{Factor risk premium})_2 + \ldots \\ &+ \beta_{ij}(\text{Factor risk premium})_j, \end{aligned} \tag{5}$$

where

β_{ij} = stock i's sensitivity to changes in the jth factor

$(\text{Factor risk premium})_j$ = expected risk premium for the jth factor

The basic idea behind these multifactor models is that the CAPM beta may not capture all the risks, especially in a global context, which include inflation, business-cycle, interest rate, exchange rate, and default risks.

There are several ways to estimate the equity risk premium, although there is no general agreement as to the best approach. The two we discuss are the historical equity risk premium approach and the survey approach.

The **historical equity risk premium approach** is a well-established approach based on the assumption that the realized equity risk premium observed over a long period of time is a good indicator of the expected equity risk premium. This approach requires compiling historical data to find the average rate of return of a country's market portfolio and the average rate of return for the risk-free rate in that country. For example, an analyst might use the historical returns to the TOPIX Index to estimate the risk premium for Japanese equities. The exceptional bull market observed during the second half of the 1990s and the bursting of the technology bubble that followed during 2000–2002 remind us that the time period for such estimates should cover complete market cycles.

Elroy Dimson, Paul Marsh, and Mike Staunton (2018) conducted an analysis of the equity risk premiums observed in markets located in 21 countries, including the United States, over the period 1900–2017. These researchers found that the annualized US equity risk premium relative to US Treasury bills was 5.6% (geometric mean) and 7.5% (arithmetic mean). They also found that the annualized US equity risk premium relative to bonds was 4.4% (geometric mean) and 6.5% (arithmetic mean). Jeremy Siegel (2005), covering the period from 1802 through 2004, observed an equity return of 6.82% and an equity risk premium in the range of 3.31%–5.36%. Note that the arithmetic mean is greater than the geometric mean as a result of the significant volatility of the observed market rate of return and the observed risk-free rate. Under the assumption of an unchanging distribution of returns over time, the arithmetic mean is the unbiased estimate of the expected single-period equity risk premium, but the geometric mean better reflects the growth rate over multiple periods. In Exhibit 2, we provide historical estimates of the equity risk premium for a few of the developed markets from Dimson et al. (2018).

Exhibit 2. Selected Equity Risk Premiums Relative to Bonds (1900–2017)

	Mean	
	Geometric	**Arithmetic**
Australia	5.0%	6.6%
Canada	3.5	5.1
France	3.1	5.4
Germany	5.1	8.4
Japan	5.1	9.1
South Africa	5.3	7.1
Switzerland	2.2	3.7
United Kingdom	3.7	5.0
United States	4.4	6.5

Note: Germany excludes 1922–1923.
Source: Dimson, Marsh, and Staunton (2018).

To illustrate the historical method as applied in the CAPM, suppose that we use the historical geometric mean for US equity of 4.4% to value Apple Computer as of early August 2018. According to Yahoo Finance, Apple had a beta of 1.14 at that time. Using a 10-year US Treasury bond yield of 3.0% to represent the risk-free rate, the estimate of the cost of equity for Apple Computer is 3.0% + 1.14(4.4%) = 8.02%.

In general, the equity risk premium can be written as

$$\text{ERP} = \overline{R}_M - \overline{R}_F,$$

where ERP is the equity risk premium, $\overline{R}_M$ is the mean return for equity, and $\overline{R}_F$ the risk-free rate.

The historical premium approach has several limitations. One limitation is that the level of risk of the stock index may change over time. Another is that the risk aversion of investors may change over time. A third limitation is that the estimates are sensitive to the method of estimation and the historical period covered.

EXAMPLE 6

Estimating the Equity Risk Premium Using Historical Rates of Return

Suppose that the arithmetic average T-bond rate observed over the last 90 years is an unbiased estimator for the risk-free rate and is 4.88%. Likewise, suppose the arithmetic average of return on the market observed over the last 90 years is an unbiased estimator for the expected return for the market. The average rate of return of the market was 9.65%. Calculate the equity risk premium.

Solution:

$$\text{ERP} = \overline{R}_M - \overline{R}_F = 9.65\% - 4.88\% = 4.77\%.$$

Another approach to estimate the equity risk premium is quite direct: Ask a panel of finance experts for their estimates, and take the mean response. This is the **survey approach**. For example, a survey of US CFOs in December 2017 found that the average expected US equity risk premium over the next 10 years was 4.42% and the median was 3.63% (Graham and Harvey 2018).

Once we have an estimate of the equity risk premium, we fine-tune this estimate for the particular company or project by adjusting it for the specific systematic risk of the project. We adjust for the specific systematic risk by multiplying the market risk premium by beta to arrive at the company's or project's risk premium, which we then add to the risk-free rate to determine the cost of equity within the framework of the CAPM.

3.3.2 *Bond Yield plus Risk Premium Approach*

For companies with publicly traded debt, the **bond yield plus risk premium approach** provides a quick estimate of the cost of equity. The BYPRP approach is based on the fundamental tenet in financial theory that the cost of capital of riskier cash flows is higher than that of less risky cash flows. In this approach, we sum the before-tax cost of debt, r_d, and a risk premium that captures the additional yield on a company's stock relative to its bonds. The estimate is, therefore,

$$r_e = r_d + \text{Risk premium}. \tag{6}$$

The risk premium compensates for the additional risk of the equity issue compared with the debt issue (recognizing that debt has a prior claim on the cash flows of the company). This risk premium is not to be confused with the equity risk premium. The equity risk premium is the difference between the cost of equity and the risk-free rate of interest. The risk premium in the bond yield plus risk premium approach is the difference between the cost of equity and the cost of debt of the company. Ideally, this risk premium is forward looking, representing the additional risk associated with the equity of the company as compared with the company's debt. However, we often estimate this premium using historical spreads between bond yields and stock yields. In developed country markets, a typical risk premium added is in the range of 3%–5%.

Looking again at Apple Computer, as of early August 2018, the yield to maturity of Apple's 3.35% coupon bonds maturing in 2027 was approximately 3.56%. Adding an arbitrary risk premium of 4.0% produces an estimate of the cost of equity of 3.56% + 4.0% = 7.56%. This estimate contrasts with the higher estimate of 8.026% from the CAPM approach. Such disparities are not uncommon and reflect the difficulty of cost of equity estimation.

4 ESTIMATING BETA

f explain and demonstrate beta estimation for public companies, thinly traded public companies, and nonpublic companies

Beta is an estimate of the company's systematic or market-related risk. It is a critical component of the CAPM, and it can be used to calculate a company's WACC. Therefore, it is essential to have a good understanding of how beta is estimated.

4.1 Estimating Beta for Public Companies

The simplest estimate of beta results from an ordinary least squares regression of the return on the stock on the return on the market. The result is often called an unadjusted or "raw" historical beta. The actual values of beta estimates are influenced by several choices:

- *The choice of the index used to represent the market portfolio*: For US equities, the S&P 500 Index and NYSE Composite have been traditional choices. In Japan, analysts would likely use the Nikkei 225 Index.
- *The length of the data period and the frequency of observations*: The most common choice is five years of monthly data, yielding 60 observations.

Researchers have observed that beta tends to regress toward 1.0. In other words, the value of a stock's beta in a future period is likely to be closer to the mean value of 1.0, the beta of an average-systematic-risk security, than to the value of the calculated raw beta. Because valuation is forward looking, it is logical to adjust the raw beta so that it more accurately predicts a future beta. The most commonly used adjustment was introduced by Blume (1971):

$$\text{Adjusted beta} = (2/3)(\text{Unadjusted beta}) + (1/3)(1.0). \quad (7)$$

For example, if the beta from a regression of an asset's returns on the market return is 1.30, adjusted beta is (2/3)(1.30) + (1/3)(1.0) = 1.20. Equation 7 acts to "smooth" raw betas by adjusting betas above and below 1.0 toward 1.0. Vendors of financial information, such as Bloomberg, often report both raw and adjusted betas.

EXAMPLE 7

Estimating the Adjusted Beta for a Public Company

Betty Lau is an analyst trying to estimate the cost of equity for Singapore Telecommunications Limited. She begins by running an ordinary least squares regression to estimate the beta. Her estimated value is 0.4, which she believes needs adjustment. What is the adjusted beta value she should use in her analysis?

Solution:

$$\text{Adjusted beta} = (2/3)(0.4) + (1/3)(1.0) = 0.6.$$

Arriving at an estimated beta for publicly traded companies is generally not a problem because of the accessibility of stock return data, the ease of use of estimating beta using simple regression, and the availability of estimated betas on publicly traded companies from financial analysis vendors.

The challenge comes in estimating a beta for a company that is thinly traded or nonpublic or for a project that is not the average or typical project of a publicly traded company. Estimating beta in these cases requires proxying for the beta by using information on the project or company combined with the beta of a publicly traded company.

4.2 Estimating Beta for Thinly Traded and Nonpublic Companies

It is not possible to run an ordinary least squares regression to estimate beta if a stock is thinly traded or a company is nonpublic. When a share issue trades infrequently, the most recent transaction price may be stale and may not reflect underlying changes in value. If beta is estimated on the basis of, for example, a monthly data series in which missing values are filled with the most recent transaction price, the estimated beta will be too small. This is because this methodology implicitly assumes that the stock's price is more stable than it really is. As a result, the required return on equity will be understated.

In these cases, a practical alternative is to base the beta estimate on the betas of comparable companies that are publicly traded. A **comparable company**, also called a **peer company**, is a company that has similar business risk. A comparable, or peer, company can be identified by using an industry classification system, such as the MSCI/Standard & Poor's Global Industry Classification Standard (GICS) or the FTSE Industry Classification Benchmark (ICB). The analyst can then indirectly estimate the beta on the basis of the betas of the peer companies.

Because financial leverage can affect beta, an adjustment must be made if the peer company has a substantially different capital structure. First, the peer company's beta must be unlevered to estimate the beta of the assets—reflecting only the systematic risk arising from the fundamentals of the industry. Then, the unlevered beta, often referred to as the **asset beta** because it reflects the business risk of the assets, must be re-levered to reflect the capital structure of the company in question.

Let β_E be the equity beta of the peer company before removing the effects of leverage. Assuming the debt of the peer company is of high quality—so that the debt's beta, or β_D, is approximately equal to zero (that is, it is assumed to have no market risk)—analysts can use the following expression to unlever the beta:

$$\beta_U = \beta_E \left[\frac{1}{1 + (1 - t)\frac{D}{E}} \right], \tag{8}$$

where β_U is the unlevered beta, t is the marginal tax rate of the peer company, and D and E are the market values of debt and equity, respectively, of the peer company.

Now we can re-lever the unlevered beta by rearranging the equation to reflect the capital structure of the thinly traded or nonpublic company in question:

$$\beta_E = \beta_U \left[1 + (1 - t)\frac{D}{E} \right]. \tag{9}$$

where β_E is now the equity beta of the thinly traded or nonpublic company, t is the marginal tax rate of the thinly traded or nonpublic company, and D and E are the debt-to-equity values, respectively, of the thinly traded or nonpublic company.

EXAMPLE 8

Estimating the Adjusted Beta for a Nonpublic Company

Raffi Azadian wants to determine the cost of equity for Elucida Oncology, a privately held company. Raffi realizes that he needs to estimate Elucida's beta before he can proceed. He determines that Merck & Co. is an appropriate publicly traded peer company. Merck has a beta of 0.7, it is 40% funded by debt, and its marginal tax rate is 21%. If Elucida is only 10% funded by debt and its marginal tax rate is also 21%, what is Elucida's beta?

Solution:

Since Merck is 40% funded by debt, it is 60% funded by equity. Therefore,

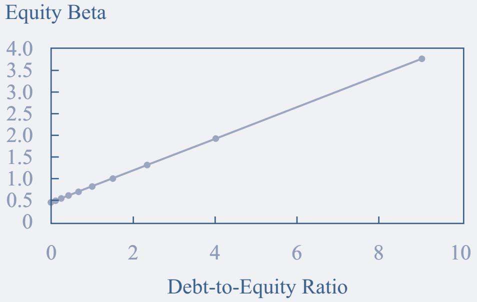
$$\text{Unlevered beta} = (0.7)\left[\frac{1}{1+(1-0.21)\left(\frac{0.4}{0.6}\right)}\right] = 0.46.$$

Now we re-lever the unlevered beta using Elucida's tax rate and capital structure:

$$\text{Elucida's beta} = 0.46\left[1+(1-0.21)\left(\frac{0.1}{0.9}\right)\right] = 0.50.$$

The following table and figure show how Elucida's beta increases as leverage rises.

Debt-to-Equity Ratio	Equity Beta
0.00	0.46
0.11	0.50
0.25	0.55
0.43	0.62
0.67	0.70
1.00	0.82
1.50	1.01
2.33	1.31
4.00	1.91
9.00	3.73



The beta estimate can then be used to determine the component cost of equity and combined with the cost of debt in a weighted average to provide an estimate of the cost of capital for the company.

EXAMPLE 9

Estimating the Weighted Average Cost of Capital

Georg Schrempp is the CFO of Bayern Chemicals KgaA, a German manufacturer of industrial, commercial, and consumer chemical products. Bayern Chemicals is privately owned, and its shares are not listed on an exchange. The CFO has appointed Markus Meier, CFA, a third-party valuator, to perform a stand-alone valuation of Bayern Chemicals. Meier has access to the following information to calculate Bayern Chemicals' weighted average cost of capital:

- The nominal risk-free rate, represented by the yield on the long-term 10-year German bund, was 4.5% at the valuation date.

- The average long-term historical equity risk premium in Germany is assumed to be 5.7%.
- Bayern Chemicals' corporate tax rate is 38%.
- Bayern Chemicals' target debt-to-equity ratio is 0.7. Its capital structure is 41% debt.
- Bayern Chemicals' cost of debt has an estimated spread of 225 bps over the 10-year bund.
- Exhibit 3 supplies additional information on comparables for Bayern Chemicals.

Exhibit 3. Information on Comparables

Comparable Companies	Country	Tax Rate (%)	Market Capitalization in Millions	Net Debt in Millions	D/E	Beta
British Chemicals Ltd.	United Kingdom	30.0	4,500	6,000	1.33	1.45
Compagnie Petrochimique S.A.	France	30.3	9,300	8,700	0.94	0.75
Rotterdam Chemie N.V.	Netherlands	30.5	7,000	7,900	1.13	1.05
Average					1.13	1.08

Based only on the information given, calculate Bayern Chemicals' WACC.

Solution:

To calculate the cost of equity, the first step is to "unlever" the betas of the comparable companies and calculate an average for a company with business risk similar to the average of these companies:

Comparable Companies	Unlevered Beta
British Chemicals Ltd.	$0.75 = \left[\dfrac{1.45}{1 + (1 - 0.30)(1.33)}\right]$
Compagnie Petrochimique S.A.	$0.45 = \left[\dfrac{0.75}{1 + (1 - 0.303)(0.94)}\right]$
Rotterdam Chemie N.V.	$0.59 = \left[\dfrac{1.05}{1 + (1 - 0.305)(1.13)}\right]$
Average*	0.60

*An analyst must apply judgment and experience to determine a representative average for the comparable companies. This example uses a simple average, but in some situations a weighted average based on some factor, such as market capitalization, may be more appropriate.

Levering the average unlevered beta for the peer group average, applying Bayern Chemicals' target debt-to-equity ratio and marginal tax rate, results in a beta of 0.86:

$$\beta_{Bayern\ Chemicals} = 0.60\{1 + [(1 - 0.38)0.7]\} = 0.86.$$

Using CAPM, the cost of equity of Bayern Chemicals (r_e) can be calculated as follows:

$$r_e = 4.5\% + (0.86)(5.7\%) = 9.4\%.$$

The weights for the cost of debt and cost of equity may be calculated as follows:

$$w_d = 0.41, \text{ and } w_e = (1 - 0.41) = 0.59.$$

The before-tax cost of debt of Bayern Chemicals (r_d) is 6.75%:

$$r_d = 4.5\% + 2.25\% = 6.75\%.$$

As a result, Bayern Chemicals' WACC is 7.26%:

$$\text{WACC} = (0.41)(0.0675)(1 - 0.38) + (0.59)(0.094)$$

$$= 0.0726, \text{ or } 7.26\%.$$

5 FLOTATION COSTS

g explain and demonstrate the correct treatment of flotation costs

When a company raises new capital, it generally seeks the assistance of investment bankers. Investment bankers charge the company a fee based on the size and type of offering. This fee is referred to as the **flotation cost**. In general, flotation costs are higher in percentage terms for equity issuances than they are for debt. They are also higher for smaller issuance amounts or for issuances that are perceived to be riskier. In the case of debt and preferred stock, we do not usually incorporate flotation costs in the estimated cost of capital because the amount of these costs is quite small, often less than 1% of the value of the offering.

However, with equity issuance, the flotation costs may be substantial, so we should consider these when estimating the cost of external equity capital. Average flotation costs for new equity have been estimated at 7.11% of the value of the offering in the United States,[1] 1.65% in Germany,[2] 5.78% in the United Kingdom,[3] and 4.53% in Switzerland.[4] A large part of the differences in costs among these studies is likely attributed to the type of offering; cash underwritten offers, typical in the United States, are generally more expensive than rights offerings, which are common in Europe.

---

[1] Inmoo Lee, Scott Lochhead, Jay R. Ritter, and Quanshui Zhao, "The Costs of Raising Capital," *Journal of Financial Research* 19 (Spring 1996): 59–71.

[2] Thomas Bühner and Christoph Kaserer, "External Financing Costs and Economies of Scale in Investment Banking: The Case of Seasoned Equity Offerings in Germany," *European Financial Management* 9 (June 2002): 249.

[3] Seth Armitage, "The Direct Costs of UK Rights Issues and Open Offers," *European Financial Management* 6 (2000): 57–68.

[4] Christoph Kaserer and Fabian Steiner, "The Cost of Raising Capital—New Evidence from Seasoned Equity Offerings in Switzerland," working paper, Technische Universität München (February 2004).

How should flotation costs be accounted for? There are two views on this topic. One view, which you can find often in textbooks, is to directly incorporate the flotation costs into the cost of capital. The other view is that flotation costs should be incorporated into the valuation analysis as an additional cost. We will argue that the second view is preferred.

Consistent with the first view, we can specify flotation costs in monetary terms as an amount per share or as a percentage of the share price. With flotation costs specified in monetary terms on a per share basis, the cost of external equity is

$$r_e = \left(\frac{D_1}{P_0 - F}\right) + g, \tag{10}$$

where

r_e is the cost of equity
D_1 is the dividend expected at the end of Period 1
P_0 is the current stock price
F is the monetary per share flotation cost
g is the growth rate

As a percentage applied against the price per share, the cost of external equity is

$$r_e = \left[\frac{D_1}{P_0(1 - f)}\right] + g, \tag{11}$$

where f is the flotation cost as a percentage of the issue price.

EXAMPLE 10

Estimating the Cost of Equity with Flotation Costs

A company has a current dividend of \$2 per share, a current price of \$40 per share, and an expected growth rate of 5%.

1 What is the cost of internally generated equity (i.e., stock is not issued and flotation costs are not incurred)?

2 What is the cost of external equity (i.e., new shares are issued and flotation costs are incurred) if the flotation costs are 4% of the issuance?

Solution to 1:

$$r_e = \left[\frac{\$2(1+0.05)}{\$40}\right] + 0.05 = 0.0525 + 0.05 = 0.1025\text{, or } 10.25\%.$$

Solution to 2:

$$r_e = \left[\frac{\$2(1+0.05)}{\$40(1-0.04)}\right] + 0.05 = 0.05469 + 0.05 = 0.1047\text{, or } 10.47\%.$$

Many experts object to this methodology. Flotation costs are a cash flow that occurs at issue and they affect the value of the project only by reducing the initial cash flow. However, by adjusting the cost of capital for flotation costs, we apply a higher cost of capital to determine the present value of the future cash flows. The result is that the calculated net present value of a project is less than its true net present value. As a result, otherwise profitable projects may get rejected when this methodology is used.

An alternative and preferred approach is to make the adjustment for flotation costs to the cash flows in the valuation computation. For example, consider a project that requires a €60,000 initial cash outlay and is expected to produce cash flows of €10,000 each year for 10 years. Suppose the company's marginal tax rate is 40%, the before-tax cost of debt is 5%, and the cost of equity is 10%. Assume the company will finance the project with 40% debt and 60% equity. Exhibit 4 summarizes the information on the component costs of capital.

Exhibit 4. After-Tax Costs of Debt and Equity

Source of Capital	Amount Raised (€)	Proportion	Marginal After-Tax Cost
Debt	24,000	0.40	0.05(1 – 0.4) = 0.03
Equity	36,000	0.60	0.10

The weighted average cost of capital is 7.2%, calculated as 0.40(3%) + 0.60(10%). Ignoring flotation costs for the moment and using a financial calculator, we find that the net present value (NPV) of this project can be expressed as

NPV = Present value of inflows – Present value of the outflows,

or

NPV = €69,591 – €60,000 = €9,591,

where €69,591 is the present value of €10,000 per year for 10 years at 7.2%. Now suppose flotation costs amount to 5% of the new equity capital: (0.05)(€36,000) = €1,800. If flotation costs are not tax deductible, the net present value considering flotation costs is

NPV = €69,591 – €60,000 – €1,800 = €7,791.

If flotation costs are tax deductible, the net present value considering flotation costs is

NPV = €69,591 – €60,000 – €1,800(1-0.4) = €8,511.

Suppose instead of considering the flotation costs as part of the cash flows, we made an adjustment to the cost of equity. Without showing the calculations, the cost of equity increases from 10% to 10.2632%, the cost of capital increases from 7.22% to 7.3579%, and the NPV decreases from €9,591 to €9,089. As you can see, we arrive at different assessments of value using these two methods.

If the preferred method is to deduct the flotation costs as part of the net present value calculation, why do many textbooks highlight the adjustment to the cost of capital? One reason is that it is often difficult to identify specific financing associated with a project. Making the adjustment for flotation costs to the cost of capital is most useful if specific project financing cannot be identified. A second reason is that by adjusting the cost of capital for the flotation costs, it is easier to demonstrate how the costs of financing a company change as a company exhausts internally generated equity (i.e., retained earnings) and switches to externally generated equity (i.e., a new stock issue).

METHODS IN USE

6

We have introduced methods that may be used to estimate the cost of capital for a company or a project, but which methods do companies actually use when making investment decisions? John Graham and Campbell Harvey (2002) surveyed a large number of CFOs to find out which methods they prefer. Their survey revealed the following:

- The most popular method for estimating the cost of equity is the capital asset pricing model.
- Few companies use the dividend discount model (which we did not cover) to estimate the cost of equity.
- Publicly traded companies are more likely to use the capital asset pricing model than are private companies.
- In evaluating projects, the majority of CFOs use a single company cost of capital, but a large portion apply some type of risk adjustment for individual projects.

Their survey also revealed that the single-factor capital asset pricing model is the most popular method for estimating the cost of equity. The second and third most popular methods, respectively, are average stock returns and multifactor return models. The lack of popularity of the dividend discount model indicates that this approach, although once favored, has lost its appeal in practice.

In a survey of publicly traded multinational European companies, Franck Bancel and Usha Mittoo (2004) provided evidence consistent with the Graham and Harvey (2002) survey. They found that over 70% of companies use the CAPM to determine the cost of equity; this finding is similar to the 73.5% of US companies that use the CAPM. In a survey of both publicly traded and private European companies, Dirk Brounen, Abe de Jong, and Kees Koedijk (2004) confirmed the result of Graham and Harvey that larger companies are more likely to use the more sophisticated methods, such as CAPM, in estimating the cost of equity. Brounen, de Jong, and Koedijk found that the use of the CAPM was less popular for their sample (ranging from 34% to 55.6%, depending on the country) than for the other two surveys, which may reflect the inclusion of smaller, private companies in their sample.

We learn from the survey evidence that the CAPM is a popular method for estimating the cost of equity capital and that it is used less often by smaller, private companies. The latter result is not surprising, because of the difficulty in estimating systematic risk in cases in which the company's equity is not publicly traded.

SUMMARY

In this reading, we provided an overview of the techniques used to calculate the cost of capital for companies and projects. We examined the weighted average cost of capital, discussing the methods commonly used to estimate the component costs of capital and the weights applied to these components.

- The weighted average cost of capital is a weighted average of the after-tax marginal costs of each source of capital: WACC = $w_d r_d (1 - t) + w_p r_p + w_e r_e$.
- The before-tax cost of debt is generally estimated by either the yield-to-maturity method or the bond rating method.

- The yield-to-maturity method of estimating the before-tax cost of debt uses the familiar bond valuation equation. Assuming semi-annual coupon payments, the equation is

  $$P_0 = \frac{PMT_1}{\left(1+\frac{r_d}{2}\right)} + \ldots + \frac{PMT_n}{\left(1+\frac{r_d}{2}\right)^n} + \frac{FV}{\left(1+\frac{r_d}{2}\right)^n} = \left[\sum_{t=1}^{n}\frac{PMT_t}{\left(1+\frac{r_d}{2}\right)^t}\right] + \frac{FV}{\left(1+\frac{r_d}{2}\right)^n}.$$

  We solve for the six-month yield ($r_d/2$) and then annualize it to arrive at the before-tax cost of debt, r_d.
- Because interest payments are generally tax deductible, the after-tax cost is the true, effective cost of debt to the company. If a yield to maturity or bond rating is not available, such as in the case of a private company without rated debt or a project, the estimate of the cost of debt becomes more challenging.
- The cost of preferred stock is the preferred stock dividend divided by the current preferred stock price:

  $$r_p = \frac{D_p}{P_p}.$$

- The cost of equity is the rate of return required by a company's common stockholders. We estimate this cost using the CAPM (or its variants).
- The CAPM is the approach most commonly used to calculate the cost of equity. The three components needed to calculate the cost of equity are the risk-free rate, the equity risk premium, and beta:

  $$\mathrm{E}(R_i) = R_F + \beta_i\left[\mathrm{E}(R_M) - R_F\right].$$

- In estimating the cost of equity, an alternative to the CAPM is the bond yield plus risk premium approach. In this approach, we estimate the before-tax cost of debt and add a risk premium that reflects the additional risk associated with the company's equity.
- When estimating the cost of equity capital using the CAPM, if we do not have publicly traded equity, we may be able to use a comparable company operating in the same business line to estimate the unlevered beta for a company with similar business risk, β_U:

  $$\beta_U = \beta_E\left[\frac{1}{1+(1-t)\frac{D}{E}}\right].$$

  Then, we lever this beta to reflect the financial risk of the project or company:

  $$\beta_E = \beta_U\left[1+(1-t)\frac{D}{E}\right].$$

- Flotation costs are costs incurred in the process of raising additional capital. The preferred method of including these costs in the analysis is as an initial cash flow in the valuation analysis.
- Survey evidence tells us that the CAPM method is the most popular method used by companies in estimating the cost of equity. The CAPM method is more popular with larger, publicly traded companies, which is understandable considering the additional analyses and assumptions required in estimating systematic risk for a private company or project.

REFERENCES

Bancel, Franck, and Usha Mittoo. 2004. "Cross-Country Determinants of Capital Structure Choice: A Survey of European Firms." *Financial Management* 33 (4): 103–32.

Blume, Marshall. 1971. "On the Assessment of Risk." *Journal of Finance* 26 (1): 1–10.

Brounen, Dirk, Abe de Jong, and Kees Koedijk. 2004. "Corporate Finance in Europe: Confronting Theory with Practice." *Financial Management* 33 (4): 71–101.

Dimson, E., P. Marsh, and M. Staunton. 2018. "*Credit Suisse Global Investment Returns Yearbook 2018*" (February).

Fernandez, Pablo, Vitaly Pershin, and Isabel Fernández Acin. 2018. "*Market Risk Premium and Risk-Free Rate Used for 59 Countries in 2018: A Survey.*" Available at SSRN: https://ssrn.com/abstract=3155709.

Graham, John R., and Campbell R. Harvey. 2002. "How Do CFOs Make Capital Budgeting and Capital Structure Decisions?" *Journal of Applied Corporate Finance* 15 (1): 8–23.

Graham, John R., and Campbell R. Harvey. 2018. "*The Equity Risk Premium in 2018.*" Working paper (27 March). Available at https://ssrn.com/abstract=3151162.

Pinto, Jerald E., Thomas R. Robinson, and John D. Stowe. 2019. "Equity Valuation: A Survey of Professional Practice." *Review of Financial Economics* 37 (2): 219–33.

Siegel, Jeremy J. 2005. "Perspectives on the Equity Risk Premium." *Financial Analysts Journal* 61 (6): 61–73.

PRACTICE PROBLEMS

1 The cost of equity is equal to the:

  A expected market return.

  B rate of return required by stockholders.

  C cost of retained earnings plus dividends.

2 Which of the following statements is correct?

  A The appropriate tax rate to use in the adjustment of the before-tax cost of debt to determine the after-tax cost of debt is the average tax rate because interest is deductible against the company's entire taxable income.

  B For a given company, the after-tax cost of debt is generally less than both the cost of preferred equity and the cost of common equity.

  C For a given company, the after-tax cost of debt is generally higher than both the cost of preferred equity and the cost of common equity.

3 An analyst's data source shows that Newmont Mining (NEM) has an estimated beta of –0.2. The risk-free rate of return is 2.5%, and the equity risk premium is estimated to be 4.5%.

  A Using the CAPM, calculate the required rate of return for investors in NEM.

  B The analyst notes that the current yield to maturity on corporate bonds with a credit rating similar to NEM is approximately 3.9%. How should this information affect the analyst's estimate?

4 An analyst assembles the following facts concerning a company's component costs of capital and capital structure. Based on the information given, calculate the company's WACC.

Facts	**(%)**
Cost of equity based on the CAPM	15.60
Pretax cost of debt	8.28
Corporate tax rate	30.00
Capital structure weight	Equity 80, Debt 20

5 Dot.Com has determined that it could issue $1,000 face value bonds with an 8% coupon paid semi-annually and a five-year maturity at $900 per bond. If Dot.Com's marginal tax rate is 38%, its after-tax cost of debt is *closest* to:

  A 6.2%.

  B 6.4%.

  C 6.6%.

6 The cost of debt can be determined using the yield-to-maturity and bond rating approaches. If the bond rating approach is used, the:

  A coupon is the yield.

  B yield is based on the interest coverage ratio.

  C company is rated and the rating can be used to assess the credit default spread of the company's debt.

© 2021 CFA Institute. All rights reserved.

7 Morgan Insurance Ltd. issued a fixed-rate perpetual preferred stock three years ago and placed it privately with institutional investors. The stock was issued at $25 per share with a $1.75 dividend. If the company were to issue preferred stock today, the yield would be 6.5%. The stock's current value is:

- **A** $25.00.
- **B** $26.92.
- **C** $37.31.

8 The Gearing Company has an after-tax cost of debt capital of 4%, a cost of preferred stock of 8%, a cost of equity capital of 10%, and a weighted average cost of capital of 7%. Gearing intends to maintain its current capital structure as it raises additional capital. In making its capital-budgeting decisions for the average-risk project, the relevant cost of capital is:

- **A** 4%.
- **B** 7%.
- **C** 8%.

9 Fran McClure, of Alba Advisers, is estimating the cost of capital of Frontier Corporation as part of her valuation analysis of Frontier. McClure will be using this estimate, along with projected cash flows from Frontier's new projects, to estimate the effect of these new projects on the value of Frontier. McClure has gathered the following information on Frontier Corporation:

	Current Year ($)	Forecasted for Next Year ($)
Book value of debt	50	50
Market value of debt	62	63
Book value of equity	55	58
Market value equity	210	220

The weights that McClure should apply in estimating Frontier's cost of capital for debt and equity are, respectively:

- **A** $w_d = 0.200$ and $w_e = 0.800$.
- **B** $w_d = 0.185$ and $w_e = 0.815$.
- **C** $w_d = 0.223$ and $w_e = 0.777$.

10 Wang Securities had a long-term stable debt-to-equity ratio of 0.65. Recent bank borrowing for expansion into South America raised the ratio to 0.75. The increased leverage has what effect on the asset beta and equity beta of the company?

- **A** The asset beta and the equity beta will both rise.
- **B** The asset beta will remain the same, and the equity beta will rise.
- **C** The asset beta will remain the same, and the equity beta will decline.

11 Brandon Wiene is a financial analyst covering the beverage industry. He is evaluating the impact of DEF Beverage's new product line of flavored waters. DEF currently has a debt-to-equity ratio of 0.6. The new product line would be financed with $50 million of debt and $100 million of equity. In estimating the valuation impact of this new product line on DEF's value, Wiene has estimated the equity beta and asset beta of comparable companies. In calculating the equity beta for the product line, Wiene is intending to use DEF's existing capital structure when converting the asset beta into a project beta. Which of the following statements is correct?

A Using DEF's debt-to-equity ratio of 0.6 is appropriate in calculating the new product line's equity beta.

B Using DEF's debt-to-equity ratio of 0.6 is not appropriate; rather, the debt-to-equity ratio of the new product, 0.5, is appropriate to use in calculating the new product line's equity beta.

C Wiene should use the new debt-to-equity ratio of DEF that would result from the additional $50 million debt and $100 million equity in calculating the new product line's equity beta.

12 Happy Resorts Company currently has 1.2 million common shares of stock outstanding, and the stock has a beta of 2.2. It also has $10 million face value of bonds that have five years remaining to maturity and an 8% coupon with semi-annual payments and are priced to yield 13.65%. If Happy issues up to $2.5 million of new bonds, the bonds will be priced at par and will have a yield of 13.65%; if it issues bonds beyond $2.5 million, the expected yield on the entire issuance will be 16%. Happy has learned that it can issue new common stock at $10 a share. The current risk-free rate of interest is 3%, and the expected market return is 10%. Happy's marginal tax rate is 30%. If Happy raises $7.5 million of new capital while maintaining the same debt-to-equity ratio, its weighted average cost of capital will be *closest* to:

A 14.5%.

B 15.5%.

C 16.5%.

The following information relates to Questions 13–17.

Jurgen Knudsen has been hired to provide industry expertise to Henrik Sandell, CFA, an analyst for a pension plan managing a global large-cap fund internally. Sandell is concerned about one of the fund's larger holdings, auto parts manufacturer Kruspa AB. Kruspa currently operates in 80 countries, with the previous year's global revenues at €5.6 billion. Recently, Kruspa's CFO announced plans for expansion into Trutan, a country with a developing economy. Sandell worries that this expansion will change the company's risk profile and wonders if he should recommend a sale of the position.

Sandell provides Knudsen with the basic information. Kruspa's global annual free cash flow to the firm is €500 million, and earnings are €400 million. Sandell estimates that cash flow will level off at a 2% rate of growth. Sandell also estimates that Kruspa's after-tax free cash flow to the firm on the Trutan project for next three years is, respectively, €48 million, €52 million, and €54.4 million. Kruspa recently announced a dividend of €4.00 per share of stock. For the initial analysis, Sandell requests that Knudsen ignore possible currency fluctuations. He expects the Trutanese plant to sell only to customers within Trutan for the first three years. Knudsen is asked to evaluate Kruspa's planned financing of the required €100 million in Sweden with an €80 million public offering of 10-year debt and the remainder with an equity offering.

Additional information:	
Equity risk premium, Sweden	4.82%
Risk-free rate of interest, Sweden	4.25%
Industry debt-to-equity ratio	0.3
Market value of Kruspa's debt	€900 million

Additional information:	
Market value of Kruspa’s equity	€2.4 billion
Kruspa’s equity beta	1.3
Kruspa’s before-tax cost of debt	9.25%
Trutan credit A2 country risk premium	1.88%
Corporate tax rate	37.5%
Interest payments each year	Level

13 Using the capital asset pricing model, Kruspa’s cost of equity capital for its typical project is *closest* to:

A 7.62%.

B 10.52%.

C 12.40%.

14 Sandell is interested in the weighted average cost of capital of Kruspa AB prior to its investing in the Trutan project. This weighted average cost of capital is *closest* to:

A 7.65%.

B 9.23%.

C 10.17%.

15 In his estimation of the project’s cost of capital, Sandell would like to use the asset beta of Kruspa as a base in his calculations. The estimated asset beta of Kruspa prior to the Trutan project is *closest* to:

A 1.053.

B 1.110.

C 1.327.

16 Sandell is performing a sensitivity analysis of the effect of the new project on the company's cost of capital. If the Trutan project has the same asset risk as Kruspa, the estimated project beta for the Trutan project, if it is financed 80% with debt, is *closest* to:

A 1.300.

B 2.635.

C 3.686.

17 As part of the sensitivity analysis of the effect of the new project on the company's cost of capital, Sandell is estimating the cost of equity of the Trutan project considering that the Trutan project requires a country equity premium to capture the risk of the project. The cost of equity for the project in this case is *closest* to:

A 10.52%.

B 19.91%.

C 28.95%.

The following information relates to Questions 18–21.

Boris Duarte, CFA, covers initial public offerings for Zellweger Analytics, an independent research firm specializing in global small-cap equities. He has been asked to evaluate the upcoming new issue of TagOn, a US-based business intelligence software company. The industry has grown at 26% per year for the previous three years. Large companies dominate the market, but sizable comparable companies, such as Relevant Ltd., ABJ Inc., and Opus Software Pvt. Ltd., also compete. Each of these competitors is domiciled in a different country, but they all have shares of stock that trade on the US NASDAQ. The debt ratio of the industry has risen slightly in recent years.

Company	Sales in Millions ($)	Market Value Equity in Millions ($)	Market Value Debt in Millions ($)	Equity Beta	Tax Rate (%)	Share Price ($)
Relevant Ltd.	752	3,800	0.0	1.702	23	42
ABJ Inc.	843	2,150	6.5	2.800	23	24
Opus Software Pvt. Ltd.	211	972	13.0	3.400	23	13

Duarte uses the information from the preliminary prospectus for TagOn's initial offering. The company intends to issue 1 million new shares. In his conversation with the investment bankers for the deal, he concludes the offering price will be between $7 and $12. The current capital structure of TagOn consists of a $2.4 million five-year noncallable bond issue and 1 million common shares. The following table includes other information that Duarte has gathered:

Currently outstanding bonds	$2.4 million five-year bonds, coupon of 12.5% paying semi-annually with a market value of $2.156 million
Risk-free rate of interest	5.25%
Estimated equity risk premium	7%
Tax rate	23%

18 The asset betas for Relevant, ABJ, and Opus, respectively, are:

A 1.70, 2.52, and 2.73.

B 1.70, 2.79, and 3.37.

C 1.70, 2.81, and 3.44.

19 The average asset beta for comparable players in this industry, Relevant, ABJ, and Opus, weighted by market value of equity is *closest* to:

A 1.67.

B 1.97.

C 2.27.

20 Using the capital asset pricing model, the cost of equity capital for a company in this industry with a debt-to-equity ratio of 0.01, an asset beta of 2.27, and a marginal tax rate of 23% is *closest* to:

A 17%.

B 21%.

C 24%.

21 The marginal cost of capital for TagOn, based on an average asset beta of 2.27 for the industry and assuming that new stock can be issued at $8 per share, is *closest* to:

A 20.5%.

B 21.0%.

C 21.5%.

22 Two years ago, a company issued $20 million in long-term bonds at par value with a coupon rate of 9%. The company has decided to issue an additional $20 million in bonds and expects the new issue to be priced at par value with a coupon rate of 7%. The company has no other debt outstanding and has a tax rate of 40%. To compute the company's weighted average cost of capital, the appropriate after-tax cost of debt is *closest* to:

A 4.2%.

B 4.8%.

C 5.4%.

23 An analyst gathered the following information about a private company and its publicly traded competitor:

Comparable Companies	Tax Rate (%)	Debt/Equity	Equity Beta
Private company	30.0	1.00	na
Public company	35.0	0.90	1.75

The estimated equity beta for the private company is *closest* to:

A 1.029.

B 1.104.

C 1.877.

24 Which of the following statements is *most accurate*? If two equity issues have the same market risk but the first issue has higher leverage, greater liquidity, and a higher required return, the higher required return *is most likely* the result of the first issue's:

A greater liquidity.

B higher leverage.

C higher leverage and greater liquidity.

25 SebCoe plc, a British firm, is evaluating an investment in a £50 million project that will be financed with 50% debt and 50% equity. Management has already determined that the NPV of this project is £5 million if it uses internally generated equity. However, if the company uses external equity, it will incur flotation costs of 5.8%. Assuming flotation costs are not tax deductible, the NPV using external equity would be:

A less than £5 million because we would discount the cash flows using a higher weighted average cost of capital that reflects the flotation costs.

B £3.55 million because flotation costs reduce NPV by $1.45 million.

C £5 million because flotation costs have no impact on NPV.

26 At the time of valuation, the estimated betas for JPMorgan Chase & Co. and the Boeing Company were 1.50 and 0.80, respectively. The risk-free rate of return was 4.35%, and the equity risk premium was 8.04%. Based on these data, calculate the required rates of return for these two stocks using the CAPM.

SOLUTIONS

1 B is correct. The cost of equity is defined as the rate of return required by stockholders.

2 B is correct. Debt is generally less costly than preferred or common stock. The cost of debt is further reduced if interest expense is tax deductible.

3 **A.** The required return is given by

$$r = 0.025 + (-0.2)(0.045) = 2.5\% - 0.9\% = 1.6\%.$$

This example indicates that Newmont Mining has a required return of 1.6%. When beta is negative, the CAPM calculation yields a required rate of return that is below the risk-free rate, which is arguably not meaningful. Cases of equities with negative betas are relatively rare.

B. The fact that the NEM's cost of debt is higher than the calculated required return on equity is another indicator that the return estimated using CAPM is not useful for valuing the company's equity.

4 The company's WACC is 13.64%, calculated as follows:

	Equity		Debt	WACC
Weight	0.80		0.20	
After-Tax Cost	15.6%		(1 – 0.30)8.28%	
Weight × Cost	12.48%	+	1.16%	= 13.64%

5 C is correct. FV = \$1,000, PMT = \$40, N = 10, and PV = \$900.

Solve for i. The six-month yield, i, is 5.3149%.

YTM = 5.3149% × 2 = 10.62985%.

$r_d(1 - t) = 10.62985\%(1 - 0.38) = 6.5905\%$.

6 C is correct. The bond rating approach depends on knowledge of the company's rating and can be compared with yields on bonds in the public market.

7 B is correct. The company can issue preferred stock at 6.5%. Therefore, the calculation of the preferred stock's current value is

$$P_p = \$1.75/0.065 = \$26.92.$$

8 B is correct. The weighted average cost of capital, using weights derived from the current capital structure, is the best estimate of the cost of capital for the average-risk project of a company.

9 C is correct. McClure should use the forecasted or target market values to calculate the weights.

$$w_d = \$63/(\$220 + 63) = 0.223.$$

$$w_e = \$220/(\$220 + 63) = 0.777.$$

10 B is correct. Asset risk does not change with a higher debt-to-equity ratio. Equity risk rises with higher debt.

11 B is correct. The debt-to-equity ratio of the new product should be used when making the adjustment from the asset beta, derived from the comparables, to the equity beta of the new product.

12 B is correct. The capital structure is as follows:

Market value of debt: FV = \$10,000,000, PMT = \$400,000, N = 10, and I/YR = 6.825%. Solving for PV gives \$7,999,688.

Market value of equity: 1.2 million shares outstanding at \$10 = \$12,000,000.

Market value of debt	\$7,999,688	40%
Market value of equity	12,000,000	60%
Total capital	\$19,999,688	100%

To raise \$7.5 million of new capital while maintaining the same capital structure, the company would issue \$7.5 million × 40% = \$3.0 million in bonds, which results in a before-tax rate of 16%.

$r_d(1 - t) = 0.16(1 - 0.3) = 0.112$, or 11.2%.

$r_e = 0.03 + 2.2(0.10 - 0.03) = 0.184$, or 18.4%.

$\text{WACC} = 0.40(0.112) + 0.6(0.184) = 0.0448 + 0.1104 = 0.1552$, or 15.52%.

13 B is correct.

$r_e = 0.0425 + (1.3)(0.0482) = 0.1052$, or 10.52%.

14 B is correct.

$$\begin{aligned}\text{WACC} &= [(€900/€3300)(0.0925)(1 - 0.375)] + [(€2{,}400/€3{,}300)(0.1052)] \\ &= 0.0923, \text{ or } 9.23\%.\end{aligned}$$

15 A is correct.

Asset beta = Unlevered beta = $1.3/\{1 + [(1 - 0.375)(€900/€2{,}400)]\} = 1.053$.

16 C is correct.

Project beta = $1.053\{1 + [(1 - 0.375)(€80/€20)]\} = 1.053(3.5) = 3.686$.

17 C is correct.

$r_e = 0.0425 + 3.686(0.0482 + 0.0188) = 0.2895$, or 28.95%.

18 B is correct.

Asset beta: $\beta_{equity}/[1 + (1 - t)(D/E)]$

Relevant = $1.702/[1 + (0.77)(0)] = 1.702$.

ABJ = $2.8/[1 + (0.77)(0.003)] = 2.7918$.

Opus = $3.4/[1 + (0.77)(0.013)] = 3.3663$.

19 C is correct.

Weights are determined on the basis of relative market values:

Comparables	Market Value of Equity in Millions	Proportion of Total
Relevant	\$3,800	0.5490
ABJ	2,150	0.3106
Opus	972	0.1404
Total	\$6,922	1.0000

Weighted average beta = (0.5490)(1.702) + (0.3106)(2.7918) + (0.1404)(3.3572) = 2.27.

20 B is correct.

Asset beta = 2.27.

Levered beta = 2.27[1 + (1 − 0.23)(0.01)] = 2.2875.

Cost of equity capital = 0.0525 + (2.2875)(0.07) = 0.2126, or 21.26%.

21 C is correct.

For debt: FV = 2,400,000; PV = 2,156,000; n = 10; PMT = 150,000.

Solve for i: i = 0.07748. YTM = 15.5%.

Before-tax cost of debt = 15.5%.

Market value of equity = 1 million shares outstanding + 1 million newly issued shares = 2 million shares at \$8 = \$16 million.

Total market capitalization = \$2.156 million + \$16 million = \$18.156 million.

Levered beta = 2.27[1 + (1 − 0.23)(2.156/16)] = 2.27(1.1038) = 2.5055.

Cost of equity = 0.0525 + 2.5055(0.07) = 0.2279, or 22.79%.

Debt weight = \$2.156/\$18.156 = 0.1187.

Equity weight = \$16/\$18.156 = 0.8813.

TagOn's MCC = (0.1187)(0.155)(1 − 0.23) + (0.8813)(0.2279)
= 0.01417 + 0.20084
= 0.2150. or 21.50%.

22 A is correct. The relevant cost is the marginal cost of debt. The before-tax marginal cost of debt can be estimated by the yield to maturity of the company's expected new issue, which is 7%. After adjusting for tax, the after-tax cost is 7%(1 − 0.4) = 7%(0.6) = 4.2%.

23 C is correct. Inferring the asset beta for the public company: Unlevered beta = 1.75/[1 + (1 − 0.35)(0.90)] = 1.104. Re-levering to reflect the target debt ratio of the private firm: Levered beta = 1.104 × [1 + (1 − 0.30)(1.00)] = 1.877.

24 B is correct. All else equal, the first issue's greater liquidity would tend to make its required return lower than the second issue's. However, the required return on equity increases as leverage increases. The first issue's higher required return must result from its higher leverage, more than offsetting the effect of its greater liquidity, given that both issues have the same market risk.

25 B is correct. Since the project will be financed with 50% equity, the company will issue £25 million of new stock. The flotation cost of external equity is (0.058 × 25,000,000) = 1,450,000. The NPV of the project using external equity is the NPV using internal equity less the flotation cost. Adjusting the cost of capital to reflect the flotation cost is not a preferred way to account for flotation costs.

26 For JPMorgan Chase, the required return is

$$r = R_F + \beta\left[\mathrm{E}\left(R_M\right) - R_F\right] = 4.35\% + 1.50\left(8.04\%\right) = 4.35\% + 12.06\%$$
$$= 16.41\%.$$

For Boeing, the required return is

$$r = R_F + \beta\left[\mathrm{E}\left(R_M\right) - R_F\right] = 4.35\% + 0.80\left(8.04\%\right) = 4.35\% + 6.43\%$$
$$= 10.78\%.$$

READING

31

Capital Structure

by Raj Aggarwal, PhD, CFA, Glen D. Campbell, MBA,
Pamela Peterson Drake, PhD, CFA, Adam Kobor, PhD, CFA, and
Gregory Noronha, PhD, CFA

Raj Aggarwal, PhD, CFA, is at the Kent State University Foundation Board (USA). Glen D. Campbell, MBA (Canada). Pamela Peterson Drake, PhD, CFA, is at James Madison University (USA). Adam Kobor, PhD, CFA, is at New York University (USA). Gregory Noronha, PhD, CFA, is at the University of Washington, Tacoma (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe how a company's capital structure may change over its life cycle;
☐	**b.** explain the Modigliani–Miller propositions regarding capital structure;
☐	**c.** describe the use of target capital structure in estimating WACC, and calculate and interpret target capital structure weights;
☐	**d.** explain factors affecting capital structure decisions;
☐	**e.** describe competing stakeholder interests in capital structure decisions.

INTRODUCTION

1

Capital structure refers to the specific mix of debt and equity used to finance a company's assets and operations. From a corporate perspective, equity represents a more expensive, permanent source of capital with greater financial flexibility. Debt, on the other hand, represents a cheaper, finite-to-maturity capital source that legally obligates the company to fixed, promised cash outflows with the need to refinance at some future date at an unknown cost.

A company's capital structure is the result of such financing decisions that may be guided by capital structure policies or targets set by management and the board. Capital structure is also the result of such factors as company size and maturity, which influence the financing options a company may have available. Besides equity and debt issuance, capital structure can also be significantly affected by merger and acquisition (M&A) activity, which can be financed by cash, borrowing, share assumption and/or

© 2021 CFA Institute. All rights reserved.

debt assumption in addition to proceeds from divestitures and asset sales. Capital structure is also affected over time by the company's operations, which might consume or generate cash, and by management decisions regarding dividends and share buybacks.

Since we are considering how a company minimizes its overall cost of capital, the focus here is on the market values of debt and equity. Therefore, capital structure is also affected by changes in the market value of a company's securities over time, particularly the share price.

We tend to think of capital structure as the result of a conscious decision by management, but it is not that simple. For example, unmanageable debt, or financial distress, can arise because a company's capital structure policy was too aggressive, but it can also occur because operating results or prospects deteriorate unexpectedly.

This reading reviews some of the key factors affecting capital structure, including the following:

- **Company life cycle:** Companies typically evolve over time from cash consumers to cash generators, with decreasing business risk and increasing debt capacity.
- **Cost of capital:** From a theoretical perspective, company management seeks to maximize shareholder value and determines an optimal capital structure to minimize the company's **weighted average cost of capital (WACC)**. "Optimal capital structure" involves a trade-off between the benefits of higher leverage, which include the tax deductibility of interest and the lower cost of debt relative to equity, and the costs of higher leverage, which include higher risk for all capital providers and the potential costs of financial distress.
- **Financing considerations:** From a practical perspective, company management may consider several factors in capital structure decisions and the use of leverage.
- **Competing stakeholder interests:** In seeking to maximize shareholder value, company management may make capital structure decisions that are not in the interests of other stakeholders, such as debtholders, suppliers, customers, or employees.

2

CAPITAL STRUCTURE AND COMPANY LIFE CYCLE

a describe how a company's capital structure may change over its life cycle

Many factors influence a company's capital structure and its ability to support debt, including its maturity, capital intensity, the strength of its market position, and the stability and nature of its business.

This section will discuss the typical changes in a company's capital structure as it evolves from a start-up, to a growth business, to a mature business. It will also highlight key variations.

2.1 Background

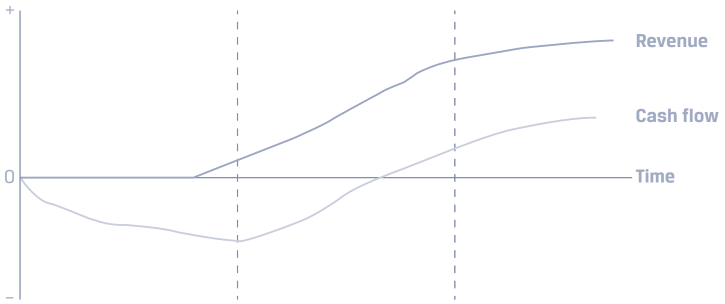
Many financial transactions can be seen in simple terms as the trading of money across time: A party with surplus cash today exchanges it with another party requiring cash in exchange for payments in the future. This is true of everyday consumer financial transactions, such as making a bank deposit, obtaining a loan, or contributing to a pension plan. It is also true of companies accessing the capital markets over their life cycle, from start-up through maturity. Typically, companies begin life as consumers of capital. As they mature, cash flows turn from negative to positive. Capital markets connect them to investors with the opposite requirement, those with cash to invest today in exchange for future payments.

The framework in Exhibit 1 describes the relationship between a company's life-cycle stage, its cash flow characteristics, and its ability to support debt. Life-cycle stage is a principal factor in determining capital structure. Capital not sourced through borrowing must come from equity (including retained earnings). The framework categorizes the stages in life-cycle development as start-up, growth, and mature and shows typical revenue and cash flow characteristics of companies in each stage. As companies mature, business risk typically declines, and their cash flow turns positive and becomes increasingly predictable, allowing for greater use of leverage at more attractive (less costly) financing terms. Debt then becomes a larger component of their capital structure.

Note that the framework references cash flow, which is net of investment. Profitable, high-growth businesses may have negative cash flow once investment is taken into account. Investment includes spending on plant equipment and other fixed assets as well as the expansion of working capital to support and grow the business over time.

Note also that the stages in a company's life cycle—start-up, growth, and mature—are like the evolutionary stages of an industry. However, start-ups and growth companies can often be found in mature industries. Examples include restaurants and apparel (with new concepts, formats, and fashions appearing regularly) and technology-driven disruption of established industries, such as advertising (by Google and Facebook) and even automobiles (by Tesla). This is a very general framework, and there are wide variations between and within sectors.

Exhibit 1 Capital Structure and Company Life Cycle



Stage in life cycle	Start-up	Growth	Mature
Financial management			
Revenue growth	Beginning	Rising	Slowing
Cash flow	Negative	Improving	Positive/Predictable
Business risk	High	Medium	Low
Debt capital/leverage			
Availability	Very limited	Limited/improving	High
Cost	High	Medium	Low
Typical cases	N/A	Secured (by receivables, fixed assets)	Unsecured (bank and public debt)
Typical % of capital structure[1]	Close to 0%	0%–20%	20% +

Note: These ratios are calculated based on the market values of equity and debt.

2.2 Start-Ups

Early in its life, a company is typically a cash consumer. Investment is required to advance concepts through the prototype stage and into commercial production. Revenues are zero or minimal, and risk of business failure is high. The company must raise capital, and since the timing and potential for cash flow generation are highly uncertain, it will generally raise equity rather than debt. Equity for start-ups is generally sourced privately, such as through venture capital rather than in public markets through an IPO, in part because many stock exchange listing requirements include minimum profitability levels, which most start-ups have yet to achieve.

At this early stage, debt capital is typically not available or available but very expensive. Most lenders require stable and positive cash flow to service debt and/or collateral to secure it. An early-stage company often has neither, making it a high-risk prospect to lenders.

From the perspective of an early-stage issuer, debt may be an attractive way to reduce or avoid the dilution associated with equity issuance. However, with cash flows that are negative and unpredictable, a typical start-up would have difficulty making regular debt payments, or "servicing the debt." Even if debt is available, the cost, inflexibility, and risk associated with borrowing are often unattractive to the start-up. As a result, debt is a negligible component in the capital structure of most start-up companies.

EXAMPLE 1

Start-Up Financing

Which of the following are limiting factors in the ability of a start-up company to take on debt?

A Lender requirement for positive cash flows

B Few to no assets that can be used as collateral

C Exchange listing requirements requiring minimum levels of profitability

D (A) and (B) only

E All of the above

Solution:

D is correct. Most lenders require positive and stable cash flows and/or collateral for servicing debt payments. Start-ups often have neither given their early stage in development. With expenditures outpacing revenues, a start-up company's cash flow is negative and unpredictable.

For equity financing, most start-ups fail to meet minimum profitability requirements for exchange listing in public markets and source their capital needs through private equity markets. Therefore, C is incorrect given an exchange listing relates to equity and is not a debt-financing limiting factor.

2.3 Growth Businesses

As a company exits the start-up stage, it is typically generating revenues, providing confirmation of the product concept and evidence of demand. Revenue growth may be rising and/or high, but investment is needed to achieve this growth and scale. This spending includes operating expenses, such as sales and marketing expense, growth related-capital expenditure, and working capital investment. Cash flow is still likely to be negative but may be improving and becoming more predictable. The company begins to establish a customer and supplier base.

As the business progresses through the growth stage, execution and competitive risks decline. Cash flow typically turns positive and then becomes more stable and predictable as the business grows. As a result, the business now becomes more attractive to lenders. Depending on the business model, there may also be assets that can be used to secure debt, such as receivables, inventory, or fixed assets. Both the availability and terms of debt financing improve for the company during this stage.

Many growth companies use debt conservatively in order to preserve operational and financial flexibility and minimize the risk of financial distress. Equity remains the predominant source of capital. Some business models are inherently "**capital light**," that is, they require little incremental investment in fixed assets or working capital to enable revenue growth. As a result, they have no need to borrow or otherwise raise capital to grow, even though they could easily support debt. Software businesses often fit this description.

EXAMPLE 2

Growth Company Financing

A growth retail business that is expected to generate positive operating cash flows in the next 3–5 years would normally be financed with:

A little or no debt.

B significant debt, to minimize equity dilution.

C significant debt, to minimize its weighted average cost of capital.

Solution:

A is correct. A growth retail business expected to generate positive operating cash flows in the next 3-5 years would normally be financed with little to no debt.

While one can often find exceptions, the standard approach to finance a business in such a highly competitive sector as retail with negative and/or unpredictable cash flows would be to rely primarily on equity. Using predominantly equity financing to meet capital needs allows management to preserve operational and financial flexibility while minimizing the risk of financial distress associated with debt.

2.4 Mature Businesses

At the maturity stage, a company's revenue growth may slow or even begin to decline. At this stage, however, a successful business generates reliable and positive cash flow and likely has an established customer and supplier base. There is typically a decline or a deceleration in growth-related investment spending. The company becomes able to support low-cost debt, often on an unsecured basis. From the company's perspective, debt financing is likely to be more attractive than higher-cost equity financing.

In practice, large, mature public companies commonly employ significant leverage, although many seek to maintain an investment-grade rating in order to preserve maximum financial flexibility. An investment-grade rating enables a company to access a very wide pool of potential investors on an unsecured basis without onerous debt covenants, and it minimizes the risk that financing might become unavailable.

Finance theory supports the idea that debt is an important component of the "optimal" capital structure, once the business can support it, due to the tax-deductibility of interest expense. "Optimal" capital structure, however, is a theoretical concept; in practice, there is wide variation in capital structures, even among otherwise similar companies. The prevalence of low interest rates increases the debt-carrying capacity of businesses and the use of corporate debt.

Mature businesses often de-leverage over time, experiencing a reduction in debt as a proportion of total capital. De-leveraging occurs due to continuing cash flow generation and because equity values commonly rise over time from share price appreciation. To offset this de-leveraging, companies may elect to do share buybacks. By using cash to buy back shares, outstanding share count, and therefore equity, is reduced in the capital structure.

Share buybacks are attractive to companies because, compared to dividends, they offer greater flexibility and, in general, are a more tax-efficient means of distributing cash to taxable investors. Buybacks can be done when cash is available and when the share price is seen by management to be attractive. Particularly in North America, investors generally expect dividend levels to be maintained once established. This is not the case with buybacks. Finally, by reducing the share count, buybacks enhance the share price, share option values, and per-share metrics, such as EPS or EPS growth, which may enhance management compensation.

EXAMPLE 3

Mature Business Capital Structure Considerations

Sleepy Mattress Company went public 10 years ago. Since then, the company has seen its revenues and cash flows increase every year. This growth is expected to continue. No dividends have ever been paid. All cash flow has gone toward debt reduction, and the company has recently become debt-free.

There is now a lively debate within the board about the company's capital structure. Board members agree that the company's capital structure should include debt in order to minimize its overall cost of capital, but they do not agree on how to achieve this. We consider each board member's recommendation, its associated scenario and likely impact on Sleepy's capital structure, and what is most likely the best course of action:

A Director TT has argued strongly that the company should initiate a regular annual cash dividend equal to 50% of after-tax cash flow.

B Director DD likes TT's idea but feels that the proposed cash dividend is too low. He recommends an annual stock dividend equal to 75% of after-tax cash flow.

C Director AH advocates making a tender offer to buy back shares at a 50% premium to market price for 50% of the company's shares and issuing debt to finance the offer. Since this would be a very large tender offer (equal to 75% of the company's equity market capitalization), it would be financed using a combination of senior and high-yield debt, carrying a relatively high 6% average interest rate. The company's cash flow will cover the projected interest payments, with operating cash flow equal to 1.1 times interest, and the surplus available for debt reduction. Projected debt/capital would be above 75%.

D Director FF favors a smaller share buyback: 20% of the company's shares at a 20% premium to market price, financed by issuing investment-grade debt with a 3% coupon. The company's cash flow will cover the projected interest payments, with operating cash flow equal to 6.9 times interest, and the surplus available for share repurchases. Projected debt/capital would be 20% to 25%.

Discussion:

Sleepy Mattress is a mature and stable business. It should therefore be able to support significant debt. The optimal capital structure for such a company would normally include debt, with the firm's value enhanced by the tax-deductibility of interest expense. But the company is now debt-free, generating cash flow, and paying no dividend. If no action is taken, cash will accumulate on the company's balance sheet. We look at each recommendation, in turn:

A TT's recommendation to pay a regular cash dividend would slow the cash accumulation on the balance sheet, since some of the company's operating cash flow will now be directed to the payment of a regular dividend. Starting to pay a dividend is probably a good idea for the company, but taking this action does not introduce any debt into the capital structure. The company would continue to have an all-equity capital structure.

B DD's recommendation to pay a stock dividend would also not introduce debt into the capital structure, and the company would continue to have an all-equity capital structure. Since paying a stock dividend does not require cash, cash would accumulate on the balance sheet as quickly as if no action were taken.

C AH's recommendation to buy back shares with debt would increase debt and reduce the equity in the capital structure. This recapitalization plan is a step in the right direction; however, the resulting interest coverage of 1.1 times is extremely low, so taking this action is likely too aggressive for the company. Any profit deterioration, and the company could easily fail to meet its interest obligations. While we are assuming that Sleepy Mattress could raise the debt, it could be challenged in doing this. At over 75%, pro forma debt as a percentage of total capital is very high. For companies with more stable and predictable cash flows, such as regulated utilities or REITS, higher percentages (i.e., above 60%) are common, but they are not common in consumer or industrial companies. Percentages above 75% often indicate that a company may be in financial distress.

D FF's recommendation is a superior course of action. It is a more balanced version of C, resulting in a manageable level of debt in the capital structure (under 25%), with a strong projected interest coverage ratio (6.9 times). It would be reasonable for the board to recommend this recapitalization to address the all-equity capital structure and, in addition, initiate a cash dividend. Doing a regular share buyback combined with a cash dividend would increase debt relative to equity in the capital structure and help counteract the de-leveraging in the business over time.

2.5 Unique Situations

In our discussion of the general relationship between company maturity and capital structure, there are important exceptions to note.

Capital Intensive Businesses with Marketable Assets Some businesses use high levels of leverage regardless of their development stage. For example, in real estate, utilities, shipping, airlines, and certain other highly capital-intensive businesses, the underlying assets can be bought and sold fairly easily, tend to retain their value regardless of who owns them, and can therefore support substantial debt secured by those assets. This would be true of a downtown office building, for example, but would not be true of a highly specialized factory in a remote location.

Note that business models have evolved in some cases to take advantage of this fact. What were once highly capital-intensive and vertically-integrated businesses (e.g., trucking, hotels, retailers, restaurants) have evolved so that many major brands are now owned by marketing or service businesses, which in turn have contractual relationships with the owners of real estate or other fixed assets used in the business. For example, Hilton Worldwide, one of the world's largest hotel companies, operates almost all its hotel rooms through long-term franchise or management agreements; the hotels themselves are owned by others.

Conversely, some relatively large and mature businesses use little debt.

Cyclical Industries In cyclical sectors, such as mining and materials, and many industrials, revenues and cash flows vary widely through the economic cycle, which limits debt capacity. As a result, businesses in cyclical sectors may have less debt in their capital structures than companies in other less cyclical industries.

"Capital-Light" Businesses Some business models, regardless of their development stage, notably software-based technology businesses, have minimal fixed investments or working capital needs. They tend to have little debt in their capital structures and in many cases have substantial net cash. This reflects several factors:

- With little or no fixed assets or capital expenditure required to support growth, these businesses are often cash flow positive from an early stage, never needing to raise large amounts of capital.
- Many companies in rapidly evolving industries see the need to accumulate cash for potential acquisitions.
- If they are rapidly growing and successful, companies may not face typical pressure to pay dividends or repurchase shares.
- A rapidly rising share price can cause the market value of the company's equity to dwarf the value of any debt that might have been raised.

EXAMPLE 4

Debt Use: Industry Variations

Green Company and Black Company are each achieving 15% annual revenue growth and have recently started to generate positive cash flow. Green Company owns and acquires renewable energy generation projects. Black Company is a cloud-based software company with a dominant market position, serving auto dealers. Which is more likely to have greater debt in its capital structure, and why?

A Black Company, because it serves a cyclical business

B Black Company, because of the strength of its market position

C Green Company, because its underlying assets can be financed with debt

Solution:

C is correct. Green Company has fixed assets, for which there is likely to be a ready and liquid market, and stable cash flows, which are supportive of debt financing. Black Company, a cloud-based software technology company, is an example of a capital "light" business, with few fixed assets. Its assets are more likely to be human capital. Additionally, servicing a cyclical industry is also likely to lead to Black Company having low debt.

MODIGLIANI–MILLER PROPOSITIONS

3

b explain the Modigliani–Miller propositions regarding capital structure

In a now-classic paper, Nobel Prize-winning economists Franco Modigliani and Merton Miller (1958) argued that given certain assumptions, a company's choice of capital structure does not affect its value. In short, managers cannot change firm value simply by changing the company's capital structure. In their work, Modigliani and Miller (MM) used simplifying assumptions to show the irrelevance of capital structure to firm value, and then they relaxed the assumptions to show the impact of taxes and financial distress costs on capital structure. Modigliani and Miller's work provides us with a starting point and allows us to examine what happens when the assumptions are relaxed to accommodate real-world considerations.

Their assumptions are shown in Exhibit 2.

Exhibit 2 Modigliani–Miller Assumptions

1. Homogenous expectations.
Investors agree on a given investment's expected cash flows.

2. Perfect capital markets.
No transaction costs, no taxes, no bankruptcy costs. Everyone has the same information.

3. Risk-free rate.
Investors can borrow and lend at the risk-free rate.

4. No agency costs.
Managers always act to maximize shareholder wealth.

5. Independent decisions.
Financing and investment decisions are independent of each other.

Let's begin by imagining a company's capital structure as a pie, with each slice of the pie representing a specific type of capital (e.g., common equity or debt) and the size of the pie representing the company's total value. We can slice the pie in any number of ways, yet the total size remains the same. This is equivalent to saying that the company's value, or present value of expected cash flows, remains the same no matter how the capital slices are allocated as long as the future cash flow stream is expected to remain the same and the risk of that cash flow stream, as reflected by the company's cost of capital, remains the same.

3.1 MM Proposition I without Taxes: Capital Structure Irrelevance

Given their assumptions, Modigliani and Miller proved that changing the capital structure does not affect firm value in part because individual investors can create any capital structure they prefer for the company by borrowing and lending in their own accounts in addition to holding shares in the company.

As an example, suppose that a company has a capital structure consisting of 50% debt and 50% equity and that an individual investor would prefer that the company's capital structure be 70% debt and 30% equity. That investor can borrow money to finance their share purchases so that their ownership of company assets reflects their preferred 70% debt financing. This would be equivalent to buying stock on margin. This action has no effect on the company's expected operating cash flows and no impact on company value.

Modigliani and Miller use the concept of arbitrage to demonstrate their point: If the value of an unlevered company (i.e., a company without any debt) is not equal to that of a levered company, investors could make a riskless arbitrage profit at no cost by selling shares of the overvalued company and using the proceeds to buy shares in the undervalued company, forcing their values to become equal.

MM Proposition I without taxes:
The market value of a company is not affected by the capital structure of the company.

In other words, the value of the levered company (V_L) is equal to the value of the unlevered company (V_U), or $V_L = V_U$. A crucial implication is that in the absence of taxes, the weighted average cost of capital is unaffected by capital structure. Note, however, that this proposition holds only under the restrictive assumptions of Modigliani and Miller.

The importance of the Modigliani and Miller result is that it demonstrates that managers cannot create firm value simply by changing the company's capital structure. In other words, the value of a company is determined solely by its cash flows, not by its relative reliance on debt and equity capital.

Adding leverage increases the risk to equityholders because greater debt increases the probability of bankruptcy. As a result, equityholders will demand a higher return as leverage increases in order to offset the increase in risk. However, overall cost of capital does not change, so there is no change in the value of the company. Note that Modigliani and Miller did not assume away the possibility of bankruptcy. They simply assumed there is no cost to bankruptcy. Given their assumptions of **perfect capital markets**—in which there are no taxes, transaction costs, or bankruptcy costs, and in which all investors have equal ("symmetric") information—capital structure does not affect the value of the company and is irrelevant. This forms the basis of the Modigliani and Miller Proposition I—and II, which follows.

3.2 MM Proposition II without Taxes: Higher Financial Leverage Raises the Cost of Equity

According to MM Proposition I, the value of the company remains unchanged or constant as the capital structure changes. Therefore, the weighted average cost of capital must also remain constant as capital structure changes. Since the cost of debt is lower than the cost of equity because debtholders have a priority in claims, it holds that as the company uses more debt, its risk goes up and the cost of equity must increase. MM Proposition II holds that the increase in the cost of equity must exactly offset the greater use of lower cost debt.

MM Proposition II without taxes:
The cost of equity is a linear function of the company's debt-to-equity ratio.

In their analysis, Modigliani and Miller prove that in the absence of corporate taxes, the cost of equity is equal to

$$r_e = r_0 + (r_0 - r_d)\frac{D}{E}, \tag{1}$$

where r_e is the cost of equity, r_o is the cost of capital for a company financed only with equity (i.e., an all-equity company), r_d is the cost of debt, D is market value of debt, and E is the market value of equity. Equation 1 is a linear function with intercept equal to r_0 and slope equal to the quantity $(r_0 - r_d)$. In short, higher leverage raises the cost of equity but does not change firm value.

Consider the example of the Leverkin Company, which currently has an all-equity capital structure. Leverkin has expected annual cash flows to equityholders (which we denote as "CF_e") of \$5,000 and a cost of equity, which is also its WACC, of 10%. For simplicity, we assume that all cash flows are perpetual. Therefore, Leverkin's value is equal to

$$V = \frac{\text{CF}_e}{r_{wacc}} = \frac{\$5{,}000}{0.10} = \$50{,}000$$

Now suppose that Leverkin plans to issue \$15,000 in debt at a cost of 5% and use the proceeds to buy back \$15,000 worth of its equity.

Under MM Proposition I, $V_L = V_U$, so the value of Leverkin must remain the same at \$50,000. Under MM Proposition II, the cost of Leverkin's equity when it has \$15,000 debt and \$35,000 equity is

$$r_e = 0.10 + (0.10 - 0.05)\frac{\$15{,}000}{\$35{,}000} \approx 0.12143 = 12.143\%$$

Furthermore, the value of Leverkin must equal the sum of the present value of cash flows to debtholders and equityholders. With \$15,000 debt at a cost of 5%, Leverkin makes annual interest payments of \$750 to debtholders, leaving \$5,000 – \$750 = \$4,250 for equityholders. Therefore, the total value of the company is:

$$V = D + E = \frac{\$750}{0.05} + \frac{\$4{,}250}{0.12143} = \$15{,}000 + \$34{,}999.59 \approx \$50{,}000$$

The more debt the company uses, the greater the cost of equity, but the total value of the company does not change and neither does the weighted average cost of capital. With its new capital structure, the company's WACC remains at 10%:

$$r_{wacc} = \left(\frac{\$15{,}000}{\$50{,}000}\right)0.05 + \left(\frac{\$35{,}000}{\$50{,}000}\right)0.12143 = 0.10 = 10\%$$

Now let's consider a company's systematic risk. Portfolio managers know that the beta of an investment portfolio is a market-value-weighted average of the betas of the investments in that portfolio. Similarly, the beta (or systematic risk) of a company's assets is a weighted average of the systematic risks of its sources of capital. This concept is explained by the Hamada equation, which does not account for default risk (see Hamada 1972):

$$\beta_a = \left(\frac{D}{V}\right)\beta_d + \left(\frac{E}{V}\right)\beta_e \tag{2}$$

where β_a is the asset's systematic risk, or **asset beta**, β_d is the beta of debt, and β_e is the equity beta. According to Modigliani and Miller, the company's cost of capital does not depend on its capital structure but rather on its business risk. As the level of debt rises, however, the risk of the company defaulting on its debt increases. This risk is borne by the equityholders. So, as the proportionate use of debt rises, the equity's beta, β_e, also rises. By reordering the terms of Equation 2, we get

$$\beta_e = \beta_a + (\beta_a - \beta_d)\left(\frac{D}{E}\right) \tag{3}$$

We have learned that under Modigliani and Miller's restrictive assumptions, when there are no taxes, leverage does not affect the value of the company or the weighted average cost of capital. However, leverage does cause the risk of equity to increase and therefore the cost of equity to increase as well.

EXAMPLE 5

Value of Firm in Perfect Capital Markets

Company A has 25% debt and 75% equity in its capital structure. Management decides to increase leverage, so it issues more debt and buys back company stock. As a result, the new capital structure is 50% debt and 50% equity. Which of the following statements is true in perfect capital markets?

A Post refinancing, the company is worth more because leverage has increased.

B Post refinancing, the company is worth less because there is a greater chance of bankruptcy.

Solution:

Neither answer is correct. In perfect capital markets, a change in capital structure has no impact on the value of the company.

EXAMPLE 6

Effect of Leverage on Equity Beta

The Chang Shou Noodle Company is financed with 10% debt and 90% equity.

1. If the asset beta is 0.8 and the beta of the debt is 0.2, what is the equity beta?
2. How does the equity beta change if management increases leverage to 30% debt?

Solution:

We use Equation 3 to solve this problem.

1 Prior to the change in capital structure, the equity beta is

$$\beta_e = 0.8 + (0.8 - 0.2)(10/90) = 0.87.$$

2 After the capital structure change, the equity beta is higher because more debt increases the risk to equityholders. The new equity beta is

$$\beta_e = 0.8 + (0.8 - 0.2)(30/70) = 1.06.$$

3.3 MM Propositions with Taxes: Taxes, Cost of Capital, and Value of the Company

Now let's explore what happens when we take a more realistic assumption, that of corporate taxes.

In most countries, interest paid on debt is deductible from income for tax purposes. In other words, debt provides a tax shield for companies that are earning profits. The money saved in taxes enhances the value of the company. Ignoring other realities, such as the costs of financial distress and bankruptcy, the value of the company increases with increasing levels of debt. The actual cost of debt is reduced by the amount of the tax benefit:

After-tax cost of debt = Before-tax cost of debt × (1 − Marginal tax rate).

Modigliani and Miller show that in the presence of corporate taxes (but not personal taxes), the value of the levered company is greater than that of the all-equity company by an amount equal to the tax rate multiplied by the value of the debt, also termed the **debt tax shield**.

> *MM Proposition I with corporate taxes:*
> *The market value of a levered company is equal to the value of an unlevered company plus the value of the debt tax shield.*

$$V_L = V_U + tD, \tag{4}$$

where t is the marginal tax rate and tD is the debt tax shield. When there are corporate taxes, a profitable company can increase its value by using debt financing.

If the value of the company (V) increases as it uses more debt, the weighted average cost of capital must decrease as it uses more debt:

$$r_{wacc} = \left(\frac{D}{V}\right) r_d (1 - t) + \left(\frac{E}{V}\right) r_e \tag{5}$$

If r_0 represents the cost of equity for an all-equity company, Modigliani and Miller also show that the cost of equity for the same company with debt is:

$$r_e = r_0 + (r_0 - r_d)(1 - t)\frac{D}{E} \tag{6}$$

which leads us to MM Proposition II with corporate taxes:

> *MM Proposition II with corporate taxes:*
> *The cost of equity is a linear function of the company's debt-to-equity ratio with an adjustment for the tax rate.*

Notice that the only difference between Equation 6 and Equation 1 (MM Proposition II with no taxes) is the presence of the term (1 – *t*). When *t* is zero, the two equations are identical. When *t* is not zero, the term (1 – *t*) is less than 1 and serves to reduce the cost of levered equity. The cost of equity rises as the company increases the amount of debt in its capital structure, but it rises at a slower rate than in the no-tax case. As debt increases, the WACC decreases and the company's value increases. This implies that when there are taxes (and no costs to financial distress or bankruptcy), debt financing is highly advantageous. Taken to an extreme, a company's optimal capital structure is to have all debt—a conclusion that is at odds with reality and a direct result of Modigliani and Miller's restrictive assumptions.

Let us return to the example of the Leverkin Company. Recall that annual cash flows to equityholders are \$5,000 and the WACC is 10%. As before, Leverkin is planning to issue \$15,000 of 5% debt in order to buy back an equivalent amount of equity. Now, however, assume that Leverkin pays corporate taxes at a rate of 25%.

Since the company does not currently have debt, the after-tax cash flows are \$5,000(1 – 0.25) or \$3,750. Because the cash flows are assumed to be perpetual, the value of the company at 10% is \$37,500, considerably less than it was when there were no taxes. Now suppose Leverkin issues \$15,000 of debt and uses the proceeds to repurchase common stock. According to MM Proposition I with corporate taxes (i.e., Equation 4), the value of the company is:

$$V_L = V_U + tD = \$37{,}500 + 0.25(\$15{,}000) = \$41{,}250.$$

Since the value of the debt is \$15,000, the value of the equity (after the buyback) must be (\$41,250 – \$15,000) = \$26,250. According to MM Proposition II with corporate taxes (Equation 6), the cost of the levered equity is:

$$r_e = 0.10 + (0.10 - 0.05)(1 - 0.25)\frac{\$15{,}000}{\$26{,}250} = 0.12143 = 12.143\%$$

Note that the value of the company must also equal the present value of cash flows to debt and to equity:

$$V_L = D + E = \frac{r_d D}{r_d} + \frac{(\text{CF}_e - r_d D)(1 - t)}{r_e}$$
$$= \frac{\$750}{0.05} + \frac{(\$5{,}000 - \$750)(1 - 0.25)}{0.12143} \approx \$41{,}250$$

which is the same result that we got from MM Proposition I with corporate taxes. As a further check, using Equation 5, the WACC for the levered Leverkin is:

$$r_{wacc} = \frac{\$15{,}000}{\$41{,}250}(0.05)(1 - 0.25) + \frac{\$26{,}250}{\$41{,}250}(0.12143)$$
$$= 0.09091 = 9.091\%$$

As expected, this is lower than the unlevered WACC of 10%. Furthermore, because debtholders and equityholders are entitled to whatever is left of the cash flows after taxes are paid, we must also have:

$$V_L = \frac{\text{CF}_e(1 - t)}{\text{WACC}} = \frac{\$5{,}000(1 - 0.25)}{0.09091} \approx \$41{,}250$$

Exhibit 3 summarizes the results:

Exhibit 3

Modigliani and Miller Propositions

	Without Taxes	With Taxes
Proposition I	$V_L = V_U$	$V_L = V_U + tD$
Proposition II	$r_e = r_0 + (r_0 - r_d)\frac{D}{E}$	$r_e = r_0 + (r_0 - r_d)(1 - t)\frac{D}{E}$

In an attempt to explain why companies are not financed entirely with debt as the Modigliani and Miller analysis with corporate taxes prescribes, Miller (1977) introduced another aspect into the analysis: personal taxes. Miller argued that the existence of a higher personal tax rate on interest income than on dividend income reduces the advantage of debt financing to the corporation. A higher tax rate on interest income causes investors to demand a higher return on debt than they otherwise would, which drives up the cost of debt financing to the company. In the Miller model, whether or not debt financing adds value to the company depends not only on the corporate tax rate but also on the personal tax rates on interest and dividend income. Depending on the levels of the various tax rates, it is possible for debt to add value, lower value, or have no effect on value. The existence of personal taxes might explain, at least in part, why companies are not financed entirely with debt as MM Proposition I with corporate taxes implies. The existence of investors with differing personal tax rates could explain why corporations often have very different capital structures, even when they have a number of other similarities.

Of course, in the real world, taxes are not the only factor that affect the value of a levered company. The analysis gets more complicated when we allow for such things as the costs of financial distress and other real-world considerations. We consider costs of financial distress next.

3.4 Costs of Financial Distress

Operating and financial leverage can magnify profits and losses. Losses can put companies into **financial distress**, the costs of which can be explicit or implicit. Financial distress refers to the heightened uncertainty regarding a company's ability to meet its various obligations because of diminished earnings power or actual current losses. Even before filing for bankruptcy, companies under financial distress may lose customers, creditors, suppliers, and valuable employees.

EXAMPLE 7

Costs of Financial Distress

The Carillion Plc was a UK-based entity with revenues of almost £4.4 billion. It had a number of contracts with the UK government for the maintenance of roads and for catering and cleaning at hundreds of schools. The share price was £308 at the end of 2015 and £236 at the end of 2016.

In mid-2017, Carillion warned that profits would fall short of expectations due to problems collecting on several construction contracts. Soon after, the share price fell by 30%. Even as it faced collection problems, the company was awarded more contracts, including those for railway and military projects.

Subsequent profit warnings sent the share price further downward. The second profit warning of that year was followed by new credit facilities and deferrals of debt repayments. The third profit warning, which included concerns over violating debt covenants, caused the share price to fall to around £21.

Discussions with creditors in December 2017 and January 2018 did not result in an agreement. Carillion was put in compulsory liquidation on 15 January 2018, and share trading was suspended.

The primary causes of Carillion's problems were its many overreaching, unprofitable projects and its inability to collect payments quickly.

The costs of financial distress included:

- the loss of all shareholder value;
- £44.2 million paid to an accounting firm to manage the insolvency;
- costs to the UK government as it grappled with nationalizing contracts; and
- creditors getting paid only a fraction of what they were owed.

The direct costs of financial distress include actual cash expenses associated with the bankruptcy process, such as legal and administrative fees. Indirect costs include forgone investment opportunities, reputational risk, impaired ability to conduct business, and costs arising from conflicts of interest between managers and debtholders, known as **agency costs of debt**, during periods in which the company is near or in bankruptcy.

The costs associated with financial distress are lower for companies whose assets have a ready secondary market. Airlines, shipping companies, and steel manufacturers typically have tangible assets that can be easily sold. High-tech growth companies, pharmaceutical companies, information technology companies, and companies in the services industry typically have fewer tangible assets that can be sold. These companies have higher costs associated with financial distress.

The probability of financial distress and bankruptcy increases when there is a greater amount of debt in the capital structure, a greater amount of business risk, and fewer reserves available to delay bankruptcy. Other factors that affect the probability of bankruptcy include the company's corporate governance structure and the quality of the management team.

4

OPTIMAL AND TARGET CAPITAL STRUCTURE

c describe the use of target capital structure in estimating WACC, and calculate and interpret target capital structure weights

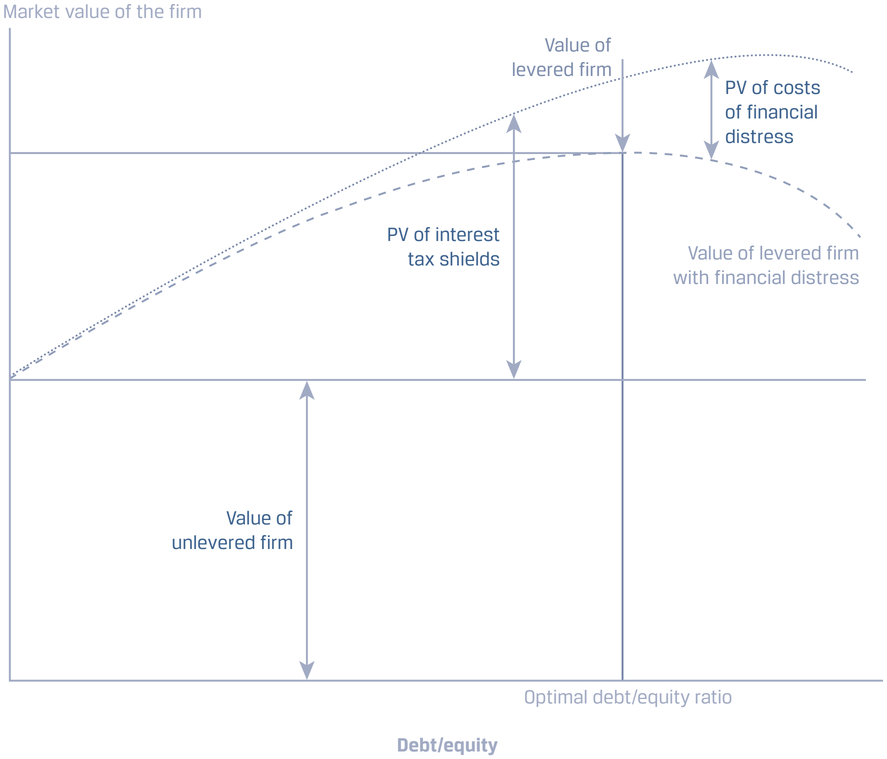
The value-enhancing effects of leverage from the tax deductibility of interest must be weighed against the value-reducing impact of the net present value of the expected (or probability weighted) costs of financial distress or bankruptcy, debt agency costs, debt structuring and issuance fees, etc. Putting the pieces together, we see that the value of the company increases as it uses more debt only up to a point. Beyond that point, the value-enhancing effects of financial leverage are offset by the value-reducing effects. The theoretical point at which the value of the company is maximized is known as the **optimal capital structure**.

Here we consider only the tax shield provided by debt and the costs of financial distress. As a result, we can write the value of a leveraged company as

$$V_L = V_U + tD - PV(\text{Costs of financial distress}), \tag{7}$$

where the last term is the present value of the costs of financial distress. Equation 7 represents the **static trade-off theory of capital structure**, which results in an optimal capital structure that is somewhere less than 100% debt. We diagram this optimum in Exhibit 4.

Exhibit 4 Trade-Off Theory with Taxes and Costs of Financial Distress. Firm Value and the Debt-to-Equity Ratio



In practice, we cannot say precisely where that point is. However, by considering the company's business risk, tax situation, corporate governance, financial accounting information, probable bankruptcy costs, and other factors, management should be able to identify a range that includes the optimal point.

When a company recognizes its optimal capital structure, it may adopt it as its **target capital structure**. This target is the company's chosen proportions of debt and equity as opposed to the actual weightings, which may be different for a number of reasons. For example, management may be able to exploit short-term opportunities in a financing source. A company issuing high-grade debt at attractive rates, for instance, may choose to increase deal size in response to strong investor demand. Or, market-value fluctuations can cause the company's actual capital structure to deviate from the target. In addition, flotation costs may make it impractical for a company to continually adjust its capital structure to keep it on target. Flotation costs are the costs incurred by a publicly traded company when it issues new debt or equity securities. In any case, while optimal capital structure may exist as a specific point in theory, it cannot be determined exactly in practice.

In practice, the optimal and target capital structures should be thought of as being within a range. For example, it would be impossible to determine with certainty that the optimal amount of debt in the capital structure should be exactly 40%. Yet, the company should be able to determine that its target capital structure has between 30% to 50% debt.

When conducting an analysis, how do we determine what WACC weights to use? Ideally, we want to use the proportion of each source of capital that the company would use in the project or company. If we assume that a company has a target capital structure and raises capital consistent with this target, we should use this target capital structure, or the capital structure that a company is striving to obtain. If we know the company's target capital structure, then of course we should use this in our analysis. Individuals outside the company, however, such as analysts, typically do not know the target capital structure and must estimate it using one of several methods:

1. Assume the company's current capital structure, at market value weights for the components, represents the company's target capital structure.
2. Examine trends in the company's capital structure or statements by management regarding capital structure policy to infer the target capital structure.
3. Use averages of comparable companies' capital structures as the target capital structure.

In the absence of knowledge of a company's target capital structure, we may take Method 1 as the baseline. Note that in applying Method 3, we use an unweighted, arithmetic average, as is often done for simplicity. An alternative is to calculate a weighted average, which would give more weight to larger companies.

Suppose we are using the company's current capital structure as a proxy for the target capital structure. In this case, we use the market value of the different capital sources in the calculation of these proportions. For example, suppose a company has the following market values for its capital:

Bonds outstanding	\$5 million
Preferred stock	1 million
Common stock	14 million
Total capital	\$20 million

The weights that we apply would be as follows:

w_d =	0.25
w_p =	0.05
w_e =	0.70

Example 8 illustrates the estimation of weights. Note that a simple way of transforming a debt-to-equity ratio (D/E) into a weight—that is, $D/(D + E)$—is to divide D/E by $1 + D/E$.

EXAMPLE 8

Estimating the Proportions of Capital

Fin Anziell is a financial analyst with Analytiker Firma. Anziell is in the process of estimating the cost of capital of Gewicht GmbH. The following information is provided:

Market value of debt: €50 million

Market value of equity: €60 million

Primary competitors and their capital structures (in millions):

Competitor	Market Value of Debt	Market Value of Equity
A	€25	€50
B	€101	€190
C	£40	£60

What are Gewicht's proportions of debt and equity that Anziell would use if estimating these proportions using the company's:

1. current capital structure?
2. competitors' capital structure?
3. Suppose Gewicht announces that a debt-to-equity ratio of 0.7 reflects its target capital structure. What weights should Anziell use in the cost of capital calculations?

Solution to 1:

Current capital structure

$$w_d = \frac{\text{€}50\text{ million}}{\text{€}50\text{ million} + \text{€}60\text{ million}} = 0.4545$$

$$w_e = \frac{\text{€}60\text{ million}}{\text{€}50\text{ million} + \text{€}60\text{ million}} = 0.5454$$

Solution to 2:

Competitors' capital structure:

$$w_d = \frac{\left(\frac{\text{€}25}{\text{€}25 + \text{€}50}\right) + \left(\frac{\text{€}101}{\text{€}101 + \text{€}190}\right) + \left(\frac{\text{£}40}{\text{£}40 + \text{£}60}\right)}{3} = 0.3601$$

$$w_e = \frac{\left(\frac{\text{€}50}{\text{€}25 + \text{€}50}\right) + \left(\frac{\text{€}190}{\text{€}101 + \text{€}190}\right) + \left(\frac{\text{£}60}{\text{£}40 + \text{£}60}\right)}{3} = 0.6399$$

Solution to 3:

A debt-to-equity ratio of 0.7 represents a weight on debt of 0.7/1.7 = 0.4118, so that $w_d = 0.4118$ and $w_e = 1 - 0.4118 = 0.5882$. These would be the preferred weights to use in a cost of capital calculation.

FACTORS AFFECTING CAPITAL STRUCTURE DECISIONS

5

d explain factors affecting capital structure decisions

In theory, capital structure is an optimal mix of equity and debt, which minimizes the firm's weighted average cost of capital and represents a trade-off between the cost-effectiveness of borrowing relative to higher cost equity and the costs of financial distress. Many practical considerations affect capital structure and the use of leverage by management. These include:

- *capital structure policies and targets*: guidelines set by management and the board that seek to establish sensible borrowing limits for the company based on the company's risk appetite and ability to support debt;
- *capital investment financing:* financing opportunities or requirements associated with acquisitions or other major investments;
- *market conditions:* current share price levels and market interest rates for a company's debt; and
- *asymmetric information:* the unequal information distribution that exists between management and other stakeholders.

We look at each of these in turn.

5.1 Capital Structure Policies and Target Capital Structures

In principle, a company's capital structure policy guides its financing and distribution decisions, with a target capital structure that achieves an optimal mix of equity and debt.

In practice, it is difficult to estimate a company's cost of equity capital or to estimate the costs of financial distress. As a result, it is common for capital structure policies to be debt-oriented, defining reasonable limits for borrowing by the company. Policy limits are often defined in straightforward terms, such as "debt/equity less than 0.5 times"; "debt to operating cash flow less than 2.0 times"; or "debt as a maximum of X% of total capital." Companies differ in the measures and thresholds used, but they frequently start with debt covenants or rating agency thresholds to which they add a cushion, or "margin of safety," to ensure that the debt limits or covenants are not breached or to avoid a rating downgrade.

Not all companies have capital structure policies. Having explicit capital structure policies is important for companies that are significant and frequent borrowers since the risks associated with too much leverage can be severe, while the risks associated with too little leverage generally are not. Companies that have little to no debt, such as start-ups or some technology companies, may not have capital structure policies or targets. In certain industries, capital structures are evaluated by regulators and thus capital structure policies will be determined accordingly.

5.1.1 *Debt Ratings*

Debt ratings are independent, third-party measures of the quality and safety of a company's debt based upon an analysis of the company's ability to pay the promised cash flows. Debt ratings are an important consideration in the practical management of leverage, and maintaining the company's rating at a certain level may also be an explicit policy target for management.

As leverage rises, rating agencies tend to lower the ratings of the company's debt to reflect the higher credit risk and probability of default arising from the increase in leverage. Lower ratings signify higher risk to both equity and debt capital providers, who demand higher returns for supplying capital to the company.

Most large companies pay one or more rating services to provide debt ratings for their bonds. Debt issues are rated for creditworthiness by these credit rating agencies after they perform a financial analysis of the company's ability to pay the promised cash flows as well as an analysis of the bond's indenture (i.e., the set of complex legal documents associated with the issuance of debt instruments).

These agencies evaluate a wealth of information about the issuer and the bond, including the bond's characteristics and indenture agreement, and provide investors with an assessment of the company's ability to pay the interest and principal on the bond as promised.

In practice, most managers consider the company's debt rating in their capital structure policies because the cost of capital is tied so closely to bond ratings. And some companies explicitly target a certain debt rating in their capital structure policy. For example, a company might target an S&P debt rating of A or higher, which is one level above BBB, the minimum investment-grade rating on S&P's scale. The cost of debt increases significantly when a bond's rating drops from investment grade to speculative grade (also called high-yield debt). In economic recessions, these spreads may widen significantly, and borrowing can become more difficult for companies with non-investment-grade ratings.

In addition to lower borrowing costs, highly rated companies have greater borrowing flexibility and better terms. They are able to borrow in debt markets for "general corporate purposes" on an unsecured basis, with very few debt covenants. Similarly, lower-rated, sub-investment-grade companies have higher borrowing costs with more restrictive debt covenants, liens on assets, and differing subordination levels.

In general, firms that are in industries with stable cash flows, greater profitability, shorter asset conversion cycles and/or more liquid assets can support more debt and tend to receive higher debt ratings and have lower borrowing costs. Such companies are often in stable and defensive sectors, such as utilities and consumer staples, and enjoy strong market positions.

On the other hand, companies that require heavy investment in inventory or specialized equipment—with long asset conversion cycles, cyclical or less stable cash flows and/or profitability—have a reduced ability to consistently meet fixed debt payments over time and therefore face lower ratings and higher borrowing costs.

5.1.2 *Market Value vs. Book Value*

While optimal capital structure is calculated using the market value of equity and debt, company capital structure targets often use book value instead for the following reasons:

1. *Market values can fluctuate substantially and seldom impact the appropriate level of borrowing.* On the contrary, a company that has seen a rapid increase in its share price might opportunistically decide to raise equity capital rather than debt to maintain a certain debt-to-equity ratio.
2. *For management, the primary concern is the amount and types of capital invested by the company, not in the company.* This perspective includes how future investment will be funded. It is different from that of a shareholder who has purchased shares at the prevailing market price, for example, and is concerned with generating a return on that investment.
3. *Capital structure policy ensures management's ability to borrow easily and at low cost.* Since lenders, debt investors, and rating agencies generally focus on the book value of debt and equity for their calculation measures, companies and their management will take this into account in their capital structure policies.

EXAMPLE 9

Capital Structure Policy: Abha Software

Abha Software's management believes the company can obtain attractive borrowing terms if debt/EBITDA (earnings before interest, taxes, depreciation, and amortization) is at 2 times or less, which for Abha is the primary criterion used by lenders to assess borrowing capacity and terms. Abha's EBITDA for the latest year is $50 million, and it has a total capitalization of $1.0 billion. Management and the board have agreed that Abha should use leverage if the company can borrow easily at attractive rates. They are seeking a capital structure policy that reflects this objective. Which of the following is *most likely* to be an appropriate capital structure policy for Abha?

A Debt should be a maximum of 25% of total capital in book value terms.

B Debt should be a maximum of 10% of total capital in market value terms.

C The debt-to-EBITDA ratio should be no more than 2 times.

Solution:

C is correct. It is the most appropriate capital structure policy since the debt-to-EBITDA ratio is the primary criterion that lenders and management are using to determine Abha's debt capacity. Note that lending criteria vary by sector, by market, and over time, including the ratios used (as well as the threshold ratios for those ratios).

B is least likely to be appropriate. While useful in calculating the cost of capital for Abha, market value weights are less practical and less commonly used in capital structure targets. Market values of equity can fluctuate a great deal without a corresponding change in the company's borrowing capacity. Increases or decreases in the company's share price would imply increases or decreases in maximum debt that are unlikely to be appropriate.

A is a better policy than B, with 10% debt/capital implying up to $100 million in debt. This ratio is unaffected by fluctuations in the share price. While the implied debt capacity of $100 million happens to equal the amount indicated by the 2.0 times debt/EBITDA that lenders seek, this policy is not the primary criterion used by lenders.

Finally, a company's cash flow generation is almost never entirely predictable and is likely to deviate from management's plan. This will affect the company's borrowing capacity and its financing plans. As management considers their financing plans, they will continually refine their views on the company's cash flow generation as well as its investment needs.

5.2 Financing Capital Investments

In practice, key financing decisions are typically tied to investment spending—and often to a specific investment or acquisition. Management will consider the specific attributes and requirements of the asset or investment at the time of financing in addition to overall capital structure of the business after the investment has been made.

The nature of the investments may be well or poorly suited to leverage, relative to the entire business. Assets suitable to leverage include those that are generally easy to market (i.e., liquid), cash-generative, and typically considered strong collateral by lenders, such as real estate or infrastructure assets. Start-up businesses—those with minimal tangible assets and that are inextricably tied to the parent's operations (for example, a sales/service operation)—are less attractive to lenders for stand-alone financing.

In matching the cash flows and maturity structures of their assets and liabilities, companies reduce risk. A company financing long-term assets with short-term obligations faces rollover risk, which may threaten profitability if short-term financing costs go up over the financing period. Similarly, a company financing short-term assets with long-term financing beyond the term needed faces the risk that the company overpays in financing cost. Asset liability misalignment increases the risk of default and cost of capital for companies as capital suppliers demand higher returns in compensation. In contrast, companies with greater asset liability alignment have lower default risk, higher ratings, and lower borrowing costs.

Some investments may have financing in place. An example would be the acquisition of an established business, with either existing debt or financing provided by the vendor.

It is common for foreign investments to be financed substantially with debt, in order to hedge currency risk by borrowing in the same (foreign) currency as revenues earned or, if tax rates are higher in the foreign jurisdiction, to reduce taxes. More specifically, borrowing in the same (foreign) currency helps to hedge against exchange rate movements on earnings from the foreign investment and on the value of the investment in local currency terms. Higher tax rates enhance the value of the interest tax deduction on new debt.

5.3 Market Conditions

Although it is common to have target capital structures expressed in terms of book values, managers pay close attention to the price of their company's shares and market interest rates for their debt in deciding when, how much, and what type of capital to raise. For this reason, financings are often opportunistic to some degree—and an important reason why capital structures often deviate from target levels defined by management.

An important consideration for corporate issuers is whether their debt or equity issuance meets index provider requirements that allow inclusion in a benchmark index, since this affects the demand for their securities by major investors. Companies reaching a certain market capitalization size in addition to meeting specific index provider requirements may qualify for index membership. Similarly, whether a specific debt issue meets requirements—including issuance size, in particular—for addition to an index is a key factor in debt-financing decisions.

A company's cost of debt is equal to a risk-free rate plus a risk premium specific to the company. The risk premium, or credit spread, above the benchmark risk-free rate reflects issuer-specific risk factors as well as overall credit market conditions. More creditworthy companies—with strong and predictable cash flows, lower risk, and higher debt ratings—have lower costs of debt. Similarly, less profitable companies—with higher cash flow volatility and/or predictability, higher risk, and lower debt ratings—have higher costs of debt.

When a company can issue new debt at a known rate, it is common practice to use that rate as the company's cost of debt. However, if leverage is increasing for the company, the cost of refinancing existing debt in the future is also likely to increase and thus should be included as part of the incremental cost of borrowing. At high levels of leverage, a company's incremental borrowing costs can climb quickly.

Macroeconomic and country-specific factors are reflected in the benchmark rate and in the overall level of credit spreads. They impact the ability to raise debt on attractive terms. Borrowing is less costly when interest rates and inflation rates are low as well as during expansionary stages in the business and credit cycle. Countries with more developed banking systems and capital markets, stronger currencies, and lender-friendly legal systems tend to have lower borrowing costs, supporting higher use of leverage by companies. Companies located in emerging markets often employ lower levels of leverage than those in developed countries.

For a typical company, debt is less expensive than equity, particularly when the tax-deductibility of interest is considered and the company is profitable. Corporate bond spreads over benchmark rates, such as US Treasuries, vary over time and by credit rating, but most investment-grade bonds have historically traded at 0.5% to 2.5% above benchmark bond rates.

In comparison, equity risk premium estimates vary widely, between 3% and 8% depending on the market and period studied. Note that such estimates reflect the returns earned by equity investors, whereas in principle, the equity risk premium refers to the return required by investors on a forward-looking basis. When a company has both debt and equity options, to maximize firm value management should choose the debt–equity combination that minimizes the weighted average cost of capital (WACC).

We next consider information asymmetries and signaling.

5.4 Information Asymmetries and Signaling

Asymmetric information (an unequal distribution of information) arises from the fact that managers have more information about a company's performance and prospects (including future investment opportunities) than do outsiders, such as owners and creditors. Whereas all companies have a certain level of asymmetric information, companies with comparatively high asymmetry in information include those with complex products. These include high-tech companies, companies with little transparency in financial accounting information, or companies with lower levels of institutional ownership. Providers of both debt and equity capital demand higher returns from companies with higher asymmetry in information because there is greater potential for conflicts of interest.

However, investors understand that management has access to more information about a business and its prospects than they do—and that management will take this information into account when raising capital. Investors will be concerned either that a company will issue shares when its shares are expensive or that it will issue debt when its creditworthiness is about to deteriorate, for example due to higher investment or acquisitions.

To address this issue, a common response by the company is to tie financing to the event that requires it or that might cause investor concern, such as an acquisition, investment program, debt retirement, or results announcement. This approach aims to reduce the extent of information asymmetry between managers and investors. In addition, it ties the timing to a specific requirement for capital rather than perceived management opportunism.

Another corporate response is to avoid equity issuance in favor of financing via the company's retained earnings or by issuing debt.

Some degree of asymmetric information always exists because investors never know as much as managers and other insiders. Consequently, investors often closely watch manager behavior for insight into insider opinions on the company's prospects. Aware of this scrutiny, managers consider how their actions might be interpreted by outsiders. The signaling model of capital structure suggests a hierarchy ("pecking order") to managers' selection of methods for financing new investments.

The **pecking order theory**, developed by Myers and Majluf (1984), suggests that managers choose methods of financing according to a hierarchy that gives first preference to methods with the least potential information content (internally generated funds) and lowest preference to the form with the greatest potential information content (public equity offerings). Public equity offerings are, in general, closely scrutinized because investors are typically skeptical that existing owners would share ownership of a company with a great future with other investors. In brief, managers prefer internal financing. If internal financing is insufficient, managers next prefer debt, then equity.

Another implication of the work of Myers and Majluf is that financial managers tend to issue equity when they believe the stock is overvalued, but they are reluctant to issue equity if they believe the stock is undervalued. Thus, additional issuance of stock is often interpreted by investors as a negative signal.

We can read the signals that managers provide in their choice of financing method. Managers can signal to the market their confidence in a company's prospects through the issuance of debt, which commits the company to future obligations (so-called "debt signaling"). For example, commitments to fixed payments, such as dividends and debt service payments, may be interpreted as the company's management having confidence in the company's future prospects of making payments. Such signals are considered too costly for poorly performing companies to afford.

Alternatively, the signal of raising money at the top of the pecking order and issuing equity at the bottom of the pecking order holds other clues. If, for instance, the company's cost of capital increases after an equity issuance, we may interpret this effect as an indication that management needed capital beyond what comes cheaply; in other words, this is a negative signal regarding the company's future prospects. Managers may hesitate to issue new equity when they believe the company's shares are underpriced, because they wish to avoid signaling that they believe the shares are overpriced and also to avoid the cost and effort involved with new equity issuance.

Information asymmetry and "signaling" about the value of a company's shares can also become important factors when there are share sales in a company that has a controlling shareholder. Investors will look for signals about the controlling shareholder's confidence in the value of the shares and reasons for selling shares. For instance, if the company is issuing new equity, is the controlling shareholder participating? If there is no new equity being issued, what is the stated reason for the sale and how convincing is that reason?

When a private equity owner takes a portfolio company public, it is typically assumed that it will eventually sell, often in stages over time. With companies controlled by founders or their families, this is less likely to be the case, although it is common for founders to seek to reduce their stake in order to reduce their exposure to a single business or for succession management reasons. In any case, public investors will assume that controlling shareholders will sell their shares at opportune times, taking advantage of favorable market conditions and/or their superior knowledge of the business.

AGENCY COSTS

6

Agency costs are the costs associated with the fact that all public companies and the larger private companies are managed by non-owners. Agency costs are the incremental costs arising from conflicts of interest when an agent makes decisions for a principal. In the context of a corporation, agency costs arise from conflicts of interest between managers, shareholders, and bondholders. "Perquisite consumption" refers to items that executives may legally authorize for themselves that have a cost to shareholders, such as subsidized dining, a corporate jet fleet, and chauffeured limousines.

The smaller the stake managers have in the company, the less their share in bearing the cost of excessive perquisite consumption—consequently, the less their desire to give their best efforts in running the company to maximize shareholder wealth. The costs arising from this conflict of interest have been called the **agency costs of equity**. Given that outside shareholders are aware of this conflict, they will take actions to monitor management and the board's activities that incur costs, such as requiring audited financial statements and holding annual meetings. Management will also take actions to assure owners they are working in the owners' best interest, such as the use of noncompete employment contracts and insurance to guarantee performance.

The better the company governance, the lower the agency costs. Good governance practices translate into higher shareholder value because managers' interests are better aligned with those of shareholders.

Additionally, agency theory predicts that a reduction in net agency costs of equity results from an increase in the use of debt versus equity. That is, there are savings in the agency costs of equity associated with the use of debt. Similarly, the more financially leveraged a company, the less freedom for managers to either take on more debt or unwisely spend cash. This is the foundation of Michael Jensen's **free cash flow hypothesis**. According to Jensen (1986), higher debt levels discipline managers by forcing them to manage the company efficiently so the company can make its interest and principal payments and by reducing the company's free cash flow and thus management's opportunities to misuse cash. This discipline can be especially important in emerging markets, in which there is a tendency to overinvest.

EXAMPLE 10

Agency Costs, Information Asymmetry, and Signaling

Cloudy Prospects plc, a small pharmaceutical company, is testing the potential for its primary drug, which is already in commercial production, to be used in a new application for which it was not intended: treatment of patients with a widespread virus. The company is managed by a team of medical scientists. Management hopes this new application for the drug could result in a big increase in sales and the value of the company. The company has been consistently profitable and debt-free.

Over the course of three months, the following occurs:

In June, while testing occurs, management monitors the results closely. Results are promising; the share price is stable.

In July, testing results continue to look promising. Anticipating that it might have to scale up production quickly, the company negotiates and announces a large increase in its credit line.

In August, Cloudy announces successful test results. The CEO receives a call from a headhunter, who thinks he might be a good candidate to head a larger competitor. The CEO takes the call and decides to meet the headhunter.

Which, if any of the following, are represented here, and by which event and when?

A Agency costs

B Information asymmetry

C Signaling

Solution:

A Agency costs. The CEO's decision in August to take the headhunter meeting about a possible opportunity with a competitor. If the CEO were acting solely in the interests of the shareholders of Cloudy Prospects, he would not take the meeting.

B Information asymmetry. Management had knowledge of the positive test results in June and July, before they were publicly announced in August.

C Signaling. The company's announcement in July that it is increasing its credit line can reasonably be taken as a signal of its confidence that test results will be positive and that it will need to expand production capacity. Note: This is true whether or not this signaling is intentional.

Next, we address competing stakeholder interests.

STAKEHOLDER INTERESTS

7

e describe competing stakeholder interests in capital structure decisions

In this coverage, we look at how and why stakeholder interests are affected when various capital structure decisions are taken by management. Stakeholder groups we consider include providers of capital, such as shareholders and debtholders, as well as others, such as customers, suppliers, employees, management, and the board.

7.1 Shareholder vs. Stakeholder Theory

In the traditional corporate governance framework, shareholders elect a board of directors, which in turn hires managers to serve the interests of shareholders. The interests of other stakeholders—such as debtholders, employees, customers, and society more broadly—are considered only to the extent that they affect shareholder value. This is referred to as the "shareholder theory" model. More recently, "stakeholder theory" advocates have argued that corporate governance should consider the interests of all stakeholders, not just shareholders. In some countries, stakeholder theory concepts have been put into law.

Stakeholder theory gives more prominence to ESG (environmental, social, and governance) considerations by making them an explicit objective for the board of directors and management. Doing this involves certain challenges: the complexity of balancing multiple objectives; a lack of clarity on how non-shareholder objectives are defined, measured, and balanced; the challenges of competing globally when competitors may not face similar constraints; as well as the direct costs in adhering to higher ESG standards.

Some of the principal stakeholder groups are shown in Exhibit 5.

Exhibit 5 Principal Stakeholder Groups

Providers of Capital

Equity shareholders

- Public (minority) common shareholders
- Controlling shareholders (founders, private equity)

Preferred shareholders

Debtholders

- Public debtholders
- Bank lenders
- Private lenders

Other Stakeholders

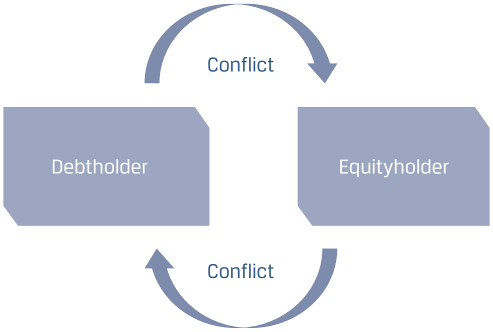
Customers
Suppliers
Employees
Managers
Directors
Regulators, Governments

Capital structure decisions impact stakeholder groups differently. In general, higher financial leverage increases risk for all stakeholders to varying degrees. However, the benefits of increased leverage, primarily through higher returns, accrue almost entirely to shareholders, often including senior management and the board. As a result, shareholders seeking higher returns from the company, and managers acting on their behalf, might prefer decisions that increase the company's financial leverage, while debtholders and other stakeholders in the company generally will not.

Next, we will look at the impact of capital structure decisions on debtholders vs. shareholders and then look more broadly at impacts on other capital providers and other stakeholder groups.

7.2 Debt vs. Equity Conflict

The conflict between equityholders and debtholders reflects their different positions in the capital structure and the different structure of risks and returns for each. Debtholders have a contractual and prior claim to cash flows and firm assets over equityholders. For a holder of debt to maturity, the return or upside is prescribed and limited, while the risk, or downside, is not. In other words, the risk–return profile for debtholders is very asymmetric. In contrast, equityholders typically have more downside risk but much higher upside return potential. As a result, debtholders will almost always favor decisions that reduce a company's leverage and financial risk, while common shareholders often prefer higher leverage levels that offer them greater return potential.



High leverage increases the potential conflict of interest between debtholders and shareholders. At high levels of leverage, the risk–return profile of shareholders becomes more asymmetric. The downside for an equity investor is never more than 100%, or the full value of their investment assuming no leverage or shorting. For highly leveraged companies, the upside reward for shareholders can be multiples of their initial investment. Meanwhile, for a holder of debt to maturity, the return, or upside, remains the face value of debt plus the coupon. However, downside risk for a debtholder increases with higher leverage and the increased probability that the company will be unable to meet its outstanding debt obligations. Similarly, lower leverage levels decrease downside risk for a debtholder as the probability that the company can make its debt obligations increases.

Let's look at a case study.

EXAMPLE 11

Acme Holdings Case Study

Acme Holdings (Acme), a hypothetical company, is planning to make an investment. Three scenarios, each with a different starting leverage for Acme, are compared for impact to Acme's shareholders and debtholders. The scenarios are:

Scenario 1. Normal leverage

Acme has a modest amount of debt, consistent with an investment-grade rating. Debt-to-equity ratio is 1:3.

Scenario 2. High leverage

Acme is highly leveraged, with debt comprising the majority of its capital structure. Debt-to-equity ratio is 9:1.

Scenario 3. Distress scenario

Acme has debt exceeding the value of its business; the intrinsic value of its equity is therefore zero. Debt-to-equity ratio is 1:0.

The case illustrates how the outcomes for shareholders and debtholders are different in each scenario and how management decisions may or may not compromise the interests of debtholders.

It is not necessary that you memorize or be able to replicate the calculations in this example. The purpose is simply to illustrate the debt/equity conflict of interest numerically.

Discussion:

Acme Holdings, with an enterprise value of $1 billion, is considering a $500 million investment to be financed using 50% new debt and 50% equity. The investment has a 50% probability of creating $1.2 billion in value, and a 50% probability of resulting in zero value, for an expected value of $600 million (i.e., 50% × $1.2 billion + 50% × $0). NPV is therefore $100 million (i.e., $600 million – $500 million). Post investment, Acme's enterprise value will either be $2.2 billion ($1.0 billion + $1.2 billion) or $1.0 billion ($1.0 billion + $0), with both outcomes having an equal probability.

Exhibit 6 shows the impact of the investment on equityholder and debtholder value. Columns A, B, C show Acme's pre- and post-investment capital. Columns D, E, F show the company's resulting enterprise value for each if the investment fails or succeeds with corresponding expected value. This value is then compared to the post-capital investment value with the difference being the impact from making the investment. The impact is either positive, negative, or neutral for each group. The last column indicates whether management is likely to proceed or not with the investment based on the expected impact to Acme's equityholders and debtholders.

In short, management is likely to make the investment when equityholders are expected to benefit, even if the impact to debtholders is negative. Management is not likely to make the investment when the impact to equityholders is negative, even if the impact to bondholders and the company is positive.

Exhibit 6 ACME Holdings: Scenarios Summary

	Capital Invested			Enterprise value, ending					
	Pre-investment	New investment	Post-investment	"Failure" (50% prob)	"Success" (50% prob)	Expected value	Impact on Value		Management will likely…
Scenario	A	B	C = A + B	D	E	F = 0.5D + 0.5E	G = F – C		
Scenario 1: Normal leverage									Proceed
Equity	750	250	1,000	500	1,700	1,100	100	Positive	
Debt	250	250	500	500	500	500	—	Neutral	
Total	1,000	500	1,500	1,000	2,200	1,600	100	Positive	
Scenario 2: High leverage									Proceed
Equity	100	250	350	—	1,050	525	175	Positive	
Debt	900	250	1,150	1,000	1,150	1,075	–75	Negative	
Total	1,000	500	1,500	1,000	2,200	1,600	100	Positive	
Scenario 3: Distress									Not proceed
Equity	—	250	250	—	450	225	–25	Negative	

(continued)

Exhibit 6 (Continued)

	Capital Invested			Enterprise value, ending					
	Pre-investment	New investment	Post-investment	“Failure” (50% prob)	“Success” (50% prob)	Expected value	Impact on Value		Management will likely…
Scenario	A	B	C = A + B	D	E	F = 0.5D + 0.5E	G = F – C		
Debt (face value 1,500)	1,000	250	1,250	1,000	1,750	1,375	125	Positive	
Total	1,000	500	1,500	1,000	2,200	1,600	100	Positive	

Scenario 1, 2, 3 “Failure”

D (Debt) = Minimum [Ending Total EV, C(Debt)]

D(Equity) = Ending Total EV – D(Debt)

Scenario 1, 2 “Success”

E(Debt) = Minimum [Ending Total EV, C(Debt)]

Scenario 3 “Success”

E (Debt) = Minimum [Ending Total EV, Face Value Debt + B (Debt)]

Scenario 1, 2, 3 “Success”

D(Equity) = Ending Total EV – D(Debt)

Now we look more closely at the numbers for each scenario:

Scenario 1: Normal Leverage

- Enterprise Value: If the investment fails, Acme is worth $1 billion. Debt capital of $500 million is covered with $500 million residual equity value. If the investment succeeds, Acme is worth $2.2 billion. Debt value remains covered at $500 million, and equityholder residual value is $1.7 billion.
- Outcome: The expected value for equity increases to $1.1 billion from $1.0 billion, so management makes the investment. The expected value of the debt remains unchanged at $500 million, and debtholders are no worse off since the enterprise value of the company ($1.6 billion) exceeds the value of the debt, regardless of whether the new investment succeeds or fails.

Scenario 2: High Leverage

- Enterprise Value: If the investment fails, debtholders lose $150 million ($1 billion – $1.15 billion) and equity value is 0. If the investment succeeds, debt value remains unchanged ($1.15 billion) and equity value increases to $1.05 billion ($2.2 billion – $1.15 billion).
- Outcome: Similar to Scenario 1, equityholders are better off with the investment, with an NPV of negative $350 million if the investment fails, which represents their full post investment value, but $700 million ($1.05 billion – $350 million) if it succeeds. In seeking to maximize equity value, management would make the investment. However, debtholders are worse off, losing $150 million if the investment fails and gaining nothing if it succeeds. In this case, there is a conflict of interest between equity investors and debtholders. Management's decision to make the investment for the benefit of equityholders compromises the interests of the company's debtholders.

Scenario 3: Distress Scenario

- Pre-Investment: Acme has debt with a face value of $1.5 billion and zero equity value.
- Enterprise Value: In this case, equity investors are worse off. If the investment fails, they lose the $250 million they contribute; if it succeeds, there is only $450 million in equity value, or $200 million after deducting the required equity investment, for an expected impact on value of negative $25 million. Debtholders, however, are better off. If the investment succeeds, they have a full recovery on their $250 million investment; if it fails, they lose $500 million ($1.5 billion – $1.0 billion). The expected value of negative $125 million ($1.375 billion – $1.5 billion) is better than without the investment (negative $250 million).
- Outcome: In this case, making the investment would create positive value for Acme, but in seeking to maximize equity value, management would not make the investment because the return to equityholders is negative. Management's decision once again compromises the interests of its bondholders in the interests of equityholders.

Distressed Debt: Observations The Acme Holdings example shows a distress scenario: The value of debt outstanding exceeds the value of the enterprise, and the company's equity has negative intrinsic value. Given this scenario, it would be difficult for Acme to raise new equity capital, and management may decline to make investments that would increase enterprise value.

In this situation, a financial restructuring is often a logical course of action. Restructuring can be involuntary, such as when the company is in default, or voluntary. In either situation, by exchanging debt for equity, the company's capital structure can be repaired, new investments can be financed, and the company can avoid liquidation—which is a very costly type of financial distress.

This situation can arise on a large scale during financial crises. It is particularly problematic in the banking sector given the critical role of banks in sustaining business investment. In such crises, the burden of bad loans on their balance sheets makes banks reluctant to make new loans, even if those new loans would be profitable on a stand-alone basis. In these situations, governments will often intervene to ensure that banks remain solvent and willing to make new loans.

When leverage and financial risk are high, the perspective of debtholders becomes more like that of equityholders. Normally, debt has low default risk and is priced to provide a modest yield to maturity. Return, or upside, is limited, unless the general level of interest rates declines. A holder of debt to maturity will be very focused on minimizing downside risk. However, with higher-risk debt, and particularly with distressed debt, the potential upside can be substantial if the company does not default or if recoveries in default are higher than expected. In these circumstances, the return profile, and thus perspective, of the debtholder becomes closer to that of an equity investor.

Other Debt Considerations

Seniority and Security Debtholder returns are influenced not only by the probability of default but also by the likely recoveries in the event of default. In a default scenario, secured lenders are likely to recover more of their principal than unsecured lenders; likewise, senior debt is likely to recover more than subordinated debt. Recall that senior debt is often secured, but not always. As a result, secured and senior lenders can better tolerate actions by management that increase the probability of default, such as new investment by the company or increasing the level of unsecured or subordinated debt.

Long-Term vs. Short-Term Debt The potential debt/equity conflict is greater in the case of long-term rather than short-term debt because the passage of time exposes debtholders to possible changes in business conditions, strategy, and management behavior.

Safeguards for Debtholders Management generally seeks to maximize equity value and thus may take actions that are unfavorable to debtholders. However, debtholders have the following safeguards to protect their interests:

- Debt covenants can be used to put contractual limits on future borrowing by the company and limit other management actions that would increase risk to debtholders. Positive covenants state what the borrower must do, such as maintaining financial ratios within certain thresholds. Ratios commonly used in debt covenants are drawn from the income and cash flow statements (cash coverage of interest or fixed charges), the balance sheet (minimum net worth, debt-to-equity ratio), or both (debt-to-EBITDA ratio). Negative covenants state what the borrower cannot do, such as additional secured borrowing or dividend payments.
- Companies generally want to preserve favorable access and terms to credit and so will act accordingly to maintain or improve their credit worthiness given they may need to borrow in the future. Management's track record—its statements and actions with respect to capital structure, investment, M&A, dividends, and strategy—will impact the company's future credit rating, banking relationships, and cost of borrowing.
- Financial distress costs can be substantial. Even with a primary focus on maximizing common share value, management will generally seek to keep default risk, and other stakeholder impacts from high leverage, at very low levels.

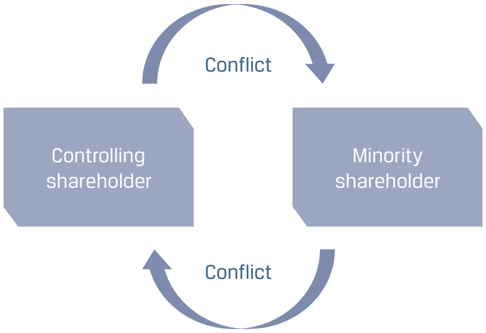
7.3 Preferred Shareholders

Preferred shares are often described as "hybrid" securities since they have both debt-like and equity-like characteristics. Holders of preferred shares rank before common equity shareholders but after debtholders in terms of security and priority of dividend payment. As with common shareholders, preferred shareholders provide long-term capital, with no maturity (repayment) date. Failure to pay a preferred dividend is not an event of default for a company, but it generally prevents the payment of dividends to common shareholders until preferred shareholders are paid. (In the case of cumulative preferred shares, any unpaid preferred dividends from prior periods must also be paid before common dividends can be paid.) The issuance of preferred equity therefore creates less risk for the company's debtholders and common shareholders than if the company were to issue debt. Whether issuing preferred equity is more attractive than issuing debt from a common shareholder perspective depends on the relative after-tax borrowing cost of the company.

Since they are providers of permanent capital and lack covenant protection, preferred shareholders are vulnerable to decisions that increase financial leverage and risk over the long term. For example, a company in a mature business might commit to a long-term policy of dividend growth and share repurchases. This strategy would result in higher financial leverage over the long term. Existing senior secured debtholders might be unconcerned, particularly for debt maturing relatively soon, since there is minimal default risk. However, preferred shareholders might be concerned that the new policy could gradually erode the company's capacity to pay preferred dividends.

7.4 Private Equity Investors/Controlling Shareholders

Not all shareholders have identical financial interests. Where there are one or more large shareholders, the issuance or retirement of common shares affects voting control of the company. Even when one shareholder is clearly in control, with a voting interest of more than 50%, that shareholder may have objectives that are at odds with those of minority shareholders. It is common for private equity majority shareholders to exercise control through the appointment of board members and senior management.



The specific circumstances are important.

- Where controlling shareholders have multiple voting shares, they can pursue expansion strategies and acquisitions that may not increase shareholder value and issue shares to do so, with no risk that minority shareholders or other stakeholders will block their actions. An example might be the acquisition of a sports team by a company in an unrelated business.
- A controlling shareholder seeking to sell its stake, such as a private equity backer, might take a very short-term view on financing, resisting any action that might result in dilution of its equity stake.
- Conversely, some companies that are founder-led and founder-controlled benefit from the fact that founders often take a very long-term perspective on their business, while minority shareholders might not. Amazon is a high-profile example since it has consistently sacrificed short-term profitability in favor of long-term growth and market share. Founder Jeff Bezos has allowed his ownership stake to be diluted as the company has grown.

7.5 Bank and Private Lenders

While the perspectives of lenders are broadly similar, there are stakeholder differences between holders of public debt, private debt, and bank lenders. These differences arise primarily from differences in investment time horizon and stakeholder access to information. Public market debtholders rely on public information and credit rating agency determinations to make their investment decisions. They may not hold debt securities to maturity, so their returns reflect changes in the price of the securities over time—not just the ability of the borrower to pay interest and principal. Bond prices are affected by many factors, including market conditions and shifting perceptions about the company and management, that may or may not impact returns for those who hold the company's debt to maturity.

Bank and private lenders generally hold company debt to maturity. They typically have direct access to company management and non-public information about the company, which in principle reduces information asymmetries. And since an individual bank or private lender can be a critical source of capital, particularly for a small- or mid-sized company, this type of lender has a degree of influence on the company. Such lenders can decide to relax debt covenants, extend more credit, or refuse to do so, which is much more difficult with public debt held by many parties. This is less true for large, "syndicated" loans involving multiple lenders.

Bank lending policies are generally conservative. Like other lenders, a bank will try to balance risk and return, but it must also meet regulatory reserve requirements that are based on the risk profile of their loan book. At the same time, banks generally have a close working relationship with the companies they finance. Banks often provide a range of services, such as payments, foreign currency transactions, and letters of credit, to the companies they finance. This close working relationship provides them with unique and immediate insights into the borrower's current and evolving financial condition and prospects. Bank loan covenants can also be tailored more easily to a borrower's situation than public debt covenants. For these reasons, banks tend to work with borrowers when the latter encounter financing difficulties.

With private lenders, like bank lenders, the general perspective of debtholders applies: Less financial leverage implies less risk and is therefore preferred. However, among private lenders, there is wide variation in their risk appetite, approach, behavior, and relationships with companies to whom they have provided capital. It is therefore difficult to generalize about their particular perspective. Some focus mainly on asset value, such as the collateral supporting their loan; some hold both debt and equity positions in the same company; some have an equity-like focus on the cash flows, value, and prospects of the business; and some lend to businesses they would be interested in owning if the borrowing company were to default on its payments.

7.6 Other Stakeholders

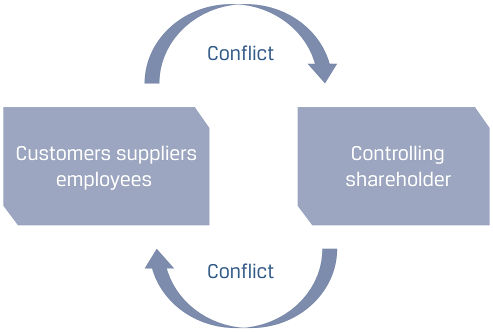
While conflicting interests with respect to capital structure decisions are most apparent with different capital providers, they also arise with other stakeholders.

7.6.1 *Customers and Suppliers*

Suppliers are typically short-term creditors. When companies are in financial distress, their financial position and their willingness to extend additional credit may be affected. However, their interest in the financial health of a customer company is often broader: Customers and suppliers make a substantial investment of time and effort in establishing vendor relationships and have a natural interest in a company's long-term stability. This might not be true for the provider of a commodity product, but it applies when products are specialized and when the customer or supplier makes a significant investment in the relationship—for example, through product design, retooling, training, or software customization. In business strategy terms, such relationships are described as having "switching barriers." Examples include insurance, software, training programs, specialized components, and equipment. In these cases, financial stability can be an important vendor selection criterion, even if there is no actual or anticipated financial distress.

7.6.2 *Employees*

Like suppliers, employees may have both a direct financial stake in their employer company through equity-oriented participation plans (such as profit-sharing, share purchase, or stock option plans) as well as a broader interest in the company's long-term stability, survival, and long-term growth. These are important considerations for companies in attracting and retaining talent. For most employees in most businesses, equity ownership is a minor part of compensation and less important than the company's stability and growth.



The difference in interests between customers, suppliers, employees, and owners is illustrated in the following example.

EXAMPLE 12

Leverage and Other Stakeholders—Mainly Metal Boats, Inc.

Mainly Metal Boats is a small, debt-free manufacturer of welded metal boats and is located in a remote small town. It sells both domestically and internationally in a market that is highly competitive and cyclical. Sales are through a network of dealers, who typically sell three or four different boat brands. Of the company's 50 employees, about half are specialized aluminum welders, with most others in sales and management. The primary purchased input for Mainly's boat production is sheet aluminum. Mainly's sole aluminum supplier is a large multinational company that has many clients. Mainly has not paid dividends previously and now has substantial retained earnings that finance the company's working capital. Mainly's owners have recently decided to borrow heavily to finance working capital so they can pay a large, one-time dividend from the retained earnings.

1. Which of the following stakeholder groups is likely to be most negatively affected by the increase in leverage?
   - **A** The welders employed by the company
   - **B** The company's dealers
   - **C** The supplier of aluminum to the business
2. Is it likely that any of these groups would be impacted positively?

Solution:

1 A is correct. As employees, the welders could face loss of employment if the company were to become financially distressed with the increase in leverage, and since their skills are very specialized, they would probably have difficulty finding another job locally. In a small and remote town, employment opportunities are likely to be limited for specialized workers.

B is incorrect. The dealers might suffer lost sales if Mainly were to fail, but they could likely replace Mainly with a competing brand.

C is incorrect. The aluminum supplier would probably suffer the least impact since it is large and Mainly is not likely a large proportion of its sales.

2 In all cases, impacts are negative. Note that for modest borrowing, these effects would be very minor.

7.6.3 *Management and Directors*

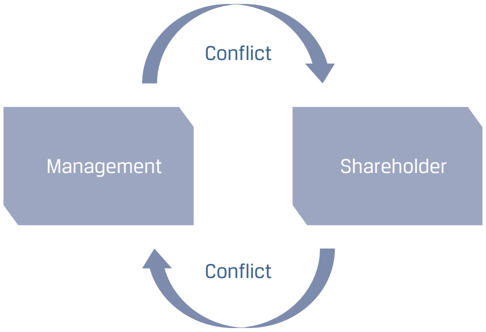
It is critical to consider the interests of management and the board since they make capital structure and business strategy decisions. In principle, management and the board are hired to maximize shareholder wealth, giving rise to the debt/equity conflict described and the risk that their actions will reduce value for debtholders to benefit shareholders.

In practice, compensation is the principal tool used to create alignment of interests between management and directors and shareholders. Equity compensation can be a substantial portion of their total compensation.

In principle, management compensation is intended to motivate managers to work hard to maximize shareholder value. However, the alignment of interests between managers and shareholders is never perfect. For example, the interests of owners are not necessarily identical to the interests of managers, who happen to receive shares and options as part of their compensation.

If we consider the typical elements of management compensation and the typical process by which compensation is determined (e.g., by a compensation committee of the board of directors), we can identify common examples of misalignment or conflicts:

- The overall level of director or executive compensation can be excessive. This is a cost to the company and can lead to "entrenchment," which is the avoidance of taking risk motivated by the desire to keep one's position. Directors, similarly motivated to keep their positions by compensation and prestige, may avoid speaking out against management in the interest of shareholders or other stakeholders.
- Where senior management compensation is high and tied to the size of the business, it can also lead to "growth for growth's sake" or "empire building," with managers motivated to pursue acquisitions and expansion that might not increase shareholder value. It is worth noting that managers two or three tiers below the CEO, whose careers and incomes are tied more to promotions than to the share price, may also be motivated to pursue questionable growth initiatives; it then falls to the CEO and the board to maintain financial and strategic discipline.
- Reliance on management stock options can motivate risk-taking behavior by their holders, particularly with respect to capital structure decisions, since optionholders participate only in upside share price moves.



The following highlights how this can occur.

EXAMPLE 13

Management Stock Options and Leverage

The company Binary Outcomes Ltd. is debt-free and has a share price of $100, with 1 million shares outstanding. Its intrinsic value is $100 million ($100 × 1 million). This is based on a 50% probability of a $110/share value and a 50% probability of a $90/share value. If management has options exercisable at $100, their current expected value is $5.

Is it in management's interest to increase leverage at the company, such as by issuing $50 million of debt to buy back shares?

Discussion:

As shown in Exhibit 7, it would be in management's interest to increase leverage, such as issuing debt to buy back 50% of the company's shares at $100 per share. This would increase debt by $50 million (purchase of 0.5 million shares at $100 per share) and reduce equity value by an equivalent amount, creating a 50% probability of a $120/share value and a 50% probability of an $80/share value. While there is no change in shareholder value, the leverage results in an increase in share value ($120) and corresponding option value ($20) in the upside scenario, while downside option value remains $0, with an average expected option value of $10. The use of leverage doubles the average expected value of management's options from $5 to $10.

Exhibit 7 Binary Outcomes Leverage Scenario

Scenario	Upside	Downside	Expected Value
Probability	50%	50%	
Firm value	$110,000,000	$90,000,000	$100,000,000
No leverage			
Debt	—	—	
Equity value	$110,000,000	$90,000,000	
Shares outstanding	1,000,000	1,000,000	
Share price	$110	$90	$100
Option exercise price	$100	$100	
Option value	$10	—	$5

(continued)

Exhibit 7 (Continued)

Scenario	Upside	Downside	Expected Value
With leverage			
Debt	$50,000,000	$50,000,000	
Equity value	$60,000,000	$40,000,000	
Shares outstanding	500,000	500,000	
Share price	$120	$80	$100
Option exercise price	$100	$100	
Option value	$20	—	$10

With typical management share option schemes, rather than paying dividends, management will be motivated to buy back shares with excess cash. This strategy benefits them by increasing the value of their share options. Similarly, management would be incentivized to retain profits rather than distribute them as a dividend to shareholders.

Continuing this example, if Binary Outcomes earns a profit of $10/share and adds $10/share in cash, retaining that profit would increase its intrinsic value from $100/share to $110/share, thus increasing the intrinsic value of management options by $5 (from $5 to $10). If the entire profit is instead distributed as a dividend, however, there would be no increase in intrinsic value for the shares or management's options.

To mitigate this issue, some companies pay executive performance bonuses in shares rather than options or have minimum share ownership requirements for senior managers and, in some cases, for directors.

In mature or stable businesses, compensation can skew heavily to salaries and bonuses tied to "easy" targets, and management complacency can be a risk. The business may not have enough upside to attract highly talented managers or enough influence to make management aggressively pursue incremental sales or cost efficiencies. In these cases, adding financial leverage is a common strategy, increasing both the incentives and pressure on management to maximize value.

When managers stay in their positions for only a few years, they may be motivated to pursue short-term rather than long-term objectives: They want to demonstrate success and keep their jobs, and they are unsure whether they will be with the company long enough to benefit from longer-term initiatives. From a shareholder perspective, the solution is not to avoid management turnover, but rather to ensure that it is based on performance (i.e., evaluated objectively and carefully) instead of on non-performance-related reasons.

Some of these misalignments offset each other. For example, stock options motivate risk-taking, while high cash compensation can create excessive caution and entrenchment.

7.6.4 *Regulators and Government*

Regulators and the government are often key stakeholders for a company. In some cases, the company's capital structure is an important consideration for regulators. Financial institutions must generally maintain certain levels of solvency or capital adequacy, as defined by regulators. Utilities may have pricing determined based on allowable rates of return, which depend on capital structure. Multinational companies may borrow or retain funds in offshore jurisdictions to minimize taxes, which governments may permit but may also view unfavorably.

Distressed companies may seek government support in order to remain in business, and those governments, in turn, may require new equity investments, block dividends, or otherwise constrain the company's financing decisions.

SUMMARY

- A company's stage in the life cycle, its cash flow characteristics, and its ability to support debt largely dictate its capital structure since capital not sourced through borrowing must come from equity (including retained earnings).
- Generally speaking, as companies mature and move from start-up, through growth, to mature, their business risk declines as operating cash flows turn positive with increasing predictability, allowing for greater use of leverage at more attractive terms.
- Modigliani and Miller's work shows us that managers cannot change firm value simply by changing the firm's capital structure. Firm value is independent of the capital structure decision.
- Given the tax deductibility of interest, adding leverage increases firm value to a point but also increases the risk of default for capital providers who demand higher returns in compensation.
- To maximize firm value, management should target the optimal capital structure that minimizes the company's weighted average cost of capital.
- In practice, financing decisions are typically tied to investment spending and based on the company's ability to support debt given the nature of its business and operating cash flows.
- Managers may provide investors with information ("signaling") through their choice of financing method. For example, commitments to fixed payments can signal management's confidence in the company's future prospects.
- Management's capital structure decisions impact various stakeholder groups differently. In seeking to maximize shareholder wealth or their own, conflicts of interest may arise in which one or more groups are favored at the expense of others.

REFERENCES

Hamada, Robert. 1972. “The Effect of the Firm’s Capital Structure on the Systematic Risk of Common Stocks.” Journal of Finance 27 (2): 435–52. doi:10.1111/j.1540-6261.1972.tb00971.x

Jensen, Michael C. 1986. “Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers.” American Economic Review 76 (2): 323–29.

Miller, Merton H. 1977. “Debt and Taxes.” Journal of Finance 32 (2): 261–75. doi:10.2307/2326758

Modigliani, Franco, and Merton H. Miller. 1958. “The Cost of Capital, Corporation Finance, and the Theory of Investment.” American Economic Review 48 (3): 261–97.

Myers, Stewart, and Nicholas S. Majluf. 1984. “Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have.” Journal of Financial Economics 13:187–221. doi:10.1016/0304-405X(84)90023-0

PRACTICE PROBLEMS

1 Which of the following is *least* likely to affect the capital structure of Longdrive Trucking Company? Longdrive has moderate leverage today.

  A The acquisition of a major competitor for shares

  B A substantial increase in share price

  C The payment of a stock dividend

2 Which of the following is *least* accurate with respect to the market value and book value of a company's equity?

  A Market value is more relevant than book value when measuring a company's cost of capital.

  B Book value is often used by lenders and in financial ratio calculations.

  C Both market value and book value fluctuate with changes in the company's share price.

3 Which of these statements is *most* accurate with respect to the use of debt by a start-up fashion retailer with negative cash flow and uncertain revenue prospects?

  A Debt financing will be unavailable or very costly.

  B The company will prefer to use equity rather than debt given its uncertain cash flow outlook.

  C Both A and B.

4 Which of the following is true of the growth stage in a company's development?

  A Cash flow is negative, by definition, with investment outlays exceeding cash flow from operations.

  B Cash flow may be negative or positive.

  C Cash flow is positive and growing quickly.

5 Which of the following mature companies is *most likely* to employ a high proportion of debt in its capital structure?

  A A mining company with a large, fixed asset base

  B A software company with very stable and predictable revenues and an asset-light business model

  C An electric utility

6 Which of the following is *most likely* to occur as a company evolves from growth stage through maturity and seeks to optimize its capital structure?

  A The company relies on equity to finance its growth.

  B Leverage increases as the company needs more capital to support organic expansion.

  C Leverage increases as the company is able to support more debt.

7 If investors have homogeneous expectations, the market is efficient, and there are no taxes, no transaction costs, and no bankruptcy costs, Modigliani and Miller's Proposition I states that:

  A bankruptcy risk rises with more leverage.

  B managers cannot change the value of the company by changing the amount of debt.

© 2020 CFA Institute. All rights reserved.

  - **C** managers cannot increase the value of the company by employing tax-saving strategies.

8 According to Modigliani and Miller's Proposition II without taxes:

  - **A** the capital structure decision has no effect on the cost of equity.
  - **B** investment and the capital structure decisions are interdependent.
  - **C** the cost of equity increases as the use of debt in the capital structure increases.

9 The current weighted average cost of capital (WACC) for Van der Welde is 10%. The company announced a debt offering that raises the WACC to 13%. The *most likely* conclusion is that for Van der Welde:

  - **A** the company's prospects are improving.
  - **B** equity financing is cheaper than debt financing.
  - **C** the company's debt/equity has moved beyond the optimal range.

10 According to the static trade-off theory:

  - **A** debt should be used only as a last resort.
  - **B** companies have an optimal level of debt.
  - **C** the capital structure decision is irrelevant.

The following information relates to Questions 11–12

Nailah Mablevi is an equity analyst who covers the entertainment industry for Kwame Capital Partners, a major global asset manager. Kwame owns a significant position with a large unrealized capital gain in Mosi Broadcast Group (MBG). On a recent conference call, MBG's management states that they plan to increase the proportion of debt in the company's capital structure. Mablevi is concerned that any changes in MBG's capital structure will negatively affect the value of Kwame's investment.

To evaluate the potential impact of such a capital structure change on Kwame's investment, she gathers the information about MBG given in Exhibit 1.

Exhibit 1 Current Selected Financial Information for MBG

Yield to maturity on debt	8.00%
Market value of debt	$100 million
Number of shares of common stock	10 million
Market price per share of common stock	$30
Cost of capital if all equity-financed	10.3%
Marginal tax rate	35%

11 MBG is *best* described as currently:

  - **A** 25% debt-financed and 75% equity-financed.
  - **B** 33% debt-financed and 66% equity-financed.
  - **C** 75% debt-financed and 25% equity-financed.

12 Holding operating earnings constant, an increase in the marginal tax rate to 40% would:

A result in a lower cost of debt capital.

B result in a higher cost of debt capital.

C not affect the company's cost of capital.

13 Which of the following is *least likely* to be true with respect to optimal capital structure?

A The optimal capital structure minimizes WACC.

B The optimal capital structure is generally close to the target capital structure.

C Debt can be a significant portion of the optimal capital structure because of the tax-deductibility of interest.

14 Fran McClure of Alba Advisers is estimating the cost of capital of Frontier Corporation as part of her valuation analysis of Frontier. McClure will be using this estimate, along with projected cash flows from Frontier's new projects, to estimate the effect of these new projects on the value of Frontier. McClure has gathered the following information on Frontier Corporation:

	Current Year ($)	Forecasted for Next Year ($)
Book value of debt	50	50
Market value of debt	62	63
Book value of shareholders' equity	55	58
Market value of shareholders' equity	210	220

The weights that McClure should apply in estimating Frontier's cost of capital for debt and equity are, respectively:

A $w_d = 0.200$; $w_e = 0.800$.

B $w_d = 0.185$; $w_e = 0.815$.

C $w_d = 0.223$; $w_e = 0.777$.

15 Which of the following is *not* a reason why target capital structure and actual capital structure tend to differ?

A Financing is often tied to a specific investment.

B Companies raise capital when the terms are attractive.

C Target capital structure is set for a particular project, while actual capital structure is measured at the consolidated company level.

16 According to the pecking order theory:

A new debt is preferable to new equity.

B new debt is preferable to internally generated funds.

C new equity is always preferable to other sources of capital.

17 Vega Company has announced that it intends to raise capital next year, but it is unsure as to the appropriate method of raising capital. White, the CFO, has concluded that Vega should apply the pecking order theory to determine the appropriate method of raising capital. Based on White's conclusion, Vega should raise capital in the following order:

A debt, internal financing, equity.

B equity, debt, internal financing.

C internal financing, debt, equity.

18 Which of the following is *least* accurate with respect to debt–equity conflicts?

- **A** Equityholders focus on potential upside and downside, while debtholders focus primarily on downside risk.
- **B** Management attempts to balance the interests of equityholders and debtholders.
- **C** Debt covenants can mitigate the conflict between debtholders and equityholders.

19 Other factors being equal, in which of the following situations are debt–equity conflicts likely to increase?

- **A** Financial leverage is low.
- **B** The company's debt is secured.
- **C** The company's debt is long-term.

20 Which of the following stakeholders are *least likely* to be positively affected by increasing the proportion of debt in the capital structure?

- **A** Senior management
- **B** Non-management employees
- **C** Shareholders

21 Which of the following is an example of agency costs? In each case, management is advocating a substantial acquisition, and management compensation is comprised heavily of stock options.

- **A** Management believes the acquisition will be positive for shareholder value but negative for the value and interests of the company's debtholders.
- **B** Management's stock options are worthless at the current share price. The acquisition has a high (50%) risk of failure (with zero value) but substantial (30%) upside if it works out.
- **C** The acquisition is positive for equityholders and does not significantly impair the position of debtholders. However, the acquisition puts the company into a new business where labor practices are harsh and the production process is environmentally damaging.

22 Which of the following is *least likely* to be true with respect to agency costs and senior management compensation?

- **A** Equity-based incentive compensation is the primary method to address the problem of agency costs.
- **B** A well-designed compensation scheme should eliminate agency costs.
- **C** High cash compensation for senior management, without significant equity-based performance incentives, can lead to excessive caution and complacency.

SOLUTIONS

1 C is correct. Stock dividends, like stock splits, have no impact on the value of a company's equity. Issuing shares to acquire a competitor would increase equity relative to debt in the capital structure. Share price appreciation would also increase the market value of equity, thus increasing equity relative to debt.

2 C is correct. Share price changes will cause the market value of the company's equity to change; book value is unaffected.

Statements A and B are accurate.

3 C is correct. For a start-up company of this nature, debt financing is likely to be unattractive to lenders—and therefore very expensive or difficult to obtain. Debt financing is also unappealing to the company, because it commits the company to interest and principal payments that might be difficult to manage given its uncertain cash flow outlook.

4 B is correct. Cash flow typically turns positive during the growth stage, but it may be negative, particularly at the beginning of this stage.

5 C is correct. An electric utility has the capacity to support substantial debt, with very stable and predictable revenues and cash flows. The software company also has these attributes, but it would have been much less likely to have raised debt during its development and may have raised equity. The mining company has fixed assets, which it would have needed to finance, but the cyclical nature of its business would limit its debt capacity.

6 C is correct. As cash flows become more predictable, the company is able to support more debt in its capital structure; the optimal capital structure includes a higher proportion of debt. While mature companies do borrow to support growth, this would typically not occur because the company is optimizing its capital structure. Likewise, while a mature company might issue equity to finance growth, this would not be the typical approach for a company optimizing its capital structure.

7 B is correct. Proposition I, or the capital structure irrelevance theorem, states that in perfect markets the level of debt versus equity in the capital structure has no effect on company value.

8 C is correct. The cost of equity rises with the use of debt in the capital structure (e.g., with increasing financial leverage).

9 C is correct. If the company's WACC increases as a result of taking on additional debt, the company has moved beyond the optimal capital range. The costs of financial distress may outweigh any tax benefits to the use of debt.

10 B is correct. The static trade-off theory indicates that there is a trade-off between the tax shield from interest on debt and the costs of financial distress, leading to an optimal amount of debt in a company's capital structure.

11 A is correct. The market value of equity is (\$30)(10,000,000) = \$300,000,000. With the market value of debt equal to \$100,000,000, the market value of the company is \$100,000,000 + \$300,000,000 = \$400,000,000. Therefore, the company is \$100,000,000/\$400,000,000 = 0.25 or 25% debt-financed.

12 A is correct. The after-tax cost of debt decreases as the marginal tax rate increases.

13 B is correct. A company's optimal and target capital structures may be very different from one another.

14 C is correct.

w_d = \$63/(\$220 + 63) = 0.223.

w_e = \$220/(\$220 + 63) = 0.777.

Market values should be used in cost of capital calculations, and forecasted market values should be used in this case given that the cost of capital will be applied to *projected* cash flows in McClure's analysis.

15 C is correct. Companies generally raise finance when capital is needed, such as for investment spending or when market pricing and terms are favorable for debt or equity issuance.

16 A is correct. According to the pecking order theory, internally generated funds are preferable to both new equity and new debt. If internal financing is insufficient, managers next prefer new debt, then new equity. Managers prefer forms of financing with the least amount of visibility to outsiders.

17 C is correct. According to the pecking order theory, managers prefer internal financing. If internal financing is insufficient, managers next prefer debt, then equity in order of increasing visibility to outsiders.

18 B is correct. Management is generally focused on maximizing the value of equity.

19 C is correct. Long-term debt is more exposed than short-term debt to the risk of a management decision that is not debtholder-friendly. Secured debt is less exposed than unsecured debt to such a risk, and with low leverage, the risk of debt–equity conflict is reduced, not increased, relative to high leverage.

20 B is correct. While leverage increases risk for all stakeholders, shareholders generally benefit through higher potential returns. Senior management typically benefits through equity-based compensation. For non-management employees, equity-based compensation is likely to be small to non-existent.

21 B is correct. Management is advocating an acquisition that is likely to be positive for the value of the company's options but negative for equityholders given the substantial risk. A is an example of the debt–equity conflict. C is an example of stakeholder interests that are not being considered by management.

22 B is correct. A well-designed management compensation scheme can reduce, but not eliminate, agency costs.

READING

32

Measures of Leverage

by Pamela Peterson Drake, PhD, CFA, Raj Aggarwal, PhD, CFA, Cynthia Harrington, CFA, and Adam Kobor, PhD, CFA

Pamela Peterson Drake, PhD, CFA, is at James Madison University (USA). Raj Aggarwal, PhD, CFA, is at the Kent State University Foundation Board (USA). Cynthia Harrington, CFA, is at teamyou.co (USA). Adam Kobor, PhD, CFA, is at New York University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** define and explain leverage, business risk, sales risk, operating risk, and financial risk and classify a risk;
☐	**b.** calculate and interpret the degree of operating leverage, the degree of financial leverage, and the degree of total leverage;
☐	**c.** analyze the effect of financial leverage on a company's net income and return on equity;
☐	**d.** calculate the breakeven quantity of sales and determine the company's net income at various sales levels;
☐	**e.** calculate and interpret the operating breakeven quantity of sales.

INTRODUCTION

1

This reading presents elementary topics in leverage. **Leverage** is the use of fixed costs in a company's cost structure. Fixed costs that are operating costs (such as depreciation or rent) create operating leverage. Fixed costs that are financial costs (such as interest expense) create financial leverage.

Analysts refer to the use of fixed costs as leverage because fixed costs act as a fulcrum for the company's earnings. Leverage can magnify earnings both up and down. The profits of highly leveraged companies might soar with small upturns in revenue. But the reverse is also true: Small downturns in revenue may lead to losses.

Analysts need to understand a company's use of leverage for three main reasons. First, the degree of leverage is an important component in assessing a company's risk and return characteristics. Second, analysts may be able to discern information about a company's business and future prospects from management's decisions about the use of operating and financial leverage. Knowing how to interpret these signals also helps

© 2019 CFA Institute. All rights reserved.

the analyst evaluate the quality of management's decisions. Third, the valuation of a company requires forecasting future cash flows and assessing the risk associated with those cash flows. Understanding a company's use of leverage should help in forecasting cash flows and in selecting an appropriate discount rate for finding their present value.

The reading is organized as follows: Section 2 introduces leverage and defines important terms. Sections 3–8 illustrate and discuss measures of operating leverage and financial leverage, which combine to define a measure of total leverage that gauges the sensitivity of net income to a given percent change in units sold. These sections also cover breakeven points in using leverage and corporate reorganization (a possible consequence of using leverage inappropriately). A summary and practice problems conclude this reading.

2

LEVERAGE

a define and explain leverage, business risk, sales risk, operating risk, and financial risk and classify a risk;

Leverage increases the volatility of a company's earnings and cash flows and increases the risk of lending to or owning a company. Additionally, the valuation of a company and its equity is affected by the degree of leverage: The greater a company's leverage, the greater its risk and, hence, the greater the discount rate that should be applied in its valuation. Further, highly leveraged (levered) companies have a greater chance of incurring significant losses during downturns, thus accelerating conditions that lead to financial distress and bankruptcy.

Consider the simple example of two companies, Impulse Robotics, Inc., and Malvey Aerospace, Inc. These companies have the following performance for the period of study:[1]

Exhibit 1 Impulse Robotics and Malvey Aerospace

	Impulse Robotics	Malvey Aerospace
Revenues	$1,000,000	$1,000,000
Operating costs	700,000	750,000
Operating income	$300,000	$250,000
Financing expense	100,000	50,000
Net income	$200,000	$200,000

These companies have the same net income, but are they identical in terms of operating and financial characteristics? Would we appraise these two companies at the same value? Not necessarily.

The risk associated with future earnings and cash flows of a company are affected by the company's cost structure. The **cost structure** of a company is the mix of variable and fixed costs. **Variable costs** fluctuate with the level of production and sales. Some examples of variable costs are the cost of goods purchased for resale, costs of materials or supplies, shipping charges, delivery charges, wages for hourly employees, sales commissions, and sales or production bonuses. **Fixed costs** are expenses that are the same regardless of the production and sales of the company. These costs include depreciation, rent, interest on debt, insurance, and wages for salaried employees.

---

[1] We are ignoring taxes for this example, but when taxes are included, the general conclusions remain the same.

Suppose that the cost structures of the companies differ in the manner shown in Exhibit 2.

Exhibit 2 Impulse Robotics and Malvey Aerospace

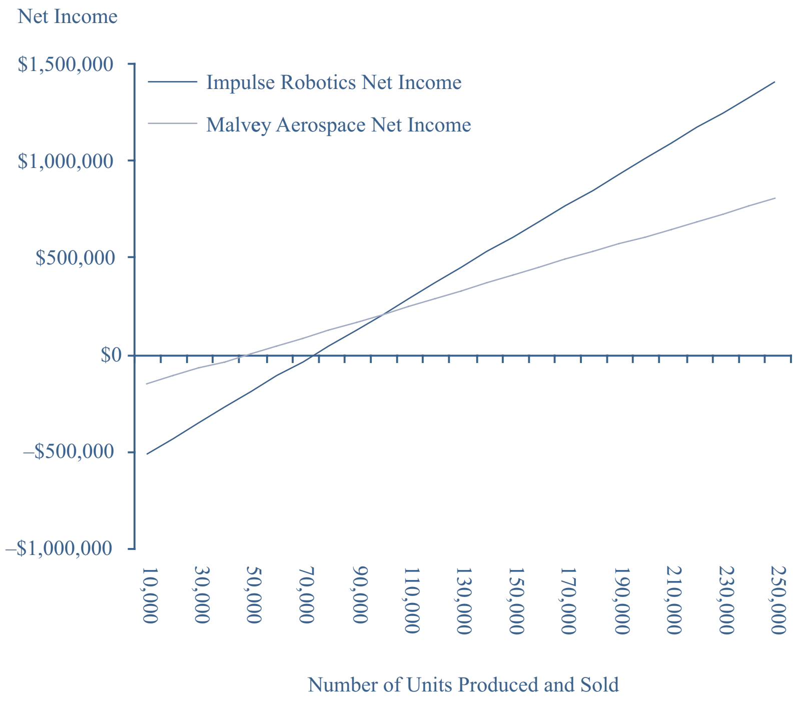
	Impulse Robotics	Malvey Aerospace
Number of units produced and sold	100,000	100,000
Sales price per unit	$10	$10
Variable cost per unit	$2	$6
Fixed operating cost	$500,000	$150,000
Fixed financing expense	$100,000	$50,000

The risk associated with these companies is different, although, as we saw in Exhibit 1, they have the same net income. They have different operating and financing cost structures, resulting in differing volatility of net income.

For example, if the number of units produced and sold is different from 100,000, the net income of the two companies diverges. If 50,000 units are produced and sold, Impulse Robotics has a loss of $200,000 and Malvey Aerospace has $0 earnings. If, on the other hand, the number of units produced and sold is 200,000, Impulse Robotics earns $1 million whereas Malvey Aerospace earns $600,000. In other words, the variability in net income is greater for Impulse Robotics, which has higher fixed costs in terms of both fixed operating costs and fixed financing costs.

Impulse Robotics' cost structure results in more leverage than that of Malvey Aerospace. We can see this effect when we plot the net income of each company against the number of units produced and sold, as in Exhibit 3. The greater leverage of Impulse Robotics is reflected in the greater slope of the line representing net income. This means that as the number of units sold changes, Impulse Robotics experiences a greater change in net income than does Malvey Aerospace for the same change in units sold.

Exhibit 3 Net Income for Different Numbers of Units Produced and Sold



Companies that have more fixed costs relative to variable costs in their cost structures have greater variation in net income as revenues fluctuate and, hence, more risk.

3 BUSINESS AND SALES RISKS

a define and explain leverage, business risk, sales risk, operating risk, and financial risk and classify a risk;

Risk arises from both the operating and financing activities of a company. In the following, we address how that happens and the measures available to the analyst to gauge the risk in each case.

3.1 Business Risk and Its Components

Business risk is the risk associated with operating earnings. Operating earnings are risky because total revenues are risky, as are the costs of producing revenues. Revenues are affected by a large number of factors, including economic conditions, industry dynamics (including the actions of competitors), government regulation, and demographics. Therefore, prices of the company's goods or services or the quantity of sales may be different from what is expected. We refer to the uncertainty with respect to the price and quantity of goods and services as **sales risk**.

Operating risk is the risk attributed to the operating cost structure, in particular the use of fixed costs in operations. The greater the fixed operating costs relative to variable operating costs, the greater the operating risk. Business risk is therefore the combination of sales risk and operating risk. Companies that operate in the same line of business generally have similar business risk.

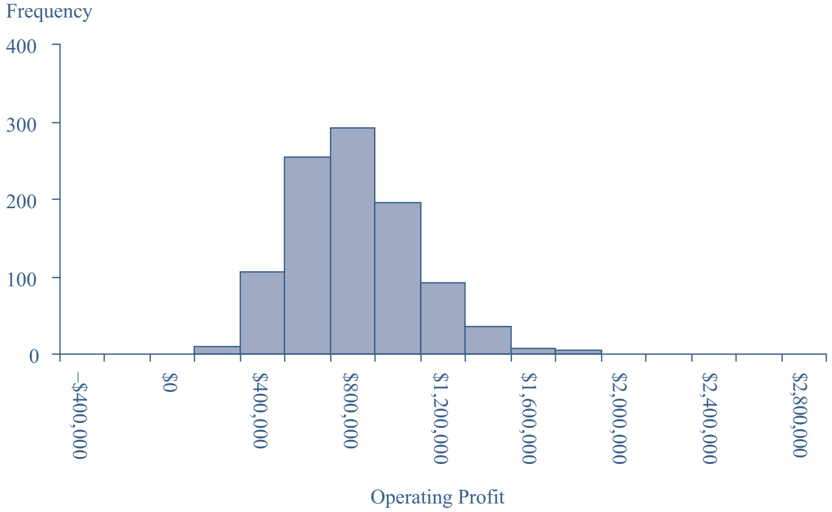
3.2 Sales Risk

Consider Impulse Robotics once again. Suppose that the forecasted number of units produced and sold in the next period is 100,000 but that the standard deviation of the number of units sold is 20,000. And suppose the price that the units sell for is expected to be $10 per unit but the standard deviation is $2. Contrast this situation with that of a company named Tolley Aerospace, Inc., which has the same cost structure but a standard deviation of units sold of 40,000 and a price standard deviation of $4.

If we assume, for simplicity's sake, that the fixed operating costs are known with certainty and that the units sold and price per unit follow a normal distribution, we can see the impact of the different risks on the operating income of the two companies through a simulation; the results are shown in Exhibit 4. Here, we see the differing distributions of operating income that result from the distributions of units sold and price per unit. So, even if the companies have the same cost structure, differing *sales risk* affects the potential variability of the company's profitability. In our example, Tolley Aerospace has a wider distribution of likely outcomes in terms of operating profit. This greater volatility in operating earnings means that Tolley Aerospace has more sales risk than Impulse Robotics.

Exhibit 4 Operating Income Simulations for Impulse Robotics and Tolley Aerospace

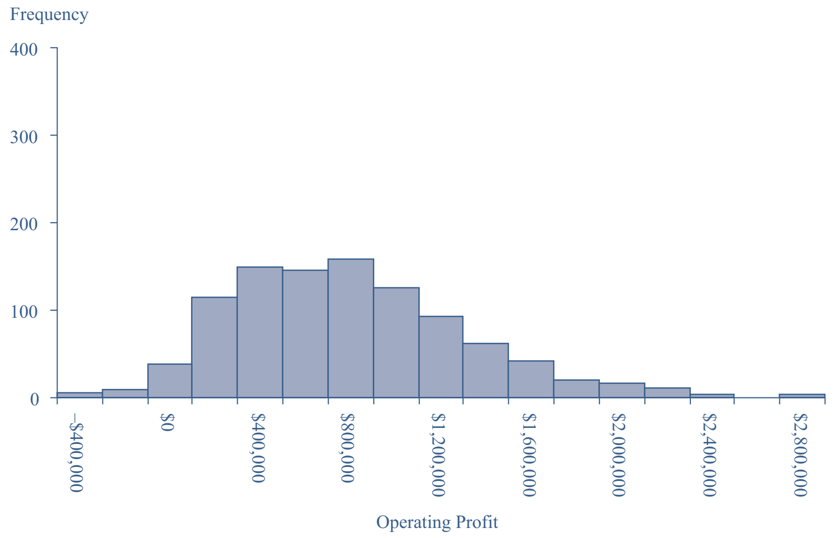
Panel A: Impulse Robotics



(continued)

Exhibit 4 (Continued)

Panel B: Tolley Aerospace



4 OPERATING RISK AND THE DEGREE OF OPERATING LEVERAGE

a define and explain leverage, business risk, sales risk, operating risk, and financial risk and classify a risk;

b calculate and interpret the degree of operating leverage, the degree of financial leverage, and the degree of total leverage;

The greater the fixed component of costs, the more difficult it is for a company to adjust its operating costs to changes in sales. The mixture of fixed and variable costs depends largely on the type of business. Even within the same line of business, companies can vary their fixed and variable costs to some degree. We refer to the risk arising from the mix of fixed and variable costs as **operating risk**. The greater the fixed operating costs relative to variable operating costs, the greater the operating risk.

Next, we look at how operating risk affects the variability of cash flows. A concept taught in microeconomics is **elasticity**, which is simply a measure of the sensitivity of changes in one item to changes in another. We can apply this concept to examine how sensitive a company's operating income is to changes in demand, as measured by unit sales. We will calculate the operating income elasticity, which we refer to as the **degree of operating leverage** (DOL). DOL is a quantitative measure of operating risk as it was defined earlier.

The degree of operating leverage is the ratio of the percentage change in operating income to the percentage change in units sold. We will simplify things and assume that the company sells all that it produces in the same period. Then,

$$\text{DOL} = \frac{\text{Percentage change in operating income}}{\text{Percentage change in units sold}} \quad (1)$$

For example, if DOL at a given level of unit sales is 2.0, a 5 percent increase in unit sales from that level would be expected to result in a (2.0)(5%) = 10 percent increase in operating income. As illustrated later in relation to Exhibit 6, a company's DOL is dependent on the level of unit sales being considered.

Returning to Impulse Robotics, the price per unit is \$10, the variable cost per unit is \$2, and the total fixed operating costs are \$500,000. If Impulse Robotics' output changes from 100,000 units to 110,000 units—an increase of 10 percent in the number of units sold—operating income changes from \$300,000 to \$380,000:[2]

Exhibit 5 Operating Leverage of Impulse Robotics

Item	Selling 100,000 Units	Selling 110,000 Units	Percentage Change
Revenues	\$1,000,000	\$1,100,000	+10.00
Less variable costs	200,000	220,000	+10.00
Less fixed costs	500,000	500,000	0.00
Operating income	\$300,000	\$380,000	+26.67

Operating income increases by 26.67 percent when units sold increases by 10 percent. What if the number of units *decreases* by 10 percent, from 100,000 to 90,000? Operating income is \$220,000, representing a *decline* of 26.67 percent.

What is happening is that for a 1 percent change in units sold, the operating income changes by 2.67 times that percentage, in the same direction. If units sold increases by 10 percent, operating income increases by 26.7 percent; if units sold decreased by 20 percent, operating income would decrease by 53.3 percent.

We can represent the degree of operating leverage as given in Equation 1 in terms of the basic elements of the price per unit, variable cost per unit, number of units sold, and fixed operating costs. Operating income is revenue minus total operating costs (with variable and fixed cost components):

$$\begin{aligned}\begin{matrix}\text{Operating}\\ \text{income}\end{matrix} &= \left[\begin{pmatrix}\text{Price}\\ \text{per unit}\end{pmatrix}\begin{pmatrix}\text{Number of}\\ \text{units sold}\end{pmatrix}\right]\\ &\quad -\left[\begin{pmatrix}\text{Variable cost}\\ \text{per unit}\end{pmatrix}\begin{pmatrix}\text{Number of}\\ \text{units sold}\end{pmatrix}\right]-\begin{bmatrix}\text{Fixed operating}\\ \text{costs}\end{bmatrix}\end{aligned}$$

or

$$\begin{matrix}\text{Operating}\\ \text{income}\end{matrix} = \underbrace{\begin{pmatrix}\text{Number of}\\ \text{units sold}\end{pmatrix}\left[\begin{pmatrix}\text{Price}\\ \text{per unit}\end{pmatrix}-\begin{pmatrix}\text{Variable cost}\\ \text{per unit}\end{pmatrix}\right]}_{\text{Contribution margin}}-\begin{bmatrix}\text{Fixed operating}\\ \text{costs}\end{bmatrix}$$

The **per unit contribution margin** is the amount that each unit sold contributes to covering fixed costs—that is, the difference between the price per unit and the variable cost per unit. That difference multiplied by the quantity sold is the **contribution margin**, which equals revenue minus variable costs.

---

[2] We provide the variable and fixed operating costs for our sample companies used in this reading to illustrate the leverage and breakeven concepts. In reality, however, the financial analyst does not have these breakdowns but rather is faced with interpreting reported account values that often combine variable and fixed costs and costs for different product lines.

How much does operating income change when the number of units sold changes? Fixed costs do not change; therefore, operating income changes by the contribution margin. The percentage change in operating income for a given change in units sold simplifies to

$$\text{DOL} = \frac{Q(P - V)}{Q(P - V) - F} \tag{2}$$

where Q is the number of units, P is the price per unit, V is the variable operating cost per unit, and F is the fixed operating cost. Therefore, $P - V$ is the per unit contribution margin and $Q(P - V)$ is the contribution margin.

Applying the formula for DOL using the data for Impulse Robotics, we can calculate the sensitivity to change in units sold from 100,000 units:

$$\begin{matrix}\text{DOL @} \\ 100{,}000\,\text{units}\end{matrix} = \frac{100{,}000(\$10 - \$2)}{100{,}000(\$10 - \$2) - \$500{,}000} = 2.67$$

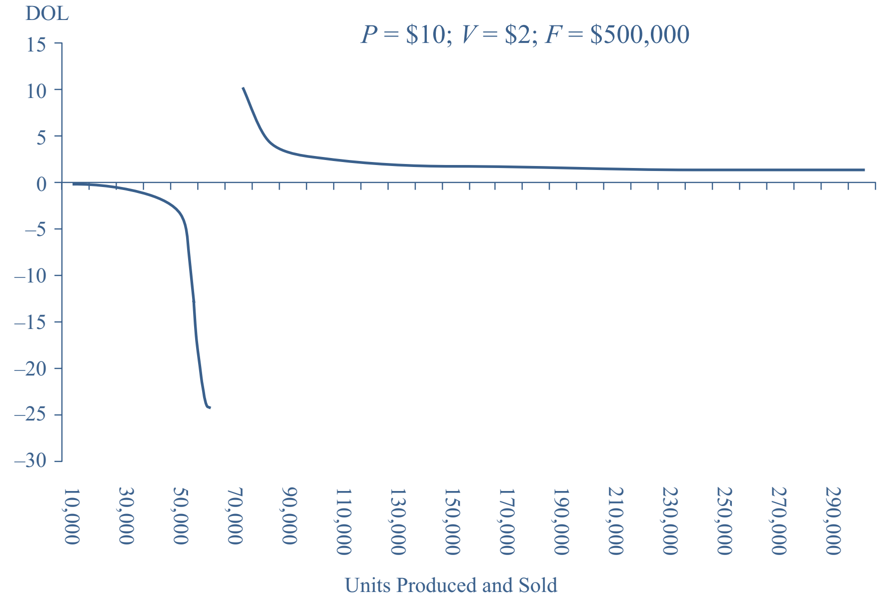
A DOL of 2.67 means that a 1 percent change in units sold results in a 1% × 2.67 = 2.67% change in operating income; a DOL of 5 means that a 1 percent change in units sold results in a 5 percent change in operating income, and so on.

Why do we specify that the DOL is at a particular quantity sold (in this case, 100,000 units)? Because the DOL is different at different numbers of units produced and sold. For example, at 200,000 units,

$$\begin{matrix}\text{DOL @} \\ 200{,}000\,\text{units}\end{matrix} = \frac{200{,}000(\$10 - \$2)}{200{,}000(\$10 - \$2) - \$500{,}000} = 1.45$$

We can see the sensitivity of the DOL for different numbers of units produced and sold in Exhibit 6. When operating profit is negative, the DOL is negative. At positions just below and just above the point where operating income is $0, operating income is at its most sensitive on a percentage basis to changes in units produced and sold. At the point at which operating income is $0 (at 62,500 units produced and sold in this example), the DOL is undefined because the denominator in the DOL calculation is $0. After this point, the DOL gradually declines as more units are produced and sold.

Exhibit 6 Impulse Robotics' Degree of Operating Leverage for Different Number of Units Produced and Sold



We will now look at a similar situation in which the company has shifted some of the operating costs away from fixed costs and into variable costs. Malvey Aerospace has a unit sales price of \$10, a variable cost of \$6 a unit, and \$150,000 in fixed costs. A change in units sold from 100,000 to 110,000 (a 10 percent change) changes operating profit from \$250,000 to \$290,000, or 16 percent. The DOL in this case is 1.6:

$$\begin{matrix}\text{DOL @}\\ 100{,}000\,\text{units}\end{matrix} = \frac{100{,}000(\$10 - \$6)}{100{,}000(\$10 - \$6) - \$150{,}000} = 1.6$$

and the change in operating income is 16 percent:

$$\begin{matrix}\text{Percentage change}\\ \text{in operating income}\end{matrix} = (\text{DOL})\begin{pmatrix}\text{Percentage change}\\ \text{in units sold}\end{pmatrix} = (1.6)(10\%) = 16\%$$

We can see the difference in leverage in the case of Impulse Robotics and Malvey Aerospace companies in Exhibit 7. In Panel A, we see that Impulse Robotics has higher operating income than Malvey Aerospace when both companies produce and sell more than 87,500 units, but lower operating income than Malvey when both companies produce and sell less than 87,500 units.[3]

Exhibit 7 Profitability and the DOL for Impulse Robotics and Malvey Aerospace

Impulse Robotics: P = \$10; V = \$2; F = \$500,000
Malvey Aerospace: P = \$10; V = \$6; F = \$150,000

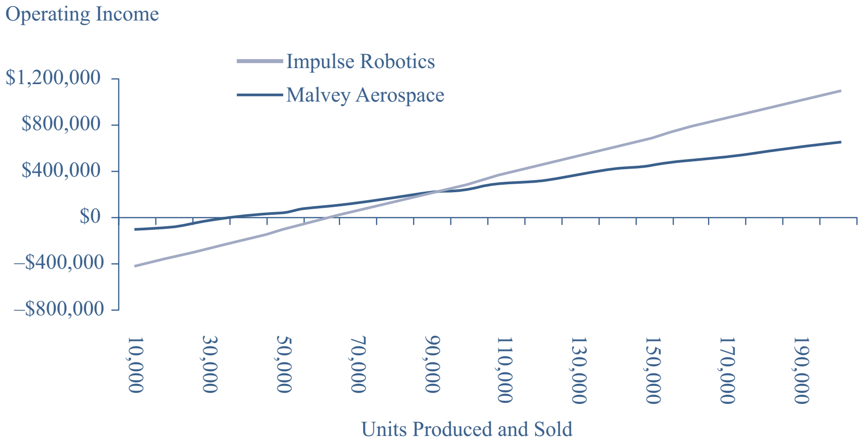
(continued)

---

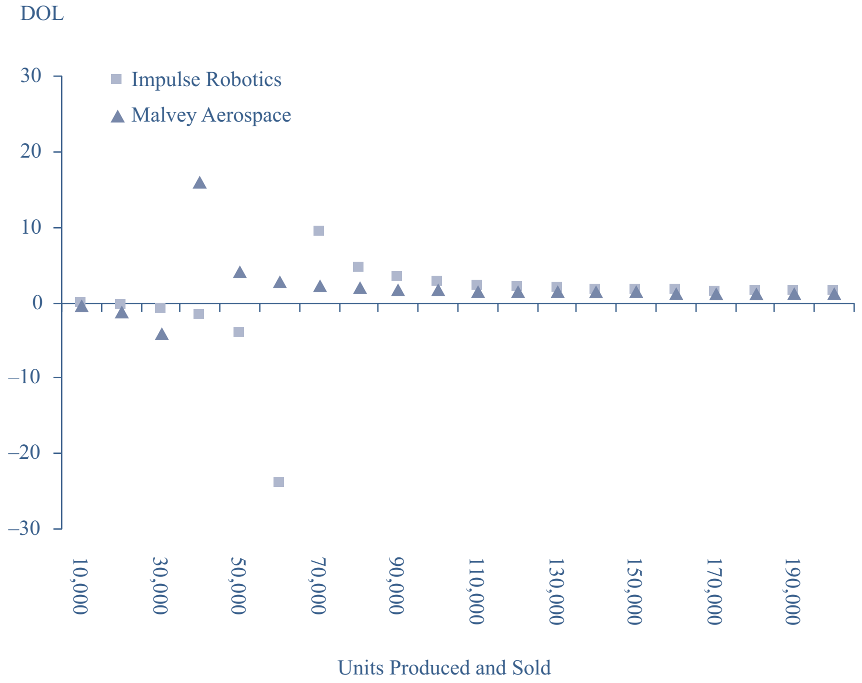
[3] We can calculate the number of units that produce the same operating income for these two companies by equating the operating incomes and solving for the number of units. Let X be the number of units. The X at which Malvey Aerospace and Impulse Robotics generate the same operating income is the X that solves the following: $10X - 2X - 500{,}000 = 10X - 6X - 150{,}000$; that is, $X = 87{,}500$.

Exhibit 7 (Continued)

Panel A: Operating Income and Number of Units Produced and Sold



Panel B: Degree of Operating Leverage (DOL)



This example confirms what we saw earlier in our reasoning of fixed and variable costs: The greater the use of fixed, relative to variable, operating costs, the more sensitive operating income is to changes in units sold and, therefore, the more operating risk. Impulse Robotics has more operating risk because it has more operating leverage. However, as Panel B of Exhibit 7 shows, the degrees of operating leverage are similar for the two companies for larger numbers of units produced and sold.

Both sales risk and operating risk influence a company's business risk. And both sales risk and operating risk are determined in large part by the type of business the company is in. But management has more opportunity to manage and control operating risk than sales risk.

Suppose a company is deciding which equipment to buy to produce a particular product. The sales risk is the same no matter what equipment is chosen to produce the product. But the available equipment may differ in terms of the fixed and variable operating costs of producing the product. Financial analysts need to consider how the operating cost structure of a company affects the company's risk.

EXAMPLE 1

Calculating the Degree of Operating Leverage

Arnaud Kenigswald is analyzing the potential impact of an improving economy on earnings at Global Auto, one of the world's largest car manufacturers. Global is headquartered in Berlin. Global Auto manufactures passenger cars and produces revenues of €168 billion. Kenigswald projects that sales will improve by 10 percent due to increased demand for cars. He wants to see how Global's earnings might respond given that level of increase in sales. He first looks at the degree of leverage at Global, starting with operating leverage.

Global sold 6 million passenger cars in 2017. The average price per car was €28,000, fixed costs associated with passenger car production total €15 billion per year, and variable costs per car are €20,500. What is the degree of operating leverage of Global Auto?

Solution:

$$\begin{array}{c}\text{DOL @}\\ \text{6 million units}\end{array} = \frac{\text{6 million (€28,000} - \text{€20,500)}}{\text{6 million (€28,000} - \text{€20,500)} - \text{€15 billion}} = 1.5$$

Operating income is [6 million × (€28,000 – €20,500)] – €15 billion = €30 billion

For a 10 percent increase in cars sold, operating income increases by 1.50 × 10% = 15.0%.

Industries that tend to have high operating leverage are those that invest up front to produce a product but spend relatively little on making and distributing it. Software developers and pharmaceutical companies fit this description. Alternatively, retailers have low operating leverage because much of the cost of goods sold is variable.

Because most companies produce more than one product, the ratio of variable to fixed costs is difficult to obtain. We can get an idea of the operating leverage of a company by looking at changes in operating income in relation to changes in sales for the entire company. This relation can be estimated by regressing changes in operating income (the variable to be explained) on changes in sales (the explanatory variable) over a recent time period.[4] Although this approach does not provide a precise measure of operating risk, it can help provide a general idea of the amount of operating leverage present. For example, compare the relation between operating earnings and revenues for Delta Air Lines, a transportation company, and Wal-Mart Stores, a discount retailer, as shown in Exhibit 8. Note that the slope of the least-squares regression line is greater for Delta Air Lines (with a slope coefficient of 0.1702) than for Wal-Mart (with a slope coefficient of 0.0493). (A visual comparison of slopes should not be relied upon because the scales of the *x*- and *y*-axes are different in diagrams for the two regressions.) We can see that operating earnings are more sensitive to changes in revenues for the higher-operating-leveraged Delta Air Lines as compared to the lower-operating-leveraged Wal-Mart Stores.

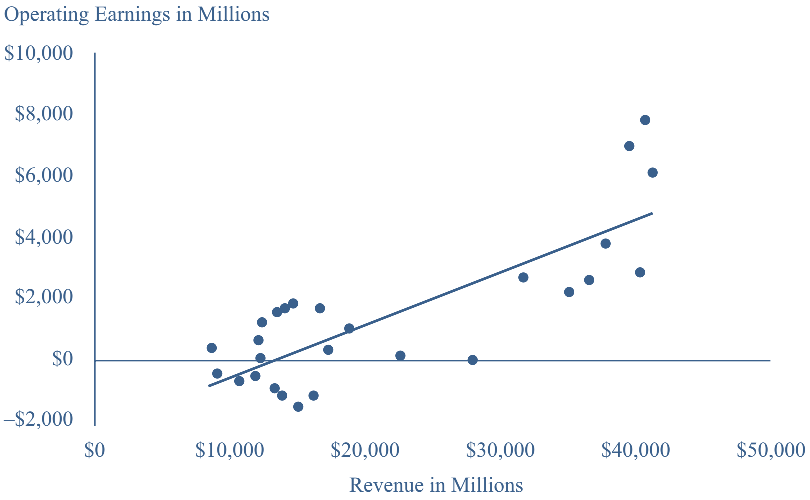
---

[4] A least-squares regression is a procedure for finding the best-fitting line (called the least squares regression line) through a set of data points by minimizing the squared deviations from the line.

Exhibit 8 Relation between Operating Earnings and Revenues

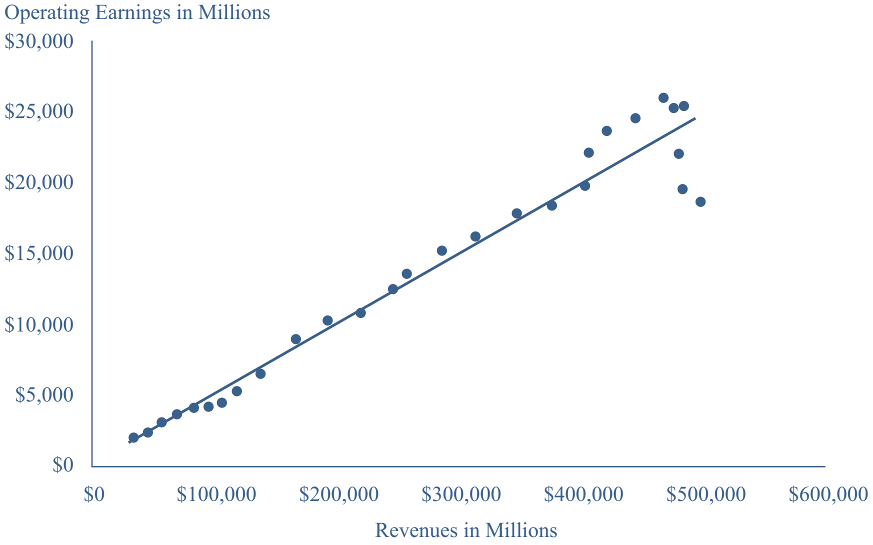
Panel A: Delta Airlines Operating Earnings and Revenues, 1990–2017

Estimated regression: Operating earnings = –$2,249 + 0.1702 Revenues
R^2 = 64.73%



Panel B: Wal-Mart Stores Operating Earnings and Revenues, 1990–2017

Estimated regression: Operating earnings = $253.16 + 0.0493 Revenues
R^2 = 94.89%



Sources: Delta Air Lines 10-K filings and Wal-Mart Stores 10-K filings, various years.

FINANCIAL RISK, THE DEGREE OF FINANCIAL LEVERAGE AND THE LEVERAGING ROLE OF DEBT

5

a define and explain leverage, business risk, sales risk, operating risk, and financial risk and classify a risk;

b calculate and interpret the degree of operating leverage, the degree of financial leverage, and the degree of total leverage;

c analyze the effect of financial leverage on a company's net income and return on equity;

We can expand on the concept of risk to accommodate the perspective of owning a security. A security represents a claim on the income and assets of a business; therefore, the risk of the security goes beyond the variability of operating earnings to include how the cash flows from those earnings are distributed among the claimants—the creditors and owners of the business. The risk of a security is therefore affected by both business risk and financial risk.

Financial risk is the risk associated with how a company finances its operations. If a company finances with debt, it is legally obligated to pay the amounts that make up its debts when due. By taking on fixed obligations, such as debt and long-term leases, the company increases its financial risk. If a company finances its business with common equity, generated either from operations (retained earnings) or from issuing new common shares, it does not incur fixed obligations. The more fixed-cost financial obligations (e.g., debt) incurred by the company, the greater its financial risk.

We can quantify this risk in the same way we did for operating risk, looking at the sensitivity of the cash flows available to owners when operating income changes. This sensitivity, which we refer to as the **degree of financial leverage** (DFL), is

$$\text{DFL} = \frac{\text{Percentage change in net income}}{\text{Percentage change in operating income}} \qquad (3)$$

For example, if DFL at a given level of operating income is 1.1, a 5 percent increase in operating income would be expected to result in a (1.1)(5%) = 5.5 percent increase in net income. A company's DFL is dependent on the level of operating income being considered.

Net income is equal to operating income, less interest and taxes.[5] If operating income changes, how does net income change? Consider Impulse Robotics. Suppose the interest payments are $100,000 and, for simplicity, the tax rate is 0 percent: If operating income changes from $300,000 to $360,000, net income changes from $200,000 to $260,000:

---

[5] More complex entities than we have been using for our examples may also need to account for other income (losses) and extraordinary income (losses) together with operating income as the basis for earnings before interest and taxes.

Exhibit 9 Financial Risk of Impulse Robotics (1)

	Operating Income of $300,000	Operating Income of $360,000	Percentage Change
Operating income	$300,000	$360,000	+20
Less interest	100,000	100,000	0
Net income	$200,000	$260,000	+30

A 20 percent increase in operating income increases net income by $60,000, or 30 percent. What if the fixed financial costs are $150,000? A 20 percent change in operating income results in a 40 percent change in the net income, from $150,000 to $210,000:

Exhibit 10 Financial Risk of Impulse Robotics (2)

	Operating Income of $300,000	Operating Income of $360,000	Percentage Change
Operating income	$300,000	$360,000	+20
Less interest	150,000	150,000	0
Net income	$150,000	$210,000	+40

Using more debt financing, which results in higher fixed costs, increases the sensitivity of net income to changes in operating income. We can represent the sensitivity of net income to a change in operating income, continuing the notation from before and including the fixed financial cost, *C*, and the tax rate, *t*, as

$$\text{DFL} = \frac{\left[Q\left(P - V\right) - F\right]\left(1 - t\right)}{\left[Q\left(P - V\right) - F - C\right]\left(1 - t\right)} = \frac{\left[Q\left(P - V\right) - F\right]}{\left[Q\left(P - V\right) - F - C\right]} \tag{4}$$

As you can see in Equation 4, the factor that adjusts for taxes, (1 – *t*), cancels out of the equation. In other words, the DFL is not affected by the tax rate.

In the case in which operating income is $300,000 and fixed financing costs are $100,000, the degree of financial leverage is

$$\begin{array}{c}\text{DFL @}\\ \text{\$300,000 operating income}\end{array} = \frac{\$300{,}000}{\$300{,}000 - \$100{,}000} = 1.5$$

If, instead, fixed financial costs are $150,000, the DFL is equal to 2.0:

$$\begin{array}{c}\text{DFL @}\\ \text{\$300,000 operating income}\end{array} = \frac{\$300{,}000}{\$300{,}000 - \$150{,}000} = 2.0$$

Again, we need to qualify our degree of leverage by the level of operating income because DFL is different at different levels of operating income.

The greater the use of financing sources that require fixed obligations, such as interest, the greater the sensitivity of net income to changes in operating income.

EXAMPLE 2

Calculating the Degree of Financial Leverage

Global Auto also employs debt financing. If Global can borrow at 8 percent, the interest cost is €18 billion. What is the degree of financial leverage of Global Auto if 6 million cars are produced and sold?

Solution:

At 6 million cars produced and sold, operating income = €30 billion. Therefore:

$$\begin{array}{c}\text{DFL @}\\ \text{€30 billion operating income}\end{array} = \frac{\text{€30 billion}}{\text{€30 billion} - \text{€18 billion}} = 2.5$$

For every 1 percent change in operating income, net income changes 2.5 percent due to financial leverage.

Unlike operating leverage, the degree of financial leverage is most often a choice by the company's management. Whereas operating costs are very similar among companies in the same industry, competitors may decide on differing capital structures.

Companies with relatively high ratios of tangible assets to total assets may be able to use higher degrees of financial leverage than companies with relatively low ratios because the claim on the tangible assets that lenders would have in the event of a default may make lenders more confident in extending larger amounts of credit. In general, businesses with plants, land, and equipment that can be used to collateralize borrowings and businesses whose revenues have below-average business cycle sensitivity may be able to use more financial leverage than businesses without such assets and with relatively high business cycle sensitivity.

Using financial leverage generally increases the variability of return on equity (net income divided by shareholders' equity). In addition, its use by a profitable company may increase the level of return on equity. Example 3 illustrates both effects.

EXAMPLE 3

The Leveraging Role of Debt

Consider the Capital Company, which is expected to generate $1,500,000 in revenues and $500,000 in operating earnings next year. Currently, the Capital Company does not use debt financing and has assets of $2,000,000.

Suppose Capital were to change its capital structure, buying back $1,000,000 of stock and issuing $1,000,000 in debt. If we assume that interest on debt is 5 percent and income is taxed at a rate of 30 percent, what is the effect of debt financing on Capital's net income and return on equity if operating earnings may vary as much as 40 percent from expected earnings?

Exhibit 11 Return on Equity of Capital Company

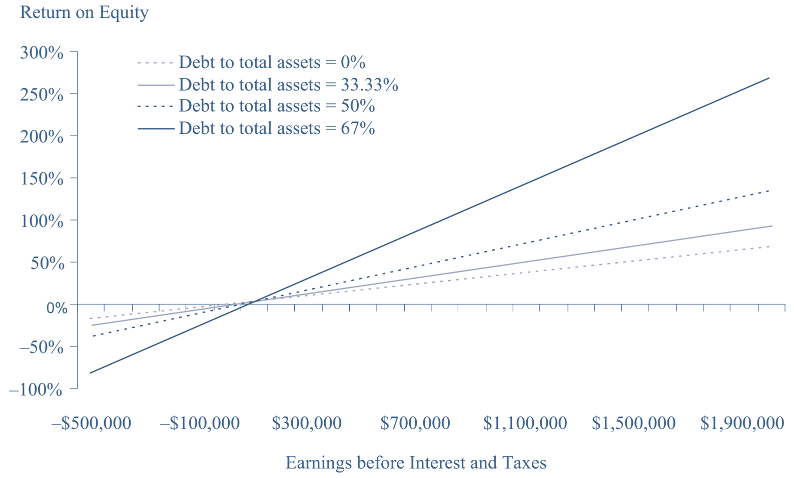
No Debt (Shareholders' Equity = $2 million)	Expected Operating Earnings, Less 40%	Expected Operating Earnings	Expected Operating Earnings, Plus 40%
Earnings before interest and taxes	$300,000	$500,000	$700,000
Interest expense	0	0	0
Earnings before taxes	$300,000	$500,000	$700,000
Taxes	90,000	150,000	210,000
Net income	$210,000	$350,000	$490,000
Return on equity[1]	10.5%	17.5%	24.5%

Debt to Total Assets = 50%; (Shareholders' Equity = $1 million)	Expected Operating Earnings, Less 40%	Expected Operating Earnings	Expected Operating Earnings, Plus 40%
Earnings before interest and taxes	$300,000	$500,000	$700,000
Interest expense	50,000	50,000	50,000
Earnings before taxes	$250,000	$450,000	$650,000
Taxes	75,000	135,000	195,000
Net income	$175,000	$315,000	$455,000
Return on equity	17.5%	31.5%	45.5%

[1] Recall that ROE is calculated as net income/shareholders' equity.

Depicting a broader array of capital structures and operating earnings, ranging from an operating loss of $500,000 to operating earnings of $2,000,000, Exhibit 12 shows the effect of leverage on the return on equity for Capital Company:

Exhibit 12 Return on Equity of Capital Company for Different Levels of Operating Earnings and Different Financing Choices



Business is generally an uncertain venture. Changes in the macroeconomic and competitive environments that influence sales and profitability are typically difficult to discern and forecast. The larger the proportion of debt in the financing mix of a business, the greater is the chance that it will face default. Similarly, the greater the proportion of debt in the capital structure, the more earnings are magnified upward in improving economic times. The bottom line? Financial leverage tends to increase the risk of ownership for shareholders.

TOTAL LEVERAGE AND THE DEGREE OF TOTAL LEVERAGE

6

a define and explain leverage, business risk, sales risk, operating risk, and financial risk and classify a risk;

b calculate and interpret the degree of operating leverage, the degree of financial leverage, and the degree of total leverage;

The degree of operating leverage gives us an idea of the sensitivity of operating income to changes in revenues. And the degree of financial leverage gives us an idea of the sensitivity of net income to changes in operating income. But often we are concerned about the combined effect of both operating leverage and financial leverage. Owners are concerned about the combined effect because both factors contribute to the risk associated with their future cash flows. And financial managers, making decisions intended to maximize owners' wealth, need to be concerned with how investment decisions (which affect the operating cost structure) and financing decisions (which affect the capital structure) affect lenders' and owners' risk.

Look back at the example of Impulse Robotics. The sensitivity of owners' cash flow to a given change in units sold is affected by both operating and financial leverage. Consider using 100,000 units as the base number produced and sold. A 10 percent increase in units sold results in a 27 percent increase in operating income and a 40 percent increase in net income; a like decrease in units sold results in a similar decrease in operating income and net income.

Exhibit 13 Total Leverage of Impulse Robotics

	Units Produced and Sold:		
	90,000	**100,000**	**110,000**
Revenues	\$900,000	\$1,000,000	\$1,100,000
Less variable costs	180,000	200,000	220,000
Less fixed costs	500,000	500,000	500,000
Operating income	\$220,000	\$300,000	\$380,000
Less interest	100,000	100,000	100,000
Net income	\$120,000	\$200,000	\$280,000
Relative to 100,000 units produced and sold			
Percentage change in units sold	–10%		+10%
Percentage change in operating profit	–27%		+27%
Percentage change in net income	–40%		+40%

Combining a company's degree of operating leverage with its degree of financial leverage results in the **degree of total leverage** (DTL), a measure of the sensitivity of net income to changes in the number of units produced and sold. We again make the simplifying assumption that a company sells all that it produces in the same period:

$$\text{DTL} = \frac{\text{Percentage change in net income}}{\text{Percentage change in the number of units sold}} \tag{5}$$

or

$$\begin{aligned}\text{DTL} &= \underbrace{\frac{Q(P-V)}{Q(P-V)-F}}_{\text{DOL}} \times \underbrace{\frac{[Q(P-V)-F]}{[Q(P-V)-F-C]}}_{\text{DFL}} \\ &= \frac{Q(P-V)}{Q(P-V)-F-C}\end{aligned} \tag{6}$$

Suppose

Number of units sold = Q = 100,000
Price per unit = P = \$10
Variable cost per unit = V = \$2
Fixed operating cost = F = \$500,000
Fixed financing cost = C = \$100,000

Then,

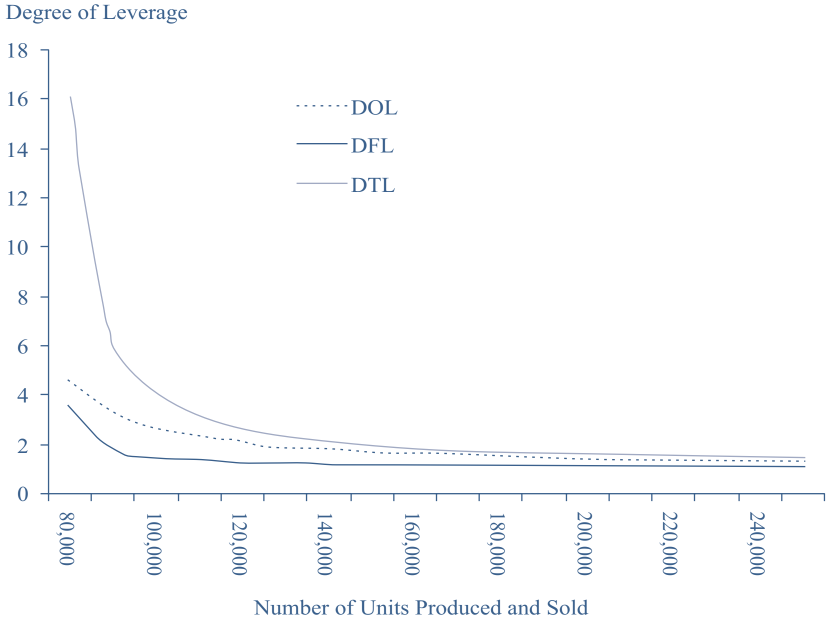
$$\text{DTL} = \frac{100{,}000(\$10-\$2)}{100{,}000(\$10-\$2)-\$500{,}000-\$100{,}000} = 4.0$$

which we could also have determined by multiplying the DOL, 2.67, by the DFL, 1.5. This means that a 1 percent increase in units sold will result in a 4 percent increase in net income; a 50 percent increase in units produced and sold results in a 200 percent increase in net income; a 5 percent decline in units sold results in a 20 percent decline in income to owners; and so on.

Because the DOL is relative to the base number of units produced and sold and the DFL is relative to the base level of operating earnings, DTL is different depending on the number of units produced and sold. We can see the DOL, DFL, and DTL for Impulse Robotics for different numbers of units produced and sold, beginning at the number of units for which the degrees are positive, in Exhibit 14.

Exhibit 14 DOL, DFL, and DTL for Different Numbers of Units Produced and Sold

P = \$10, V = \$2, F = \$500,000, C = \$100,000



In the case of operating leverage, the fixed operating costs act as a fulcrum. The greater the proportion of operating costs that are fixed, the more sensitive operating income is to changes in sales. In the case of financial leverage, the fixed financial costs, such as interest, act as a fulcrum. The greater the proportion of financing with fixed cost sources, such as debt, the more sensitive cash flows available to owners are to changes in operating income. Combining the effects of both types of leverage, we see that fixed operating and financial costs together increase the sensitivity of earnings to owners.

EXAMPLE 4

Calculating the Degree of Total Leverage

Continuing from Examples 1 and 2, Global Auto's total leverage is

$$\begin{matrix}\text{DTL @}\\ \text{6 million units}\end{matrix} = \begin{matrix}\text{DOL @}\\ \text{6 million units}\end{matrix} \times \begin{matrix}\text{DFL @}\\ \text{€30 billion}\end{matrix}$$

$$\begin{matrix}\text{DTL @}\\ \text{6 million units}\end{matrix} = \frac{6\text{ million}\left(€28{,}000 - €20{,}500\right)}{\left[6\text{ million}\left(€28{,}000 - €20{,}500\right) - €15\text{ billion}\right] - €18\text{ billion}}$$

$$= \frac{€45\text{ billion}}{€12\text{ billion}} = 3.75$$

$$\begin{matrix}\text{DTL @}\\ \text{6 million units}\end{matrix} = 1.5 \times 2.5 = 3.75$$

Given Global Auto's operating and financial leverage, a 1 percent change in unit sales changes net income by 3.75 percent.

7 BREAKEVEN POINTS AND OPERATING BREAKEVEN POINTS

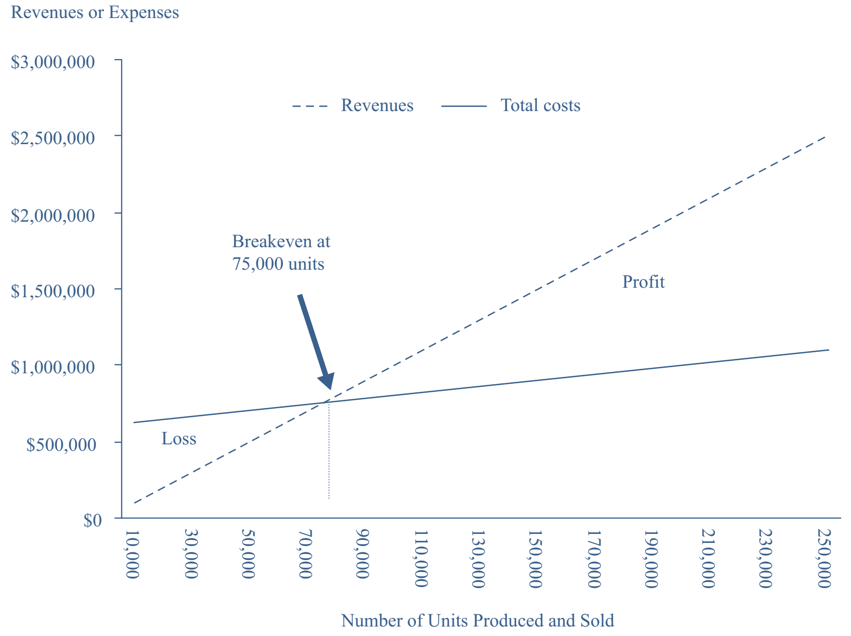
d calculate the breakeven quantity of sales and determine the company's net income at various sales levels;

e calculate and interpret the operating breakeven quantity of sales.

Looking back at Exhibit 3, we see that there is a number of units at which the company goes from being unprofitable to being profitable—that is, the number of units at which the net income is zero. This number is referred to as the breakeven point. The **breakeven point**, Q_{BE}, is the number of units produced and sold at which the company's net income is zero—the point at which revenues are equal to costs.

Plotting revenues and total costs against the number of units produced and sold, as in Exhibit 15, indicates that the breakeven is at 75,000 units. At this number of units produced and sold, revenues are equal to costs and, hence, profit is zero.

Exhibit 15 Impulse Robotics Breakeven



We can calculate this breakeven point for Impulse Robotics and Malvey Aerospace. Consider that net income is zero when the revenues are equal to the expenses. We can represent this equality of revenues and costs (summing variable operating costs, fixed operating costs, and fixed financing costs) by the following equation:

$$PQ = VQ + F + C$$

where

P = the price per unit
Q = the number of units produced and sold
V = the variable cost per unit
F = the fixed operating costs
C = the fixed financial cost

Therefore,

$$PQ_{BE} = VQ_{BE} + F + C$$

and the breakeven number of units, Q_{BE}, is[6]

$$Q_{BE} = \frac{F + C}{P - V} \tag{7}$$

In the case of Impulse Robotics and Malvey Aerospace, Impulse Robotics has a higher breakeven point. Using numbers taken from Exhibit 2:

$$\text{Impulse Robotics: } Q_{BE} = \frac{\$500{,}000 + \$100{,}000}{\$10 - \$2} = 75{,}000 \text{ units}$$

$$\text{Malvey Aerospace: } Q_{BE} = \frac{\$150{,}000 + \$50{,}000}{\$10 - \$6} = 50{,}000 \text{ units}$$

---

[6] You will notice that we did not consider taxes in our calculation of the breakeven point. This is because at the point of breakeven, taxable income is zero.

This means that Impulse Robotics must produce and sell more units to achieve a profit. So, while the higher-leveraged Impulse Robotics has a greater breakeven point relative to Malvey Aerospace, the profit that Impulse Robotics generates beyond this breakeven point is greater than that of Malvey Aerospace. Therefore, leverage has its rewards in terms of potentially greater profit, but it also increases risk.

In addition to the breakeven point specified in terms of net income, Q_{BE}, we can also specify the breakeven point in terms of operating profit, which we refer to as the **operating breakeven** point, Q_{OBE}. Revenues at the operating breakeven point are set equal to operating costs at the operating breakeven point to solve for the operating breakeven number of units, Q_{OBE}. The expression shows Q_{OBE} as equal to fixed operating costs divided by the difference between price per unit and variable cost per unit:

$$PQ_{OBE} = VQ_{OBE} + F$$

$$Q_{OBE} = \frac{F}{P - V}$$

For the two companies in our example, Impulse Robotics and Malvey Aerospace, the operating breakevens are 62,500 and 37,500 units, respectively:

$$\text{Impulse Robotics: } Q_{OBE} = \frac{\$500{,}000}{\$10 - \$2} = 62{,}500 \text{ units}$$

$$\text{Malvey Aerospace: } Q_{OBE} = \frac{\$150{,}000}{\$10 - \$6} = 37{,}500 \text{ units}$$

Impulse Robotics has a higher operating breakeven point in terms of the number of units produced and sold.

EXAMPLE 5

Calculating Operating Breakeven and Breakeven Points

Continuing with his analysis, Kenigswald considers the effect of a possible downturn on Global Auto's earnings. He divides the fixed operating costs of €15 billion by the per unit contribution margin:

$$Q_{OBE} = \frac{\text{€15 billion}}{\text{€28,000} - \text{€20,500}} = 2 \text{ million cars}$$

The operating breakeven for Global is 2,000,000 cars, or €56 billion in revenues. We calculate the breakeven point by dividing fixed operating costs, plus interest costs, by the contribution margin:

$$Q_{BE} = \frac{\text{€15 billion} + \text{€18 billion}}{\text{€28,000} - \text{€20,500}} = 4{,}400{,}000$$

Considering the degree of total leverage, Global's breakeven is 4.4 million cars, or revenues of €123.2 billion.

We can verify these calculations by constructing an income statement for the breakeven sales (in € billions):

	2,000,000 Cars	**4,400,000 Cars**
Revenues (= $P \times Q$)	€56.0	€123.2
Variable operating costs (= $V \times Q$)	41.0	90.2
Fixed operating costs (F)	15.0	15.0
Operating income	**€0**	€18.0
Fixed financial costs (C)	18.0	18.0
Net income	–€18.0	**€0**

As business expands or contracts beyond or below breakeven points, fixed costs do not change. The breakeven points for companies with low operating and financial leverage are less important than those for companies with high leverage. Companies with greater total leverage must generate more revenue to cover fixed operating and financing costs. The farther unit sales are from the breakeven point for high-leverage companies, the greater the magnifying effect of this leverage.

THE RISKS OF CREDITORS AND OWNERS

8

As we discussed earlier, business risk refers to the effect of economic conditions as well as the level of operating leverage. Uncertainty about demand, output prices, and costs are among the many factors that affect business risk. When conditions change for any of these factors, companies with higher business risk experience more volatile earnings. Financial risk is the additional risk that results from the use of debt and preferred stock. The degree of financial risk grows with greater use of debt. Who bears this risk?

The risk for providers of equity and debt capital differs because of the relative rights and responsibilities associated with the use of borrowed money in a business. Lenders have a prior claim on assets relative to shareholders, so they have greater security. In return for lending money to a business, lenders require the payment of interest and principal when due. These contractual payments to lenders must be made regardless of the profitability of the business. A business must satisfy these claims in a timely fashion or face the pain of bankruptcy should it default. In return for their higher priority in claims, lenders get predefined yet limited returns.

In contrast, equity providers claim whatever is left over after all expenses, including debt service, have been paid. So, unlike the fixed and known commitments to the lenders, what is left over for the owners may be a great deal or may be nothing. In exchange for this risk, providers of equity capital exercise the decision-making power over the business, including the right to hire, guide, and if necessary, fire managers. In public companies, ownership rights are usually exercised through an elected board of directors. They undertake the decisions over what portion of the business's earnings should be paid out as dividends for common shareholders.

Legal codes in most countries provide for these rights, as well as conditions for companies to file for bankruptcy (with reference to businesses, often called insolvency). A number of bankruptcy codes provide in some form for two categories of bankruptcies. One form provides for a temporary protection from creditors so that a viable business may reorganize. In the United States, the US Bankruptcy Code sets the terms for the form of negotiated **reorganization** of a company's capital structure that allows it to remain a going concern in Chapter 11.[7] For businesses that are not viable, the second form of bankruptcy process allows for the orderly satisfaction of the creditors' claims. In the United States, this form of bankruptcy is referred to as **liquidation**.[8] Whereas both types of bankruptcy lead to major dislocations in the rights and privileges of owners, lenders, employees, and managers, it is in this latter category of bankruptcy that the original business ceases to exist.

---

[7] US Code, Title 11—Bankruptcy, Chapter 11—Reorganization. Companies filing for bankruptcy under this code are referred to as having filed for Chapter 11 bankruptcy.

[8] US Code, Title 11—Bankruptcy, Chapter 7—Liquidation.

The difference between a company that reorganizes and emerges from bankruptcy and one that is liquidated is often the difference between operating and financial leverage. Companies with high operating leverage have less flexibility in making changes, and bankruptcy protection does little to help reduce operating costs. Companies with high financial leverage use bankruptcy laws and protection to change their capital structure and, once the restructuring is complete, can emerge as ongoing concerns.

EXAMPLE 6

Dow Corning: In and Out of Chapter 11 Bankruptcy

Dow Corning, a leading silicone producer, was a joint venture of Dow Chemical and Corning Inc. Dow Chemical filed for bankruptcy protection using Chapter 11 in 1995 as a result of the lawsuits related to silicone implants. The company was profitable at the time of filing, with more than $500 million in net income in 1994, but the potential liability from lawsuits, initially estimated around $2 billion, was significant when compared to its $4 billion in total assets. In 2004, Dow Corning emerged from bankruptcy, and in 2016 Dow Chemical completed the acquisition of 100 percent of Dow Corning, which currently operates as a wholly owned subsidiary of Dow Chemical Company.

EXAMPLE 7

Retailers Do Not Deliver

Traditional brick-and-mortar retail stores have been challenged with competition from online retailers. Whereas some retailers successfully added online access, others have struggled to compete effectively. A number of US retail stores filed for bankruptcy protection in 2017 and 2018, including Radio Shack, Nine West, Toys R Us, Brookstone, Payless, and hhgregg. Though some of these retailers were “reborn” online (e.g., hhgregg.com), others liquidated (e.g., Toys R Us), and many are closing stores and working closely with creditors to stave off liquidation.

Whereas the ability to file for bankruptcy is important to the economy, the goal of most investors is to avoid ownership of companies that are heading toward this extreme step, as well as to be able to evaluate opportunities among companies already in bankruptcy. Under both Chapter 7 and Chapter 11, providers of equity capital generally lose all value during the bankruptcy. On the other hand, debtholders typically receive at least a portion of their capital, but the payments of principal and interest are delayed during the period of bankruptcy protection.

SUMMARY

In this reading, we have reviewed the fundamentals of business risk, financial risk, and measures of leverage.

- Leverage is the use of fixed costs in a company’s cost structure. Business risk is the risk associated with operating earnings and reflects both sales risk (uncertainty with respect to the price and quantity of sales) and operating risk (the risk related to the use of fixed costs in operations). Financial risk is the risk associated with how a company finances its operations (i.e., the split between equity and debt financing of the business).

- The degree of operating leverage (DOL) is the ratio of the percentage change in operating income to the percentage change in units sold. We can use the following formula to measure the degree of operating leverage:

$$\text{DOL} = \frac{Q(P-V)}{Q(P-V)-F}$$

- The degree of financial leverage (DFL) is the percentage change in net income for a one percent change in operating income. We can use the following formula to measure the degree of financial leverage:

$$\text{DFL} = \frac{\left[Q(P-V)-F\right](1-t)}{\left[Q(P-V)-F-C\right](1-t)} = \frac{\left[Q(P-V)-F\right]}{\left[Q(P-V)-F-C\right]}$$

- The degree of total leverage (DTL) is a measure of the sensitivity of net income to changes in unit sales, which is equivalent to DTL = DOL × DFL.
- The breakeven point, Q_{BE}, is the number of units produced and sold at which the company's net income is zero, which we calculate as

$$Q_{\text{BE}} = \frac{F+C}{P-V}$$

- The operating breakeven point, Q_{OBE}, is the number of units produced and sold at which the company's operating income is zero, which we calculate as

$$Q_{\text{OBE}} = \frac{F}{P-V}$$

PRACTICE PROBLEMS

1 If two companies have identical unit sales volume and operating risk, they are *most likely* to also have identical:

A sales risk.

B business risk.

C sensitivity of operating earnings to changes in the number of units produced and sold.

2 Degree of operating leverage is *best* described as a measure of the sensitivity of:

A net earnings to changes in sales.

B fixed operating costs to changes in variable costs.

C operating earnings to changes in the number of units produced and sold.

3 The Fulcrum Company produces decorative swivel platforms for home televisions. If Fulcrum produces 40 million units, it estimates that it can sell them for $100 each. Variable production costs are $65 per unit and fixed production costs are $1.05 billion. Which of the following statements is *most accurate*? Holding all else constant, the Fulcrum Company would:

A generate positive operating income if unit sales were 25 million.

B have less operating leverage if fixed production costs were 10 percent greater than $1.05 billion.

C generate 20 percent more operating income if unit sales were 5 percent greater than 40 million.

4 The business risk of a particular company is *most accurately* measured by the company's:

A debt-to-equity ratio.

B efficiency in using assets to generate sales.

C operating leverage and level of uncertainty about demand, output prices, and competition.

5 Consider two companies that operate in the same line of business and have the same degree of operating leverage: the Basic Company and the Grundlegend Company. The Basic Company and the Grundlegend Company have, respectively, no debt and 50 percent debt in their capital structure. Which of the following statements is *most accurate*? Compared to the Basic Company, the Grundlegend Company has:

A a lower sensitivity of net income to changes in unit sales.

B the same sensitivity of operating income to changes in unit sales.

C the same sensitivity of net income to changes in operating income.

6 Myundia Motors now sells 1 million units at ¥3,529 per unit. Fixed operating costs are ¥1,290 million and variable operating costs are ¥1,500 per unit. If the company pays ¥410 million in interest, the levels of sales at the operating breakeven and breakeven points are, respectively:

A ¥1,500,000,000 and ¥2,257,612,900.

B ¥2,243,671,760 and ¥2,956,776,737.

C ¥2,975,148,800 and ¥3,529,000,000.

© 2011 CFA Institute. All rights reserved.

7 Juan Alavanca is evaluating the risk of two companies in the machinery industry: The Gearing Company and Hebelkraft, Inc. Alavanca used the latest fiscal year's financial statements and interviews with managers of the respective companies to gather the following information:

	The Gearing Company	**Hebelkraft, Inc.**
Number of units produced and sold	1 million	1.5 million
Sales price per unit	\$200	\$200
Variable cost per unit	\$120	\$100
Fixed operating cost	\$40 million	\$90 million
Fixed financing expense	\$20 million	\$20 million

Based on this information, the breakeven points for The Gearing Company and Hebelkraft, Inc. are:

A 0.75 million and 1.1 million units, respectively.

B 1 million and 1.5 million units, respectively.

C 1.5 million and 0.75 million units, respectively.

The following information relates to Questions 8–16

Mary Benn, CFA, is a financial analyst for Twin Fields Investments, located in Storrs, Connecticut, USA. She has been asked by her supervisor, Bill Cho, to examine two small Japanese cell phone component manufacturers: 4G, Inc. and Qphone Corp. Cho indicates that his clients are most interested in the use of leverage by 4G and Qphone. Benn states, "I will have to specifically analyze each company's respective business risk, sales risk, operating risk, and financial risk." "Fine, I'll check back with you shortly," Cho, answers.

Benn begins her analysis by examining the sales prospects of the two firms. The results of her sales analysis appear in Exhibit 1. She also expects very little price variability for these cell phones. She next gathers more data on these two companies to assist her analysis of their operating and financial risk.

When Cho inquires as to her progress Benn responds, "I have calculated Qphone's degree of operating leverage (DOL) and degree of financial leverage (DFL) at Qphone's 2009 level of unit sales. I have also calculated Qphone's breakeven level for unit sales. I will have 4G's leverage results shortly."

Cho responds, "Good, I will call a meeting of some potential investors for tomorrow. Please help me explain these concepts to them, and the differences in use of leverage by these two companies. In preparation for the meeting, I have a number of questions":

- "You mentioned business risk; what is included in that?"
- "How would you classify the risk due to the varying mix of variable and fixed costs?"
- "Could you conduct an analysis and tell me how the two companies will fare relative to each other in terms of net income if their unit sales increased by 10 percent above their 2009 unit sales levels?"
- "Finally, what would be an accurate verbal description of the degree of total leverage?"

The relevant data for analysis of 4G is contained in Exhibit 2, and Benn's analysis of the Qphone data appears in Exhibit 3:

Exhibit 1 Benn's Unit Sales Estimates for 4G, Inc. and Qphone Corp.

Company	2009 Unit Sales	Standard Deviation of Unit Sales	2010 Expected Unit Sales Growth Rate (%)
4G, Inc.	1,000,000	25,000	15
Qphone Corp.	1,500,000	10,000	15

Exhibit 2 Sales, Cost, and Expense Data for 4G, Inc. (At Unit Sales of 1,000,000)

Number of units produced and sold	1,000,000
Sales price per unit	¥108
Variable cost per unit	¥72
Fixed operating cost	¥22,500,000
Fixed financing expense	¥9,000,000

Exhibit 3 Benn's Analysis of Qphone (At Unit Sales of 1,500,000)

Degree of operating leverage	1.40
Degree of financial leverage	1.15
Breakeven quantity (units)	571,429

8 Based on Benn's analysis, 4G's sales risk relative to Qphone's is *most likely* to be:

A lower.

B equal.

C higher.

9 What is the *most appropriate* response to Cho's question regarding the components of business risk?

A Sales risk and financial risk.

B Operating risk and sales risk.

C Financial risk and operating risk.

10 The *most appropriate* response to Cho's question regarding the classification of risk arising from the mixture of variable and fixed costs is:

A sales risk.

B financial risk.

C operating risk.

11 Based on the information in Exhibit 2, the degree of operating leverage (DOL) of 4G, Inc., at unit sales of 1,000,000, is *closest* to:

A 1.60.

B 2.67.

C 3.20.

12 Based on the information in Exhibit 2, 4G, Inc.'s degree of financial leverage (DFL), at unit sales of 1,000,000, is *closest* to:

A 1.33.

B 2.67.

C 3.00.

13 Based on the information in Exhibit 1 and Exhibit 3, Qphone's expected percentage change in operating income for 2010 is *closest* to:

A 17.25%.

B 21.00%.

C 24.30%.

14 4G's breakeven quantity of unit sales is *closest* to:

A 437,500 units.

B 625,000 units.

C 875,000 units.

15 In response to Cho's question regarding an increase in unit sales above 2009 unit sales levels, it is *most likely* that 4G's net income will increase at:

A a slower rate than Qphone's.

B the same rate as Qphone's.

C a faster rate than Qphone's.

16 The *most appropriate* response to Cho's question regarding a description of the degree of total leverage is that degree of total leverage is:

A the percentage change in net income divided by the percentage change in units sold.

B the percentage change in operating income divided by the percentage change in units sold.

C the percentage change in net income divided by the percentage change in operating income.

SOLUTIONS

1 C is correct. The companies' degree of operating leverage should be the same, consistent with C. Sales risk refers to the uncertainty of the number of units produced and sold and the price at which units are sold. Business risk is the joint effect of sales risk and operating risk.

2 C is correct. The degree of operating leverage is the elasticity of operating earnings with respect to the number of units produced and sold. As an elasticity, the degree of operating leverage measures the sensitivity of operating earnings to a change in the number of units produced and sold.

3 C is correct. Because DOL is 4, if unit sales increase by 5 percent, Fulcrum's operating earnings are expected to increase by 4 × 5% = 20%. The calculation for DOL is:

$$\begin{aligned}\text{DOL} &= \frac{(40\text{ million})(\$100-\$65)}{\left[(40\text{ million})(\$100-\$65)\right]-\$1.05\text{ billion}}\\ &= \frac{\$1.400\text{ billion}}{\$1.400\text{ billion}-\$1.05\text{ billion}} = \frac{\$1.4}{\$0.35} = 4\end{aligned}$$

4 C is correct. Business risk reflects operating leverage and factors that affect sales (such as those given).

5 B is correct. Grundlegend's degree of operating leverage is the same as Basic Company's, whereas Grundlegend's degree of total leverage and degree of financial leverage are higher.

6 B is correct.

$$\text{Operating breakeven units} = \frac{¥1{,}290\text{ million}}{(¥3{,}529-¥1{,}500)} = 635{,}781.173\text{ units}$$

$$\text{Operating breakeven sales} = ¥3{,}529 \times 635{,}781.173\text{ units} = ¥2{,}243{,}671{,}760$$

or

$$\text{Operating breakeven sales} = \frac{¥1{,}290\text{ million}}{1-(¥1{,}500/¥3{,}529)} = ¥2{,}243{,}671{,}760$$

$$\begin{aligned}\text{Total breakeven} &= \frac{¥1{,}290\text{ million}+¥410\text{ million}}{(¥3{,}529-¥1{,}500)} = \frac{¥1{,}700\text{ million}}{¥2{,}029}\\ &= 837{,}851.1582\text{ units}\end{aligned}$$

$$\text{Breakeven sales} = ¥3{,}529 \times 837{,}851.1582\text{ units} = ¥2{,}956{,}776{,}737$$

or

$$\text{Breakeven sales} = \frac{¥1{,}700\text{ million}}{1-(¥1{,}500/¥3{,}529)} = ¥2{,}956{,}776{,}737$$

7 A is correct. For The Gearing Company,

$$Q_{\text{BE}} = \frac{F+C}{P-V} = \frac{\$40\text{ million}+\$20\text{ million}}{\$200-\$120} = 750{,}000$$

For Hebelkraft, Inc.,

$$Q_{\text{BE}} = \frac{F+C}{P-V} = \frac{\$90\text{ million}+\$20\text{ million}}{\$200-\$100} = 1{,}100{,}000$$

8 C is correct. Sales risk is defined as uncertainty with respect to the price or quantity of goods and services sold. 4G has a higher standard deviation of unit sales than Qphone; in addition, 4G's standard deviation of unit sales stated as a fraction of its level of unit sales, at 25,000/1,000,000 = 0.025, is greater than the comparable ratio for Qphone, 10,000/1,500,000 = 0.0067.

9 B is correct. Business risk is associated with operating earnings. Operating earnings are affected by sales risk (uncertainty with respect to price and quantity), and operating risk (the operating cost structure and the level of fixed costs).

10 C is correct. Operating risk refers to the risk arising from the mix of fixed and variable costs.

11 B is correct. $\text{DOL} = \dfrac{Q(P-V)}{Q(P-V)-F}$

$$\begin{array}{c}\text{DOL @}\\ \text{1,000,000 units}\end{array} = \frac{1{,}000{,}000(¥108-¥72)}{1{,}000{,}000(¥108-¥72)-¥22{,}500{,}000} = 2.67$$

12 C is correct. Degree of financial leverage is

$$\begin{aligned}\text{DFL} &= \frac{[Q(P-V)-F]}{[Q(P-V)-F-C]}\\ &= \frac{1{,}000{,}000(¥108-¥72)-¥22{,}500{,}000}{1{,}000{,}000(¥108-¥72)-¥22{,}500{,}000-¥9{,}000{,}000} = 3.00\end{aligned}$$

13 B is correct. The degree of operating leverage of Qphone is 1.4. The percentage change in operating income is equal to the DOL times the percentage change in units sold, therefore:

$$\begin{array}{c}\text{Percentage change}\\ \text{in operating income}\end{array} = (\text{DOL})\begin{pmatrix}\text{Percentage change}\\ \text{in units sold}\end{pmatrix} = (1.4)(15\%) = 21\%$$

14 C is correct. The breakeven quantity is computed

$$Q_{\text{BE}} = \frac{F+C}{P-V} = \frac{(¥22{,}500{,}000+¥9{,}000{,}000)}{(¥108-¥72)} = 875{,}000$$

15 C is correct. 4G, Inc.'s degree of total leverage can be shown to equal 8, whereas Qphone Corp.'s degree of total leverage is only DOL × DFL = 1.4 × 1.15 = 1.61. Therefore, a 10 percent increase in unit sales will mean an 80 percent increase in net income for 4G, but only a 16.1 percent increase in net income for Qphone Corp. The calculation for 4G, Inc.'s DTL is

$$\begin{aligned}\text{DTL} &= \frac{Q(P-V)}{Q(P-V)-F-C}\\ &= \frac{1{,}000{,}000(¥108-¥72)}{1{,}000{,}000(¥108-¥72)-¥22{,}500{,}000-¥9{,}000{,}000} = 8.00\end{aligned}$$

16 A is correct. Degree of total leverage is defined as the percentage change in net income divided by the percentage change in units sold.

Equity Investments

STUDY SESSIONS

Study Session 11	Equity Investments (1)
Study Session 12	Equity Investments (2)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to describe characteristics of equity investments, security markets, and indexes. The candidate should also be able to analyze industries, companies, and equity securities and to describe and demonstrate the use of basic equity valuation models.

Global equities are an important asset class for meeting longer term growth and diversification objectives. Global equities also represent a substantial share of capital markets that has been expanding in breadth and depth as developing economies come to market for equity capital. As developed and emerging economies continue to open their markets to investment, their activity is expected to significantly change the composition of world equity markets.

© 2021 CFA Institute. All rights reserved.

EQUITY INVESTMENTS

STUDY SESSION

11

Equity Investments (1)

This study session provides a structural overview of financial markets and their operating characteristics. Overview markets include equities, fixed income, derivatives, and alternative investments. Various asset types, market participants, and how assets trade within these markets and ecosystems are described. Coverage of these core asset classes continues in subsequent Level I study sessions, laying the foundation for further study in Levels II and III. The study session then turns to the calculation, construction, and use of security market indexes. A discussion of market efficiency and the degree to which market prices may reflect available information concludes the session.

READING ASSIGNMENTS

Reading 33	Market Organization and Structure by Larry Harris, PhD, CFA
Reading 34	Security Market Indexes by Paul D. Kaplan, PhD, CFA, and Dorothy C. Kelly, CFA
Reading 35	Market Efficiency by Sean Cleary, PhD, CFA, Howard J. Atkinson, CIMA, ICD.D, CFA, and Pamela Peterson Drake, PhD, CFA

© 2021 CFA Institute. All rights reserved.

READING

33

Market Organization and Structure

by Larry Harris, PhD, CFA

Larry Harris, PhD, CFA, is at the USC Marshall School of Business (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** explain the main functions of the financial system;
☐	**b.** describe classifications of assets and markets;
☐	**c.** describe the major types of securities, currencies, contracts, commodities, and real assets that trade in organized markets, including their distinguishing characteristics and major subtypes;
☐	**d.** describe types of financial intermediaries and services that they provide;
☐	**e.** compare positions an investor can take in an asset;
☐	**f.** calculate and interpret the leverage ratio, the rate of return on a margin transaction, and the security price at which the investor would receive a margin call;
☐	**g.** compare execution, validity, and clearing instructions;
☐	**h.** compare market orders with limit orders;
☐	**i.** define primary and secondary markets and explain how secondary markets support primary markets;
☐	**j.** describe how securities, contracts, and currencies are traded in quote-driven, order-driven, and brokered markets;
☐	**k.** describe characteristics of a well-functioning financial system;
☐	**l.** describe objectives of market regulation.

© 2019 CFA Institute. All rights reserved.

1

INTRODUCTION

Financial analysts gather and process information to make investment decisions, including those related to buying and selling assets. Generally, the decisions involve trading securities, currencies, contracts, commodities, and real assets such as real estate. Consider several examples:

- Fixed income analysts evaluate issuer credit-worthiness and macroeconomic prospects to determine which bonds and notes to buy or sell to preserve capital while obtaining a fair rate of return.
- Stock analysts study corporate values to determine which stocks to buy or sell to maximize the value of their stock portfolios.
- Corporate treasurers analyze exchange rates, interest rates, and credit conditions to determine which currencies to trade and which notes to buy or sell to have funds available in a needed currency.
- Risk managers work for producers or users of commodities to calculate how many commodity futures contracts to buy or sell to manage inventory risks.

Financial analysts must understand the characteristics of the markets in which their decisions will be executed. This reading, by examining those markets from the analyst's perspective, provides that understanding.

This reading is organized as follows. Section 2 examines the functions of the financial system. Sections 3–6 introduce assets that investors, information-motivated traders, and risk managers use to advance their financial objectives and presents ways practitioners classify these assets into markets. These assets include such financial instruments as securities, currencies, and some contracts; certain commodities; and real assets. Financial analysts must know the distinctive characteristics of these trading assets.

Sections 7–9 is an overview of financial intermediaries (entities that facilitate the functioning of the financial system). Sections 10–11 discuss the positions that can be obtained while trading assets. You will learn about the benefits and risks of long and short positions, how these positions can be financed, and how the financing affects their risks. Section2 12–13 discuss how market participants order trades and how markets process those orders. These processes must be understood to achieve trading objectives while controlling transaction costs.

Section 14 focuses on describing primary markets. Section 15 describes the structures of secondary markets in securities. Sections 16 and 16 close the reading with discussions of the characteristics of a well-functioning financial system and of how regulation helps make financial markets function better. A summary reviews the reading's major ideas and points, and practice problems conclude.

2

THE FUNCTIONS OF THE FINANCIAL SYSTEM

a explain the main functions of the financial system;

The financial system includes markets and various financial intermediaries that help transfer financial assets, real assets, and financial risks in various forms from one entity to another, from one place to another, and from one point in time to another. These transfers take place whenever someone exchanges one asset or financial contract for another. The assets and contracts that people (people act on behalf of themselves, companies, charities, governments, etc., so the term "people" has a broad definition in this reading) trade include notes, bonds, stocks, exchange-traded funds, currencies, forward contracts, futures contracts, option contracts, swap contracts, and certain commodities. When the buyer and seller voluntarily arrange their trades, as is usually the case, the buyer and the seller both expect to be better off.

People use the financial system for six main purposes:

1. to save money for the future;
2. to borrow money for current use;
3. to raise equity capital;
4. to manage risks;
5. to exchange assets for immediate and future deliveries; and
6. to trade on information.

The main functions of the financial system are to facilitate:

1. the achievement of the purposes for which people use the financial system;
2. the discovery of the rates of return that equate aggregate savings with aggregate borrowings; and
3. the allocation of capital to the best uses.

These functions are extremely important to economic welfare. In a well-functioning financial system, transaction costs are low, analysts can value savings and investments, and scarce capital resources are used well.

Sections 2.1 through 2.3 expand on these three functions. The six subsections of Section 2.1 cover the six main purposes for which people use the financial system and how the financial system facilitates the achievement of those purposes. Sections 2.2 and 2.3 discuss determining rates of return and capital allocation efficiency, respectively.

2.1 Helping People Achieve Their Purposes in Using the Financial System

People often arrange transactions to achieve more than one purpose when using the financial system. For example, an investor who buys the stock of an oil producer may do so to move her wealth from the present to the future, to hedge the risk that she will have to pay more for energy in the future, and to exploit insightful research that she conducted that suggests the company's stock is undervalued in the marketplace. If the investment proves to be successful, she will have saved money for the future, managed her energy risk exposure, and obtained a return on her research.

The separate discussions of each of the six main uses of the financial system by people will help you better identify the reasons why people trade. Your ability to identify the various uses of the financial system will help you avoid confusion that often leads to poor financial decisions. The financial intermediaries that are mentioned in these discussions are explained further in Sections 7–9.

2.1.1 *Saving*

People often have money that they choose not to spend now and that they want available in the future. For example, workers who save for their retirements need to move some of their current earnings into the future. When they retire, they will use their savings to replace the wages that they will no longer be earning. Similarly, companies save money from their sales revenue so that they can pay vendors when their bills come due, repay debt, or acquire assets (for example, other companies or machinery) in the future.

To move money from the present to the future, savers buy notes, certificates of deposit, bonds, stocks, mutual funds, or real assets such as real estate. These alternatives generally provide a better expected rate of return than simply storing money. Savers then sell these assets in the future to fund their future expenditures. When savers commit money to earn a financial return, they commonly are called investors. They invest when they purchase assets, and they divest when they sell them.

Investors require a fair rate of return while their money is invested. The required fair rate of return compensates them for the use of their money and for the risk that they may lose money if the investment fails or if inflation reduces the real value of their investments.

The financial system facilitates savings when institutions create investment vehicles, such as bank deposits, notes, stocks, and mutual funds, that investors can acquire and sell without paying substantial transaction costs. When these instruments are fairly priced and easy to trade, investors will use them to save more.

2.1.2 *Borrowing*

People, companies, and governments often want to spend money now that they do not have. They can obtain money to fund projects that they wish to undertake now by borrowing it. Companies can also obtain funds by selling ownership or equity interests (covered in Section 2.1.3). Banks and other investors provide those requiring funds with money because they expect to be repaid with interest or because they expect to be compensated with future disbursements, such as dividends and capital gains, as the ownership interest appreciates in value.

People may borrow to pay for such items as vacations, homes, cars, or education. They generally borrow through mortgages and personal loans, or by using credit cards. People typically repay these loans with money they earn later.

Companies often require money to fund current operations or to engage in new capital projects. They may borrow the needed funds in a variety of ways, such as arranging a loan or a line of credit with a bank, or selling fixed income securities to investors. Companies typically repay their borrowing with income generated in the future. In addition to borrowing, companies may raise funds by selling ownership interests.

Governments may borrow money to pay salaries and other expenses, to fund projects, to provide welfare benefits to their citizens and residents, and to subsidize various activities. Governments borrow by selling bills, notes, or bonds. Governments repay their debt using future revenues from taxes and in some instances from the projects funded by these debts.

Borrowers can borrow from lenders only if the lenders believe that they will be repaid. If the lenders believe, however, that repayment in full with interest may not occur, they will demand higher rates of interest to cover their expected losses and to compensate them for the discomfit they experience wondering whether they will lose their money. To lower the costs of borrowing, borrowers often pledge assets as collateral for their loans. The assets pledged as collateral often include those that will be purchased by the proceeds of the loan. If the borrowers do not repay their loans, the lenders can sell the collateral and use the proceeds to settle the loans.

Lenders often will not loan to borrowers who intend to invest in risky projects, especially if the borrowers cannot pledge other collateral. Investors may still be willing to supply capital for these risky projects if they believe that the projects will likely produce valuable future cash flows. Rather than lending money, however, they will contribute capital in exchange for equity in the projects.

The financial system facilitates borrowing. Lenders aggregate from savers the funds that borrowers require. Borrowers must convince lenders that they can repay their loans, and that, in the event they cannot, lenders can recover most of the funds lent. Credit bureaus, credit rating agencies, and governments promote borrowing; credit bureaus and credit rating agencies do so by collecting and disseminating information that lenders need to analyze credit prospects and governments do so by establishing bankruptcy codes and courts that define and enforce the rights of borrowers and lenders. When the transaction costs of loans (i.e., the costs of arranging, monitoring, and collecting them) are low, borrowers can borrow more to fund current expenditures with credible promises to return the money in the future.

2.1.3 *Raising Equity Capital*

Companies often raise money for projects by selling (issuing) ownership interests (e.g., corporate common stock or partnership interests). Although these equity instruments legally represent ownership in companies rather than loans to the companies, selling equity to raise capital is simply another mechanism for moving money from the future to the present. When shareholders or partners contribute capital to a company, the company obtains money in the present in exchange for equity instruments that will be entitled to distributions in the future. Although the repayment of the money is not scheduled as it would be for loans, equity instruments also represent potential claims on money in the future.

The financial system facilitates raising equity capital. Investment banks help companies issue equities, analysts value the securities that companies sell, and regulatory reporting requirements and accounting standards attempt to ensure the production of meaningful financial disclosures. The financial system helps promote capital formation by producing the financial information needed to determine fair prices for equity. Liquid markets help companies raise capital. In these markets, shareholders can easily divest their equities as desired. When investors can easily value and trade equities, they are more willing to fund reasonable projects that companies wish to undertake.

EXAMPLE 1

Financing Capital Projects

As a chief financial officer (CFO) of a large industrial firm, you need to raise cash within a few months to pay for a project to expand existing and acquire new manufacturing facilities. What are the primary options available to you?

Solution:

Your primary options are to borrow the funds or to raise the funds by selling ownership interests. If the company borrows the funds, you may have the company pledge some or all of the project as collateral to reduce the cost of borrowing.

2.1.4 *Managing Risks*

Many people, companies, and governments face financial risks that concern them. These risks include default risk and the risk of changes in interest rates, exchange rates, raw material prices, and sale prices, among many other risks. These risks are often managed by trading contracts that serve as hedges for the risks.

For example, a farmer and a food processor both face risks related to the price of grain. The farmer fears that prices will be lower than expected when his grain is ready for sale whereas the food processor fears that prices will be higher than expected when she has to buy grain in the future. They both can eliminate their exposures to these risks if they enter into a binding forward contract for the farmer to sell a specified quantity of grain to the food processor at a future date at a mutually agreed upon price. By entering into a forward contract that sets the future trade price, they both eliminate their exposure to changing grain prices.

In general, hedgers trade to offset or insure against risks that concern them. In addition to forward contracts, they may use futures contracts, option contracts, or insurance contracts to transfer risk to other entities more willing to bear the risks (these contracts will be covered in Section 6). Often the hedger and the other entity face exactly the opposite risks, so the transfer makes both more secure, as in the grain example.

The financial system facilitates risk management when liquid markets exist in which risk managers can trade instruments that are correlated (or inversely correlated) with the risks that concern them without incurring substantial transaction costs. Investment banks, exchanges, and insurance companies devote substantial resources to designing such contracts and to ensuring that they will trade in liquid markets. When such markets exist, people are better able to manage the risks that they face and often are more willing to undertake risky activities that they expect will be profitable.

2.1.5 *Exchanging Assets for Immediate Delivery (Spot Market Trading)*

People and companies often trade one asset for another that they rate more highly or, equivalently, that is more useful to them. They may trade one currency for another currency, or money for a needed commodity or right. Following are some examples that illustrate these trades:

- Volkswagen pays its German workers in euros, but the company receives dollars when it sells cars in the United States. To convert money from dollars to euros, Volkswagen trades in the foreign exchange markets.
- A Mexican investor who is worried about the prospects for peso inflation or a potential devaluation of the peso may buy gold in the spot gold market. (This transaction may hedge against the risk of devaluation of the peso because the value of gold may increase with inflation.)
- A plastic producer must buy carbon credits to emit carbon dioxide when burning fuel to comply with environmental regulations. The carbon credit is a legal right that the producer must have to engage in activities that emit carbon dioxide.

In each of these cases, the trades are considered spot market trades because the instruments trade for immediate delivery. The financial system facilitates these exchanges when liquid spot markets exist in which people can arrange and settle trades without substantial transaction costs.

2.1.6 *Information-Motivated Trading*

Information-motivated traders trade to profit from information that they believe allows them to predict future prices. Like all other traders, they hope to buy at low prices and sell at higher prices. Unlike pure investors, however, they expect to earn a return on their information in addition to the normal return expected for bearing risk through time.

Active investment managers are information-motivated traders who collect and analyze information to identify securities, contracts, and other assets that their analyses indicate are under- or overvalued. They then buy those that they consider undervalued and sell those that they consider overvalued. If successful, they obtain a greater return than the unconditional return that would be expected for bearing the risk in their positions. The return that they expect to obtain is a conditional return earned on the basis of the information in their analyses. Practitioners often call this process active portfolio management.

Note that the distinction between pure investors and information-motivated traders depends on their motives for trading and not on the risks that they take or their expected holding periods. Investors trade to move wealth from the present to the future whereas information-motivated traders trade to profit from superior information about future values. When trading to move wealth forward, the time period may be short or long. For example, a bank treasurer may only need to move money overnight and might use money market instruments trading in an interbank funds market to accomplish that. A pension fund, however, may need to move money 30 years forward and might do that by using shares trading in a stock market. Both are investors although their expected holding periods and the risks in the instruments that they trade are vastly different.

In contrast, information-motivated traders trade because their information-based analyses suggest to them that prices of various instruments will increase or decrease in the future at a rate faster than others without their information or analytical models would expect. After establishing their positions, they hope that prices will change quickly in their favor so that they can close their positions, realize their profits, and redeploy their capital. These price changes may occur almost instantaneously, or they may take years to occur if information about the mispricing is difficult to obtain or understand.

The two categories of traders are not mutually exclusive. Investors also are often information-motivated traders. Many investors who want to move wealth forward through time collect and analyze information to select securities that will allow them to obtain conditional returns that are greater than the unconditional returns expected for securities in their asset classes. If they have rational reasons to expect that their efforts will indeed produce superior returns, they are information-motivated traders. If they consistently fail to produce such returns, their efforts will be futile, and they would have been better off simply buying and holding well-diversified portfolios.

EXAMPLE 2

Investing versus Information-Motivated Trading

The head of a large labor union with a pension fund asks you, a pension consultant, to distinguish between investing and information-motivated trading. You are expected to provide an explanation that addresses the financial problems that she faces. How would you respond?

Solution:

The object of investing for the pension fund is to move the union's pension assets from the present to the future when they will be needed to pay the union's retired pensioners. The pension fund managers will typically do this by buying stocks, bonds, and perhaps other assets. The pension fund managers expect to receive a fair rate of return on the pension fund's assets without paying excessive transaction costs and management fees. The return should compensate the fund for the risks that it bears and for the time that other people are using the fund's money.

The object of information-motivated trading is to earn a return in excess of the fair rate of return. Information-motivated traders analyze information that they collect with the hope that their analyses will allow them to predict better than others where prices will be in the future. They then buy assets that they think will produce excess returns and sell those that they think will underperform. Active investment managers are information-motivated traders.

The characteristic that most distinguishes investors from information-motivated traders is the return that they expect. Although both types of traders hope to obtain extraordinary returns, investors rationally expect to receive only fair returns during the periods of their investments. In contrast, information-motivated traders expect to make returns in excess of required fair rates of return. Of course, not all investing or information-motivated trading is successful (in other words, the actual returns may not equal or exceed the expected returns).

The financial system facilitates information-motivated trading when liquid markets allow active managers to trade without significant transaction costs. Accounting standards and reporting requirements that produce meaningful financial disclosures reduce the costs of being well informed, but do not necessarily help informed traders profit because they often compete with each other. The most profitable well-informed traders are often those that have the most unique insights into future values.

2.1.7 *Summary*

People use the financial system for many purposes, the most important of which are saving, borrowing, raising equity capital, managing risk, exchanging assets in spot markets, and information-motivated trading. The financial system best facilitates these uses when people can trade instruments that interest them in liquid markets, when institutions provide financial services at low cost, when information about assets and about credit risks is readily available, and when regulation helps ensure that everyone faithfully honors their contracts.

2.2 Determining Rates of Return

Saving, borrowing, and selling equity are all means of moving money through time. Savers move money from the present to the future whereas borrowers and equity issuers move money from the future to the present.

Because time machines do not exist, money can travel forward in time only if an equal amount of money is travelling in the other direction. This equality always occurs because borrowers and equity sellers create the securities in which savers invest. For example, the bond sold by a company that needs to move money from the future to the present is the same bond bought by a saver who needs to move money from the present to the future.

The aggregate amount of money that savers will move from the present to the future is related to the expected rate of return on their investments. If the expected return is high, they will forgo current consumption and move more money to the future. Similarly, the aggregate amount of money that borrowers and equity sellers will move from the future to the present depends on the costs of borrowing funds or of giving up ownership. These costs can be expressed as the rate of return that borrowers and equity sellers are expected to deliver in exchange for obtaining current funds. It is the same rate that savers expect to receive when delivering current funds. If this rate is low, borrowers and equity sellers will want to move more money to the present from the future. In other words, they will want to raise more funds.

Because the total money saved must equal the total money borrowed and received in exchange for equity, the expected rate of return depends on the aggregate supply of funds through savings and the aggregate demand for funds. If the rate is too high, savers will want to move more money to the future than borrowers and equity issuers will want to move to the present. The expected rate will have to be lower to discourage the savers and to encourage the borrowers and equity issuers. Conversely, if the rate is too low, savers will want to move less money forward than borrowers and equity issuers will want to move to the present. The expected rate will have to be higher to encourage the savers and to discourage the borrowers and equity issuers. Between rates too high and too low, an expected rate of return exists, in theory, in which the aggregate supply of funds for investing (supply of funds saved) and the aggregate demand for funds through borrowing and equity issuing are equal.

Economists call this rate the equilibrium interest rate. It is the price for moving money through time. Determining this rate is one of the most important functions of the financial system. The equilibrium interest rate is the only interest rate that would exist if all securities were equally risky, had equal terms, and were equally liquid. In fact, the required rates of return for securities vary by their risk characteristics, terms, and liquidity. For a given issuer, investors generally require higher rates of return for equity than for debt, for long-term securities than for short-term securities, and for illiquid securities than for liquid ones. Financial analysts recognize that all required rates of return depend on a common equilibrium interest rate plus adjustments for risk.

EXAMPLE 3

Interest Rates

For a presentation to private wealth clients by your firm's chief economist, you are asked to prepare the audience by explaining the most fundamental facts concerning the role of interest rates in the economy. You agree. What main points should you try to convey?

Solution:

Savers have money now that they will want to use in the future. Borrowers want to use money now that they do not have, but they expect that they will have money in the future. Borrowers are loaned money by savers and promise to repay it in the future.

The interest rate is the return that lenders, the savers, expect to receive from borrowers for allowing borrowers to use the savers' money. The interest rate is the price of using money.

Interest rates depend on the total amount of money that people want to borrow and the total amount of money that people are willing to lend. Interest rates are high when, in aggregate, people value having money now substantially more than they value having money in the future. In contrast, if many people with money want to use it in the future and few people presently need more money than they have, interest rates will be low.

2.3 Capital Allocation Efficiency

Primary capital markets (primary markets) are the markets in which companies and governments raise capital (funds). Companies may raise funds by borrowing money or by issuing equity. Governments may raise funds by borrowing money.

Economies are said to be **allocationally efficient** when their financial systems allocate capital (funds) to those uses that are most productive. Although companies may be interested in getting funding for many potential projects, not all projects are worth funding. One of the most important functions of the financial system is to ensure that only the best projects obtain scarce capital funds; the funds available from savers should be allocated to the most productive uses.

In market-based economies, savers determine, directly or indirectly, which projects obtain capital. Savers determine capital allocations directly by choosing which securities they will invest in. Savers determine capital allocations indirectly by giving funds to financial intermediaries that then invest the funds. Because investors fear the loss of their money, they will lend at lower interest rates to borrowers with the best credit prospects or the best collateral, and they will lend at higher rates to other borrowers with less secure prospects. Similarly, they will buy only those equities that they believe have the best prospects relative to their prices and risks.

To avoid losses, investors carefully study the prospects of the various investment opportunities available to them. The decisions that they make tend to be well informed, which helps ensure that capital is allocated efficiently. The fear of losses by investors and by those raising funds to invest in projects ensures that only the best projects tend to be funded. The process works best when investors are well informed about the prospects of the various projects.

In general, investors will fund an equity project if they expect that the value of the project is greater than its cost, and they will not fund projects otherwise. If the investor expectations are accurate, only projects that should be undertaken will be funded and all such projects will be funded. Accurate market information thus leads to efficient capital allocation.

EXAMPLE 4

Primary Market Capital Allocation

How can poor information about the value of a project result in poor capital allocation decisions?

Solution:

Projects should be undertaken only if their value is greater than their cost. If investors have poor information and overestimate the value of a project in which its true value is less than its cost, a wealth-diminishing project may be undertaken. Alternatively, if investors have poor information and underestimate the value of a project in which its true value is greater than its cost, a wealth-enhancing project may not be undertaken.

3 ASSETS AND CONTRACTS

b describe classifications of assets and markets;

People, companies, and governments use many different assets and contracts to further their financial goals and to manage their risks. The most common assets include financial assets (such as bank deposits, certificates of deposit, loans, mortgages, corporate and government bonds and notes, common and preferred stocks, real estate investment trusts, master limited partnership interests, pooled investment products, and exchange-traded funds), currencies, certain commodities (such as gold and oil), and real assets (such as real estate). The most common contracts are option, futures, forward, swap, and insurance contracts. People, companies, and governments use these assets and contracts to raise funds, to invest, to profit from information-motivated trading, to hedge risks, and/or to transfer money from one form to another.

3.1 Classifications of Assets and Markets

Practitioners often classify assets and the markets in which they trade by various common characteristics to facilitate communications with their clients, with each other, and with regulators.

The most actively traded assets are securities, currencies, contracts, and commodities. In addition, real assets are traded. Securities generally include debt instruments, equities, and shares in pooled investment vehicles. **Currencies** are monies issued by national monetary authorities. Contracts are agreements to exchange securities, currencies, commodities or other contracts in the future. Commodities include precious metals, energy products, industrial metals, and agricultural products. Real assets are tangible properties such as real estate, airplanes, or machinery. Securities, currencies, and contracts are classified as financial assets whereas commodities and real assets are classified as physical assets.

Securities are further classified as debt or equity. Debt instruments (also called fixed-income instruments) are promises to repay borrowed money. Equities represent ownership in companies. Pooled investment vehicle shares represent ownership of an undivided interest in an investment portfolio. The portfolio may include securities, currencies, contracts, commodities, or real assets. Pooled investment vehicles, such as exchange-traded funds, which exclusively own shares in other companies, generally are also considered equities.

Securities are also classified by whether they are public or private securities. Public securities are those registered to trade in public markets, such as on exchanges or through dealers. In most jurisdictions, issuers must meet stringent minimum regulatory standards, including reporting and corporate governance standards, to issue publicly traded securities.

Private securities are all other securities. Often, only specially qualified investors can purchase private equities and private debt instruments. Investors may purchase them directly from the issuer or indirectly through an investment vehicle specifically formed to hold such securities. Issuers often issue private securities when they find public reporting standards too burdensome or when they do not want to conform to the regulatory standards associated with public equity. Venture capital is private equity that investors supply to companies when or shortly after they are founded. Private securities generally are illiquid. In contrast, many public securities trade in liquid markets in which sellers can easily find buyers for their securities.

Contracts are derivative contracts if their values depend on the prices of other underlying assets. Derivative contracts may be classified as physical or financial depending on whether the underlying instruments are physical products or financial securities. Equity derivatives are contracts whose values depend on equities or indexes of equities. Fixed-income derivatives are contracts whose values depend on debt securities or indexes of debt securities.

Practitioners classify markets by whether the markets trade instruments for immediate delivery or for future delivery. Markets that trade contracts that call for delivery in the future are forward or futures markets. Those that trade for immediate delivery are called **spot markets** to distinguish them from forward markets that trade contracts on the same underlying instruments. Options markets trade contracts that deliver in the future, but delivery takes place only if the holders of the options choose to exercise them.

When issuers sell securities to investors, practitioners say that they trade in the **primary market**. When investors sell those securities to others, they trade in the **secondary market**. In the primary market, funds flow to the issuer of the security from the purchaser. In the secondary market, funds flow between traders.

Practitioners classify financial markets as money markets or capital markets. **Money markets** trade debt instruments maturing in one year or less. The most common such instruments are repurchase agreements (defined in Section 4), negotiable certificates of deposit, government bills, and commercial paper. In contrast, **capital markets** trade instruments of longer duration, such as bonds and equities, whose values depend on the credit-worthiness of the issuers and on payments of interest or dividends that will be made in the future and may be uncertain. Corporations generally finance their operations in the capital markets, but some also finance a portion of their operations by issuing short-term securities, such as commercial paper.

Finally, practitioners distinguish between **traditional investment markets** and **alternative investment markets**. Traditional investments include all publicly traded debts and equities and shares in pooled investment vehicles that hold publicly traded debts and/or equities. Alternative investments include **hedge funds**, private equities (including venture capital), commodities, real estate securities and real estate properties, securitized debts, operating leases, machinery, collectibles, and precious gems. Because these investments are often hard to trade and hard to value, they may sometimes trade at substantial deviations from their intrinsic values. The discounts compensate investors for the research that they must do to value these assets and for their inability to easily sell the assets if they need to liquidate a portion of their portfolios.

The remainder of this section describes the most common assets and contracts that people, companies, and governments trade.

EXAMPLE 5

Asset and Market Classification

The investment policy of a mutual fund only permits the fund to invest in public equities traded in secondary markets. Would the fund be able to purchase:

1. Common stock of a company that trades on a large stock exchange?
2. Common stock of a public company that trades only through dealers?
3. A government bond?
4. A single stock futures contract?
5. Common stock sold for the first time by a properly registered public company?
6. Shares in a privately held bank with €10 billion of capital?

Solution to 1:

Yes. Common stock is equity. Those common stocks that trade on large exchanges invariably are public equities that trade in secondary markets.

Solution to 2:

Yes. Dealer markets are secondary markets and the security is a public equity.

Solution to 3:

No. Although government bonds are public securities, they are not equities. They are debt securities.

Solution to 4:

No. Although the underlying instruments for single stock futures are invariably public equities, single stock futures are derivative contracts, not equities.

Solution to 5:

No. The fund would not be able to buy these shares because a purchase from the issuer would be in the primary market. The fund would have to wait until it could buy the shares from someone other than the issuer.

Solution to 6:

No. These shares are private equities, not public equities. The public prominence of the company does not make its securities public securities unless they have been properly registered as public securities.

SECURITIES

4

c describe the major types of securities, currencies, contracts, commodities, and real assets that trade in organized markets, including their distinguishing characteristics and major subtypes;

People, companies, and governments sell securities to raise money. Securities include bonds, notes, commercial paper, mortgages, common stocks, preferred stocks, warrants, mutual fund shares, unit trusts, and depository receipts. These can be classified broadly as fixed-income instruments, equities, and shares in pooled investment vehicles. Note that the legal definition of a security varies by country and may or may not coincide with the usage here. Securities that are sold to the public or that can be resold to the public are called issues. Companies and governments are the most common issuers.

4.1 Fixed Income

Fixed-income instruments contractually include predetermined payment schedules that usually include interest and principal payments. Fixed-income instruments generally are promises to repay borrowed money but may include other instruments with payment schedules, such as settlements of legal cases or prizes from lotteries. The payment amounts may be pre-specified or they may vary according to a fixed formula that depends on the future values of an interest rate or a commodity price. Bonds, notes, bills, certificates of deposit, commercial paper, repurchase agreements, loan agreements, and mortgages are examples of promises to repay money in the future. People, companies, and governments create fixed-income instruments when they borrow money.

Corporations and governments issue bonds and notes. Fixed-income securities with shorter maturities are called "notes," those with longer maturities are called "bonds." The cutoff is usually at 10 years. In practice, however, the terms are generally used interchangeably. Both become short-term instruments when the remaining time until maturity is short, usually taken to be one year or less.

Some corporations issue convertible bonds, which are typically convertible into stock, usually at the option of the holder after some period. If stock prices are high so that conversion is likely, convertibles are valued like stock. Conversely, if stock prices are low so that conversion is unlikely, convertibles are valued like bonds.

Bills, certificates of deposit, and commercial paper are respectively issued by governments, banks, and corporations. They usually mature within a year of being issued; certificates of deposit sometimes have longer initial maturities.

Repurchase agreements (repos) are short-term lending instruments. The term can be as short as overnight. A borrower seeking funds will sell an instrument—typically a high-quality bond—to a lender with an agreement to repurchase it later at a slightly higher price based on an agreed upon interest rate.

Practitioners distinguish between short-term, intermediate-term, and long-term fixed-income securities. No general consensus exists about the definition of short-term, intermediate-term, and long-term. Instruments that mature in less than one to two years are considered short-term instruments whereas those that mature in more than five to ten years are considered long-term instruments. In the middle are intermediate-term instruments.

Instruments trading in money markets are called money market instruments. Such instruments are traded debt instruments maturing in one year or less. Money market funds and corporations seeking a return on their short-term cash balances typically hold money market instruments.

4.2 Equities

Equities represent ownership rights in companies. These include common and preferred shares. Common shareholders own residual rights to the assets of the company. They have the right to receive any dividends declared by the boards of directors, and in the event of liquidation, any assets remaining after all other claims are paid. Acting through the boards of directors that they elect, common shareholders usually can select the managers who run the corporations.

Preferred shares are equities that have preferred rights (relative to common shares) to the cash flows and assets of the company. Preferred shareholders generally have the right to receive a specific dividend on a regular basis. If the preferred share is a cumulative preferred equity, the company must pay the preferred shareholders any previously omitted dividends before it can pay dividends to the common shareholders. Preferred shareholders also have higher claims to assets relative to common shareholders in the event of corporate liquidation. For valuation purposes, financial analysts generally treat preferred stocks as fixed-income securities when the issuers will clearly be able to pay their promised dividends in the foreseeable future.

Warrants are securities issued by a corporation that allow the warrant holders to buy a security issued by that corporation, if they so desire, usually at any time before the warrants expire or, if not, upon expiration. The security that warrant holders can buy usually is the issuer's common stock, in which case the warrants are considered equities because the warrant holders can obtain equity in the company by exercising their warrants. The warrant **exercise price** is the price that the warrant holder must pay to buy the security.

EXAMPLE 6

Securities

What factors distinguish fixed-income securities from equities?

Solution:

Fixed-income securities generate income on a regular schedule. They derive their value from the promise to pay a scheduled cash flow. The most common fixed-income securities are promises made by people, companies, and governments to repay loans.

Equities represent residual ownership in companies after all other claims—including any fixed-income liabilities of the company—have been satisfied. For corporations, the claims of preferred equities typically have priority over the claims of common equities. Common equities have the residual ownership in corporations.

4.3 Pooled Investments

Pooled investment vehicles are mutual funds, trusts, depositories, and hedge funds, that issue securities that represent shared ownership in the assets that these entities hold. The securities created by mutual funds, trusts, depositories, and hedge fund are respectively called *shares*, *units*, *depository receipts*, and *limited partnership interests* but practitioners often use these terms interchangeably. People invest in pooled investment vehicles to benefit from the investment management services of their managers and from diversification opportunities that are not readily available to them on an individual basis.

Mutual funds are investment vehicles that pool money from many investors for investment in a portfolio of securities. They are often legally organized as investment trusts or as corporate investment companies. Pooled investment vehicles may be open-ended or closed-ended. Open-ended funds issue new shares and redeem existing shares on demand, usually on a daily basis. The price at which a fund redeems and sells the fund's shares is based on the net asset value of the fund's portfolio, which is the difference between the fund's assets and liabilities, expressed on a per share basis. Investors generally buy and sell open-ended mutual funds by trading with the mutual fund.

In contrast, closed-end funds issue shares in primary market offerings that the fund or its investment bankers arrange. Once issued, investors cannot sell their shares of the fund back to the fund by demanding redemption. Instead, investors in closed-end funds must sell their shares to other investors in the secondary market. The secondary market prices of closed-end funds may differ—sometimes quite significantly—from their net asset values. Closed-end funds generally trade at a discount to their net asset values. The discount reflects the expenses of running the fund and sometimes investor concerns about the quality of the management. Closed-end funds may also trade at a discount or a premium to net asset value when investors believe that the portfolio securities are overvalued or undervalued. Many financial analysts thus believe that discounts and premiums on closed-end funds measure market sentiment.

Exchange-traded funds (ETFs) and exchange-traded notes (ETNs) are open-ended funds that investors can trade among themselves in secondary markets. The prices at which ETFs trade rarely differ much from net asset values because a class of investors, known as authorized participants (APs), has the option of trading directly with the ETF. If the market price of an equity ETF is sufficiently below its net asset value, APs will buy shares in the secondary market at market price and redeem shares at net asset value with the fund. Conversely, if the price of an ETF is sufficiently above its net asset value, APs will buy shares from the fund at net asset value and sell shares in the secondary market at market price. As a result, the market price and net asset values of ETFs tend to converge.

Many ETFs permit only in-kind deposits and redemptions. Buyers who buy directly from such a fund pay for their shares with a portfolio of securities rather than with cash. Similarly, sellers receive a portfolio of securities. The transaction portfolio generally is very similar—often essentially identical—to the portfolio held by the fund. Practitioners sometimes call such funds "depositories" because they issue depository receipts for the portfolios that traders deposit with them. The traders then trade the receipts in the secondary market. Some warehouses holding industrial materials and precious metals also issue tradable warehouse receipts.

Asset-backed securities are securities whose values and income payments are derived from a pool of assets, such as mortgage bonds, credit card debt, or car loans. These securities typically pass interest and principal payments received from the pool of assets through to their holders on a monthly basis. These payments may depend on formulas that give some classes of securities—called tranches—backed by the pool more value than other classes.

Hedge funds are investment funds that generally organize as limited partnerships. The hedge fund managers are the general partners. The limited partners are qualified investors who are wealthy enough and well informed enough to tolerate and accept substantial losses, should they occur. The regulatory requirements to participate in a hedge fund and the regulatory restrictions on hedge funds vary by jurisdiction. Most hedge funds follow only one investment strategy, but no single investment strategy characterizes hedge funds as a group. Hedge funds exist that follow almost every imaginable strategy ranging from long–short arbitrage in the stock markets to direct investments in exotic alternative assets.

The primary distinguishing characteristic of hedge funds is their management compensation scheme. Almost all funds pay their managers with an annual fee that is proportional to their assets and with an additional performance fee that depends on the wealth that the funds generate for their shareholders. A secondary distinguishing characteristic of many hedge funds is the use of leverage to increase risk exposure and to hopefully increase returns.

5 CURRENCIES, COMMODITIES, AND REAL ASSETS

c describe the major types of securities, currencies, contracts, commodities, and real assets that trade in organized markets, including their distinguishing characteristics and major subtypes;

Currencies are monies issued by national monetary authorities. Approximately 180 currencies are currently in use throughout the world. Some of these currencies are regarded as reserve currencies. Reserve currencies are currencies that national central banks and other monetary authorities hold in significant quantities. The primary reserve currencies are the US dollar and the euro. Secondary reserve currencies include the British pound, the Japanese yen, and the Swiss franc.

Currencies trade in foreign exchange markets. In spot currency transactions, one currency is immediately or almost immediately exchanged for another. The rate of exchange is called the spot exchange rate. Traders typically negotiate institutional trades in multiples of large quantities, such as US$1 million or ¥100 million. Institutional trades generally settle in two business days.

Retail currency trades most commonly take place through commercial banks when their customers exchange currencies at a location of the bank, use ATM machines when travelling to withdraw a different currency than the currency in which their bank accounts are denominated, or use credit cards to buy items priced in different currencies. Retail currency trades also take place at airport kiosks, at store front currency exchanges, or on the street.

5.1 Commodities

Commodities include precious metals, energy products, industrial metals, agricultural products, and carbon credits. Spot commodity markets trade commodities for immediate delivery whereas the forward and futures markets trade commodities for future delivery. Managers seeking positions in commodities can acquire them directly by trading in the spot markets or indirectly by trading forward and futures contracts.

The producers and processors of industrial metals and agricultural products are the primary users of the spot commodity markets because they generally are best able to take and make delivery and to store physical products. They undertake these activities in the normal course of operating their businesses. Their ability to handle physical products and the information that they gather operating businesses also gives them substantial advantages as information-motivated traders in these markets. Many producers employ financial analysts to help them analyze commodity market conditions so that they can best manage their inventories to hedge their operational risks and to speculate on future price changes.

Commodities also interest information-motivated traders and investment managers because they can use them as hedges against risks that they hold in their portfolios or as vehicles to speculate on future price changes. Most such traders take positions in the futures markets because they usually do not have facilities to handle most physical products nor can they easily obtain them. They also cannot easily cope with the normal variation in qualities that characterizes many commodities. Information-motivated traders and investment managers also prefer to trade in futures markets because most futures markets are more liquid than their associated spot markets and forward markets. The liquidity allows them to easily close their positions before delivery so that they can avoid handling physical products.

Some information-motivated traders and investment managers, however, trade in the spot commodity markets, especially when they can easily contract for low-cost storage. Commodities for which delivery and storage costs are lowest are nonperishable products for which the ratio of value to weight is high and variation in quality is low. These generally include precious metals, industrial diamonds, such high-value industrial metals as copper, aluminum, and mercury, and carbon credits.

5.2 Real Assets

Real assets include such tangible properties as real estate, airplanes, machinery, or lumber stands. These assets normally are held by operating companies, such as real estate developers, airplane leasing companies, manufacturers, or loggers. Many institutional investment managers, however, have been adding real assets to their portfolios as direct investments (involving direct ownership of the real assets) and indirect investments (involving indirect ownership, for example, purchase of securities of companies that invest in real assets or real estate investment trusts). Investments in real assets are attractive to them because of the income and tax benefits that they often generate, and because changes in their values may have a low correlation with other investments that the managers hold.

Direct investments in real assets generally require substantial management to ensure that the assets are maintained and used efficiently. Investment managers investing in such assets must either hire personnel to manage them or hire outside management companies. Either way, management of real assets is quite costly.

Real assets are unique properties in the sense that no two assets are alike. An example of a unique property is a real estate parcel. No two parcels are the same because, if nothing else, they are located in different places. Real assets generally differ in their conditions, remaining useful lives, locations, and suitability for various purposes. These differences are very important to the people who use them, so the market for a given real asset may be very limited. Thus, real assets tend to trade in very illiquid markets.

The heterogeneity of real assets, their illiquidity, and the substantial costs of managing them are all factors that complicate the valuation of real assets and generally make them unsuitable for most investment portfolios. These same problems, however, often cause real assets to be misvalued in the market, so astute information-motivated traders may occasionally identify significantly undervalued assets. The benefits from purchasing such assets, however, are often offset by the substantial costs of searching for them and by the substantial costs of managing them.

Many financial intermediaries create entities, such as real estate investment trusts (REITs) and master limited partnerships (MLPs), to securitize real assets and to facilitate indirect investment in real assets. The financial intermediaries manage the assets and pass through the net benefits after management costs to the investors who hold these securities. Because these securities are much more homogenous and divisible than the real assets that they represent, they tend to trade in much more liquid markets. Thus, they are much more suitable as investments than the real assets themselves.

Of course, investors seeking exposure to real assets can also buy shares in corporations that hold and operate real assets. Although almost all corporations hold and operate real assets, many specialize in assets that particularly interest investors seeking exposure to specific real asset classes. For example, investors interested in owning aircraft can buy an aircraft leasing company such as Waha Capital (Abu Dhabi Securities Exchange) and Aircastle Limited (NYSE).

EXAMPLE 7

Assets and Contracts

Consider the following assets and contracts:

Bank deposits
Certificates of deposit
Common stocks
Corporate bonds
Currencies
Exchange-traded funds
Lumber forward contracts
Crude oil futures contracts
Gold
Hedge funds
Master limited partnership interests
Mortgages
Mutual funds
Stock option contracts
Preferred stocks
Real estate parcels
Interest rate swaps
Treasury notes

1. Which of these represent ownership in corporations?
2. Which of these are debt instruments?
3. Which of these are created by traders rather than by issuers?
4. Which of these are pooled investment vehicles?
5. Which of these are real assets?
6. Which of these would a home builder most likely use to hedge construction costs?
7. Which of these would a corporation trade when moving cash balances among various countries?

Solution to 1:

Common and preferred stocks represent ownership in corporations.

Solution to 2:

Bank deposits, certificates of deposit, corporate bonds, mortgages, and Treasury notes are all debt instruments. They respectively represent loans made to banks, corporations, mortgagees (typically real estate owners), and the Treasury.

Solution to 3:

Lumber forward contracts, crude oil futures contracts, stock option contracts, and interest rate swaps are created when the seller sells them to a buyer.

Solution to 4:

Exchange-traded funds, hedge funds, and mutual funds are pooled investment vehicles. They represent shared ownership in a portfolio of other assets.

Solution to 5:

Real estate parcels are real assets.

Solution to 6:

A builder would buy lumber forward contracts to lock in the price of lumber needed to build homes.

Solution to 7:

Corporations often trade currencies when moving cash from one country to another.

CONTRACTS

6

c describe the major types of securities, currencies, contracts, commodities, and real assets that trade in organized markets, including their distinguishing characteristics and major subtypes;

A contract is an agreement among traders to do something in the future. Contracts include forward, futures, swap, option, and insurance contracts. The values of most contracts depend on the value of an **underlying** asset. The underlying asset may be a commodity, a security, an index representing the values of other instruments, a currency pair or basket, or other contracts.

Contracts provide for some physical or cash settlement in the future. In a physically settled contract, settlement occurs when the parties to the contract physically exchange some item, such as avocados, pork bellies, or gold bars. Physical settlement also includes the delivery of such financial instruments as bonds, equities, or futures contracts even though the delivery is electronic. In contrast, cash settled contracts settle through cash payments. The amount of the payment depends on formulas specified in the contracts.

Financial analysts classify contracts by whether they are physical or financial based on the nature of the underlying asset. If the underlying asset is a physical product, the contract is a physical; otherwise, the contract is a financial. Examples of assets classified as physical include contracts for the delivery of petroleum, lumber, and gold. Examples of assets classified as financial include option contracts, and contracts on interest rates, stock indexes, currencies, and credit default swaps.

Contracts that call for immediate delivery are called spot contracts, and they trade in spot markets. Immediate delivery generally is three days or less, but depends on each market. All other contracts involve what practitioners call futurity. They derive their values from events that will take place in the future.

EXAMPLE 8

Contracts for Difference

Contracts for difference (CFD) allow people to speculate on price changes for an underlying asset, such as a common stock or an index. Dealers generally sell CFDs to their clients. When the clients sell the CFDs back to their dealer, they receive any appreciation in the underlying asset's price between the time of purchase and sale (open and close) of the contract. If the underlying asset's price drops over this interval, the client pays the dealer the difference.

1 Are contracts for difference derivative contracts?

2 Are contracts for difference based on copper prices cash settled or physically settled?

Solution to 1:

Contracts for difference are derivative contracts because their values are derived from changes in the prices of the underlying asset on which they are based.

Solution to 2:

All contracts for difference are cash settled contracts regardless of the underlying asset on which they are based because they settle in cash and not in the underlying asset.

6.1 Forward Contracts

A **forward contract** is an agreement to trade the underlying asset in the future at a price agreed upon today. For example, a contract for the sale of wheat after the harvest is a forward contract. People often use forward contracts to reduce risk. Before planting wheat, farmers like to know the price at which they will sell their crop. Similarly, before committing to sell flour to bakers in the future, millers like to know the prices that they will pay for wheat. The farmer and the miller both reduce their operating risks by agreeing to trade wheat forward.

Practitioners call such traders hedgers because they use their contractual commitments to hedge their risks. If the price of wheat falls, the wheat farmer's crop will drop in value on the spot market but he has a contract to sell wheat in the future at a higher fixed price. The forward contract has become more valuable to the farmer. Conversely, if the price of wheat rises, the miller's future obligation to sell flour will become more burdensome because of the high price he would have to pay for wheat on the spot market, but the miller has a contract to buy wheat at a lower fixed price. The forward contract has become more valuable to the miller. In both cases, fluctuations in the spot price are hedged by the forward contract. The forward contract offsets the operating risks that the hedgers face.

Consider a simple example of hedging. An avocado farmer in Mexico expects to harvest 15,000 kilograms of avocados and that the price at harvest will be 60 pesos per kilogram. That price, however, could fluctuate significantly before the harvest. If the price of avocados drops to 50 pesos, the farmer would lose 10 pesos per kilogram (60 pesos − 50 pesos) relative to his expectations, or a total of 150,000 pesos. Now, suppose that the farmer can sell avocados forward to Del Rey Avocado at 58 pesos for delivery at the harvest. If the farmer sells 15,000 kilograms forward, and the price of avocados drops to 50 pesos, the farmer would still be able to sell his avocados for 58 pesos, and thus would not suffer from the drop in the price of avocados below this level.

EXAMPLE 9

Hedging Gold Production

A Zimbabwean gold producer invests in a mine expansion project on the expectation that gold prices will remain at or above $1,200 USD per ounce when the new project starts producing ore.

1 What risks does the gold producer face with respect to the price of gold?
2 How might the gold producer hedge its gold price risk?

Solution to 1:

The gold producer faces the risk that the price of gold could fall below $1,200 USD before it can sell its new production. If so, the investment in the expansion project will be less profitable than expected, and may even generate losses for the mine.

Solution to 2:

The gold producer could hedge the gold price risk by selling gold forward, hopefully at a price near $1,200 USD. Even if the price of gold falls, the gold producer would get paid the contract price.

Forward contracts are very common, but two problems limit their usefulness for many market participants. The first problem is counterparty risk. **Counterparty risk** is the risk that the other party to a contract will fail to honor the terms of the contract. Concerns about counterparty risk ensure that generally only parties who have long-standing relationships with each other execute forward contracts. Trustworthiness is critical when prices are volatile because, after a large price change, one side or the other may prefer not to settle the contract.

The second problem is liquidity. Trading out of a forward contract is very difficult because it can only be done with the consent of the other party. The liquidity problem ensures that forward contracts tend to be executed only among participants for whom delivery is economically efficient and quite certain at the time of contracting so that both parties will want to arrange for delivery.

The counterparty risk problem and the liquidity problem often make it difficult for market participants to obtain the hedging benefits associated with forward contracting. Fortunately, futures contracts have been developed to mitigate these problems.

6.2 Futures Contracts

A **futures contract** is a standardized forward contract for which a clearinghouse guarantees the performance of all traders. The buyer of a futures contract is the side that will take physical delivery or its cash equivalent. The seller of a futures contract is the side that is liable for the delivery or its cash equivalent. A **clearinghouse** is an organization that ensures that no trader is harmed if another trader fails to honor the contract. In effect, the clearinghouse acts as the buyer for every seller and as the seller for every buyer. Buyers and sellers, therefore, can trade futures without worrying whether their counterparties are creditworthy. Because futures contracts are standardized, a buyer can eliminate his obligation to buy by selling his contract to anyone. A seller similarly can eliminate her obligation to deliver by buying a contact from anyone. In either case, the clearinghouse will release the trader from all future obligations if his or her long and short positions exactly offset each other.

To protect against defaults, futures clearinghouses require that all participants post with the clearinghouse an amount of money known as **initial margin** when they enter a contract. The clearinghouse then settles the margin accounts on a daily basis. All participants who have lost on their contracts that day will have the amount of their losses deducted from their margin by the clearinghouse. The clearinghouse similarly increases margins for all participants who gained on that day. Participants whose margins drop below the required **maintenance margin** must replenish their accounts. If a participant does not provide sufficient additional margin when required, the participant's broker will immediately trade to offset the participant's position. These **variation margin** payments ensure that the liabilities associated with futures contracts do not grow large.

EXAMPLE 10

Futures Margin

NYMEX's Light Sweet Crude Oil futures contract specifies the delivery of 1,000 barrels of West Texas Intermediate Crude Oil when the contract finally settles. A broker requires that its clients post an initial overnight margin of $7,763 per contract and an overnight maintenance margin of $5,750 per contract. A client buys ten contracts at $75 per barrel through this broker. On the next day, the contract settles for $72 per barrel. How much additional margin will the client have to provide to his broker?

Solution:

The client lost three dollars per barrel (he is the side committed to take delivery or its cash equivalent at $75 per barrel). This results in a $3,000 loss on each of his 10 contracts, and a total loss of $30,000. His initial margin of $77,630 is reduced by $30,000 leaving $47,630 in his margin account. Because his account has dropped below the maintenance margin requirement of $57,500, the client will get a margin call. The client must provide an additional $30,000 = $77,630 – $47,630 to replenish his margin account; the account is replenished to the amount of the initial margin. The client will only receive another margin call if his account drops to below $57,500 again.

Futures contracts have vastly improved the efficiency of forward contracting markets. Traders can trade standardized futures contracts with anyone without worrying about counterparty risk, and they can close their positions by arranging offsetting trades. Hedgers for whom the terms of the standard contract are not ideal generally still use the futures markets because the contracts embody most of the price risk that concerns them. They simply offset (close out) their futures positions, at the same time they enter spot contracts on which they make or take ultimate delivery.

EXAMPLE 11

Forward and Futures Contracts

What feature most distinguishes futures contracts from forward contracts?

Solution:

A futures contract is a standardized forward contract for which a clearinghouse guarantees the performance of all buyers and sellers. The clearinghouse reduces the counterparty risk problem. The clearinghouse allows a buyer who has bought a contract from one person and sold the same contract to another person to net out the two obligations so that she is no longer liable for either side of the contract; the positions are closed. The ability to trade futures contracts provides liquidity in futures contracts compared with forward contracts.

6.3 Swap Contracts

A **swap contract** is an agreement to exchange payments of periodic cash flows that depend on future asset prices or interest rates. For example, in a typical **interest rate swap**, at periodic intervals, one party makes fixed cash payments to the counterparty in exchange for variable cash payments from the counterparty. The variable payments are based on a pre-specified variable interest rate such as the London Interbank Offered Rate (Libor). This swap effectively exchanges fixed interest payments for variable interest payments. Because the variable rate is set in the future, the cash flows for this contract are uncertain when the parties enter the contract.

Investment managers often enter interest rate swaps when they own a fixed long-term income stream that they want to convert to a cash flow that varies with current short-term interest rates, or vice versa. The conversion may allow them to substantially reduce the total interest rate risk to which they are exposed. Hedgers often use swap contracts to manage risks.

In a **commodity swap**, one party typically makes fixed payments in exchange for payments that depend on future prices of a commodity such as oil. In a **currency swap**, the parties exchange payments denominated in different currencies. The payments may be fixed, or they may vary depending on future interest rates in the two countries. In an **equity swap**, the parties exchange fixed cash payments for payments that depend on the returns to a stock or a stock index.

EXAMPLE 12

Swap and Forward Contracts

What feature most distinguishes a swap contract from a cash-settled forward contract?

Solution:

Both contracts provide for the exchange of cash payments in the future. A forward contract only has a single cash payment at the end that depends on an underlying price or index at the end. In contrast, a swap contract has several scheduled periodic payments, each of which depends on an underlying price or index at the time of the payment.

6.4 Option Contracts

An **option contract** allows the holder (the purchaser) of the **option** to buy or sell, depending on the type of option, an underlying instrument at a specified price at or before a specified date in the future. Those that do buy or sell are said to **exercise** their contracts. An option to buy is a **call option**, and an option to sell is a **put** option. The specified price is called the strike price (exercise price). If the holders can exercise their contracts only when they mature, they are **European-style** contracts. If they can exercise the contracts earlier, they are **American-style** contracts. Many exchanges list standardized option contracts on individual stocks, stock indexes, futures contracts, currencies, swaps, and precious metals. Institutions also trade many customized option contracts with dealers in the over-the-counter derivative market.

Option holders generally will exercise call options if the strike price is below the market price of the underlying instrument, in which case, they will be able to buy at a lower price than the market price. Similarly, they will exercise put options if the strike price is above the underlying instrument price so that they sell at a higher price than the market price. Otherwise, option holders allow their options to expire as worthless.

The price that traders pay for an option is the option premium. Options can be quite expensive because, unlike forward and futures contracts, they do not impose any liability on the holder. The premium compensates the sellers of options—called option writers—for giving the call option holders the right to potentially buy below market prices and put option holders the right to potentially sell above market prices. Because the writers must trade if the holders exercise their options, option contracts may impose substantial liabilities on the writers.

EXAMPLE 13

Option and Forward Contracts

What feature most distinguishes option contracts from forward contracts?

Solution:

The holder of an option contract has the right, but not the obligation, to buy (for a call option) or sell (for a put option) the underlying instrument at some time in the future. The writer of an option contract must trade the underlying instrument if the holder exercises the option.

In contrast, the two parties to a forward contract must trade the underlying instrument (or its equivalent value for a cash-settled contract) at some time in the future if either party wants to settle the contract.

6.5 Other Contracts

Insurance contracts pay their beneficiaries a cash benefit if some event occurs. Life, liability, and automobile insurance are examples of insurance contracts sold to retail clients. People generally use insurance contracts to compensate for losses that they will experience if bad things happen unexpectedly. Insurance contracts allow them to hedge risks that they face.

Credit default swaps (CDS) are insurance contracts that promise payment of principal in the event that a company defaults on its bonds. Bondholders use credit default swaps to convert risky bonds into more secure investments. Other creditors of the company may also buy them to hedge against the risk they will not be paid if the company goes bankrupt.

Well-informed traders who believe that a corporation will default on its bonds may buy credit default swaps written on the corporation's bonds if the swap prices are sufficiently low. If they are correct, the traders will profit if the payoff to the swap is more than the cost of buying and maintaining the swap position.

People sometimes also buy insurance contracts as investments, especially in jurisdictions where payouts from insurance contracts are not subject to as much taxation as are payouts to other investment vehicles. They may buy these contracts directly from insurance companies, or they may buy already issued contracts from their owners. For example, the life settlements market trades life insurance contracts that people sell to investors when they need cash.

FINANCIAL INTERMEDIARIES

7

d describe types of financial intermediaries and services that they provide;

Financial intermediaries help entities achieve their financial goals. These intermediaries include commercial, mortgage, and investment banks; credit unions, credit card companies, and various other finance corporations; brokers and exchanges; dealers and arbitrageurs; clearinghouses and depositories; mutual funds and hedge funds; and insurance companies. The services and products that financial intermediaries provide allow their clients to solve the financial problems that they face more efficiently than they could do so by themselves. Financial intermediaries are essential to well-functioning financial systems.

Financial intermediaries are called intermediaries because the services and products that they provide help connect buyers to sellers in various ways. Whether the connections are easy to identify or involve complex financial structures, financial intermediaries stand between one or more buyers and one or more sellers and help them transfer capital and risk between them. Financial intermediaries' activities allow buyers and sellers to benefit from trading, often without any knowledge of the other.

This section introduces the main financial intermediaries that provide services and products in well-developed financial markets. The discussion starts with those intermediaries whose services most obviously connect buyers to sellers and then proceeds to those intermediaries whose services create more subtle connections. Because many financial intermediaries provide many different types of services, some are mentioned more than once. The section concludes with a general characterization of the various ways in which financial intermediaries add value to the financial system.

7.1 Brokers, Exchanges, and Alternative Trading Systems

Brokers are agents who fill orders for their clients. They do not trade with their clients. Instead, they search for traders who are willing to take the other side of their clients' orders. Individual brokers may work for large brokerage firms, the brokerage arm of banks, or at exchanges. Some brokers match clients to clients personally. Others use specialized computer systems to identify potential trades and help their clients fill their orders. Brokers help their clients trade by reducing the costs of finding counterparties for their trades.

Block brokers provide brokerage service to large traders. Large orders are hard to fill because finding a counterparty willing to do a large trade is often quite difficult. A large buy order generally will trade at a premium to the current market price, and a large sell order generally will trade at a discount to the current market price. These price concessions encourage other traders to trade with the large traders. They also make large traders reluctant, however, to expose their orders to the public before their trades are arranged because they do not want to move the market. Block brokers, therefore, carefully manage the exposure of the orders entrusted to them, which makes filling them difficult.

Investment banks provide advice to their mostly corporate clients and help them arrange transactions such as initial and seasoned securities offerings. Their corporate finance divisions help corporations finance their business by issuing securities, such as common and preferred shares, notes, and bonds. Another function of corporate finance divisions is to help companies identify and acquire other companies (i.e., in mergers and acquisitions).

Exchanges provide places where traders can meet to arrange their trades. Historically, brokers and dealers met on an exchange floor to negotiate trades. Increasingly, exchanges arrange trades for traders based on orders that brokers and dealers submit to them. Such exchanges essentially act as brokers. The distinction between exchanges and brokers has become quite blurred. Exchanges and brokers that use electronic order matching systems to arrange trades among their clients are functionally indistinguishable in this respect. Examples of exchanges include the NYSE, Eurex, Frankfurt Stock Exchange, the Chicago Mercantile Exchange, the Tokyo Stock Exchange, and the Singapore Exchange.

Exchanges are easily distinguished from brokers by their regulatory operations. Most exchanges regulate their members' behavior when trading on the exchange, and sometimes away from the exchange.

Many securities exchanges regulate the issuers that list their securities on the exchange. These regulations generally require timely financial disclosure. Financial analysts use this information to value the securities traded at the exchange. Without such disclosure, valuing securities could be very difficult and market prices might not reflect the fundamental values of the securities. In such situations, well-informed participants may profit from less-informed participants. To avoid such losses, the less-informed participants may withdraw from the market, which can greatly increase corporate costs of capital.

Some exchanges also prohibit issuers from creating capital structures that would concentrate voting rights in the hands of a few owners who do not own a commensurate share of the equity. These regulations attempt to ensure that corporations are run for the benefit of all shareholders and not to promote the interests of controlling shareholders who do not have significant economic stakes in the company.

Exchanges derive their regulatory authority from their national or regional governments, or through the voluntary agreements of their members and issuers to subject themselves to the exchange regulations. In most countries, government regulators oversee the exchange rules and the regulatory operations. Most countries also impose financial disclosure standards on public issuers. Examples of government regulatory bodies include the Japanese Financial Services Agency, the British Financial Conduct Authority, the German Federal Financial Supervisory Authority (BaFin), the US Securities and Exchange Commission, the Ontario Securities Commission, and the Argentine National Securities Commission (CNV).

Alternative trading systems (ATSs), also known as **electronic communications networks** (ECNs) or **multilateral trading facilities** (MTFs) are trading venues that function like exchanges but that do not exercise regulatory authority over their subscribers except with respect to the conduct of their trading in their trading systems. Some ATSs operate electronic trading systems that are otherwise indistinguishable from the trading systems operated by exchanges. Others operate innovative trading systems that suggest trades to their customers based on information that their customers share with them or that they obtain through research into their customers' preferences. Many ATSs are known as **dark pools** because they do not display the orders that their clients send to them. Large investment managers especially like these systems because market prices often move to their disadvantage when other traders know about their large orders. ATSs may be owned and operated by broker–dealers, exchanges, banks, or by companies organized solely for this purpose, many of which may be owned by a consortia of brokers–dealers and banks. Examples of ATSs include MATCHNow (Canada), BATS (United States), POSIT (United States), Liquidnet (United States), Baxter-FX (Ireland), and Turquoise (Europe). Many of these ATSs provide services in many markets besides the ones in which they are domiciled.

7.2 Dealers

Dealers fill their clients' orders by trading with them. When their clients want to sell securities or contracts, dealers buy the instruments for their own accounts. If their clients want to buy securities, dealers sell securities that they own or have borrowed.

After completing a transaction, dealers hope to reverse the transaction by trading with another client on the other side of the market. When they are successful, they effectively connect a buyer who arrived at one point in time with a seller who arrived at another point in time.

The service that dealers provide is liquidity. **Liquidity** is the ability to buy or sell with low transactions costs when you want to trade. By allowing their clients to trade when they want to trade, dealers provide liquidity to them. In over-the-counter markets, dealers offer liquidity when their clients ask them to trade with them. In exchange markets, dealers offer liquidity to anyone who is willing to trade at the prices that the dealers offer at the exchange. Dealers profit when they can buy at prices that on average are lower than the prices at which they sell.

Dealers may organize their operations within proprietary trading houses, investment banks, and hedge funds, or as sole proprietorships. Some dealers are traditional dealers in the sense that individuals make trading decisions. Others use computerized trading to make all trading decisions. Examples of companies with large dealing operations include Deutsche Bank (Germany), RBC Capital Markets (Canada), Nomura Securities (Japan), Timber Hill (United States), Goldman Sachs (United States), and IG Group (United Kingdom). Almost all investment banks have large dealing operations.

Most dealers also broker orders, and many brokers deal to their customers. Accordingly, practitioners often use the term **broker–dealer** to refer to dealers and brokers. Broker–dealers have a conflict of interest with respect to how they fill their customers' orders. When acting as a broker, they must seek the best price for their customers' orders. When acting as dealers, however, they profit most when they sell to their customers at high prices or buy from their customers at low prices. The problem is most serious when the customer allows the broker–dealer to decide whether to trade the order with another trader or to fill it as a dealer. Consequently, when trading with a broker–dealer, some customers specify how they want their orders filled. They may also trade only with pure agency brokers who do not also deal.

Primary dealers are dealers with whom central banks trade when conducting monetary policy. They buy bills, notes, and bonds when the central banks sell them to decrease the money supply. The dealers then sell these instruments to their clients. Similarly, when the central banks want to increase the money supply, the primary dealers buy these instruments from their clients and sell them to the central banks.

EXAMPLE 14

Brokers and Dealers

What characteristic *most likely* distinguishes brokers from dealers?

Solution:

Brokers are agents that arrange trades on behalf of their clients. They do not trade with their clients. In contrast, dealers are proprietary traders who trade with their clients.

7.3 Arbitrageurs

Arbitrageurs trade when they can identify opportunities to buy and sell identical or essentially similar instruments at different prices in different markets. They profit when they can buy in one market for less than they sell in another market. Arbitrageurs are financial intermediaries because they connect buyers in one market to sellers in another market.

The purest form of arbitrage involves buying and selling the same instrument in two different markets. Arbitrageurs who do such trades sell to buyers in one market and buy from sellers in the other market. They provide liquidity to the markets because they make it easier for buyers and sellers to trade when and where they want to trade.

Because dealers and arbitrageurs both provide liquidity to other traders, they compete with each other. The dealers connect buyers and sellers who arrive in the same market at different times whereas the arbitrageurs connect buyers and sellers who arrive at the same time in different markets. In practice, traders who profit from offering liquidity rarely are purely dealers or purely arbitrageurs. Instead, most traders attempt to identify and exploit every opportunity they can to manage their inventories profitably.

If information about prices is readily available to market participants, pure arbitrages involving the same instrument will be quite rare. Traders who are well informed about market conditions usually route their orders to the market offering the best price so that arbitrageurs will have few opportunities to match traders across markets when they want to trade the exact same instrument.

Arbitrageurs often trade securities or contracts whose values depend on the same underlying factors. For example, dealers in equity option contracts often sell call options in the contract market and buy the underlying shares in the stock market. Because the values of the call options and of the underlying shares are closely correlated (the value of the call increases with the value of the shares), the long stock position hedges the risk in the short call position so that the dealer's net position is not too risky.

Similar to the pure arbitrage that involves the same instrument in different markets, these arbitrage trades connect buyers in one market to sellers in another market. In this case, however, the buyers and sellers are interested in different instruments whose values are closely related. In the example, the buyer is interested in buying a call options contract, the value of which is a nonlinear function of the value of the underlying stock; the seller is interested in selling the underlying stock.

Options dealers buy stock and sell calls when calls are overpriced relative to the underlying stocks. They use complicated financial models to value options in relation to underlying stock values, and they use financial engineering techniques to control the risk of their portfolios. Successful arbitrageurs must know valuation relations well and they must manage the risk in their portfolios well to trade profitably. They profit by buying the relatively undervalued instrument and selling the relatively overvalued instrument.

Buying a risk in one form and selling it another form involves a process called replication. Arbitrageurs use various trading strategies to replicate the returns to securities and contracts. If they can substantially replicate those returns, they can use the replication trading strategy to offset the risk of buying or selling the actual securities and contracts. The combined effect of their trading is to transform risk from one form to another. This process allows them to create or eliminate contracts in response to the excess demand for, and supply of, contracts.

For example, when traders want to buy more call contracts than are presently available, they push the call contract prices up so that calls become overvalued relative to the underlying stock. The arbitrageurs replicate calls by using a particular financial engineering strategy to buy the underlying stock, and then create the desired call option contracts by selling them short. In contrast, if more calls have been created than traders want to hold, call prices will fall so that calls become undervalued relative to the underlying stock. The arbitrageurs will trade stocks and contracts to absorb the excess contracts. Arbitrageurs who use these strategies are financial intermediaries because they connect buyers and sellers who want to trade the same underlying risks but in different forms.

EXAMPLE 15

Dealers and Arbitrageurs

With respect to providing liquidity to market participants, what characteristics most clearly distinguish dealers from arbitrageurs?

Solution:

Dealers provide liquidity to buyers and sellers who arrive at the same market at different times. They move liquidity through time. Arbitrageurs provide liquidity to buyers and sellers who arrive at different markets at the same time. They move liquidity across markets.

SECURITIZERS, DEPOSITORY INSTITUTIONS AND INSURANCE COMPANIES

8

d describe types of financial intermediaries and services that they provide;

Banks and investment companies create new financial products when they buy and repackage securities or other assets. For example, mortgage banks commonly originate hundreds or thousands of residential mortgages by lending money to homeowners. They then place the mortgages in a pool and sell shares of the pool to investors as mortgage pass-through securities, which are also known as mortgage-backed securities. All payments of principal and interest are passed through to the investors each month, after deducting the costs of servicing the mortgages. Investors who purchase these pass-through securities obtain securities that in aggregate have the same net cash flows and associated risks as the pool of mortgages.

The process of buying assets, placing them in a pool, and then selling securities that represent ownership of the pool is called securitization.

Mortgage-backed securities have the advantage that default losses and early repayments are much more predictable for a diversified portfolio of mortgages than they are for individual mortgages. They are also attractive to investors who cannot efficiently service mortgages but wish to invest in mortgages. By securitizing mortgage pools, the mortgage banks allow investors who are not large enough to buy hundreds of mortgages to obtain the benefits of diversification and economies of scale in loan servicing.

Securitization greatly improves liquidity in the mortgage markets because it allows investors in the pass-through securities to buy mortgages indirectly that they otherwise would not buy. Because the financial risks associated with mortgage-backed securities (debt securities with specified claims on the cash flows of a portfolio of mortgages) are much more predictable than those of individual mortgages, mortgage-backed securities are easier to price and thus easier to sell when investors need to raise cash. These characteristics make the market for mortgage-backed securities much more liquid than the market for individual mortgages. Because investors value liquidity—the ability to sell when they want to—they will pay more for securitized mortgages than for individual mortgages. The homeowners benefit because higher mortgage prices imply lower interest rates.

The mortgage bank is a financial intermediary because it connects investors who want to buy mortgages to homeowners who want to borrow money. The homeowners sell mortgages to the bank when the bank lends them money.

Some mortgage banks form mortgage pools from mortgages that they buy from other banks that originate the loans. These mortgage banks are also financial intermediaries because they connect sellers of mortgages to buyers of mortgage-backed securities. Although the sellers of the mortgages are the originating lenders and not the borrowers, the benefits of creating liquid mortgage-backed securities ultimately flow back to the borrowers.

The creation of the pass-through securities generally takes place on the accounts of the mortgage bank. The bank buys mortgages and sells pass-through securities whose values depend on the mortgage pool. The mortgages appear on the bank's accounts as assets and the mortgage-backed securities appear as liabilities.

In many securitizations, the financial intermediary avoids placing the assets and liabilities on its balance sheet by setting up a special corporation or trust that buys the assets and issues the securities. That corporation or trust is called a **special purpose vehicle** (SPV) or alternatively a **special purpose entity** (SPE). Conducting a securitization through a special purpose vehicle is advantageous to investors because their interests in the asset pool are better protected in an SPV than they would be on the balance sheet of the financial intermediary if the financial intermediary were to go bankrupt.

Financial intermediaries securitize many assets. Besides mortgages, banks securitize car loans, credit card receivables, bank loans, and airplane leases, to name just a few assets. As a class, these securities are called asset-backed securities.

When financial intermediaries securitize assets, they often create several classes of securities, called tranches, that have different rights to the cash flows from the asset pool. The tranches are structured so that some produce more predictable cash flows than do others. The senior tranches have first rights to the cash flow from the asset pool. Because the overall risk of a given asset pool cannot be changed, the more junior tranches bear a disproportionate share of the risk of the pool. Practitioners often call the most junior tranche toxic waste because it is so risky. The complexity associated with slicing asset pools into tranches can make the resulting securities difficult to value. Mistakes in valuing these securities contributed to the financial crisis that started in 2007.

Investment companies also create pass-through securities based on investment pools. For example, an exchange-traded fund is an asset-backed security that represents ownership in the securities and contracts held by the fund. The shareholders benefit from the securitization because they can buy or sell an entire portfolio in a single transaction. Because the transaction cost savings are quite substantial, exchange-traded funds often trade in very liquid markets. The investment companies (and sometimes the arbitrageurs) that create exchange-traded funds are financial intermediaries because they connect the buyers of the funds to the sellers of the assets that make up the fund portfolios.

More generally, the creators of all pooled investment vehicles are financial intermediaries that transform portfolios of securities and contracts into securities that represent undivided ownership of the portfolios. The investors in these funds thus indirectly invest in the securities held by the fund. They benefit from the expertise of the investment manager and from obtaining a portfolio that may be more diversified than one they might otherwise be able to hold.

8.1 Depository Institutions and Other Financial Corporations

Depository institutions include commercial banks, savings and loan banks, credit unions, and similar institutions that raise funds from depositors and other investors and lend it to borrowers. The banks give their depositors interest and transaction services, such as check writing and check cashing, in exchange for using their money. They may also raise funds by selling bonds or equity in the bank.

These banks are financial intermediaries because they transfer funds from their depositors and investors to their borrowers. The depositors and investors benefit because they obtain a return (in interest, transaction services, dividends, or capital appreciation) on their funds without having to contract with the borrowers and manage their loans. The borrowers benefit because they obtain the funds that they need without having to search for investors who will trust them to repay their loans.

Many other financial corporations provide credit services. For example, acceptance corporations, discount corporations, payday advance corporations, and factors provide credit to borrowers by lending them money secured by such assets as consumer loans, machinery, future paychecks, or accounts receivables. They finance these loans by selling commercial paper, bonds, and shares to investors. These corporations are intermediaries because they connect investors to borrowers. The investors obtain investments secured by a diversified portfolio of loans while the borrowers obtain funds without having to search for investors.

Brokers also act as financial intermediaries when they lend funds to clients who want to buy securities on margin. They generally obtain the funds from other clients who deposit them in their accounts. Brokers who provide these services to hedge funds and other similar institutions are called prime brokers.

Banks, financial corporations, and brokers can only raise money from depositors and other lenders because their equity owners retain residual interests in the performance of the loans that they make. If the borrowers default, the depositors and other lenders have priority claims over the equity owners. If insufficient money is collected from the borrowers, shareholders' equity is used to pay their depositors and other lenders. The risk of losing capital focuses the equity owners' and management's attention so that credit is not offered foolishly.

Because the ability of these companies to cover their credit losses is limited by the capital that their owners invest in them, the depositors and other investors who lend them money pay close attention to how much money the owners have at risk. For example, if a finance corporation is poorly capitalized, its shareholders will lose little if its clients default on the loans that the finance corporation makes to them. In that case, the finance corporation will have little incentive to lend only to creditworthy borrowers and to effectively manage collection on those loans once they have been made. Worse, it may even choose to lend to borrowers with poor credit because the interest rates that they can charge such borrowers are higher. Until those loans default, the higher income will make the corporation appear to be more profitable than it actually is. Depositors and other investors are aware of these problems and generally pay close attention to them. Accordingly, poorly capitalized financial institutions cannot easily borrow money to finance their operations at favorable rates.

Depository banks and financial corporations are similar to securitized asset pools that issue pass-through securities. Their depositors and investors own securities that ultimately are backed by an asset pool consisting of their loan portfolios. The depositors generally hold the most senior tranche, followed by the other creditors. The shareholders hold the most junior tranche. In the event of bankruptcy, they are paid only if everyone else is paid.

EXAMPLE 16

Commercial Banks

What services do commercial banks provide that make them financial intermediaries?

Solution:

Commercial banks collect deposits from investors and lend them to borrowers. They are intermediaries because they connect lenders to borrowers. Commercial banks also provide transaction services that make it easier for the banks' depository customers to pay bills and collect funds from their own customers.

8.2 Insurance Companies

Insurance companies help people and companies offset risks that concern them. They do this by creating insurance contracts (policies) that provide a payment in the event that some loss occurs. The insured buy these contracts to hedge against potential losses. Common examples of insurance contracts include auto, fire, life, liability, medical, theft, and disaster insurance contracts.

Credit default swaps are also insurance contracts, but historically they have not been subject to the same reserve requirements that most governments apply to more traditional insurance contracts. They may be sold by insurance companies or by other financial entities, such as investment banks or hedge funds.

Insurance contracts transfer risk from those who buy the contracts to those who sell them. Although insurance companies occasionally broker trades between the insured and the insurer, they more commonly provide the insurance themselves. In that case, the insurance company's owners and creditors become the indirect insurers of the risks that the insurance company assumes. Insurance companies also often transfer risks that they do not wish to bear by buying reinsurance policies from reinsurers.

Insurers are financial intermediaries because they connect the buyers of their insurance contracts with investors, creditors, and reinsurers who are willing to bear the insured risks. The buyers benefit because they can easily obtain the risk transfers that they seek without searching for entities that would be willing to assume those risks.

The owners, creditors, and reinsurers of the insurance company benefit because the company allows them to sell their tolerance for risk easily without having to manage the insurance contracts. Instead, the company manages the relationships with the insured—primarily collections and claims—and hopefully controls the various problems—fraud, moral hazard, and adverse selection—that often plague insurance markets. Fraud occurs when people deliberately cause or falsely report losses to collect on insurance. Moral hazard occurs when people are less careful about avoiding insured losses than they would be if they were not insured so that losses occur more often than they would otherwise. Adverse selection occurs when only those who are most at risk buy insurance so that insured losses tend to be greater than average.

Everyone benefits because insurance companies hold large diversified portfolios of policies. Loss rates for well-diversified portfolios of insurance contracts are much more predictable than for single contracts. For such contracts as auto insurance in which losses are almost uncorrelated across policies, diversification ensures that the financial performance of a large portfolio of contracts will be quite predictable and so holding the portfolio will not be very risky. The insured benefit because they do not have to pay the insurers much to compensate them for bearing risk (the expected loss is quite predictable so the risk is relatively low). Instead, their insurance premiums primarily reflect the expected loss rate in the portfolio plus the costs of running and financing the company.

SETTLEMENT AND CUSTODIAL SERVICES AND SUMMARY

9

d describe types of financial intermediaries and services that they provide;

In addition to connecting buyers to sellers through a variety of direct and indirect means, financial intermediaries also help their customers settle their trades and ensure that the resulting positions are not stolen or pledged more than once as collateral.

Clearinghouses arrange for final settlement of trades. In futures markets, they guarantee contract performance. In other markets, they may act only as escrow agents, transferring money from the buyer to the seller while transferring securities from the seller to the buyer.

The members of a clearinghouse are the only traders for whom the clearinghouse will settle trades. To ensure that their members settle the trades that they present to the clearinghouse, clearinghouses require that their members have adequate capital and post-performance bonds (margins). Clearinghouses also limit the aggregate net (buy minus sell) quantities that their members can settle.

Brokers and dealers who are not members of the clearinghouse must arrange to have a clearinghouse member settle their trades. To ensure that the non-member brokers and dealers can settle their trades, clearinghouse members require that their customers (the non-member brokers and dealers) have adequate capital and post-margins. They also limit the aggregate net quantities that their customers can settle and they monitor their customers' trading to ensure that they do not arrange trades that they cannot settle.

Brokers and dealers similarly monitor the trades made by their retail and institutional customers, and regulate their customers to ensure that they do not arrange trades that they cannot settle.

This hierarchical system of responsibility generally ensures that traders settle their trades. The brokers and dealers guarantee settlement of the trades they arrange for their retail and institutional customers. The clearinghouse members guarantee settlement of the trades that their customers present to them, and clearinghouses guarantee settlement of all trades presented to them by their members. If a clearinghouse member fails to settle a trade, the clearinghouse settles the trade using its own capital or capital drafted from the other members.

Reliable settlement of all trades is extremely important to a well-functioning financial system because it allows strangers to confidently contract with each other without worrying too much about **counterparty risk**, the risk that their counterparties will not settle their trades. A secure clearinghouse system thus greatly increases liquidity because it greatly increases the number of counterparties with whom a trader can safely arrange a trade.

In many national markets, clearinghouses clear all securities trades so that traders can trade securities through any exchange, broker, alternative trading system, or dealer. These clearinghouse systems promote competition among these exchange service providers.

In contrast, most futures exchanges have their own clearinghouses. These clearinghouses usually will not accept trades arranged away from their exchanges so that a competing exchange cannot trade another exchange's contracts. Competing exchanges may create similar contracts, but moving traders from one established market to a new market is extraordinarily difficult because traders prefer to trade where other traders trade.

Depositories or custodians hold securities on behalf of their clients. These services, which are often offered by banks, help prevent the loss of securities through fraud, oversight, or natural disaster. Broker–dealers also often hold securities on behalf of their customers so that the customers do not have to hold the securities in certificate form. To avoid problems with lost certificates, securities increasingly are issued only in electronic form.

EXAMPLE 17

Financial Intermediaries

As a relatively new member of the business community, you decide it would be advantageous to join the local lunch club to network with businessmen. Upon learning that you are a financial analyst, club members soon enlist you to give a lunch speech. During the question and answer session afterwards, a member of the audience asks, “I keep reading in the newspaper about the need to regulate ‘financial intermediaries,’ but really don’t understand exactly what they are. Can you tell me?” How do you answer?

Solution:

Financial intermediaries are companies that help their clients achieve their financial goals. They are called intermediaries because, in some way or another, they stand between two or more people who would like to trade with each other, but for various reasons find it difficult to do so directly. The intermediary arranges the trade for them, or more often, trades with both sides.

For example, a commercial bank is an intermediary that connects investors with money to borrowers who need money. The investors buy certificates of deposit from the bank, buy bonds or stock issued by the bank, or simply are depositors in the bank. The borrowers borrow this money from the bank when they arrange loans. Without the bank’s intermediation, the investors would have to find trustworthy borrowers themselves, which would be difficult, and the borrowers would have to find trusting lenders, which would also be difficult.

Similarly, an insurance company is an intermediary because it connects customers who want to insure risks with investors who are willing to bear those risks. The investors own shares or bonds issued by the insurance company, or they have sold reinsurance contracts to the insurance company. The insured benefit because they can more easily buy a policy from an insurance company than they can find counterparties who would be willing to bear their risks. The investors benefit because the insurance company creates a diversified portfolio of risks by selling insurance to thousands or millions of customers. Diversification ensures that the net risk borne by the insurance company and its investors will be predictable and thus financially manageable.

In both cases, the financial intermediary also manages the relationships with its customers and investors so that neither side has to worry about the credit-worthiness or trust-worthiness of its counterparties. For example, the bank manages credit quality and collections on its loans and the insurance company manages risk exposure and collections on its policies. These services benefit both sides by reducing the costs of connecting investors to borrowers or of insured to insurers.

These are only two examples of financial intermediation. Many others involve firms engaged in brokerage, dealing, arbitrage, securitization, investment management, and the clearing and settlement of trades. In all cases, the financial intermediary stands between a buyer and a seller, offering them services that allow them to better achieve their financial goals in a cost effective and efficient manner.

9.1 Summary

By facilitating transactions among buyers and sellers, financial intermediaries provide services essential to a well-functioning financial system. They facilitate transactions the following ways:

1. Brokers, exchanges, and various alternative trading systems match buyers and sellers interested in trading the same instrument at the same place and time. These financial intermediaries specialize in discovering and organizing information about who wants to trade.
2. Dealers and arbitrageurs connect buyers to sellers interested in trading the same instrument but who are not present at the same place and time. Dealers connect buyers to sellers who are present at the same place but at different times whereas arbitrageurs connect buyers to sellers who are present at the same time but in different places. These financial intermediaries trade for their own accounts when providing these services. Dealers buy or sell with one client and hope to do the offsetting transaction later with another client. Arbitrageurs buy from a seller in one market while simultaneously selling to a buyer in another market.
3. Many financial intermediaries create new instruments that depend on the cash flows and associated financial risks of other instruments. The intermediaries provide these services when they securitize assets, manage investment funds, operate banks and other finance corporations that offer investments to investors and loans to borrowers, and operate insurance companies that pool risks. The instruments that they create generally are more attractive to their clients than the instruments on which they are based. The new instruments also may be differentiated to appeal to diverse clienteles. Their efforts connect buyers of one or more instruments to sellers of other instruments, all of which in aggregate provide the same cash flows and risk exposures. Financial intermediaries thus effectively arrange trades among traders who otherwise would not trade with each other.
4. Arbitrageurs who conduct arbitrage among securities and contracts whose values depend on common factors convert risk from one form to another. Their trading connects buyers and sellers who want to trade similar risks expressed in different forms.
5. Banks, clearinghouses, and depositories provide services that ensure traders settle their trades and that the resulting positions are not stolen or pledged more than once as collateral.

POSITIONS AND SHORT POSITIONS

10

e compare positions an investor can take in an asset;

People generally solve their financial and risk management problems by taking positions in various assets or contracts. A **position** in an asset is the quantity of the instrument that an entity owns or owes. A portfolio consists of a set of positions.

People have **long positions** when they own assets or contracts. Examples of long positions include ownership of stocks, bonds, currencies, contracts, commodities, or real assets. Long positions benefit from an appreciation in the prices of the assets or contracts owned.

People have **short positions** when they have sold assets that they do not own, or when they write and sell contracts. Short positions benefit from a decrease in the prices of the assets or contracts sold. Short sellers profit by selling at high prices and repurchasing at lower prices. Information-motivated traders sell assets and contracts short positions when they believe that prices will fall.

Hedgers also often sell instruments short. They short securities and contracts when the financial risks inherent in the instruments are positively correlated with the risks to which they are exposed. For example, to hedge the risk associated with holding copper inventories, a wire manufacturer would sell short copper futures. If the price of copper falls, the manufacturer will lose on his copper inventories but gain on his short futures position. (If the risk in an instrument is inversely correlated with a risk to which hedgers are exposed, the hedgers will hedge with long positions.)

Contracts have long sides and short sides. The long side of a forward or futures contract is the side that will take physical delivery or its cash equivalent. The short side of such contracts is the side that is liable for the delivery. The long side of a futures contract increases in value when the value of the underlying asset increases in value.

The identification of the two sides can be confusing for option contracts. The long side of an option contract is the side that holds the right to exercise the option. The short side is the side that must satisfy the obligation. Practitioners say that that the long side *holds* the option and the short side *writes* the option, so the long side is the holder and the short side is the writer. The put contracts are the source of the potential confusion. The put contract holder has the right to sell the underlying to the writer. The holder will benefit if the price of the underlying falls, in which case the price of the put contract will rise. The holder is long the put contract and has an indirect short position in the underlying instrument. Analysts call the indirect short position short exposure to the underlying. The put contract holders have long exposure to their option contract and short exposure to the underlying instrument.

Exhibit 1 Option Positions and Their Associated Underlying Risk Exposures

Type of Option	Option Position	Exposure to Underlying Risk
Call	Long	Long
Call	Short	Short
Put	Long	Short
Put	Short	Long

The identification of the long side in a swap contract is often arbitrary because swap contracts call for the exchange of contractually determined cash flows rather than for the purchase (or the cash equivalent) of some underlying instrument. In general, the side that benefits from an increase in the quoted price is the long side.

The identification of the long side in currency contracts also may be confusing. In this case, the confusion stems from symmetry in the contracts. The buyer of one currency is the seller of the other currency, and vice versa for the seller. Thus, a long forward position in one currency is a short forward position in the other currency. When practitioners describe a position, they generally will say, “I’m long the dollar against the yen,” which means they have bought dollars and sold yen.

10.1 Short Positions

Short sellers create short positions in contracts by selling contracts that they do not own. In a sense, they become the issuers of the contract when they create the liabilities associated with their contracts. This analogy will also help you better understand risk when you study corporate finance: Corporations create short positions in their bonds when they issue bonds in exchange for cash. Although bonds are generally considered to be securities, they are also contracts between the issuer and the bondholder.

Short sellers create short positions in securities by borrowing securities from security lenders who are long holders. The short sellers then sell the borrowed securities to other traders. Short sellers close their positions by repurchasing the securities and returning them to the security lenders. If the securities drop in value, the short sellers profit because they repurchase the securities at lower prices than the prices at which they sold the securities. If the securities rise in value, they will lose. Short sellers who buy to close their positions are said to cover their positions.

The potential gains in a long position generally are unbounded. For example, the stock prices of such highly successful companies as Yahoo! have increased more than 50-fold since they were first publicly traded. The potential losses on long positions, however, are limited to no more than 100 percent—a complete loss—for long positions without any associated liabilities.

In contrast, the potential gains on a short position are limited to no more than 100 percent whereas the potential losses are unbounded. The unbounded potential losses on short positions make short positions very risky in volatile instruments. As an extreme example of this, if you had shorted 100 shares of Yahoo! in July 1996 at $20 and kept the position open for four years, you would have lost $148,000 on your $2,000 initial short position. During this period, Yahoo! rose 75-fold to $1,500 on a split-adjusted equivalent basis.

Although security lenders generally believe that they are long the securities that they lend, in fact, they do not actually own the securities during the periods of their loans. Instead, they own promises made by the short sellers to return the securities. These promises are memorialized in security lending agreements. These agreements specify that the short sellers will pay the long sellers all dividends or interest that they otherwise would have received had they not lent their securities. These payments are called payments-in-lieu of dividends (or of interest), and they may have different tax treatments than actual dividends and interest. The security lending agreements also protect the lenders in the event of a stock split.

To secure the security loans, lenders require that the short seller leave the proceeds of the short sale on deposit with them as collateral for the stock loan. They invest the collateral in short-term securities, and they rebate the interest to the short sellers at rates called short rebate rates. The short rebate rates are determined in the market and generally are available only to institutional short-sellers and some large retail traders. If a security is hard to borrow, the rebate rate may be very small or even negative. Such securities are said to be “on special”. Most security lending agreements require various margin payments to keep the credit risk among the parties from growing when prices change.

Securities lenders lend their securities because the short rebate rates they pay on the collateral are lower than the interest rates they receive from investing the collateral. The difference is because of the implicit loan fees that they receive from the borrowers for borrowing the stock. The difference also compensates lenders for risks that the lenders take when investing the collateral and for the risk that the borrowers will default if prices rise significantly.

EXAMPLE 18

Short Positions in Securities and Contracts

How is the process of short selling shares of Siemens different from that of short selling a Siemens equity call option contract?

Solution:

To short sell shares of Siemens, the seller (or his broker) must borrow the shares from a long holder so that he can deliver them to the buyer. To short sell a Siemens equity call option contract, the seller simply creates the contract when he sells it to the buyer.

11 LEVERAGED POSITIONS

f calculate and interpret the leverage ratio, the rate of return on a margin transaction, and the security price at which the investor would receive a margin call;

In many markets, traders can buy securities by borrowing some of the purchase price. They usually borrow the money from their brokers. The borrowed money is called the **margin loan**, and they are said to buy on margin. The interest rate that the buyers pay for their margin loan is called the **call money rate**. The call money rate is above the government bill rate and is negotiable. Large buyers generally obtain more favorable rates than do retail buyers. For institutional-size buyers, the call money rate is quite low because the loans are generally well secured by securities held as collateral by the lender.

Trader's equity is that portion of the security price that the buyer must supply. Traders who buy securities on margin are subject to minimum margin requirements. The **initial margin requirement** is the minimum fraction of the purchase price that must be trader's equity. This requirement may be set by the government, the exchange, or the exchange clearinghouse. For example, in the United States, the Federal Reserve Board sets the initial margin requirement through Regulation T. In Hong Kong SAR, the Securities and Futures Commission sets the margin requirements. In all markets, brokers often require more equity than the government-required minimum from their clients when lending to them.

Many markets allow brokers to lend their clients more money if the brokers use risk models to measure and control the overall risk of their clients' portfolios. This system is called portfolio margining.

Buying securities on margin can greatly increase the potential gains or losses for a given amount of equity in a position because the trader can buy more securities on margin than he could otherwise. The buyer thus earns greater profits when prices rise and suffers greater losses when prices fall. The relation between risk and borrowing is called **financial leverage** (often simply called leverage). Traders leverage their positions when they borrow to buy more securities. A highly leveraged position is large relative to the equity that supports it.

The leverage ratio is the ratio of the value of the position to the value of the equity investment in it. The leverage ratio indicates how many times larger a position is than the equity that supports it. The maximum leverage ratio associated with a position financed by the minimum margin requirement is one divided by the minimum margin requirement. If the requirement is 40 percent, then the maximum leverage ratio is 2.5 = 100% position ÷ 40% equity.

The leverage ratio indicates how much more risky a leveraged position is relative to an unleveraged position. For example, if a stock bought on 40 percent margin rises 10 percent, the buyer will experience a 25 percent (2.5 × 10%) return on the equity investment in her leveraged position. But if the stock falls by 10 percent, the return on the equity investment will be −25 percent (before the interest on the margin loan and before payment of commissions).

Financial analysts must be able to compute the total return to the equity investment in a leveraged position. The total return depends on the price change of the purchased security, the dividends or interest paid by the security, the interest paid on the margin loan, and the commissions paid to buy and sell the security. The following example illustrates the computation of the total return to a leveraged purchase of stock that pays a dividend.

EXAMPLE 19

Computing Total Return to a Leveraged Stock Purchase

A buyer buys stock on margin and holds the position for exactly one year, during which time the stock pays a dividend. For simplicity, assume that the interest on the loan and the dividend are both paid at the end of the year.

Purchase price	$20/share
Sale price	$15/share
Shares purchased	1,000
Leverage ratio	2.5
Call money rate	5%
Dividend	$0.10/share
Commission	$0.01/share

1 What is the total return on this investment?

2 Why is the loss greater than the 25 percent decrease in the market price?

Solution to 1:

To find the return on this investment, first determine the initial equity and then determine the equity remaining after the sale. The total purchase price is $20,000. The leverage ratio of 2.5 indicates that the buyer's equity financed 40 percent = (1 ÷ 2.5) of the purchase price. Thus, the equity investment is $8,000 = 40% of $20,000. The $12,000 remainder is borrowed. The actual investment is slightly higher because the buyer must pay a commission of $10 = $0.01/share × 1,000 shares to buy the stock. The total initial investment is $8,010.

At the end of the year, the stock price has declined by $5/share. The buyer lost $5,000 = $5/share × 1,000 shares as a result of the price change. In addition, the buyer has to pay interest at 5 percent on the $12,000 loan, or $600. The buyer also receives a dividend of $0.10/share, or $100. The trader's equity remaining after the sale is computed from the initial equity investment as follows:

Initial investment	$8,010
Purchase commission	−10
Trading gains/losses	−5,000
Margin interest paid	−600
Dividends received	100

(continued)

Sales commission paid	–10
Remaining equity	$2,490

or

Proceeds on sale	$15,000
Payoff loan	–12,000
Margin interest paid	–600
Dividends received	100
Sales commission paid	–10
Remaining equity	$2,490

so that the return on the initial investment of $8,010 is (2,490 – 8,010)/8,010 = –68.9%.

Solution to 2:

The realized loss is substantially greater than the stock price return of ($15 – $20)/$20 = –25%. Most of the difference is because of the leverage with the remainder primarily the result of the interest paid on the loan. Based on the leverage alone and ignoring the other cash flows, we would expect that the return on the equity would be –62.5% = 2.5 leverage times the –25% stock price return.

In the above example, if the stock dropped more than the buyer's original 40 percent margin (ignoring commissions, interest, and dividends), the trader's equity would have become negative. In that case, the investor would owe his broker more than the stock is worth. Brokers often lose money in such situations if the buyer does not repay the loan out of other funds.

To prevent such losses, brokers require that margin buyers always have a minimum amount of equity in their positions. This minimum is called the **maintenance margin requirement**. It is usually 25 percent of the current value of the position, but it may be higher or lower depending on the volatility of the instrument and the policies of the broker.

If the value of the equity falls below the maintenance margin requirement, the buyer will receive a **margin call**, or request for additional equity. If the buyer does not deposit additional equity with the broker in a timely manner, the broker will close the position to prevent further losses and thereby secure repayment of the margin loan.

When you buy securities on margin, you must know the price at which you will receive a margin call if prices drop. The answer to this question depends on your initial equity and on the maintenance margin requirement.

EXAMPLE 20

Margin Call Price

A trader buys stock on margin posting 40 percent of the initial stock price of $20 as equity. The maintenance margin requirement for the position is 25 percent. Below what price will a margin call occur?

Solution:

The trader's initial equity is 40 percent of the initial stock price of $20, or $8 per share. Subsequent changes in equity per share are equal to the share price change so that equity per share is equal to $8 + ($P$ – 20) where P is the current share price. The margin call takes place when equity drops below the 25 percent maintenance margin requirement. The price below which a margin call will take place is the solution to the following equation:

$$\frac{\text{Equity/share}}{\text{Price/share}} = \frac{\$8 + P - 20}{P} = 25\%$$

which occurs at P = 16. When the price drops below \$16, the equity will be under \$4/share, which is less than 25 percent of the price.

Traders who sell securities short are also subject to margin requirements because they have borrowed securities. Initially, the trader's equity supporting the short position must be at least equal to the margin requirement times the initial value of the short position. If prices rise, equity will be lost. At some point, the short seller will have to contribute additional equity to meet the maintenance margin requirement. Otherwise, the broker will buy the security back to cover the short position to prevent further losses and thereby secure repayment of the stock loan.

ORDERS AND EXECUTION INSTRUCTIONS

12

g compare execution, validity, and clearing instructions;

h compare market orders with limit orders;

Buyers and sellers communicate with the brokers, exchanges, and dealers that arrange their trades by issuing **orders**. All orders specify what instrument to trade, how much to trade, and whether to buy or sell. Most orders also have other instructions attached to them. These additional instructions may include execution instructions, validity instructions, and clearing instructions. **Execution instructions** indicate how to fill the order, **validity instructions** indicate when the order may be filled, and **clearing instructions** indicate how to arrange the final settlement of the trade.

In this section, we introduce various order instructions and explain how traders use them to achieve their objectives. We discuss execution mechanisms—how exchanges, brokers and dealers fill orders—in the next section. To understand the concepts in this section, however, you need to know a little about order execution mechanisms.

In most markets, dealers and various other proprietary traders often are willing to buy from, or sell to, other traders seeking to sell or buy. The prices at which they are willing to buy are called **bid** prices and those at which they are willing to sell are called **ask** prices, or sometimes **offer** prices. The ask prices are invariably higher than the bid prices.

The traders who are willing to trade at various prices may also indicate the quantities that they will trade at those prices. These quantities are called **bid sizes** and **ask sizes** depending on whether they are attached to bids or offers.

Practitioners say that the traders who offer to trade make a market. Those who trade with them take the market.

The highest bid in the market is the **best bid**, and the lowest ask in the market is the **best offer**. The difference between the best bid and the best offer is the **market bid–ask spread**. When traders ask, "What's the market?" they want to know the best bid and ask prices and their associated sizes. Bid–ask spreads are an implicit cost of trading. Markets with small bid–ask spreads are markets in which the costs of trading are small, at least for the sizes quoted. Dealers often quote both bid and ask prices, and in that case, practitioners say that they quote a two-sided market. The market spread is never more than any dealer spread.

12.1 Execution Instructions

Market and limit orders convey the most common execution instructions. A **market order** instructs the broker or exchange to obtain the best price immediately available when filling the order. A **limit order** conveys almost the same instruction: Obtain the best price immediately available, but in no event accept a price higher than a specified limit price when buying or accept a price lower than a specified limit price when selling.

Many people mistakenly believe that limit orders specify the prices at which the orders will trade. Although limit orders do often trade at their limit prices, remember that the first instruction is to obtain the best price available. If better prices are available than the limit price, brokers and exchanges should obtain those prices for their clients.

Market orders generally execute immediately if other traders are willing to take the other side of the trade. The main drawback with market orders is that they can be expensive to execute, especially when the order is placed in a market for a thinly traded security, or more generally, when the order is large relative to the normal trading activity in the market. In that case, a market buy order may fill at a high price, or a market sell order may fill at a low price if no traders are willing to trade at better prices. High purchase prices and low sale prices represent price concessions given to other traders to encourage them to take the other side of the trade. Because the sizes of price concessions can be difficult to predict, and because prices often change between when a trader submits an order and when the order finally fills, the execution prices for market orders are often uncertain.

Buyers and sellers who are concerned about the possibility of trading at unacceptable prices add limit price instructions to their orders. The main problem with limit orders is that they may not execute. Limit orders do not execute if the limit price on a buy order is too low, or if the limit price on a sell order is too high. For example, if an investment manager submits a limit order to buy at the limit price of 20 (buy limit 20) and nobody is willing to sell at or below 20, the order will not trade. If prices never drop to 20, the manager will never buy. If the price subsequently rises, the manager will have lost the opportunity to profit from the price rise.

Whether traders use market orders or limit orders when trying to arrange trades depends on their concerns about price, trading quickly, and failing to trade. On average, limit orders trade at better prices than do market orders, but they often do not trade. Traders generally regret when their limit orders fail to trade because they usually would have profited if they had traded. Limit buy orders do not fill when prices are rising, and limit sell orders do not fill when prices are falling. In both cases, traders would be better off if their orders had filled.

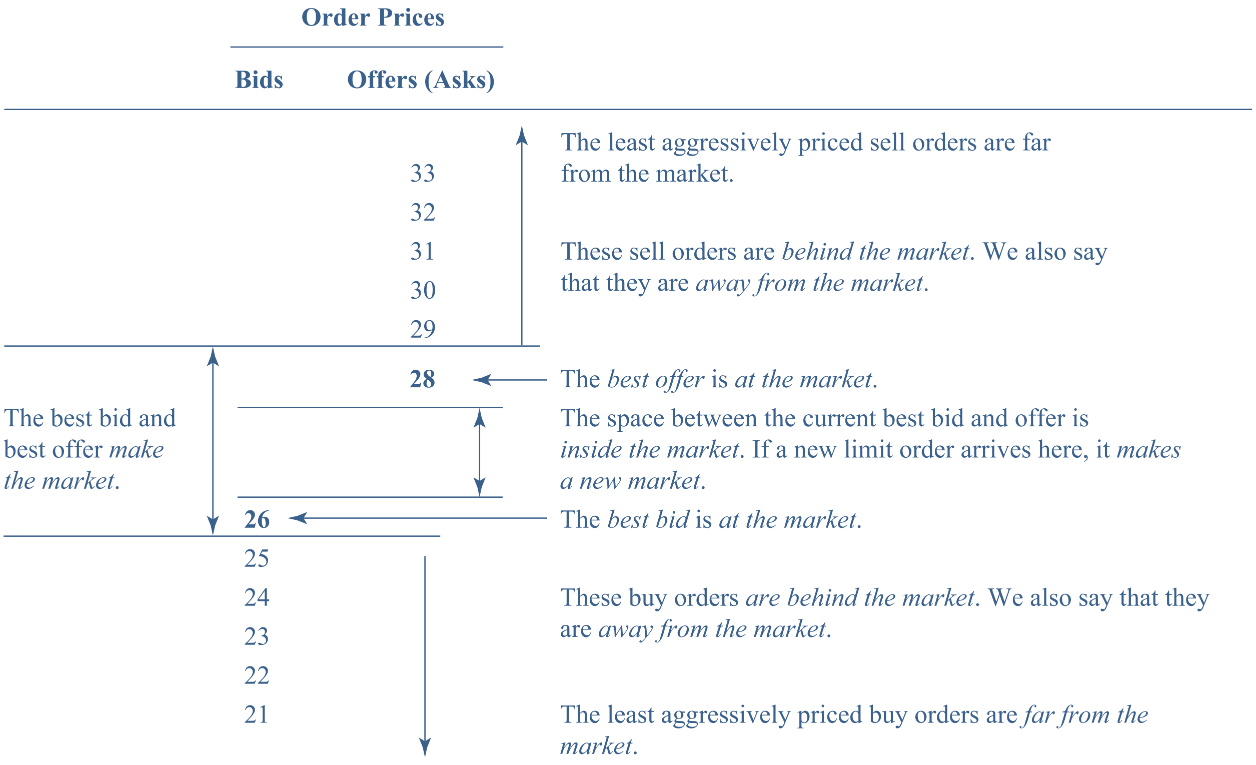
The probability that a limit order will execute depends on where the order is placed relative to market prices. An aggressively priced order is more likely to trade than is a less aggressively priced order. A limit buy order is aggressively priced when the limit price is high relative to the market bid and ask prices. If the limit price is placed above the best offer, the buy order generally will partially or completely fill at the best offer price, depending on the size available at the best offer. Such limit orders are called **marketable limit orders** because at least part of the order can trade immediately. A limit buy order with a very high price relative to the market is essentially a market order.

If the buy order is placed above the best bid but below the best offer, traders say the order makes a new market because it becomes the new best bid. Such orders generally will not immediately trade, but they may attract sellers who are interested in trading. A buy order placed at the best bid is said to make market. It may have to wait until all other buy orders at that price trade first. Finally, a buy order placed below the best bid is **behind the market**. It will not execute unless market prices drop. Traders call limit orders that are waiting to trade **standing limit orders**.

Sell limit orders are aggressively priced if the limit price is low relative to market prices. The limit price of a marketable sell limit order is below the best bid. A limit sell order placed between the best bid and the best offer makes a new market on the sell side, one placed at the best offer makes market, and one placed above the best offer is behind the market.

Exhibit 2 presents a simplified **limit order book** in which orders are presented ranked by their limit prices for a hypothetical market. The market is "26 bid, offered at 28" because the best bid is 26 and the best offer (ask) is 28.

Exhibit 2 Terms Traders Use to Describe Standing Limit Orders



The least aggressively priced sell orders are far from the market.

These sell orders are *behind the market*. We also say that they are *away from the market*.

The *best offer* is *at the market*.

The best bid and best offer *make the market*.

The space between the current best bid and offer is *inside the market*. If a new limit order arrives here, it *makes a new market*.

The *best bid* is *at the market*.

These buy orders *are behind the market*. We also say that they are *away from the market*.

The least aggressively priced buy orders are *far from the market*.

Source: Trading and Exchanges.
Harris, Larry. 2003. *Trading and Exchanges: Market Microstructure for Practitioners*. New York: Oxford University Press.

EXAMPLE 21

Making and Taking

1. What is the difference between making a market and taking a market?
2. What order types are most likely associated with making a market and taking a market?

Solution to 1:

A trader makes a market when the trader offers to trade. A trader takes a market when the trader accepts an offer to trade.

Solution to 2:

Traders place standing limit orders to give other traders opportunities to trade. Standing limit orders thus make markets. In contrast, traders use market orders or marketable limit orders to take offers to trade. These marketable orders take the market.

A trade-off exists between how aggressively priced an order is and the ultimate trade price. Although aggressively priced orders fill faster and with more certainty then do less aggressively priced limit orders, the prices at which they execute are inferior. Buyers seeking to trade quickly must pay higher prices to increase the probability of trading quickly. Similarly, sellers seeking to trade quickly must accept lower prices to increase the probability of trading quickly.

Some order execution instructions specify conditions on size. For example, **all-or-nothing (AON) orders** can only trade if their entire sizes can be traded. Traders can similarly specify minimum fill sizes. This specification is common when settlement costs depend on the number of trades made to fill an order and not on the aggregate size of the order.

Exposure instructions indicate whether, how, and perhaps to whom orders should be exposed. **Hidden orders** are exposed only to the brokers or exchanges that receive them. These agencies cannot disclose hidden orders to other traders until they can fill them. Traders use hidden orders when they are afraid that other traders might behave strategically if they knew that a large order was in the market. Traders can discover hidden size only by submitting orders that will trade with that size. Thus, traders can only learn about hidden size after they have committed to trading with it.

Traders also often indicate a specific **display size** for their orders. Brokers and exchanges then expose only the display size for these orders. Any additional size is hidden from the public but can be filled if a suitably large order arrives. Traders sometimes call such orders **iceberg orders** because most of the order is hidden. Traders specify display sizes when they do not want to display their full sizes, but still want other traders to know that someone is willing to trade at the displayed price. Traders on the opposite side who wish to trade additional size at that price can discover the hidden size only if they trade the displayed size, at which point the broker or exchange will display any remaining size up to the display size. They also can discover the hidden size by submitting large orders that will trade with that size.

EXAMPLE 22

Market versus Limit and Hidden versus Displayed Orders

You are the buy-side trader for a very clever investment manager. The manager has hired a commercial satellite firm to take regular pictures of the parking lots in which new car dealers store their inventories. It has also hired some part-time workers to count the cars on the lots. With this information and some econometric analyses, the manager can predict weekly new car sale announcements more accurately than can most analysts. The manager typically makes a quarter percent each week on this strategy. Once a week, a day before the announcements are made, the manager gives you large orders to buy or sell car manufacturers based on his insights into their dealers' sales. What primary issues should you consider when deciding whether to:

1 use market or limit orders to fill his orders?

2 display the orders or hide them?

Solution to 1:

The manager's information is quite perishable. If his orders are not filled before the weekly sales are reported to the public, the manager will lose the opportunity to profit from the information as prices immediately adjust to the news. The manager, therefore, needs to get the orders filled quickly. This consideration suggests that the orders should be submitted as market orders. If submitted as limit orders, the orders might not execute and the firm would lose the opportunity to profit.

Large market orders, however, can be very expensive to execute, especially if few people are willing to trade significant size on the other side of the market. Because transaction costs can easily exceed the expected quarter percent return, you should submit limit orders to limit the execution prices that you are willing to accept. It is better to fail to trade than to trade at losing prices.

Solution to 2:

Your large orders could easily move the market if many people were aware of them, and even more so if others were aware that you are trading on behalf of a successful information-motivated trader. You thus should consider submitting hidden orders. The disadvantage of hidden orders is that they do not let people know that they can trade the other side if they want to.

VALIDITY INSTRUCTIONS AND CLEARING INSTRUCTIONS

13

g compare execution, validity, and clearing instructions;

Validity instructions indicate when an order may be filled. The most common validity instruction is the **day order**. A day order is good for the day on which it is submitted. If it has not been filled by the close of business, the order expires unfilled.

Good-till-cancelled orders (GTC) are just that. In practice, most brokers limit how long they will manage an order to ensure that they do not fill orders that their clients have forgotten. Such brokers may limit their GTC orders to a few months.

Immediate or cancel orders (IOC) are good only upon receipt by the broker or exchange. If they cannot be filled in part or in whole, they cancel immediately. In some markets these orders are also known as **fill or kill** orders. When searching for hidden liquidity, electronic algorithmic trading systems often submit thousands of these IOC orders for every order that they fill.

Good-on-close orders can only be filled at the close of trading. These orders often are market orders, so traders call them **market-on-close** orders. Traders often use on-close orders when they want to trade at the same prices that will be published as the closing prices of the day. Mutual funds often like to trade at such prices because they value their portfolios at closing prices. Many traders also use **good-on-open** orders.

13.1 Stop Orders

A **stop order** is an order in which a trader has specified a stop price condition. The stop order may not be filled until the stop price condition has been satisfied. For a sell order, the stop price condition suspends execution of the order until a trade occurs at or below the stop price. After that trade, the stop condition is satisfied and the order becomes valid for execution, subject to all other execution instructions attached to it. If the market price subsequently rises above the sell order's stop price before the order trades, the order remains valid. Similarly, a buy order with a stop condition becomes valid only after a price rises above the specified stop price.

Traders often call stop orders **stop-loss orders** because many traders use them with the hope of stopping losses on positions that they have established. For example, a trader who has bought stock at 40 may want to sell the stock if the price falls below 30. In that case, the trader might submit a "GTC, stop 30, market sell" order. If the price falls to or below 30, the market order becomes valid and it should immediately execute at the best price then available in the market. That price may be substantially lower than 30 if the market is falling quickly. The stop-loss order thus does not guarantee a stop to losses at the stop price. If potential sellers are worried about trading at too low of a price, they can attach stop instructions to limit orders instead of market orders. In this example, if the trader is unwilling to sell below 25, the trader would submit a "GTC, stop 30, limit 25 sell" order.

If a trader wants to guarantee that he can sell at 30, the trader would buy a put option contract struck at 30. The purchase price of the option would include a premium for the insurance that the trader is buying. Option contracts can be viewed as limit orders for which execution is guaranteed at the strike price. A trader similarly might use a stop-buy order or a call option to limit losses on a short position.

A portfolio manager might use a stop-buy order when the manager believes that a security is undervalued but is unwilling to trade without market confirmation. For example, suppose that a stock currently trades for 50 RMB and a manager believes that it should be worth 100 RMB. Further, the manager believes that the stock will much more likely be worth 100 RMB if other traders are willing to buy it above 65 RMB. To best take advantage of this information, the manager would consider issuing a "GTC, stop 65 RMB, limit 100 RMB buy" order. Note that if the manager relies too much on the market when making this trading decision, however, he may violate CFA Standard of Professional Conduct V.A.2, which requires that all investment actions have a reasonable and adequate basis supported by appropriate research and investigation.

Because stop-sell orders become valid when prices are falling and stop-buy orders become valid when prices are rising, traders using stop orders contribute to market momentum as their sell orders push prices down further and their buy orders push prices up. Execution prices for stop orders thus are often quite poor.

EXAMPLE 23

Limit and Stop Instructions

In what ways do limit and stop instructions differ?

Solution:

Although both limit and stop instructions specify prices, the role that these prices play in the arrangement of a trade are completely different. A limit price places a limit on what trade prices will be acceptable to the trader. A buyer will accept prices only at or lower than the limit price whereas a seller will accept prices only at or above the limit price.

In contrast, a stop price indicates when an order can be filled. A buy order can only be filled once the market has traded at a price at or above the stop price. A sell order can only be filled once the market has traded at a price at or below the stop price.

Both order instructions may delay or prevent the execution of an order. A buy limit order will not execute until someone is willing to sell at or below the limit price. Similarly, a sell limit order will not execute until someone is willing to buy at or above the limit sell price. In contrast, a stop-buy order will not execute if the market price never rises to the stop price. Similarly, a stop-sell order will not execute if the market price never falls to the stop price.

13.2 Clearing Instructions

Clearing instructions tell brokers and exchanges how to arrange final settlement of trades. Traders generally do not attach these instructions to each order—instead they provide them as standing instructions. These instructions indicate what entity is responsible for clearing and settling the trade. For retail trades, that entity is the customer's broker. For institutional trades, that entity may be a custodian or another broker. When a client uses one broker to arrange trades and another broker to settle trades, traders say that the first broker gives up the trade to the other broker, who is often known as the prime broker. Institutional traders provide these instructions so they can obtain specialized execution services from different brokers while maintaining a single account for custodial services and other prime brokerage services, such as margin loans.

An important clearing instruction that must appear on security sale orders is an indication of whether the sale is a long sale or a short sale. In either case, the broker representing the sell order must ensure that the trader can deliver securities for settlement. For a long sale, the broker must confirm that the securities held are available for delivery. For a short sale, the broker must either borrow the security on behalf of the client or confirm that the client can borrow the security.

PRIMARY SECURITY MARKETS

14

i. define primary and secondary markets and explain how secondary markets support primary markets;

When issuers first sell their securities to investors, practitioners say that the trades take place in the **primary markets**. An issuer makes an **initial public offering** (IPO)—sometimes called a placing—of a security issue when it sells the security to the public for the first time. A seasoned security is a security that an issuer has already issued. If the issuer wants to sell additional units of a previously issued security, it makes a **seasoned offering** (sometimes called a secondary offering). Both types of offerings occur in the **primary market** where issuers sell their securities to investors. Later, if investors trade these securities among themselves, they trade in **secondary markets**. This section discusses primary markets and the procedures that issuers use to offer their securities to the public.

14.1 Public Offerings

Corporations generally contract with an investment bank to help them sell their securities to the public. The investment bank then lines up subscribers who will buy the security. Investment bankers call this process **book building**. In London, the book builder is called the book runner. The bank tries to build a book of orders to which they can sell the offering. Investment banks often support their book building by providing investment information and opinion about the issuer to their clients and to the public. Before the offering, the issuer generally makes a very detailed disclosure of its business, of the risks inherent in it, and of the uses to which the new funds will be placed.

When time is of the essence, issuers in Europe may issue securities through an **accelerated book build**, in which the investment bank arranges the offering in only one or two days. Such sales often occur at discounted prices.

The first public offering of common stock in a company consists of newly issued shares to be sold by the company. It may also include shares that the founders and other early investors in the company seek to sell. The initial public offering provides these investors with a means of liquidating their investments.

In an **underwritten offering**—the most common type of offering—the investment bank guarantees the sale of the issue at an offering price that it negotiates with the issuer. If the issue is undersubscribed, the bank will buy whatever securities it cannot sell at the offering price. In the case of an IPO, the underwriter usually also promises to make a market in the security for about a month to ensure that the secondary market will be liquid and to provide price support, if necessary. For large issues, a syndicate of investment banks and broker–dealers helps the **lead underwriter** build the book. The issuer usually pays an underwriting fee of about 7 percent for these various services. The underwriting fee is a placement cost of the offering.

In a **best effort offering**, the investment bank acts only as broker. If the offering is undersubscribed, the issuer will not sell as much as it hoped to sell.

For both types of offerings, the issuer and the bank usually jointly set the offering price following a negotiation. If they set a price that buyers consider too high, the offering will be undersubscribed, and they will fail to sell the entire issue. If they set the price too low, the offering will be oversubscribed, in which case the securities are often allocated to preferred clients or on a pro-rata basis.

(Note that CFA Standard of Professional Conduct III.B—fair dealing—requires that the allocation be based on a written policy disclosed to clients and suggests that the securities be offered on a pro-rata basis among all clients who have comparable relationships with their broker–dealers.)

Investment banks have a conflict of interest with respect to the offering price in underwritten offerings. As agents for the issuers, they generally are supposed to select the offering price that will raise the most money. But as underwriters, they have strong incentives to choose a low price. If the price is low, the banks can allocate valuable shares to benefit their clients and thereby indirectly benefit the banks. If the price is too high, the underwriters will have to buy overvalued shares in the offering and perhaps also during the following month if they must support the price in the secondary market, which directly costs the banks. These considerations tend to lower initial offering prices so that prices in the secondary market often rise immediately following an IPO. They are less important in a seasoned offering because trading in the secondary market helps identify the proper price for the offering.

First time issuers generally accept lower offering prices because they and many others believe that an undersubscribed IPO conveys very unfavorable information to the market about the company's prospects at a time when it is most vulnerable to public opinion about its prospects. They fear that an undersubscribed initial public offering will make it substantially harder to raise additional capital in subsequent seasoned offerings.

EXAMPLE 24

The Healthybots Initial Public Offering

Healthybots is a health care company that treats diseases using artificial intelligence–based solutions. Healthybots raised approximately £265 million gross through an initial public offering of 103,142,466 ordinary shares at £2.57 per ordinary share. After the initial public offering, Healthybots had 213,333,333 ordinary shares issued and outstanding.

Healthybots received gross proceeds of approximately £34.3 million and net proceeds of £31.8 million. The ordinary shares that were sold to the public represented approximately 48 percent of Healthybots' total issued ordinary shares.

The shares commenced trading at 8:00 a.m. on the AIM market of the London Stock Exchange, where Healthybots opened at £2.74, traded 37 million shares between £2.68 and £2.74, and closed at £2.73.

1. Approximately how many new shares were issued by the company and how many shares were sold by the company's founders? What fraction of their holdings in the company did the founders sell?
2. Approximately what return did the subscribers who participated in the IPO make on the first day it traded?
3. Approximately how much did Healthybots pay in placement costs as a percentage of the new funds raised?

Solution to 1:

Healthybots received gross proceeds of £34.3 million at £2.57 per share, so the company issued and sold 13,346,304 shares (= £34.3 million/£2.57 per share). The total placement was for 103,142,466 shares, so the founders sold 89,796,162 shares (= 103,142,466 shares − 13,346,304 shares). Because approximately 200 million shares (= 213.3 million shares – 13.3 million shares) were outstanding before the placement, the founders sold approximately 45 percent (= 90 million shares/200 million shares) of the company.

Solution to 2:

The subscribers bought the stock for £2.57 per share, and it closed at £2.73. The first day return thus was $6.2\% = \frac{2.73 - 2.57}{2.57} \times 100$.

Solution to 3:

Healthybots obtained gross proceeds of £34.3 million but only raised net proceeds of £31.8 million. The £2.5 million difference was the total cost of the placement to the firm, which is 7.9 percent of net proceeds, or new funds raised (£2.5 million/£31.8 million).

14.2 Private Placements and Other Primary Market Transactions

Corporations sometimes issue their securities in private placements. In a **private placement**, corporations sell securities directly to a small group of qualified investors, usually with the assistance of an investment bank. Qualified investors have sufficient knowledge and experience to recognize the risks that they assume, and sufficient wealth to assume those risks responsibly. Most countries allow corporations to do private placements without nearly as much public disclosure as is required for public offerings. Private placements, therefore, may be cheaper than public offerings, but the buyers generally require higher returns (lower purchase prices) because they cannot subsequently trade the securities in an organized secondary market.

Corporations sometimes sell new issues of seasoned securities directly to the public on a piecemeal basis via a shelf registration. In a **shelf registration**, the corporation makes all public disclosures that it would for a regular offering, but it does not sell the shares in a single transaction. Instead, it sells the shares directly into the secondary market over time, generally when it needs additional capital. Shelf registrations provide corporations with flexibility in the timing of their capital transactions, and they can alleviate the downward price pressures often associated with large secondary offerings.

Many corporations may also issue shares via dividend reinvestment plans (DRPs or DRIPs, for short) that allow their shareholders to reinvest their dividends in newly issued shares of the corporation (in particular, DRPs specify that the corporation issue new shares for the plan rather than purchase them on the open market). These plans sometimes also allow existing shareholders and other investors to buy additional stock at a slight discount to current prices.

Finally, corporations can issue new stock via a rights offering. In a rights offering, the corporation distributes rights to buy stock at a fixed price to existing shareholders in proportion to their holdings. Because the rights need not be exercised, they are options. The exercise price, however, is set below the current market price of the stock so that buying stock with the rights is immediately profitable. Consequently, shareholders will experience dilution in the value of their existing shares. They can offset the dilution loss by exercising their rights or by selling the rights to others who will exercise them. Shareholders generally do not like rights offerings because they must provide additional capital (or sell their rights) to avoid losses through dilution. Financial analysts recognize that these securities, although called rights, are actually short-term stock warrants and value them accordingly.

The national governments of financially strong countries generally issue their bonds, notes, and bills in public auctions organized by a government agency (usually associated with the finance ministry). They may also sell them directly to dealers.

Smaller and less financially secure national governments and most regional governments often contract with investment banks to help them sell and distribute their securities. The laws of many governments, however, require that they auction their securities.

EXAMPLE 25

Private and Public Placements

In what ways do private placements differ from public placements?

Solution:

Issuers make private placements to a limited number of investors that generally are financially sophisticated and well informed about risk. The investors generally have some relationship to the issuer. Issuers make public placements when they sell securities to the general public. Public placements generally require substantially more financial disclosure than do private placements.

14.3 Importance of Secondary Markets to Primary Markets

Corporations and governments can raise money in the primary markets at lower cost when their securities will trade in liquid secondary markets. In a **liquid market**, traders can buy or sell with low transaction costs and small price concessions when they want to trade. Buyers value liquidity because they may need to sell their securities to meet liquidity needs. Investors thus will pay more for securities that they can easily sell than for those that they cannot easily sell. Higher prices translate into lower costs of capital for the issuers.

SECONDARY SECURITY MARKET AND CONTRACT MARKET STRUCTURES

15

i. define primary and secondary markets and explain how secondary markets support primary markets;

j. describe how securities, contracts, and currencies are traded in quote-driven, order-driven, and brokered markets;

Trading is the successful outcome to a bilateral search in which buyers look for sellers and sellers look for buyers. Many market structures have developed to reduce the costs of this search. Markets are liquid when the costs of finding a suitable counterparty to a trade are low.

Trading in securities and contracts takes place in a variety of market structures. The structures differ by when trades can be arranged, who arranges the trades, how they do so, and how traders learn about possible trading opportunities and executed trades. This section introduces the various market structures used to trade securities and contracts. We first consider trading sessions, then execution mechanisms, and finally market information systems.

15.1 Trading Sessions

Markets are organized as call markets or as continuous trading markets. In a **call market**, trades can be arranged only when the market is called at a particular time and place. In contrast in a **continuous trading market**, trades can be arranged and executed anytime the market is open.

Buyers can easily find sellers and vice versa in call markets because all traders interested in trading (or orders representing their interests) are present at the same time and place. Call markets thus have the potential to be very liquid when they are called. But they are completely illiquid between trading sessions. In contrast, traders can arrange and execute their trades at any time in continuous trading markets, but doing so can be difficult if the buyers and sellers (or their orders) are not both present at the same time.

Most call markets use single price auctions to match buyers to sellers. In these auctions, the market constructs order books representing all buy orders and all seller orders. The market then chooses a single trade price that will maximize the total volume of trade. The order books are supply and demand schedules, and the point at which they cross determines the trade price.

Call markets usually are organized just once a day, but some markets organize calls at more frequent intervals.

Many continuous trading markets start their trading with a call market auction. During a pre-opening period, traders submit their orders for the market call. At the opening, any possible trades are arranged and then trading continues in the continuous trading session. Some continuous trading markets also close their trading with a call. In these markets, traders who are only interested in trading in the closing call submit market- or limit-on-close orders.

EXAMPLE 26

Call Markets and Continuous Trading Markets

1 What is the main advantage of a call market compared with a continuous trading market?

2 What is the main advantage of a continuous trading market compared with a call market?

Solution to 1:

By gathering all traders to the same place at the same time, a call market makes it easier for buyers to find sellers and vice versa. In contrast, if buyers and sellers (or their orders) are not present at the same time in a continuous market, they cannot trade.

Solution to 2:

In a continuous trading market, a willing buyer and seller can trade at any time the market is open. In contrast, in a call market trading can take place only when the market is called.

15.2 Execution Mechanisms

The three main types of market structures are quote-driven markets (sometimes called price-driven or dealer markets), order-driven markets, and brokered markets. In **quote-driven markets**, customers trade with dealers. In **order-driven markets**, an order matching system run by an exchange, a broker, or an alternative trading system uses rules to arrange trades based on the orders that traders submit. Most exchanges and ECNs organize order-driven markets. In **brokered markets**, brokers arrange trades between their customers. Brokered markets are common for transactions of unique instruments, such as real estate properties, intellectual properties, or large blocks of securities. Many trading systems use more than one type of market structure.

15.2.1 *Quote-Driven Markets*

Worldwide, most trading, other than in stocks, takes place in quote-driven markets. Almost all bonds and currencies and most spot commodities trade in quote-driven markets. Traders call them quote-driven (or price-driven or dealer) because customers trade at the prices quoted by dealers. Depending on the instrument traded, the dealers work for commercial banks, for investment banks, for broker–dealers, or for proprietary trading houses.

Quote-driven markets also often are called over-the-counter (OTC) markets because securities used to be literally traded over the dealer's counter in the dealer's office. Now, most trades in OTC markets are conducted over proprietary computer communications networks, by telephone, or sometimes over instant messaging systems.

15.2.2 *Order-Driven Markets*

Order-driven markets arrange trades using rules to match buy orders to sell orders. The orders may be submitted by customers or by dealers. Almost all exchanges use order-driven trading systems, and every automated trading system is an order-driven system.

Because rules match buyers to sellers, traders often trade with complete strangers. Order-driven markets thus must have procedures to ensure that buyers and sellers perform on their trade contracts. Otherwise, dishonest traders would enter contracts that they would not settle if a change in market conditions made settlement unprofitable.

Two sets of rules characterize order-driven market mechanisms: Order matching rules and trade pricing rules. The order matching rules match buy orders to sell orders. The trade pricing rules determine the prices at which the matched trades take place.

15.2.2.1 Order Matching Rules Order-driven trading systems match buyers to sellers using rules that rank the buy orders and the sell orders based on price, and often along with other secondary criteria. The systems then match the highest-ranking buy order with the highest-ranking sell order. If the buyer is willing to pay at least as much as the seller is willing to receive, the system will arrange a trade for the minimum of the buy and sell quantities. The remaining size, if any, is then matched with the next order on the other side and the process continues until no further trades can be arranged.

The **order precedence hierarchy** determines which orders go first. The first rule is **price priority**: The highest priced buy orders and the lowest priced sell orders go first. They are the most aggressively priced orders. **Secondary precedence rules** determine how to rank orders at the same price. Most trading systems use time precedence to rank orders at the same price. The first order to arrive has precedence over other orders. In trading systems that permit hidden and partially hidden orders, displayed quantities at a given price generally have precedence over the undisplayed quantities. So the complete precedence hierarchy is given by price priority, display precedence at a given price, and finally time precedence among all orders with the same display status at a given price. These rules give traders incentives to improve price, display their orders, and arrive early if they want to trade quickly. These incentives increase market liquidity.

15.2.2.2 Trade Pricing Rules After the orders are matched, the trading system then uses its trade pricing rule to determine the trade price. The three rules that various order-driven markets use to price their trades are the uniform pricing rule, the discriminatory pricing rule, and the derivative pricing rule.

Call markets commonly use the uniform pricing rule. Under this rule, all trades execute at the same price. The market chooses the price that maximizes the total quantity traded.

Continuous trading markets use the **discriminatory pricing rule**. Under this rule, the limit price of the order or quote that first arrived—the standing order—determines the trade price. This rule allows a large arriving trader to discriminate among standing limit orders by filling the most aggressively priced orders first at their limit prices and then filling less aggressively priced orders at their less favorable (from the point of view of the arriving trader) limit prices. If trading systems did not use this pricing rule, large traders would break their orders into pieces to price discriminate on their own.

EXAMPLE 27

Filling a Large Order in a Continuous Trading Market

Before the arrival of a large order, the Tokyo Stock Exchange has the following limit orders standing on its book:

Buyer	Bid Size	Limit Price(¥)	Offer Size	Seller
Takumi	15	100.1		
Hiroto	8	100.2		
Shou	10	100.3		
		100.4	4	Hina
		100.5	6	Sakura
		100.6	12	Miku

Tsubasa submits a day order to buy 15 contracts, limit ¥100.5. With whom does he trade, what is his average trade price, and what does the limit order book look like afterward?

Solution:

Tsubasa's buy order first fills with the most aggressively priced sell order, which is Hina's order for four contracts. A trade takes place at ¥100.4 for four contracts, Hina's order fills completely, and Tsubasa still has 11 more contracts remaining.

The next most aggressively priced sell order is Sakura's order for six contracts. A second trade takes place at ¥100.5 for six contracts, Sakura's order fills completely, and Tsubasa still has five more contracts remaining.

The next most aggressively priced sell order is Miku's order at ¥100.6. No further trade is possible, however, because her limit sell price is above Tsubasa's limit buy price. Tsubasa's average trade price is $¥100.46 = \frac{4 \times ¥100.4 + 6 \times ¥100.5}{4 + 6}$.

Because Tsubasa issued a day order, the remainder of his order is placed on the book on the buy side at ¥100.5. The following orders are then on the book:

Buyer	Bid Size	Limit Price (¥)	Offer Size	Seller
Takumi	15	100.1		
Hiroto	8	100.2		
Shou	10	100.3		
		100.4		
Tsubasa	5	100.5		
		100.6	12	Miku

If Tsubasa had issued an immediate-or-cancel order, the remaining five contracts would have been cancelled.

Crossing networks use the derivative pricing rule. **Crossing networks** are trading systems that match buyers and sellers who are willing to trade at prices obtained from other markets. Most systems cross their trades at the midpoint of the best bid and ask quotes published by the exchange at which the security primarily trades. This pricing rule is called a **derivative pricing rule** because the price is derived from another market. In particular, the price does not depend on the orders submitted to the crossing network. Some crossing networks are organized as call markets and others as continuously trading markets. The most important crossing market is the equity trading system POSIT.

15.2.3 *Brokered Markets*

The third execution mechanism is the **brokered market**, in which brokers arrange trades among their clients. Brokers organize markets for instruments for which finding a buyer or a seller willing to trade is difficult because the instruments are unique and thus of interest only to a limited number of people or institutions. These instruments generally are also infrequently traded and expensive to carry in inventory. Examples of such instruments include very large blocks of stock, real estate properties, fine art masterpieces, intellectual properties, operating companies, liquor licenses, and taxi medallions. Because dealers generally are unable or unwilling to hold these assets in their inventories, they will not make markets in them. Organizing order-driven markets for these instruments is not sensible because too few traders would submit orders to them.

Successful brokers in these markets try to know everyone who might now or in the future be willing to trade. They spend most of their time on the telephone and in meetings building their networks.

EXAMPLE 28

Quote-Driven, Order-Driven, and Brokered Markets

What are the primary advantages of quote-driven, order-driven, and brokered markets?

Solution:

In a quote-driven market, dealers generally are available to supply liquidity. In an order-driven market, traders can supply liquidity to each other. In a brokered market, brokers help find traders who are willing to trade when dealers would not be willing to make markets and when traders would not be willing to post orders.

15.3 Market Information Systems

Markets vary in the type and quantity of data that they disseminate to the public. Traders say that a market is pre-trade transparent if the market publishes real-time data about quotes and orders. Markets are post-trade transparent if the market publishes trade prices and sizes soon after trades occur.

Buy-side traders value transparency because it allows them to better manage their trading, understand market values, and estimate their prospective and actual transaction costs. In contrast, dealers prefer to trade in opaque markets because, as frequent traders, they have an information advantage over those who know less than they do. Bid–ask spreads tend to be wider and transaction costs tend to be higher in opaque markets because finding the best available price is harder for traders in such markets.

WELL-FUNCTIONING FINANCIAL SYSTEMS

16

k describe characteristics of a well-functioning financial system;

The financial system allows traders to solve financing and risk management problems. In a well-functioning financial system:

- investors can easily move money from the present to the future while obtaining a fair rate of return for the risks that they bear;
- borrowers can easily obtain funds that they need to undertake current projects if they can credibly promise to repay the funds in the future;
- hedgers can easily trade away or offset the risks that concern them; and
- traders can easily trade currencies for other currencies or commodities that they need.

If the assets or contracts needed to solve these problems are available to trade, the financial system has **complete markets**. If the costs of arranging these trades are low, the financial system is **operationally efficient**. If the prices of the assets and contracts reflect all available information related to fundamental values, the financial system is informationally efficient.

Well-functioning financial systems are characterized by:

- the existence of well-developed markets that trade instruments that help people solve their financial problems (complete markets);
- liquid markets in which the costs of trading—commissions, bid–ask spreads, and order price impacts—are low (operationally efficient markets);
- timely financial disclosures by corporations and governments that allow market participants to estimate the fundamental values of securities (support **informationally efficient markets**); and
- prices that reflect fundamental values so that prices vary primarily in response to changes in fundamental values and not to demands for liquidity made by uninformed traders (informationally efficient markets).

Such complete and operationally efficient markets are produced by financial intermediaries who:

- organize exchanges, brokerages, and alternative trading systems that match buyers to sellers;
- provide liquidity on demand to traders;
- securitize assets to produce investment instruments that are attractive to investors and thereby lower the costs of funds for borrowers;
- run banks that match investors to borrowers by taking deposits and making loans;
- run insurance companies that pool uncorrelated risks;
- provide investment advisory services that help investors manage and grow their assets at low cost;
- organize clearinghouses that ensure everyone settles their trades and contracts; and
- organize depositories that ensure nobody loses their assets.

The benefits of a well-functioning financial system are huge. In such systems, investors who need to move money to the future can easily connect with entrepreneurs who need money now to develop new products and services. Similarly, producers who would otherwise avoid valuable projects because they are too risky can easily transfer those risks to others who can better bear them. Most importantly, these transactions generally can take place among strangers so that the benefits from trading can be derived from an enormous number of potential matches.

In contrast, economies that have poorly functioning financial systems have great difficulties allocating capital among the many companies who could use it. Financial transactions tend to be limited to arrangements within families when people cannot easily find trustworthy counterparties who will honor their contracts. In such economies, capital is allocated inefficiently, risks are not easily shared, and production is inefficient.

An extraordinarily important byproduct of an operationally efficient financial system is the production of informationally efficient prices. Prices are informationally efficient when they reflect all available information about fundamental values. Informative prices are crucially important to the welfare of an economy because they help ensure that resources go where they are most valuable. Economies that use resources where they are most valuable are allocationally efficient. Economies that do not use resources where they are most valuable waste their resources and consequently often are quite poor.

Well-informed traders make prices informationally efficient. When they buy assets and contracts that they think are undervalued, they tend to push the assets' prices up. Similarly, when they sell assets and contracts that they think are overvalued, they tend to push the assets' prices down. The effect of their trading thus causes prices to reflect their information about values.

How accurately prices reflect fundamental information depends on the costs of obtaining fundamental information and on the liquidity available to well-informed traders. Accounting standards and reporting requirements that produce meaningful and timely financial disclosures reduce the costs of obtaining fundamental information and thereby allow analysts to form more accurate estimates of fundamental values. Liquid markets allow well-informed traders to fill their orders at low cost. If filling orders is very costly, informed trading may not be profitable. In that case, information-motivated traders will not commit resources to collect and analyze data and they will not trade. Without their research and their associated trading, prices would be less informative.

EXAMPLE 29

Well-Functioning Financial Systems

As a financial analyst specializing in emerging market equities, you understand that a well-functioning financial system contributes to the economic prosperity of a country. You are asked to start covering a new small market country. What factors will you consider when characterizing the quality of its financial markets?

Solution:

In general, you will consider whether:

- the country has markets that allow its companies and residents to finance projects, save for the future, and exchange risk;
- the costs of trading in those markets is low; and
- prices reflect fundamental values.

You may specifically check to see whether:

- fixed income and stock markets allow borrowers to easily obtain capital from investors;
- corporations disclose financial and operating data on a timely basis in conformity to widely respected reporting standards, such as IFRS;
- forward, futures, and options markets trade instruments that companies need to hedge their risks;
- dealers and arbitrageurs allow traders to trade when they want to;
- bid–ask spreads are small;
- trades and contracts invariably settle as expected;
- investment managers provide high-quality management services for reasonable fees;
- banks and other financing companies are well capitalized and thus able to help investors provide capital to borrowers;
- securitized assets are available and represent reasonable credit risks;

- insurance companies are well capitalized and thus able to help those exposed to risks insure against them; and
- price volatility appears consistent with changes in fundamental values.

17 MARKET REGULATION

l. describe objectives of market regulation.

Government agencies and practitioner organizations regulate many markets and the financial intermediaries that participate in them. The regulators generally seek to promote fair and orderly markets in which traders can trade at prices that accurately reflect fundamental values without incurring excessive transaction costs. This section identifies the problems that financial regulators hope to solve and the objectives of their regulations.

Regrettably, some people will steal from each other if given a chance, especially if the probability of detection is low or if the penalty for being caught is low. The number of ways that people can steal or misappropriate wealth generally increases with the complexity of their relationships and with asymmetries in their knowledge. Because financial markets tend to be complex, and because customers are often much less sophisticated than the professionals that serve them, the potential for losses through various frauds can be unacceptably high in unregulated markets.

Regulators thus ensure that systems are in place to protect customers from fraud. In principle, the customers themselves would demand such systems as a condition of doing business. When customers are unsophisticated or poorly informed, however, they may not know how to protect themselves. When the costs of learning are large—as they often are in complex financial markets—having regulators look out for the public interest can be economically efficient.

More customer money is probably lost in financial markets through negligence than through outright fraud. Most customers in financial markets use various agents to help them solve problems that they do not understand well. These agents include securities brokers, financial advisers, investment managers, and insurance agents. Because customers generally do not have much information about market conditions, they find it extremely difficult to measure the added value they obtain from their agents. This problem is especially challenging when performance has a strong random component. In that case, determining whether agents are skilled or lucky is very difficult. Moreover, if the agent is a good salesman, the customer may not critically evaluate their agent's performance. These conditions, which characterize most financial markets, ensure that customers cannot easily determine whether their agents are working faithfully for them. They tend to lose if their agents are unqualified or lazy, or if they unconsciously favor themselves and their friends over their clients, as is natural for even the most honest people.

Regulators help solve these agency problems by setting minimum standards of competence for agents and by defining and enforcing minimum standards of practice. CFA Institute provides significant standard setting leadership in the areas of investment management and investment performance reporting through its Chartered Financial Analyst Program, in which you are studying, and its Global Investment Performance Standards. In principle, regulation would not be necessary if customers could identify competent agents and effectively measure their performance. In the financial markets, doing so is very difficult.

Regulators often act to level the playing field for market participants. For example, in many jurisdictions, insider trading in securities is illegal. The rule prevents corporate insiders and others with access to corporate information from trading on material information that has not been released to the public. The purpose of the rule is to reduce the profits that insiders could extract from the markets. These profits would come from other traders who would lose when they trade with well-informed insiders. Because traders tend to withdraw from markets when they lose, rules against insider trading help keep markets liquid. They also keep corporate insiders from hoarding information.

Many situations arise in financial markets in which common standards benefit everyone involved. For example, having all companies report financial results on a common basis allows financial analysts to easily compare companies. Accordingly, the International Accounting Standards Board (IASB) and the US-based Financial Accounting Standards Board (FASB), among many others, promulgate common financial standards to which all companies must report. The benefits of having common reporting standards has led to a very successful and continuing effort to converge all accounting standards to a single worldwide standard. Without such regulations, investors might eventually refuse to invest in companies that do not report to a common standard, but such market-based discipline is a very slow regulator of behavior, and it would have little effect on companies that do not need to raise new capital.

Regulators generally require that financial firms maintain minimum levels of capital. These capital requirements serve two purposes. First, they ensure that the companies will be able to honor their contractual commitments when unexpected market movements or poor decisions cause them to lose money. Second, they ensure that the owners of financial firms have substantial interest in the decisions that they make. Without a substantial financial interest in the decisions that they make, companies often take too many risks and exercise poor judgment about extending credit to others. When such companies fail, they impose significant costs on others. Minimum capital requirements reduce the probability that financial firms will fail and they reduce the disruptions associated with those failures that do occur. In principle, a firm's customers and counterparties could require minimum capital levels as a condition of doing business with the firm, but they have more difficulty enforcing their contracts than do governments who can imprison people.

Regulators similarly regulate insurance companies and pension funds that make long-term promises to their clients. Such entities need to maintain adequate reserves to ensure that they can fund their liabilities. Unfortunately, their managers have a tendency to underestimate these reserves if they will not be around when the liabilities come due. Again, in principle, policyholders and employees could regulate the behavior of their insurance funds and their employers by refusing to contract with them if they do not promise to adequately fund their liabilities. In practice, however, the sophistication, information, and time necessary to write and enforce contracts that control these problems are beyond the reach of most people. The government thus is a sensible regulator of such problems.

Many regulators are self-regulating organizations (SROs) that regulate their members. Exchanges, clearinghouses, and dealer trade organizations are examples of self-regulating organizations. In some cases, the members of these organizations voluntarily subject themselves to the SRO's regulations to promote the common good. In other cases, governments delegate regulatory and enforcement authorities to SROs, usually subject to the supervision of a government agency, such as a national securities and exchange authority. Exchanges, dealer associations, and clearing agencies often regulate their members with these delegated powers.

By setting high standards of behavior, SROs help their members obtain the confidence of their customers. They also reduce the chance that members of the SRO will incur losses when dealing with other members of the SRO.

When regulators fail to solve the problems discussed here, the financial system does not function well. People who lose money stop saving and borrowers with good ideas cannot fund their projects. Similarly, hedgers withdraw from markets when the costs of hedging are high. Without the ability to hedge, producers become reluctant to specialize because specialization generally increases risk. Because specialization also decreases costs, however, production becomes less efficient as producers chose safer technologies. Economies that cannot solve the regulatory problems described in this section tend to operate less efficiently than do better regulated economies, and they tend to be less wealthy.

To summarize, the objectives of market regulation are to:

1 control fraud;
2 control agency problems;
3 promote fairness;
4 set mutually beneficial standards;
5 prevent undercapitalized financial firms from exploiting their investors by making excessively risky investments; and
6 ensure that long-term liabilities are funded.

Regulation is necessary because regulating certain behaviors through market-based mechanisms is too costly for people who are unsophisticated and uninformed. Effectively regulated markets allow people to better achieve their financial goals.

EXAMPLE 30

Bankrupt Traders

You are the chief executive officer of a brokerage that is a member of a clearinghouse. A trader who clears through your firm is bankrupt at midday, but you do not yet know it even though your clearing agreement with him explicitly requires that he immediately report significant losses. The trader knows that if he takes a large position, prices might move in his favor so that he will no longer be bankrupt. The trader attempts to do so and succeeds. You find out about this later in the evening.

1 Why does the clearinghouse regulate its members?
2 What should you do about the trader?
3 Why would the clearinghouse allow you to keep his trading profits?

Solution to 1:

The clearinghouse regulates its members to ensure that no member imposes costs on another member by failing to settle a trade.

Solution to 2:

You should immediately end your clearing relationship with the trader and confiscate his trading profits. The trader was trading with your firm's capital after he became bankrupt. Had he lost, your firm would have borne the loss.

Solution to 3:

If the clearinghouse did not permit you to keep his trading profits, other traders similarly situated might attempt the same strategy.

SUMMARY

This reading introduces how the financial system operates and explains how well-functioning financial systems lead to wealthy economies. Financial analysts need to understand how the financial system works because their analyses often lead to trading decisions.

The financial system consists of markets and the financial intermediaries that operate in them. These institutions allow buyers to connect with sellers. They may trade directly with each other when they trade the same instrument or they only may trade indirectly when a financial intermediary connects the buyer to the seller through transactions with each that appear on the intermediary's balance sheet. The buyer and seller may exchange instruments, cash flows, or risks.

The following points, among others, were made in this reading:

- The financial system consists of mechanisms that allow strangers to contract with each other to move money through time, to hedge risks, and to exchange assets that they value less for those that they value more.
- Investors move money from the present to the future when they save. They expect a normal rate of return for bearing risk through time. Borrowers move money from the future to the present to fund current projects and expenditures. Hedgers trade to reduce their exposure to risks they prefer not to take. Information-motivated traders are active investment managers who try to identify under- and overvalued instruments.
- Securities are first sold in primary markets by their issuers. They then trade in secondary markets.
- People invest in pooled investment vehicles to benefit from the investment management services of their managers.
- Forward contracts allow buyers and sellers to arrange for future sales at predetermined prices. Futures contracts are forward contracts guaranteed by clearinghouses. The guarantee ensures that strangers are willing to trade with each other and that traders can offset their positions by trading with anybody. These features of futures contract markets make them highly attractive to hedgers and information-motivated traders.
- Many financial intermediaries connect buyers to sellers in a given instrument, acting directly as brokers and exchanges or indirectly as dealers and arbitrageurs.
- Financial intermediaries create instruments when they conduct arbitrage, securitize assets, borrow to lend, manage investment funds, or pool insurance contracts. These activities all transform cash flows and risks from one form to another. Their services allow buyers and sellers to connect with each other through instruments that meet their specific needs.
- Financial markets work best when strangers can contract with each other without worrying about whether their counterparts are able and willing to honor their contract. Clearinghouses, variation margins, maintenance margins, and settlement guarantees made by creditworthy brokers on behalf of their clients help manage credit risk and ultimately allow strangers to contract with each other.
- Information-motivated traders short sell when they expect that prices will fall. Hedgers short sell to reduce the risks of a long position in a related contract or commodity.

- Margin loans allow people to buy more securities than their equity would otherwise permit them to buy. The larger positions expose them to more risk so that gains and losses for a given amount of equity will be larger. The leverage ratio is the value of a position divided by the value of the equity supporting it. The returns to the equity in a position are equal to the leverage ratio times the returns to the unleveraged position.
- To protect against credit losses, brokers demand maintenance margin payments from their customers who have borrowed cash or securities when adverse price changes cause their customer's equity to drop below the maintenance margin ratio. Brokers close positions for customers who do not satisfy these margin calls.
- Orders are instructions to trade. They always specify instrument, side (buy or sell), and quantity. They usually also provide several other instructions.
- Market orders tend to fill quickly but often at inferior prices. Limit orders generally fill at better prices if they fill, but they may not fill. Traders choose order submission strategies on the basis of how quickly they want to trade, the prices they are willing to accept, and the consequences of failing to trade.
- Stop instructions are attached to other orders to delay efforts to fill them until the stop condition is satisfied. Although stop orders are often used to stop losses, they are not always effective.
- Issuers sell their securities using underwritten public offerings, best efforts public offerings, private placements, shelf registrations, dividend reinvestment programs, and rights offerings. Investment banks have a conflict of interests when setting the initial offering price in an IPO.
- Well-functioning secondary markets are essential to raising capital in the primary markets because investors value the ability to sell their securities if they no longer want to hold them or if they need to disinvest to raise cash. If they cannot trade their securities in a liquid market, they will not pay as much for them.
- Matching buyers and sellers in call markets is easy because the traders (or their orders) come together at the same time and place.
- Dealers provide liquidity in quote-driven markets. Public traders as well as dealers provide liquidity in order-driven markets.
- Order-driven markets arrange trades by ranking orders using precedence rules. The rules generally ensure that traders who provide the best prices, display the most size, and arrive early trade first. Continuous order-driven markets price orders using the discriminatory pricing rule. Under this rule, standing limit orders determine trade prices.
- Brokers help people trade unique instruments or positions for which finding a buyer or a seller is difficult.
- Transaction costs are lower in transparent markets than in opaque markets because traders can more easily determine market value and more easily manage their trading in transparent markets.
- A well-functioning financial system allows people to trade instruments that best solve their wealth and risk management problems with low transaction costs. Complete and liquid markets characterize a well-functioning financial system. Complete markets are markets in which the instruments needed to solve investment and risk management problems are available to trade. Liquid markets are markets in which traders can trade when they want to trade at low cost.
- The financial system is operationally efficient when its markets are liquid. Liquid markets lower the costs of raising capital.

- A well-functioning financial system promotes wealth by ensuring that capital allocation decisions are well made. A well-functioning financial system also promotes wealth by allowing people to share the risks associated with valuable products that would otherwise not be undertaken.
- Prices are informationally efficient when they reflect all available information about fundamental values. Information-motivated traders make prices informationally efficient. Prices will be most informative in liquid markets because information-motivated traders will not invest in information and research if establishing positions based on their analyses is too costly.
- Regulators generally seek to promote fair and orderly markets in which traders can trade at prices that accurately reflect fundamental values without incurring excessive transaction costs. Governmental agencies and self-regulating organizations of practitioners provide regulatory services that attempt to make markets safer and more efficient.
- Mandated financial disclosure programs for the issuers of publicly traded securities ensure that information necessary to estimate security values is available to financial analysts on a consistent basis.

PRACTICE PROBLEMS

1 Akihiko Takabe has designed a sophisticated forecasting model, which predicts the movements in the overall stock market, in the hope of earning a return in excess of a fair return for the risk involved. He uses the predictions of the model to decide whether to buy, hold, or sell the shares of an index fund that aims to replicate the movements of the stock market. Takabe would *best* be characterized as a(n):

A hedger.

B investor.

C information-motivated trader.

2 James Beach is young and has substantial wealth. A significant proportion of his stock portfolio consists of emerging market stocks that offer relatively high expected returns at the cost of relatively high risk. Beach believes that investment in emerging market stocks is appropriate for him given his ability and willingness to take risk. Which of the following labels *most appropriately* describes Beach?

A Hedger.

B Investor.

C Information-motivated trader.

3 Lisa Smith owns a manufacturing company in the United States. Her company has sold goods to a customer in Brazil and will be paid in Brazilian real (BRL) in three months. Smith is concerned about the possibility of the BRL depreciating more than expected against the US dollar (USD). Therefore, she is planning to sell three-month futures contracts on the BRL. The seller of such contracts generally gains when the BRL depreciates against the USD. If Smith were to sell these future contracts, she would *most appropriately* be described as a(n):

A hedger.

B investor.

C information-motivated trader.

4 Which of the following is *not* a function of the financial system?

A To regulate arbitrageurs' profits (excess returns).

B To help the economy achieve allocational efficiency.

C To facilitate borrowing by businesses to fund current operations.

5 An investor primarily invests in stocks of publicly traded companies. The investor wants to increase the diversification of his portfolio. A friend has recommended investing in real estate properties. The purchase of real estate would *best* be characterized as a transaction in the:

A derivative investment market.

B traditional investment market.

C alternative investment market.

6 A hedge fund holds its excess cash in 90-day commercial paper and negotiable certificates of deposit. The cash management policy of the hedge fund is *best described* as using:

A capital market instruments.

© 2019 CFA Institute. All rights reserved.

B money market instruments.

C intermediate-term debt instruments.

7 An oil and gas exploration and production company announces that it is offering 30 million shares to the public at $45.50 each. This transaction is *most likely* a sale in the:

A futures market.

B primary market.

C secondary market.

8 Consider a mutual fund that invests primarily in fixed-income securities that have been determined to be appropriate given the fund's investment goal. Which of the following is *least likely* to be a part of this fund?

A Warrants.

B Commercial paper.

C Repurchase agreements.

9 A friend has asked you to explain the differences between open-end and closed-end funds. Which of the following will you *most likely* include in your explanation?

A Closed-end funds are unavailable to new investors.

B When investors sell the shares of an open-end fund, they can receive a discount or a premium to the fund's net asset value.

C When selling shares, investors in an open-end fund sell the shares back to the fund whereas investors in a closed-end fund sell the shares to others in the secondary market.

10 The usefulness of a forward contract is limited by some problems. Which of the following is *most likely* one of those problems?

A Once you have entered into a forward contract, it is difficult to exit from the contract.

B Entering into a forward contract requires the long party to deposit an initial amount with the short party.

C If the price of the underlying asset moves adversely from the perspective of the long party, periodic payments must be made to the short party.

11 Tony Harris is planning to start trading in commodities. He has heard about the use of futures contracts on commodities and is learning more about them. Which of the following is Harris *least likely* to find associated with a futures contract?

A Existence of counterparty risk.

B Standardized contractual terms.

C Payment of an initial margin to enter into a contract.

12 A German company that exports machinery is expecting to receive $10 million in three months. The firm converts all its foreign currency receipts into euros. The chief financial officer of the company wishes to lock in a minimum fixed rate for converting the $10 million to euro but also wants to keep the flexibility to use the future spot rate if it is favorable. What hedging transaction is *most likely* to achieve this objective?

A Selling dollars forward.

B Buying put options on the dollar.

C Selling futures contracts on dollars.

13 A book publisher requires substantial quantities of paper. The publisher and a paper producer have entered into an agreement for the publisher to buy and the producer to supply a given quantity of paper four months later at a price agreed upon today. This agreement is a:

- **A** futures contract.
- **B** forward contract.
- **C** commodity swap.

14 The Standard & Poor's Depositary Receipts (SPDRs) is an investment that tracks the S&P 500 stock market index. Purchases and sales of SPDRs during an average trading day are *best* described as:

- **A** primary market transactions in a pooled investment.
- **B** secondary market transactions in a pooled investment.
- **C** secondary market transactions in an actively managed investment.

15 The Standard & Poor's Depositary Receipts (SPDRs) is an exchange-traded fund in the United States that is designed to track the S&P 500 stock market index. The latest price of a share of SPDRs is $290. A trader has just bought call options on shares of SPDRs for a premium of $3 per share. The call options expire in six months and have an exercise price of $305 per share. On the expiration date, the trader will exercise the call options (ignore any transaction costs) if and only if the shares of SPDRs are trading:

- **A** below $305 per share.
- **B** above $305 per share.
- **C** above $308 per share.

16 Which of the following statements about exchange-traded funds is *most correct*?

- **A** Exchange-traded funds are not backed by any assets.
- **B** The investment companies that create exchange-traded funds are financial intermediaries.
- **C** The transaction costs of trading shares of exchange-traded funds are substantially greater than the combined costs of trading the underlying assets of the fund.

17 Jason Schmidt works for a hedge fund and he specializes in finding profit opportunities that are the result of inefficiencies in the market for convertible bonds—bonds that can be converted into a predetermined amount of a company's common stock. Schmidt tries to find convertibles that are priced inefficiently relative to the underlying stock. The trading strategy involves the simultaneous purchase of the convertible bond and the short sale of the underlying common stock. The above process could best be described as:

- **A** hedging.
- **B** arbitrage.
- **C** securitization.

18 Pierre-Louis Robert just purchased a call option on shares of the Michelin Group. A few days ago he wrote a put option on Michelin shares. The call and put options have the same exercise price, expiration date, and number of shares underlying. Considering both positions, Robert's exposure to the risk of the stock of the Michelin Group is:

- **A** long.
- **B** short.
- **C** neutral.

19 An online brokerage firm has set the minimum margin requirement at 55 percent. What is the maximum leverage ratio associated with a position financed by this minimum margin requirement?

A 1.55.

B 1.82.

C 2.22.

20 A trader has purchased 200 shares of a non-dividend-paying firm on margin at a price of $50 per share. The leverage ratio is 2.5. Six months later, the trader sells these shares at $60 per share. Ignoring the interest paid on the borrowed amount and the transaction costs, what was the return to the trader during the six-month period?

A 20 percent.

B 33.33 percent.

C 50 percent.

21 Jason Williams purchased 500 shares of a company at $32 per share. The stock was bought on 75 percent margin. One month later, Williams had to pay interest on the amount borrowed at a rate of 2 percent per month. At that time, Williams received a dividend of $0.50 per share. Immediately after that he sold the shares at $28 per share. He paid commissions of $10 on the purchase and $10 on the sale of the stock. What was the rate of return on this investment for the one-month period?

A –12.5 percent.

B –15.4 percent.

C –50.1 percent.

22 Caroline Rogers believes the price of Gamma Corp. stock will go down in the near future. She has decided to sell short 200 shares of Gamma Corp. at the current market price of €47. The initial margin requirement is 40 percent. Which of the following is an appropriate statement regarding the margin requirement that Rogers is subject to on this short sale?

A She will need to contribute €3,760 as margin.

B She will need to contribute €5,640 as margin.

C She will only need to leave the proceeds from the short sale as deposit and does not need to contribute any additional funds.

23 The current price of a stock is $25 per share. You have $10,000 to invest. You borrow an additional $10,000 from your broker and invest $20,000 in the stock. If the maintenance margin is 30 percent, at what price will a margin call first occur?

A $9.62.

B $17.86.

C $19.71.

24 You have placed a sell market-on-open order—a market order that would automatically be submitted at the market's open tomorrow and would fill at the market price. Your instruction, to sell the shares at the market open, is a(n):

A execution instruction.

B validity instruction.

C clearing instruction.

25 A market has the following limit orders standing on its book for a particular stock. The bid and ask sizes are number of shares in hundreds.

Bid Size	Limit Price (€)	Offer Size
5	9.73	
12	9.81	
4	9.84	
6	9.95	
	10.02	5
	10.10	12
	10.14	8

What is the market?

A 9.73 bid, offered at 10.14.

B 9.81 bid, offered at 10.10.

C 9.95 bid, offered at 10.02.

26 Consider the following limit order book for a stock. The bid and ask sizes are number of shares in hundreds.

Bid Size	Limit Price (¥)	Offer Size
3	122.80	
8	123.00	
4	123.35	
	123.80	7
	124.10	6
	124.50	7

A new buy limit order is placed for 300 shares at ¥123.40. This limit order is said to:

A take the market.

B make the market.

C make a new market.

27 Currently, the market in a stock is "$54.62 bid, offered at $54.71." A new sell limit order is placed at $54.62. This limit order is said to:

A take the market.

B make the market.

C make a new market.

28 Jim White has sold short 100 shares of Super Stores at a price of $42 per share. He has also simultaneously placed a "good-till-cancelled, stop 50, limit 55 buy" order. Assume that if the stop condition specified by White is satisfied and the order becomes valid, it will get executed. Excluding transaction costs, what is the maximum possible loss that White can have?

A $800.

B $1,300.

C Unlimited.

29 You own shares of a company that are currently trading at $30 a share. Your technical analysis of the shares indicates a support level of $27.50. That is, if the price of the shares is going down, it is more likely to stay above this level rather than fall below it. If the price does fall below this level, however, you believe that the price may continue to decline. You have no immediate intent to sell the shares but are concerned about the possibility of a huge loss if the share price declines below the support level. Which of the following types of orders could you place to most appropriately address your concern?

A Short sell order.

B Good-till-cancelled stop sell order.

C Good-till-cancelled stop buy order.

30 In an underwritten offering, the risk that the entire issue may not be sold to the public at the stipulated offering price is borne by the:

A issuer.

B investment bank.

C buyers of the part of the issue that is sold.

31 A British company listed on AIM (formerly the Alternative Investment Market) of the London Stock Exchange announced the sale of 6,686,665 shares to a small group of qualified investors at £0.025 per share. Which of the following *best describes* this sale?

A Shelf registration.

B Private placement.

C Initial public offering.

32 A German publicly traded company, to raise new capital, gave its existing shareholders the opportunity to subscribe for new shares. The existing shareholders could purchase two new shares at a subscription price of €4.58 per share for every 15 shares held. This is an example of a(n):

A rights offering.

B private placement.

C initial public offering.

33 Consider an order-driven system that allows hidden orders. The following four sell orders on a particular stock are currently in the system's limit order book. Based on the commonly used order precedence hierarchy, which of these orders will have precedence over others?

Order	Time of Arrival (HH:MM:SS)	Limit Price (€)	Special Instruction (If any)
I	9:52:01	20.33	
II	9:52:08	20.29	Hidden order
III	9:53:04	20.29	
IV	9:53:49	20.29	

A Order I (time of arrival of 9:52:01).

B Order II (time of arrival of 9:52:08).

C Order III (time of arrival of 9:53:04).

34 Zhenhu Li has submitted an immediate-or-cancel buy order for 500 shares of a company at a limit price of CNY 74.25. There are two sell limit orders standing in that stock's order book at that time. One is for 300 shares at a limit price of CNY 74.30 and the other is for 400 shares at a limit price of CNY 74.35. How many shares in Li's order would get cancelled?

A None (the order would remain open but unfilled).

B 200 (300 shares would get filled).

C 500 (there would be no fill).

35 A market has the following limit orders standing on its book for a particular stock:

Buyer	Bid Size (Number of Shares)	Limit Price (£)	Offer Size (Number of Shares)	Seller
Keith	1,000	19.70		
Paul	200	19.84		
Ann	400	19.89		
Mary	300	20.02		
		20.03	800	Jack
		20.11	1,100	Margaret
		20.16	400	Jeff

Ian submits a day order to sell 1,000 shares, limit £19.83. Assuming that no more buy orders are submitted on that day after Ian submits his order, what would be Ian's average trade price?

A £19.70.

B £19.92.

C £20.05.

36 A financial analyst is examining whether a country's financial market is well functioning. She finds that the transaction costs in this market are low and trading volumes are high. She concludes that the market is quite liquid. In such a market:

A traders will find it hard to make use of their information.

B traders will find it easy to trade and their trading will make the market less informationally efficient.

C traders will find it easy to trade and their trading will make the market more informationally efficient.

37 The government of a country whose financial markets are in an early stage of development has hired you as a consultant on financial market regulation. Your first task is to prepare a list of the objectives of market regulation. Which of the following is *least likely* to be included in this list of objectives?

A Minimize agency problems in the financial markets.

B Ensure that financial markets are fair and orderly.

C Ensure that investors in the stock market achieve a rate of return that is at least equal to the risk-free rate of return.

SOLUTIONS

1 C is correct. Takabe is best characterized as an information-motivated trader. Takabe believes that his model provides him superior information about the movements in the stock market and his motive for trading is to profit from this information.

2 B is correct. Beach is an investor. He is simply investing in risky assets consistent with his level of risk aversion. Beach is not hedging any existing risk or using information to identify and trade mispriced securities. Therefore, he is not a hedger or an information-motivated trader.

3 A is correct. Smith is a hedger. The short position on the BRL futures contract offsets the BRL long position in three months. She is hedging the risk of the BRL depreciating against the USD. If the BRL depreciates, the value of the cash inflow goes down in USD terms but there is a gain on the futures contracts.

4 A is correct. Regulation of arbitrageurs' profits is not a function of the financial system. The financial system facilitates the allocation of capital to the best uses and the purposes for which people use the financial system, including borrowing money.

5 C is correct. The purchase of real estate properties is a transaction in the alternative investment market.

6 B is correct. The 90-day commercial paper and negotiable certificates of deposit are money market instruments.

7 B is correct. This transaction is a sale in the primary market. It is a sale of shares from the issuer to the investor and funds flow to the issuer of the security from the purchaser.

8 A is correct. Warrants are least likely to be part of the fund. Warrant holders have the right to buy the issuer's common stock. Thus, warrants are typically classified as equity and are least likely to be a part of a fixed-income mutual fund. Commercial paper and repurchase agreements are short-term fixed-income securities.

9 C is correct. When investors want to sell their shares, investors of an open-end fund sell the shares back to the fund whereas investors of a closed-end fund sell the shares to others in the secondary market. Closed-end funds are available to new investors but they must purchase shares in the fund in the secondary market. The shares of a closed-end fund trade at a premium or discount to net asset value.

10 A is correct. Once you have entered into a forward contract, it is difficult to exit from the contract. As opposed to a futures contract, trading out of a forward contract is quite difficult. There is no exchange of cash at the origination of a forward contract. There is no exchange on a forward contract until the maturity of the contract.

11 A is correct. Harris is least likely to find counterparty risk associated with a futures contract. There is limited counterparty risk in a futures contract because the clearinghouse is on the other side of every contract.

12 B is correct. Buying a put option on the dollar will ensure a minimum exchange rate but does not have to be exercised if the exchange rate moves in a favorable direction. Forward and futures contracts would lock in a fixed rate but would not allow for the possibility to profit in case the value of the dollar three months later in the spot market turns out to be greater than the value in the forward or futures contract.

13 B is correct. The agreement between the publisher and the paper supplier to respectively buy and supply paper in the future at a price agreed upon today is a forward contract.

14 B is correct. SPDRs trade in the secondary market and are a pooled investment vehicle.

15 B is correct. The holder of the call option will exercise the call options if the price is above the exercise price of \$305 per share. Note that if the stock price is above \$305 but less than \$308, the option would be exercised even though the net result for the option buyer after considering the premium is a loss. For example, if the stock price is \$307, the option buyer would exercise the option to make \$2 = \$307 – \$305 per share, resulting in a loss of \$1 = \$3 – \$2 after considering the premium. It is better to exercise and have a loss of only \$1, however, rather than not exercise and lose the entire \$3 premium.

16 B is correct. The investment companies that create exchange-traded funds (ETFs) are financial intermediaries. ETFs are securities that represent ownership in the assets held by the fund. The transaction costs of trading shares of ETFs are substantially lower than the combined costs of trading the underlying assets of the ETF.

17 B is correct. The process can best be described as arbitrage because it involves buying and selling instruments, whose values are closely related, at different prices in different markets.

18 A is correct. Robert's exposure to the risk of the stock of the Michelin Group is long. The exposure as a result of the long call position is long. The exposure as a result of the short put position is also long. Therefore, the combined exposure is long.

19 B is correct. The maximum leverage ratio is 1.82 = 100% position ÷ 55% equity. The maximum leverage ratio associated with a position financed by the minimum margin requirement is one divided by the minimum margin requirement.

20 C is correct. The return is 50 percent. If the position had been unleveraged, the return would be 20% = (60 – 50)/50. Because of leverage, the return is 50% = 2.5 × 20%.

Another way to look at this problem is that the equity contributed by the trader (the minimum margin requirement) is 40% = 100% ÷ 2.5. The trader contributed \$20 = 40% of \$50 per share. The gain is \$10 per share, resulting in a return of 50% = 10/20.

21 B is correct. The return is –15.4 percent.

Total cost of the purchase = \$16,000 = 500 × \$32

Equity invested = \$12,000 = 0.75 × \$16,000

Amount borrowed = \$4,000 = 16,000 – 12,000

Interest paid at month end = \$80 = 0.02 × \$4,000

Dividend received at month end = \$250 = 500 × \$0.50

Proceeds on stock sale = \$14,000 = 500 × \$28

Total commissions paid = \$20 = \$10 + \$10

Net gain/loss = –\$1,850 = –16,000 – 80 + 250 + 14,000 – 20

Initial investment including commission on purchase = \$12,010

Return = –15.4% = –\$1,850/\$12,010

22 A is correct. She will need to contribute €3,760 as margin. In view of the possibility of a loss, if the stock price goes up, she will need to contribute €3,760 = 40% of €9,400 as the initial margin. Rogers will need to leave the proceeds from the short sale (€9,400 = 200 × €47) on deposit.

23 B is correct. A margin call will first occur at a price of $17.86. Because you have contributed half and borrowed the remaining half, your initial equity is 50 percent of the initial stock price, or $12.50 = 0.50 × $25. If P is the subsequent price, your equity would change by an amount equal to the change in price. So, your equity at price P would be 12.50 + (P – 25). A margin call will occur when the percentage margin drops to 30 percent. So, the price at which a margin call will occur is the solution to the following equation.

$$\frac{\text{Equity/Share}}{\text{Price/Share}} = \frac{12.50 + P - 25}{P} = 30\%$$

The solution is P = $17.86.

24 B is correct. An instruction regarding when to fill an order is considered a validity instruction.

25 C is correct. The market is 9.95 bid, offered at 10.02. The best bid is at €9.95 and the best offer is €10.02.

26 C is correct. This order is said to make a new market. The new buy order is at ¥123.40, which is better than the current best bid of ¥123.35. Therefore, the buy order is making a new market. Had the new order been at ¥123.35, it would be said to make the market. Because the new buy limit order is at a price less than the best offer of ¥123.80, it will not immediately execute and is not taking the market.

27 A is correct. This order is said to take the market. The new sell order is at $54.62, which is at the current best bid. Therefore, the new sell order will immediately trade with the current best bid and is taking the market.

28 B is correct. The maximum possible loss is $1,300. If the stock price crosses $50, the stop buy order will become valid and will get executed at a maximum limit price of $55. The maximum loss per share is $13 = $55 – $42, or $1,300 for 100 shares.

29 B is correct. The most appropriate order is a good-till-cancelled stop sell order. This order will be acted on if the stock price declines below a specified price (in this case, $27.50). This order is sometimes referred to as a good-till-cancelled stop loss sell order. You are generally bullish about the stock, as indicated by no immediate intent to sell, and would expect a loss on short selling the stock. A stop buy order is placed to buy a stock when the stock is going up.

30 B is correct. The investment bank bears the risk that the issue may be undersubscribed at the offering price. If the entire issue is not sold, the investment bank underwriting the issue will buy the unsold securities at the offering price.

31 B is correct. This sale is a private placement. As the company is already publicly traded, the share sale is clearly not an initial public offering. The sale also does not involve a shelf registration because the company is not selling shares to the public on a piecemeal basis.

32 A is correct. This offering is a rights offering. The company is distributing rights to buy stock at a fixed price to existing shareholders in proportion to their holdings.

33 C is correct. Order III (time of arrival of 9:53:04) has precedence. In the order precedence hierarchy, the first rule is price priority. Based on this rule, sell orders II, III, and IV get precedence over order I. The next rule is display precedence at a given price. Because order II is a hidden order, orders III and IV get precedence. Finally, order III gets precedence over order IV based on time priority at same price and same display status.

34 C is correct. The order for 500 shares would get cancelled; there would be no fill. Li is willing to buy at CNY 74.25 or less but the minimum offer price in the book is CNY 74.30; therefore, no part of the order would be filled. Because Li's order is immediate-or-cancel, it would be cancelled.

35 B is correct. Ian's average trade price is:

$$\text{£}19.92 = \frac{300 \times \text{£}20.02 + 400 \times \text{£}19.89 \ + \ 200 \times \text{£}19.84}{300 + 400 + 200}$$

Ian's sell order first fills with the most aggressively priced buy order, which is Mary's order for 300 shares at £20.02. Ian still has 700 shares for sale. The next most aggressively priced buy order is Ann's order for 400 shares at £19.89. This order is filled. Ian still has 300 shares for sale. The next most aggressively priced buy order is Paul's order for 200 shares at £19.84. A third trade takes place. Ian still has 100 shares for sale.

The next buy order is Keith's order for 1,000 shares at £19.70. However, this price is below Ian's limit price of £19.83. Therefore, no more trade is possible.

36 C is correct. In such a market, well-informed traders will find it easy to trade and their trading will make the market more informationally efficient. In a liquid market, it is easier for informed traders to fill their orders. Their trading will cause prices to incorporate their information and the prices will be more in line with the fundamental values.

37 C is correct. Ensure that investors in the stock market achieve a rate of return that is at least equal to the risk-free rate of return is least likely to be included as an objective of market regulation. Stocks are risky investments and there would be occasions when a stock market investment would not only have a return less than the risk-free rate but also a negative return. Minimizing agency costs and ensuring that financial markets are fair and orderly are objectives of market regulation.

READING

34

Security Market Indexes

by Paul D. Kaplan, PhD, CFA, and Dorothy C. Kelly, CFA

Paul D. Kaplan, PhD, CFA, is at Morningstar (Canada). Dorothy C. Kelly, CFA, is at McIntire School of Commerce, University of Virginia (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe a security market index;
☐	**b.** calculate and interpret the value, price return, and total return of an index;
☐	**c.** describe the choices and issues in index construction and management;
☐	**d.** compare the different weighting methods used in index construction;
☐	**e.** calculate and analyze the value and return of an index given its weighting method;
☐	**f.** describe rebalancing and reconstitution of an index;
☐	**g.** describe uses of security market indexes;
☐	**h.** describe types of equity indexes;
☐	**i.** describe types of fixed-income indexes;
☐	**j.** describe indexes representing alternative investments;
☐	**k.** compare types of security market indexes.

INTRODUCTION

1

Investors gather and analyze vast amounts of information about security markets on a continual basis. Because this work can be both time consuming and data intensive, investors often use a single measure that consolidates this information and reflects the performance of an entire security market.

© 2019 CFA Institute. All rights reserved.

Security market indexes were first introduced as a simple measure to reflect the performance of the US stock market. Since then, security market indexes have evolved into important multi-purpose tools that help investors track the performance of various security markets, estimate risk, and evaluate the performance of investment managers. They also form the basis for new investment products.

in·dex, *noun* (*pl.* **in·dex·es** *or* **in·di·ces**) Latin *indic-, index,* from *indicare* to indicate: an indicator, sign, or measure of something.

ORIGIN OF MARKET INDEXES

Investors had access to regularly published data on individual security prices in London as early as 1698, but nearly 200 years passed before they had access to a simple indicator to reflect security market information. To give readers a sense of how the US stock market in general performed on a given day, publishers Charles H. Dow and Edward D. Jones introduced the Dow Jones Average, the world's first security market index, in 1884. The index, which appeared in *The Customers' Afternoon Letter*, consisted of the stocks of nine railroads and two industrial companies. It eventually became the Dow Jones Transportation Average. Convinced that industrial companies, rather than railroads, would be "the great speculative market" of the future, Dow and Jones introduced a second index in May 1896—the Dow Jones Industrial Average (DJIA). It had an initial value of 40.94 and consisted of 12 stocks from major US industries. Today, investors can choose from among thousands of indexes to measure and monitor different security markets and asset classes.

This reading is organized as follows. Section 2 defines a security market index and explains how to calculate the price return and total return of an index for a single period and over multiple periods. Sections 3–4 describe how indexes are constructed and managed. Section 5 discusses the use of market indexes. Sections 6, 7, and 8 discuss various types of indexes, and the final section summarizes the reading. Practice problems follow the conclusions and summary.

2

INDEX DEFINITION AND CALCULATIONS OF VALUE AND RETURNS

a describe a security market index;

b calculate and interpret the value, price return, and total return of an index;

A **security market index** represents a given security market, market segment, or asset class. Most indexes are constructed as portfolios of marketable securities.

The value of an index is calculated on a regular basis using either the actual or estimated market prices of the individual securities, known as **constituent securities**, within the index. For each security market index, investors may encounter two versions of the same index (i.e., an index with identical constituent securities and weights): one version based on price return and one version based on total return. As the name suggests, a **price return index**, also known as a **price index**, reflects *only* the prices of the constituent securities within the index. A **total return index**, in contrast, reflects not only the prices of the constituent securities but also the reinvestment of all income received since inception.

At inception, the values of the price and total return versions of an index are equal. As time passes, however, the value of the total return index, which includes the reinvestment of all dividends and/or interest received, will exceed the value of the price return index by an increasing amount. A look at how the values of each version are calculated over multiple periods illustrates why.

The value of a price return index is calculated as:

$$V_{PRI} = \frac{\sum_{i=1}^{N} n_i P_i}{D} \tag{1}$$

where

V_{PRI} = the value of the price return index

n_i = the number of units of constituent security i held in the index portfolio

N = the number of constituent securities in the index

P_i = the unit price of constituent security i

D = the value of the divisor

The **divisor** is a number initially chosen at inception. It is frequently chosen so that the price index has a convenient initial value, such as 1,000. The index provider then adjusts the value of the divisor as necessary to avoid changes in the index value that are unrelated to changes in the prices of its constituent securities. For example, when changing index constituents, the index provider may adjust the divisor so that the value of the index with the new constituents equals the value of the index prior to the changes.

Index return calculations, like calculations of investment portfolio returns, may measure price return or total return. **Price return** measures only price appreciation or percentage change in price. **Total return** measures price appreciation plus interest, dividends, and other distributions.

2.1 Calculation of Single-Period Returns

For a security market index, price return can be calculated in two ways: either the percentage change in value of the price return index, or the weighted average of price returns of the constituent securities. The price return of an index can be expressed as:

$$\text{PR}_I = \frac{V_{PRI1} - V_{PRI0}}{V_{PRI0}} \tag{2}$$

where

PR_I = the price return of the index portfolio (as a decimal number, i.e., 12 percent is 0.12)

V_{PRI1} = the value of the price return index at the end of the period

V_{PRI0} = the value of the price return index at the beginning of the period

Similarly, the price return of each constituent security can be expressed as:

$$\text{PR}_i = \frac{P_{i1} - P_{i0}}{P_{i0}} \tag{3}$$

where

PR_i = the price return of constituent security i (as a decimal number)
P_{i1} = the price of constituent security i at the end of the period
P_{i0} = the price of constituent security i at the beginning of the period

Because the price return of the index equals the weighted average of price returns of the individual securities, we can write:

$$\text{PR}_I = \sum_{i=1}^{N} \text{w}_i \text{PR}_i = \sum_{i=1}^{N} \text{w}_i \left(\frac{P_{i1} - P_{i0}}{P_{i0}} \right) \tag{4}$$

where:

PR_I = the price return of index portfolio (as a decimal number)
PR_i = the price return of constituent security i (as a decimal number)
N = the number of individual securities in the index
w_i = the weight of security i (the fraction of the index portfolio allocated to security i)
P_{i1} = the price of constituent security i at the end of the period
P_{i0} = the price of constituent security i at the beginning of the period

Equation 4 can be rewritten simply as:

$$\text{PR}_I = \text{w}_1\text{PR}_1 + \text{w}_2\text{PR}_2 + \ldots + \text{w}_N\text{PR}_N \tag{5}$$

where

PR_I = the price return of index portfolio (as a decimal number)
PR_i = the price return of constituent security i (as a decimal number)
w_i = the weight of security i (the fraction of the index portfolio allocated to security i)
N = the number of securities in the index

Total return measures price appreciation plus interest, dividends, and other distributions. Thus, the **total return** of an index is the price appreciation, or change in the value of the price return index, plus income (dividends and/or interest) over the period, expressed as a percentage of the beginning value of the price return index. The total return of an index can be expressed as:

$$\text{TR}_I = \frac{V_{PRI1} - V_{PRI0} + Inc_I}{V_{PRI0}} \tag{6}$$

where

TR_I = the total return of the index portfolio (as a decimal number)
V_{PRI1} = the value of the price return index at the end of the period
V_{PRI0} = the value of the price return index at the beginning of the period
Inc_I = the total income (dividends and/or interest) from all securities in the index held over the period

The total return of an index can also be calculated as the weighted average of total returns of the constituent securities. The total return of each constituent security in the index is calculated as:

$$\text{TR}_i = \frac{P_{1i} - P_{0i} + Inc_i}{P_{0i}} \tag{7}$$

where

TR_i = the total return of constituent security i (as a decimal number)
P_{1i} = the price of constituent security i at the end of the period
P_{0i} = the price of constituent security i at the beginning of the period
Inc_i = the total income (dividends and/or interest) from security i over the period

Because the total return of an index can be calculated as the weighted average of total returns of the constituent securities, we can express total return as:

$$\mathrm{TR}_I = \sum_{i=1}^{N} \mathrm{w}_i \mathrm{TR}_i = \sum_{i=1}^{N} \mathrm{w}_i \left(\frac{P_{1i} - P_{0i} + Inc_i}{P_{0i}} \right) \quad (8)$$

Equation 8 can be rewritten simply as

$$\mathrm{TR}_I = \mathrm{w}_1\mathrm{TR}_1 + \mathrm{w}_2\mathrm{TR}_2 + \ldots + \mathrm{w}_N\mathrm{TR}_N \quad (9)$$

where

TR_I = the total return of the index portfolio (as a decimal number)
TR_i = the total return of constituent security i (as a decimal number)
w_i = the weight of security i (the fraction of the index portfolio allocated to security i)
N = the number of securities in the index

2.2 Calculation of Index Values over Multiple Time Periods

The calculation of index values over multiple time periods requires geometrically linking the series of index returns. With a series of price returns for an index, we can calculate the value of the price return index with the following equation:

$$\mathrm{V}_{PRIT} = \mathrm{V}_{PRI0}(1 + \mathrm{PR}_{I1})(1 + \mathrm{PR}_{I2})\ldots(1 + \mathrm{PR}_{IT}) \quad (10)$$

where

V_{PRI0} = the value of the price return index at inception
V_{PRIT} = the value of the price return index at time t
PR_{IT} = the price return (as a decimal number) on the index over period t, t = 1, 2, …, T

For an index with an inception value set to 1,000 and price returns of 5 percent and 3 percent for Periods 1 and 2 respectively, the values of the price return index would be calculated as follows:

Period	Return (%)	Calculation	Ending Value
0		1,000(1.00)	1,000.00
1	5.00	1,000(1.05)	1,050.00
2	3.00	1,000(1.05)(1.03)	1,081.50

Similarly, the series of total returns for an index is used to calculate the value of the total return index with the following equation:

$$\mathrm{V}_{TRIT} = \mathrm{V}_{TRI0}(1 + \mathrm{TR}_{I1})(1 + \mathrm{TR}_{I2})\ldots(1 + \mathrm{TR}_{IT}) \quad (11)$$

where

V_{TRI0} = the value of the index at inception
V_{TRIT} = the value of the total return index at time t
TR_{IT} = the total return (as a decimal number) on the index over period t, t = 1, 2, …, T

Suppose that the same index yields an additional 1.5 percent return from income in Period 1 and an additional 2.0 percent return from income in Period 2, bringing the total returns for Periods 1 and 2, respectively, to 6.5 percent and 5 percent. The values of the total return index would be calculated as follows:

Period	Return (%)	Calculation	Ending Value
0		1,000(1.00)	1,000.00
1	6.50	1,000(1.065)	1,065.00
2	5.00	1,000(1.065)(1.05)	1,118.25

As illustrated above, as time passes, the value of the total return index, which includes the reinvestment of all dividends and/or interest received, exceeds the value of the price return index by an increasing amount.

3

INDEX CONSTRUCTION

c describe the choices and issues in index construction and management;

d compare the different weighting methods used in index construction;

e calculate and analyze the value and return of an index given its weighting method;

Constructing and managing a security market index is similar to constructing and managing a portfolio of securities. Index providers must decide the following:

1. Which target market should the index represent?
2. Which securities should be selected from that target market?
3. How much weight should be allocated to each security in the index?
4. When should the index be rebalanced?
5. When should the security selection and weighting decision be re-examined?

3.1 Target Market and Security Selection

The first decision in index construction is identifying the target market, market segment, or asset class that the index is intended to represent. The target market may be defined very broadly or narrowly. It may be based on asset class (e.g., equities, fixed income, real estate, commodities, hedge funds); geographic region (e.g., Japan, South Africa, Latin America, Europe); the exchange on which the securities are traded (e.g., Shanghai, Toronto, Tokyo), and/or other characteristics (e.g., economic sector, company size, investment style, duration, or credit quality).

The target market determines the investment universe and the securities available for inclusion in the index. Once the investment universe is identified, the number of securities and the specific securities to include in the index must be determined. The constituent securities could be nearly all those in the target market or a representative sample of the target market. Some equity indexes, such as the S&P 500 Index and the FTSE 100, fix the number of securities included in the index and indicate this number in the name of the index. Other indexes allow the number of securities to vary to reflect changes in the target market or to maintain a certain percentage of the target market. For example, the Tokyo Stock Price Index (TOPIX) represents and includes all of the largest stocks, known as the First Section, listed on the Tokyo Stock Exchange. To be included in the First Section—and thus the TOPIX—stocks must meet certain criteria, such as the number of shares outstanding, the number of shareholders, and market capitalization. Stocks that no longer meet the criteria are removed from the First Section and also the TOPIX. Objective or mechanical rules determine the constituent securities of most, but not all, indexes. The S&P Bombay Stock Exchange Sensitive Index, also called the S&P BSE SENSEX and the S&P 500, for example, use a selection committee and more subjective decision-making rules to determine constituent securities.

3.2 Index Weighting

The weighting decision determines how much of each security to include in the index and has a substantial impact on an index's value. Index providers use a number of methods to weight the constituent securities in an index. Indexes can be price weighted, equal weighted, market-capitalization weighted, or fundamentally weighted. Each weighting method has its advantages and disadvantages.

3.2.1 *Price Weighting*

The simplest method to weight an index and the one used by Charles Dow to construct the Dow Jones Industrial Average is **price weighting**. In price weighting, the weight on each constituent security is determined by dividing its price by the sum of all the prices of the constituent securities. The weight is calculated using the following formula:

$$w_i^P = \frac{P_i}{\sum_{i=1}^{N} P_i} \tag{12}$$

Exhibit 1 illustrates the values, weights, and single-period returns following inception of a price-weighted equity index with five constituent securities. The value of the price-weighted index is determined by dividing the sum of the security values (101.50) by the divisor, which is typically set at inception to equal the initial number of securities in the index. Thus, in our example, the divisor is 5 and the initial value of the index is calculated as 101.50 ÷ 5 = 20.30.

As illustrated in this exhibit, Security A, which has the highest price, also has the highest weighting and thus will have the greatest impact on the return of the index. Note how both the price return and the total return of the index are calculated on the basis of the corresponding returns on the constituent securities.

A property unique to price-weighted indexes is that a stock split on one constituent security changes the weights on all the securities in the index.[1] To prevent the stock split and the resulting new weights from changing the value of the index, the index provider must adjust the value of the divisor as illustrated in Exhibit 2. Given a 2-for-1 split in Security A, the divisor is adjusted by dividing the sum of the constituent prices *after* the split (77.50) by the value of the index *before* the split (21.00). This adjustment results in changing the divisor from 5 to 3.69 so that the index value is maintained at 21.00.

---

[1] A stock split is an increase in the number of shares outstanding and a proportionate decrease in the price per share such that the total market value of equity, as well as investors' proportionate ownership in the company, does not change.

Exhibit 1 Example of a Price-Weighted Equity Index

Security	Shares in Index	BOP Price	Value (Shares × BOP Price)	BOP Weight (%)	EOP Price	Dividends Per Share	Value (Shares × EOP Price)	Total Dividends	Price Return (%)	Total Return (%)	BOP Weight × Price Return (%)	BOP Weight × Total Return (%)	EOP Weight (%)
A	1	50.00	50.00	49.26	55.00	0.75	55.00	0.75	10.00	11.50	4.93	5.66	52.38
B	1	25.00	25.00	24.63	22.00	0.10	22.00	0.10	−12.00	−11.60	−2.96	−2.86	20.95
C	1	12.50	12.50	12.32	8.00	0.00	8.00	0.00	−36.00	−36.00	−4.43	−4.43	7.62
D	1	10.00	10.00	9.85	14.00	0.05	14.00	0.05	40.00	40.50	3.94	3.99	13.33
E	1	4.00	4.00	3.94	6.00	0.00	6.00	0.00	50.00	50.00	1.97	1.97	5.72
Total			**101.50**	**100.00**			**105.00**	**0.90**			**3.45**	**4.33**	**100.00**
Index Value			**20.30**				**21.00**	**0.18**	**3.45**	**4.33**			

Divisor = 5

BOP = Beginning of period

EOP = End of period

Type of Index	BOP Value	Return (%)	EOP Value
Price Return	20.30	3.45	21.00
Total Return	20.30	4.33	21.18

The primary advantage of price weighting is its simplicity. The main disadvantage of price weighting is that it results in arbitrary weights for each security. In particular, a stock split in any one security causes arbitrary changes in the weights of all the constituents' securities.

Exhibit 2 Impact of 2-for-1 Split in Security A

Security	Price before Split	Weight before Split (%)	Price after Split	Weight after Split (%)
A	55.00	52.38	27.50	35.48
B	22.00	20.95	22.00	28.39
C	8.00	7.62	8.00	10.32
D	14.00	13.33	14.00	18.07
E	6.00	5.72	6.00	7.74
Total	105.00	100.00	77.50	100.00
Divisor	5.00		3.69	
Index Value	21.00		21.00	

3.2.2 *Equal Weighting*

Another simple index weighting method is **equal weighting**. This method assigns an equal weight to each constituent security at inception. The weights are calculated as:

$$w_i^E = \frac{1}{N} \tag{13}$$

where

w_i = fraction of the portfolio that is allocated to security *i* or weight of security *i*

N = number of securities in the index

To construct an equal-weighted index from the five securities in Exhibit 1, the index provider allocates one-fifth (20 percent) of the value of the index (at the beginning of the period) to each security. Dividing the value allocated to each security by each security's individual share price determines the number of shares of each security to include in the index. Unlike a price-weighted index, where the weights are arbitrarily determined by the market prices, the weights in an equal-weighted index are arbitrarily assigned by the index provider.

Exhibit 3 illustrates the values, weights, and single-period returns following inception of an equal-weighted index with the same constituent securities as those in Exhibit 1. This example assumes a beginning index portfolio value of 10,000 (i.e., an investment of 2,000 in each security). To set the initial value of the index to 1,000, the divisor is set to 10 (10,000 ÷ 10 = 1,000).

Exhibits 1 and 3 demonstrate how different weighting methods result in different returns. The 10.4 percent price return of the equal-weighted index shown in Exhibit 3 differs significantly from the 3.45 percent price return of the price-weighted index in Exhibit 1.

Like price weighting, the primary advantage of equal weighting is its simplicity. Equal weighting, however, has a number of disadvantages. First, securities that constitute the largest fraction of the target market value are underrepresented, and securities that constitute a small fraction of the target market value are overrepresented.

Exhibit 3 Example of an Equal-Weighted Equity Index

Security	Shares in Index	BOP Price	Value (Shares × BOP Price)	Weight (%)	EOP Price	Dividends Per Share	Value (Shares × EOP Price)	Total Dividends	Price Return (%)	Total Return (%)	Weight × Price Return (%)	Weight × Total Return (%)	EOP Weight (%)
A	40	50.00	2,000	20.00	55.00	0.75	2,200	30	10.00	11.50	2.00	2.30	19.93
B	80	25.00	2,000	20.00	22.00	0.10	1,760	8	−12.00	−11.60	−2.40	−2.32	15.94
C	160	12.50	2,000	20.00	8.00	0.00	1,280	0	−36.00	−36.00	−7.20	−7.20	11.60
D	200	10.00	2,000	20.00	14.00	0.05	2,800	10	40.00	40.50	8.00	8.10	25.36
E	500	4.00	2,000	20.00	6.00	0.00	3,000	0	50.00	50.00	10.00	10.00	27.17
Total			**10,000**	**100.00**			**11,040**	**48**			**10.40**	**10.88**	**100.00**
Index Value			**1,000**				**1,104**	**4.80**	**10.40**	**10.88**			

Divisor = 10

BOP = Beginning of period

EOP = End of period

Type of Index	BOP Value	Return (%)	EOP Value
Price Return	1,000.00	10.40	1,104.00
Total Return	1,000.00	10.88	1,108.80

Second, after the index is constructed and the prices of constituent securities change, the index is no longer equally weighted. Therefore, maintaining equal weights requires frequent adjustments (rebalancing) to the index.

3.2.3 *Market-Capitalization Weighting*

In **market-capitalization weighting**, or value weighting, the weight on each constituent security is determined by dividing its market capitalization by the total market capitalization (the sum of the market capitalization) of all the securities in the index. Market capitalization or value is calculated by multiplying the number of shares outstanding by the market price per share.

The market-capitalization weight of security i is:

$$w_i^M = \frac{Q_i P_i}{\sum_{j=1}^{N} Q_j P_j} \tag{14}$$

where

w_i = fraction of the portfolio that is allocated to security i or weight of security i

Q_i = number of shares outstanding of security i

P_i = share price of security i

N = number of securities in the index

Exhibit 4 illustrates the values, weights, and single-period returns following inception of a market-capitalization-weighted index for the same five-security market. Security A, with 3,000 shares outstanding and a price of 50 per share, has a market capitalization of 150,000 or 26.29 percent (150,000/570,500) of the entire index portfolio. The resulting index weights in the exhibit reflect the relative value of each security as measured by its market capitalization.

As shown in Exhibits 1, 3, and 4, the weighting method affects the index's returns. The price and total returns of the market-capitalization index in Exhibit 4 (1.49 percent and 2.13 percent, respectively) differ significantly from those of the price-weighted (3.45 percent and 4.33 percent, respectively) and equal-weighted (10.40 percent and 10.88 percent respectively) indexes. To understand the source and magnitude of the difference, compare the weights and returns of each security under each of the weighting methods. The weight of Security A, for example, ranges from 49.26 percent in the price-weighted index to 20 percent in the equal-weighted index. With a price return of 10 percent, Security A contributes 4.93 percent to the price return of the price-weighted index, 2.00 percent to the price return of the equal-weighted index, and 2.63 percent to the price return of the market-capitalization-weighted index. With a total return of 11.50 percent, Security A contributes 5.66 percent to the total return of the price-weighted index, 2.30 percent to the total return of the equal-weighted index, and 3.02 percent to the total return of the market-capitalization-weighted index.

3.2.3.1 Float-Adjusted Market-Capitalization Weighting

In **float-adjusted market-capitalization weighting**, the weight on each constituent security is determined by adjusting its market capitalization for its **market float**. Typically, market float is the number of shares of the constituent security that are available to the investing public. For companies that are closely held, only a portion of the shares outstanding are available to the investing public (the rest are held by a small group of controlling investors). In addition to excluding shares held by controlling shareholders, most float-adjusted market-capitalization-weighted indexes also exclude shares held by other corporations and governments. Some providers of indexes that are designed to represent the investment opportunities of global investors further reduce the number of shares included

Exhibit 4 Example of a Market-Capitalization-Weighted Equity Index

Stock	Shares Out-standing	BOP Price	BOP Market Cap	BOP Weight (%)	EOP Price	Dividends Per Share	EOP Market Cap	Total Dividends	Price Return (%)	Total Return (%)	BOP Weight × Price Return (%)	BOP Weight × Total Return (%)	EOP Weight (%)
A	3,000	50.00	150,000	26.29	55.00	0.75	165,000	2,250	10.00	11.50	2.63	3.02	28.50
B	10,000	25.00	250,000	43.82	22.00	0.10	220,000	1,000	−12.00	−11.60	−5.26	−5.08	38.00
C	5,000	12.50	62,500	10.96	8.00	0.00	40,000	0	−36.00	−36.00	−3.95	−3.95	6.91
D	8,000	10.00	80,000	14.02	14.00	0.05	112,000	400	40.00	40.50	5.61	5.68	19.34
E	7,000	4.00	28,000	4.91	6.00	0.00	42,000	0	50.00	50.00	2.46	2.46	7.25
Total			**570,500**	**100.00**			**579,000**	**3,650**			**1.49**	**2.13**	**100.00**
Index Value			**1,000**				**1,014.90**	**6.40**	**1.49**	**2.13**			

Divisor = 570.50

BOP = Beginning of period

EOP = End of period

Type of Index	BOP Value	Return (%)	EOP Value
Price Return	1,000.00	1.49	1,014.90
Total Return	1,000.00	2.13	1,021.30

in the index by excluding shares that are not available to foreigner investors. The index providers may refer to these indexes as "free-float-adjusted market-capitalization-weighted indexes."

Float-adjusted market-capitalization-weighted indexes reflect the shares available for public trading by multiplying the market price per share by the number of shares available to the investing public (i.e., the float-adjusted market capitalization) rather than the total number of shares outstanding (total market capitalization). Currently, most market-capitalization-weighted indexes are float adjusted. Therefore, unless otherwise indicated, for the remainder of this reading, "market-capitalization" weighting refers to float-adjusted market-capitalization weighting.

The float-adjusted market-capitalization weight of security *i* is calculated as:

$$w_i^M = \frac{f_i Q_i P_i}{\sum_{j=1}^{N} f_j Q_j P_j} \tag{15}$$

where

f_i = fraction of shares outstanding in the market float

w_i = fraction of the portfolio that is allocated to security *i* or weight of security *i*

Q_i = number of shares outstanding of security *i*

P_i = share price of security *i*

N = number of securities in the index

Exhibit 5 illustrates the values, weights, and single-period returns following inception of a float-adjusted market-capitalization-weighted equity index using the same five securities as before. The low percentage of shares of Security D in the market float compared with the number of shares outstanding indicates that the security is closely held.

The primary advantage of market-capitalization weighting (including float adjusted) is that constituent securities are held in proportion to their value in the target market. The primary disadvantage is that constituent securities whose prices have risen the most (or fallen the most) have a greater (or lower) weight in the index (i.e., as a security's price rises relative to other securities in the index, its weight increases; and as its price decreases in value relative to other securities in the index, its weight decreases). This weighting method leads to overweighting stocks that have risen in price (and may be overvalued) and underweighting stocks that have declined in price (and may be undervalued). The effect of this weighting method is similar to a momentum investment strategy in that over time, the securities that have risen in price the most will have the largest weights in the index.

3.2.4 *Fundamental Weighting*

Fundamental weighting attempts to address the disadvantages of market-capitalization weighting by using measures of a company's size that are independent of its security price to determine the weight on each constituent security. These measures include book value, cash flow, revenues, earnings, dividends, and number of employees.

Some fundamental indexes use a single measure, such as total dividends, to weight the constituent securities, whereas others combine the weights from several measures to form a composite value that is used for weighting.

Exhibit 5 Example of Float-Adjusted Market-Capitalization-Weighted Equity Index

Stock	Shares Out-standing	% Shares in Market Float	Shares in Index	BOP Price	BOP Float-Adjusted Market Cap	BOP Weight (%)	EOP Price	Dividends Per Share	Ending Float-Adjusted Market Cap	Total Dividends	Price Return (%)	Total Return (%)	BOP Weight × Price Return (%)	BOP Weight × Total Return (%)	EOP Weight (%)
A	3,000	100	3,000	50.00	150,000	35.40	55.00	0.75	165,000	2,250	10.00	11.50	3.54	4.07	39.61
B	10,000	70	7,000	25.00	175,000	41.31	22.00	0.10	154,000	700	−12.00	−11.60	−4.96	−4.79	36.97
C	5,000	90	4,500	12.50	56,250	13.28	8.00	0.00	36,000	0	−36.00	−36.00	−4.78	−4.78	8.64
D	8,000	25	2,000	10.00	20,000	4.72	14.00	0.05	28,000	100	40.00	40.50	1.89	1.91	6.72
E	7,000	80	5,600	4.00	22,400	5.29	6.00	0.00	33,600	0	50.00	50.00	2.65	2.65	8.06
Total					**423,650**	**100.00**			**416,600**	**3,050**			**−1.66**	**−0.94**	**100.00**
Index Value					**1,000**				**983.36**	**7.20**	**−1.66**	**−0.94**			

Divisor = 423.65

BOP = Beginning of period

EOP = End of period

Type of Index	Initial Value	Return (%)	Ending Value
Price Return	1,000.00	−1.66	983.36
Total Return	1,000.00	−0.94	990.56

Letting F_i denote a given fundamental size measure of company i, the fundamental weight on security i is:

$$w_i^F = \frac{F_i}{\sum_{j=1}^{N} F_j} \tag{16}$$

Relative to a market-capitalization-weighted index, a fundamental index with weights based on such an item as earnings will result in greater weights on constituent securities with earnings yields (earnings divided by price) that are higher than the earnings yield of the overall market-weighted portfolio. Similarly, stocks with earnings yields less than the yield on the overall market-weighted portfolio will have lower weights. For example, suppose there are two stocks in an index. Stock A has a market capitalization of €200 million, Stock B has a market capitalization of €800 million, and their aggregate market capitalization is €1 billion (€1,000 million). Both companies have earnings of €20 million and aggregate earnings of €40 million. Thus, Stock A has an earnings yield of 10 percent (20/200) and Stock B has an earnings yield of 2.5 percent (20/800). The earnings weight of Stock A is 50 percent (20/40), which is higher than its market-capitalization weight of 20 percent (200/1,000). The earnings weight of Stock B is 50 percent (20/40), which is less than its market-capitalization weight of 80 percent (800/1,000). Relative to the market-cap-weighted index, the earnings-weighted index over-weights the high-yield Stock A and under-weights the low-yield Stock B.

The most important property of fundamental weighting is that it leads to indexes that have a “value” tilt. That is, a fundamentally weighted index has ratios of book value, earnings, dividends, etc. to market value that are higher than its market-capitalization-weighted counterpart. Also, in contrast to the momentum “effect” of market-capitalization-weighted indexes, fundamentally weighted indexes generally will have a contrarian “effect” in that the portfolio weights will shift away from securities that have increased in relative value and toward securities that have fallen in relative value whenever the portfolio is rebalanced.

INDEX MANAGEMENT: REBALANCING AND RECONSTITUTION

4

f describe rebalancing and reconstitution of an index;

So far, we have discussed index construction. Index management entails the two remaining questions:

- When should the index be rebalanced?
- When should the security selection and weighting decisions be re-examined?

4.1 Rebalancing

Rebalancing refers to adjusting the weights of the constituent securities in the index. To maintain the weight of each security consistent with the index’s weighting method, the index provider rebalances the index by adjusting the weights of the constituent securities on a regularly scheduled basis (rebalancing dates)—usually quarterly. Rebalancing is necessary because the weights of the constituent securities change as their market prices change. Note, for example, that the weights of the securities in the equal-weighted index (Exhibit 3) at the end of the period are no longer equal (i.e., 20 percent):

Security A	19.93%
Security B	15.94
Security C	11.60
Security D	25.36
Security E	27.17

In rebalancing the index, the weights of Securities D and E (which had the highest returns) would be decreased and the weights of Securities A, B, and C (which had the lowest returns) would be increased. Thus, rebalancing creates turnover within an index.

Price-weighted indexes are not rebalanced because the weight of each constituent security is determined by its price. For market-capitalization-weighted indexes, rebalancing is less of a concern because the indexes largely rebalance themselves. In our market-capitalization index, for example, the weight of Security C automatically declined from 10.96 percent to 6.91 percent, reflecting the 36 percent decline in its market price. Market-capitalization weights are only adjusted to reflect mergers, acquisitions, liquidations, and other corporate actions between rebalancing dates.

4.2 Reconstitution

Reconstitution is the process of changing the constituent securities in an index. It is similar to a portfolio manager deciding to change the securities in his or her portfolio. Reconstitution is part of the rebalancing cycle. The reconstitution date is the date on which index providers review the constituent securities, re-apply the initial criteria for inclusion in the index, and select which securities to retain, remove, or add. Constituent securities that no longer meet the criteria are replaced with securities that do meet the criteria. Once the revised list of constituent securities is determined, the weighting method is re-applied. Indexes are reconstituted to reflect changes in the target market (bankruptcies, de-listings, mergers, acquisitions, etc.) and/or to reflect the judgment of the selection committee.

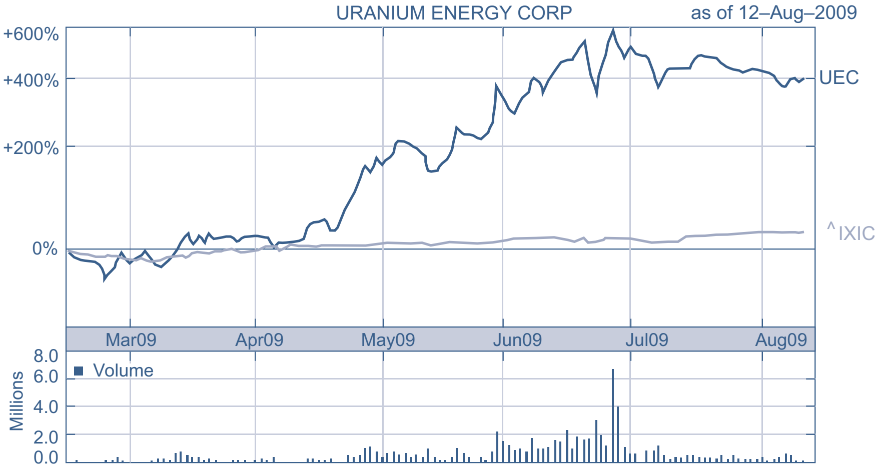
Reconstitution creates turnover in a number of different ways, particularly for market-capitalization-weighted indexes. When one security is removed and another is added, the index provider has to change the weights of the other securities in order to maintain the market-capitalization weighting of the index.

The frequency of reconstitution is a major issue for widely used indexes and their constituent securities. The Russell 2000 Index, for example, reconstitutes annually. It is used as a benchmark by numerous investment funds, and each year, prior to the index's reconstitution, the managers of these funds buy stocks they think will be added to the index—driving those stocks' prices up—and sell stocks they think will be deleted from the index—driving those stocks' prices down. Exhibit 6 illustrates a historical example of the potential impact of these decisions. Beginning in late April 2009, some managers began acquiring and bidding up the price of Uranium Energy Corporation (UEC) because they believed that it would be included in the reconstituted Russell 2000 Index. On 12 June, Russell listed UEC as a preliminary addition to the Russell 2000 Index and the Russell 3000 Index.[2] By that time, the stock value had increased by more than 300 percent. Investors continued to bid up the stock price in the weeks following the announcement, and the stock closed on the reconstitution date of 30 June at USD2.90, up nearly 400 percent for the quarter.

---

[2] According to the press release, final membership in the index would be published after market close on Friday, 26 June.

Exhibit 6 **Three-Month Performance of Uranium Energy Corporation and NASDAQ April through June 2009**



Source: Yahoo! Finance and Capital IQ.

USES OF MARKET INDEXES

5

g describe uses of security market indexes;

Indexes were initially created to give a sense of how a particular security market performed on a given day. With the development of modern financial theory, their uses in investment management have expanded significantly. Some of the major uses of indexes include:

- gauges of market sentiment;
- proxies for measuring and modeling returns, systematic risk, and risk-adjusted performance;
- proxies for asset classes in asset allocation models;
- benchmarks for actively managed portfolios; and
- model portfolios for such investment products as index funds and exchange-traded funds (ETFs).

Investors using security market indexes must be familiar with how various indexes are constructed in order to select the index or indexes most appropriate for their needs.

5.1 Gauges of Market Sentiment

The original purpose of stock market indexes was to provide a gauge of investor confidence or market sentiment. As indicators of the collective opinion of market participants, indexes reflect investor attitudes and behavior. The Dow Jones Industrial Average has a long history, is frequently quoted in the media, and remains a popular gauge of market sentiment. It may not accurately reflect the overall attitude of investors or the "market," however, because the index consists of only 30 of the thousands of US stocks traded each day.

5.2 Proxies for Measuring and Modeling Returns, Systematic Risk, and Risk-Adjusted Performance

The capital asset pricing model (CAPM) defines beta as the systematic risk of a security with respect to the entire market. The market portfolio in the CAPM consists of all risky securities. To represent the performance of the market portfolio, investors use a broad index. For example, the Tokyo Price Index (TOPIX) and the S&P 500 often serve as proxies for the market portfolio in Japan and the United States, respectively, and are used for measuring and modeling systematic risk and market returns.

Security market indexes also serve as market proxies when measuring risk-adjusted performance. The beta of an actively managed portfolio allows investors to form a passive alternative with the same level of systematic risk. For example, if the beta of an actively managed portfolio of global stocks is 0.95 with respect to the MSCI World Index, investors can create a passive portfolio with the same systematic risk by investing 95 percent of their portfolio in a MSCI World Index fund and holding the remaining 5 percent in cash. Alpha, the difference between the return of the actively managed portfolio and the return of the passive portfolio, is a measure of risk-adjusted return or investment performance. Alpha can be the result of manager skill (or lack thereof), transaction costs, and fees.

5.3 Proxies for Asset Classes in Asset Allocation Models

Because indexes exhibit the risk and return profiles of select groups of securities, they play a critical role as proxies for asset classes in asset allocation models. They provide the historical data used to model the risks and returns of different asset classes.

5.4 Benchmarks for Actively Managed Portfolios

Investors often use indexes as benchmarks to evaluate the performance of active portfolio managers. The index selected as the benchmark should reflect the investment strategy used by the manager. For example, an active manager investing in global small-capitalization stocks should be evaluated using a benchmark index, such as the FTSE Global Small Cap Index, which includes approximately 4,400 liquid small-capitalization stocks across 47 countries as of August 2018.

The choice of an index to use as a benchmark is important because an inappropriate index could lead to incorrect conclusions regarding an active manager's investment performance. Suppose that the small-cap manager underperformed the small-cap index but outperformed a broad equity market index. If investors use the broad market index as a benchmark, they might conclude that the small-cap manager is earning his or her fees and should be retained or given additional assets to invest. Using the small-cap index as a benchmark might lead to a very different conclusion.

5.5 Model Portfolios for Investment Products

Indexes also serve as the basis for the development of new investment products. Using indexes as benchmarks for actively managed portfolios has led some investors to conclude that they should invest in the benchmarks instead. Based on the CAPM's conclusion that investors should hold the market portfolio, broad market index funds have been developed to function as proxies for the market portfolio.

Investment management firms initially developed and managed index portfolios for institutional investors. Eventually, mutual fund companies introduced index funds for individual investors. Subsequently, investment management firms introduced exchange-traded funds, which are managed the same way as index mutual funds but trade like stocks.

The first ETFs were based on existing indexes. As the popularity of ETFs increased, index providers created new indexes for the specific purpose of forming ETFs, leading to the creation of numerous narrowly defined indexes with corresponding ETFs. The VanEck Vectors Vietnam ETF, for example, allows investors to invest in the equity market of Vietnam.

The choice of indexes to meet the needs of investors is extensive. Index providers are constantly looking for opportunities to develop indexes to meet the needs of investors.

EQUITY INDEXES

6

h describe types of equity indexes;

k compare types of security market indexes.

A wide variety of equity indexes exist, including broad market, multi-market, sector, and style indexes.

6.1 Broad Market Indexes

A broad equity market index, as its name suggests, represents an entire given equity market and typically includes securities representing more than 90 percent of the selected market. For example, the Shanghai Stock Exchange Composite Index (SSE) is a market-capitalization-weighted index of all shares that trade on the Shanghai Stock Exchange. In the United States, the Wilshire 5000 Total Market Index is a market-capitalization-weighted index that includes all US equities with readily available prices and is designed to represent the entire US equity market.[3] The Russell 3000, consisting of the largest 3,000 stocks by market capitalization, represents approximately 98 percent of the US equity market.

6.2 Multi-Market Indexes

Multi-market indexes usually comprise indexes from different countries and regions and are designed to represent multiple security markets. Multi-market indexes may represent multiple national markets, geographic regions, economic development groups, and, in some cases, the entire world. World indexes are of importance to investors who take a global approach to equity investing without any particular bias toward a particular country or region. A number of index providers publish families of multi-market equity indexes.

---

[3] Despite its name, the Wilshire 5000 has no constraint on the number of securities that can be included. It included approximately 5,000 securities at inception.

MSCI offers a number of multi-market indexes. As shown in Exhibit 7, MSCI classifies countries and regions along two dimensions: level of economic development and geographic region. Developmental groups, which MSCI refers to as market classifications, include developed markets, emerging markets, and frontier markets. The geographic regions are largely divided by longitudinal lines of the globe: the Americas, Europe with Africa, and Asia with the Pacific. MSCI provides country- and region-specific indexes for each of the developed and emerging markets within its multi-market indexes. MSCI periodically reviews the classifications of markets in its indexes for movement from frontier markets to emerging markets and from emerging markets to developed markets and reconstitutes the indexes accordingly.

Exhibit 7 MSCI Global Investable Market Indexes (as of October 2018)

Developed Markets

Americas	Europe and Middle East	Pacific
Canada, United States	Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Israel, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom	Australia, Hong Kong SAR, Japan, New Zealand, Singapore

Emerging Markets

Americas	Europe, Middle East, Africa	Asia
Brazil, Chile, Colombia, Mexico, Peru	Czech Republic, Egypt, Greece, Hungary, Poland, Qatar, Russia, South Africa, Turkey, United Arab Emirates	Chinese mainland, India, Indonesia, South Korea, Malaysia, Pakistan, Philippines, Taiwan region, Thailand

Frontier Markets

Americas	Europe & CIS	Africa	Middle East	Asia
Argentina	Croatia, Estonia, Lithuania, Kazakhstan, Romania, Serbia, Slovenia	Kenya, Mauritius, Morocco, Nigeria, Tunisia, WAEMU[1]	Bahrain, Jordan, Kuwait, Lebanon, Oman	Bangladesh, Sri Lanka, Vietnam

MSCI Standalone Market Indexes[2]

Europe, Middle East, and Africa	Americas	Europe and CIS	Africa	Middle East
Saudi Arabia	Jamaica, Panama,[3] Trinidad & Tobago	Bosnia Herzegovina, Bulgaria, Ukraine	Botswana, Ghana, Zimbabwe	Palestine

[1] The West African Economic and Monetary Union (WAEMU) consists of the following countries: Benin, Burkina Faso, Ivory Coast, Guinea-Bissau, Mali, Niger, Senegal, and Togo. Currently the MSCI WAEMU Indexes include securities classified in Senegal, Ivory Coast, and Burkina Faso.

[2] The MSCI Standalone Market Indexes are not included in the MSCI Emerging Markets Index or MSCI Frontier Markets Index. However, these indexes use either the Emerging Markets or the Frontier Markets methodological criteria concerning size and liquidity.

[3] MSCI Panama Index has been launched as a Standalone Market Index.

Source: adapted from MSCI (https://www.msci.com/en/market-cap-weighted-indexes), October 2018.

6.2.1 *Fundamental Weighting in Multi-Market Indexes*

Some index providers weight the securities within each country/region by market capitalization and then weight each country/region in the overall index in proportion to its relative GDP, effectively creating fundamental weighting in multi-market indexes. GDP-weighted indexes were some of the first fundamentally weighted indexes created. Introduced in 1987 by MSCI to address the 60 percent weight of Japanese equities in the market-capitalization-weighted MSCI EAFE Index at the time, GDP-weighted indexes reduced the allocation to Japanese equities by half.[4]

6.3 Sector Indexes

Sector indexes represent and track different economic sectors—such as consumer goods, energy, finance, health care, and technology—on either a national, regional, or global basis. Because different sectors of the economy behave differently over the course of the business cycle, some investors may seek to overweight or underweight their exposure to particular sectors.

Sector indexes are organized as families; each index within the family represents an economic sector. Typically, the aggregation of a sector index family is equivalent to a broad market index. Economic sector classification can be applied on a global, regional, or country-specific basis, but no universally agreed upon sector classification method exists.

Sector indexes play an important role in performance analysis because they provide a means to determine whether a portfolio manager is more successful at stock selection or sector allocation. Sector indexes also serve as model portfolios for sector-specific ETFs and other investment products.

6.4 Style Indexes

Style indexes represent groups of securities classified according to market capitalization, value, growth, or a combination of these characteristics. They are intended to reflect the investing styles of certain investors, such as the growth investor, value investor, and small-cap investor.

6.4.1 *Market Capitalization*

Market-capitalization indexes represent securities categorized according to the major capitalization categories: large cap, midcap, and small cap. With no universal definition of these categories, the indexes differ on the distinctions between large cap and midcap and between midcap and small cap, as well as the minimum market-capitalization size required to be included in a small-cap index. Classification into categories can be based on absolute market capitalization (e.g., below €100 million) or relative market capitalization (e.g., the smallest 2,500 stocks).

6.4.2 *Value/Growth Classification*

Some indexes represent categories of stocks based on their classifications as either value or growth stocks. Different index providers use different factors and valuation ratios (low price-to-book ratios, low price-to-earnings ratios, high dividend yields, etc.) to distinguish between value and growth equities.

---

[4] Steven A. Schoenfeld, *Active Index Investing* (Hoboken, NJ: John Wiley & Sons, 2004):220.

6.4.3 *Market Capitalization and Value/Growth Classification*

Combining the three market-capitalization groups with value and growth classifications results in six basic style index categories:

- Large-Cap Value
- Mid-Cap Value
- Small-Cap Value
- Large-Cap Growth
- Mid-Cap Growth
- Small-Cap Growth

Because indexes use different size and valuation classifications, the constituents of indexes designed to represent a given style, such as small-cap value, may differ—sometimes substantially.

Because valuation ratios and market capitalizations change over time, stocks frequently migrate from one style index category to another on reconstitution dates. As a result, style indexes generally have much higher turnover than do broad market indexes.

7

FIXED-INCOME INDEXES

i. describe types of fixed-income indexes;

k. compare types of security market indexes.

A wide variety of fixed-income indexes exists, but the nature of the fixed-income markets and fixed-income securities leads to some very important challenges to fixed-income index construction and replication. These challenges are the number of securities in the fixed-income universe, the availability of pricing data, and the liquidity of the securities.

7.1 Construction

The fixed-income universe includes securities issued by governments, government agencies, and corporations. Each of these entities may issue a variety of fixed-income securities with different characteristics. As a result, the number of fixed-income securities is many times larger than the number of equity securities. To represent a specific fixed-income market or segment, indexes may include thousands of different securities. Over time, these fixed-income securities mature, and issuers offer new securities to meet their financing needs, leading to turnover in fixed-income indexes.

Another challenge in index construction is that fixed-income markets are predominantly dealer markets. This means that firms (dealers) are assigned to specific securities and are responsible for creating liquid markets for those securities by purchasing and selling them from their inventory. In addition, many securities do not trade frequently and, as a result, are relatively illiquid. As a result, index providers must contact dealers to obtain current prices on constituent securities to update the index or they must estimate the prices of constituent securities using the prices of traded fixed-income securities with similar characteristics.

These challenges can result in indexes with dissimilar numbers of bonds representing the same markets. The large number of fixed-income securities—combined with the lack of liquidity of some securities—has made it more costly and difficult, compared with equity indexes, for investors to replicate fixed-income indexes and duplicate their performance.

7.2 Types of Fixed-Income Indexes

The wide variety of fixed-income securities, ranging from zero-coupon bonds to bonds with embedded options (i.e., callable or putable bonds), results in a number of different types of fixed-income indexes. Similar to equities, fixed-income securities can be categorized according to the issuer's economic sector, the issuer's geographic region, or the economic development of the issuer's geographic region. Fixed-income securities can also be classified along the following dimensions:

- type of issuer (government, government agency, corporation);
- type of financing (general obligation, collateralized);
- currency of payments;
- maturity;
- credit quality (investment grade, high yield, credit agency ratings); and
- absence or presence of inflation protection.

Fixed-income indexes are based on these various dimensions and can be categorized as follows:

- aggregate or broad market indexes;
- market sector indexes;
- style indexes;
- economic sector indexes; and
- specialized indexes such as high-yield, inflation-linked, and emerging market indexes.

The first fixed-income index created, the Bloomberg Barclays US Aggregate Bond Index (formerly the Barclays Capital Aggregate Bond Index), is an example of a single-country aggregate index. Designed to represent the broad market of US fixed-income securities, it comprises approximately 8,000 securities, including US Treasury, government-related, corporate, mortgage-backed, asset-backed, and commercial mortgage-backed securities.

Aggregate indexes can be subdivided by market sector (government, government agency, collateralized, corporate); style (maturity, credit quality); economic sector, or some other characteristic to create more narrowly defined indexes. A common distinction reflected in indexes is between investment grade (e.g., those with a Standard & Poor's credit rating of BBB– or better) and high-yield securities. Investment-grade indexes are typically further subdivided by maturity (i.e., short, intermediate, or long) and by credit rating (e.g., AAA, BBB, etc.).[5] The wide variety of fixed-income indexes reflects the partitioning of fixed-income securities on the basis of a variety of dimensions.

Exhibit 8 illustrates how the major types of fixed-income indexes can be organized on the basis of various dimensions.

Exhibit 8 Dimensions of Fixed-Income Indexes

Market	Global
	Regional
	Country or currency zone

(continued)

---

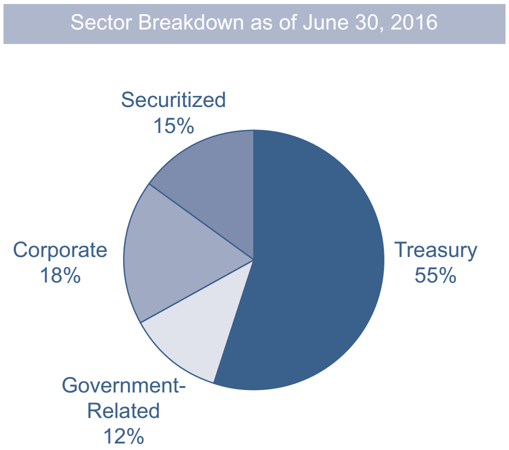
[5] Credit ratings are discussed in depth in the Level I CFA Program reading "Fundamentals of Credit Analysis."

Exhibit 8 (Continued)

Type	Corporate	Collateralized *Securitized* *Mortgage-backed*	Government agency	Government
Maturity	For example, 1–3, 3–5, 5–7, 7–10, 10+ years; short-term, medium-term, or long-term			
Credit quality	For example, AAA, AA, A, BBB, etc.; Aaa, Aa, A, Baa, etc.; investment grade, high yield			

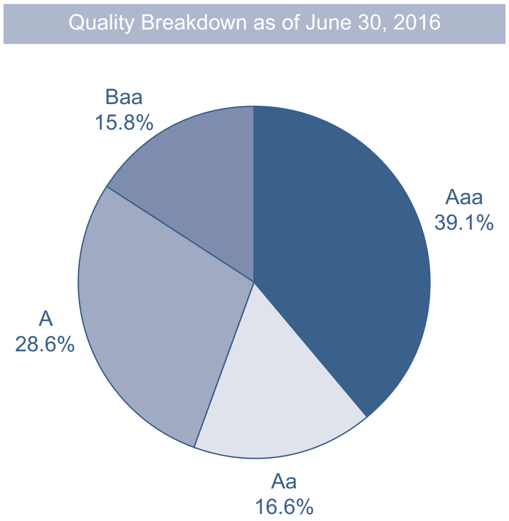
All aggregate indexes include a variety of market sectors and credit ratings. The breakdown of the Bloomberg Barclays Global Aggregate Bond Index by market sectors and by credit rating is shown in Exhibit 9 and Exhibit 10, respectively.

Exhibit 9 Market Sector Breakdown of the Bloomberg Barclays Global Aggregate Bond Index



Source: Bloomberg Barclays Indices, Global Aggregate Index Factsheet, August 24, 2016. https://data.bloomberglp.com/indices/sites/2/2016/08/Factsheet-Global-Aggregate.pdf.

Exhibit 10 Credit Breakdown of the Bloomberg Barclays Global Aggregate Bond Index



Source: Bloomberg Barclays Indices, Global Aggregate Index Factsheet, August 24, 2016. https://data.bloombergl p.com/indices/sites/2/2016/08/Factsheet-Global-Aggregate.pdf.

INDEXES FOR ALTERNATIVE INVESTMENTS

8

j describe indexes representing alternative investments;

k compare types of security market indexes.

Many investors seek to lower the risk or enhance the performance of their portfolios by investing in assets classes other than equities and fixed income. Interest in alternative assets and investment strategies has led to the creation of indexes designed to represent broad classes of alternative investments. Three of the most widely followed alternative investment classes are commodities, real estate, and hedge funds.

8.1 Commodity Indexes

Commodity indexes consist of futures contracts on one or more commodities, such as agricultural products (rice, wheat, sugar), livestock (cattle, hogs), precious and common metals (gold, silver, copper), and energy commodities (crude oil, natural gas).

Although some commodity indexes may include the same commodities, the returns of these indexes may differ because each index may use a different weighting method. Because commodity indexes do not have an obvious weighting mechanism, such as market capitalization, commodity index providers create their own weighting methods. Some indexes, such as the Thomson Reuters/Core Commodity CRB Index (TR/CC CRB Index), formerly known as the Commodity Research Bureau (CRB) Index, contain a fixed number of commodities that are weighted equally. The S&P GSCI uses a combination of liquidity measures and world production values in its weighting scheme and allocates more weight to commodities that have risen in price. Other indexes have fixed weights that are determined by a committee.

The different weighting methods can also lead to large differences in exposure to specific commodities. The S&P GSCI in 2018, for example, weights the energy-sector approximately 50% higher and the agriculture sector 40% lower than the CRB Index. These differences result in indexes with very different risk and return profiles. Unlike commodity indexes, broad equity and fixed-income indexes that target the same markets share similar risk and return profiles.

The performance of commodity indexes can also be quite different from their underlying commodities because the indexes consist of futures contracts on the commodities rather than the actual commodities. Index returns are affected by factors other than changes in the prices of the underlying commodities because futures contracts must be continually "rolled over" (i.e., replacing a contract nearing expiration with a new contract). Commodity index returns reflect the risk-free interest rate, the changes in future prices, and the roll yield. Therefore, a commodity index return can be quite different from the return based on changes in the prices of the underlying commodities.

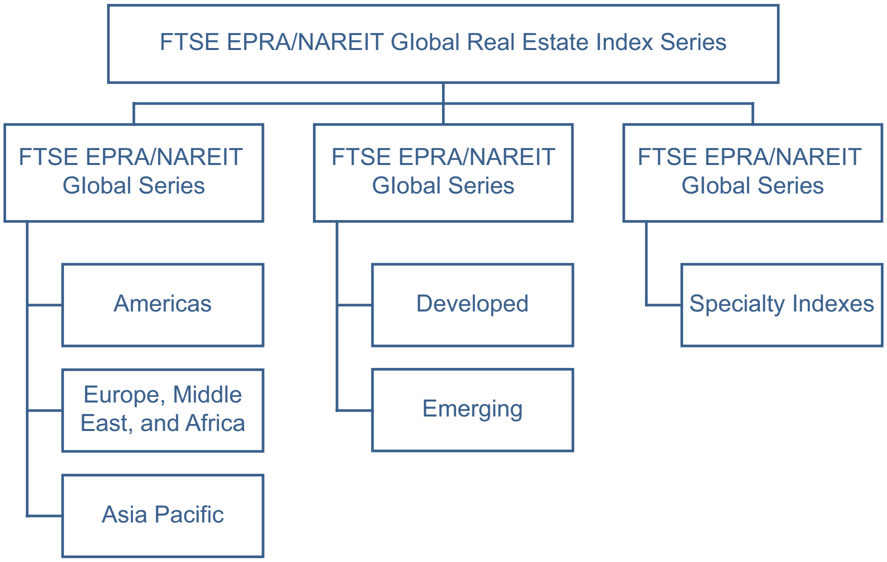
8.2 Real Estate Investment Trust Indexes

Real estate indexes represent not only the market for real estate securities but also the market for real estate—a highly illiquid market and asset class with infrequent transactions and pricing information. Real estate indexes can be categorized as appraisal indexes, repeat sales indexes, and real estate investment trust (REIT) indexes.

REIT indexes consist of shares of publicly traded REITs. REITS are public or private corporations organized specifically to invest in real estate, either through ownership of properties or investment in mortgages. Shares of public REITs are traded on the world's various stock exchanges and are a popular choice for investing in commercial real estate properties. Because REIT indexes are based on publicly traded REITs with continuous market pricing, the value of REIT indexes is calculated continuously.

The FTSE EPRA/NAREIT global family of REIT indexes shown in Exhibit 11 seeks to represent trends in real estate stocks worldwide and includes representation from the European Public Real Estate Association (EPRA) and the National Association of Real Estate Investment Trusts (NAREIT).

Exhibit 11 The FTSE EPRA/NAREIT Global REIT Index Family


FTSE EPRA/NAREIT Global Real Estate Index Series
FTSE EPRA/NAREIT Global Series
FTSE EPRA/NAREIT Global Series
FTSE EPRA/NAREIT Global Series
Americas
Europe, Middle East, and Africa
Asia Pacific
Developed
Emerging
Specialty Indexes

Source: FTSE International, "FTSE EPRA/NAREIT Global & Global Ex US Indices" Factsheet 2009). "FTSE®" is a trade mark of the London Stock Exchange Plc, "NAREIT®" is a trade mark of the National Association of Real Estate Investment Trusts ("NAREIT") and "EPRA®" is a trade mark of the European Public Real Estate Association ("EPRA") and all are used by FTSE International Limited ("FTSE") under license.

8.3 Hedge Fund Indexes

Hedge fund indexes reflect the returns on hedge funds. **Hedge funds** are private investment vehicles that typically use leverage and long and short investment strategies.

A number of research organizations maintain databases of hedge fund returns and summarize these returns into indexes. These database indexes are designed to represent the performance of the hedge funds on a very broad global level (hedge funds in general) or the strategy level. Most of these indexes are equal weighted and represent the performance of the hedge funds within a particular database.

Most research organizations rely on the voluntary cooperation of hedge funds to compile performance data. As unregulated entities, however, hedge funds are not required to report their performance to any party other than their investors. Therefore, each hedge fund decides to which database(s) it will report its performance. As a result, rather than index providers determining the constituents, the constituents determine the index.

Frequently, a hedge fund reports its performance to only one database. The result is little overlap of funds covered by the different indexes. With little overlap between their constituents, different global hedge fund indexes may reflect very different performance for the hedge fund industry over the same period of time.

Another consequence of the voluntary performance reporting is the potential for survivorship bias and, therefore, inaccurate performance representation. This means that hedge funds with poor performance may be less likely to report their performance to the database or may stop reporting to the database, so their returns may be excluded when measuring the return of the index. As a result, the index may not accurately reflect actual hedge fund performance so much as the performance of hedge funds that are performing well.

REPRESENTATIVE INDEXES WORLDWIDE

As indicated in this reading, the choice of indexes to meet the needs of investors is extensive. Investors using security market indexes must be careful in their selection of the index or indexes most appropriate for their needs. The following table illustrates the variety of indexes reflecting different asset classes, markets, and weighting methods.

Index	**Representing**	**Number of Securities**	**Weighting Method**	**Comments**
Dow Jones Industrial Average	US blue chip companies	30	Price	The oldest and most widely known US equity index. *Wall Street Journal* editors choose 30 stocks from among large, mature blue-chip companies.
Nikkei Stock Average	Japanese blue chip companies	225	Modified price	Known as the Nikkei 225 and originally formulated by Dow Jones & Company. Because of extreme variation in price levels of component securities, some high-priced shares are weighted as a fraction of share price.
TOPIX	All companies listed on the Tokyo Stock Exchange First Section	Varies	Float-adjusted market cap	Represents about 93 percent of the market value of all Japanese equities. Contains a large number of very small, illiquid stocks, making exact replication difficult.
MSCI All Country World Index	Stocks of 23 developed and 24 emerging markets	Varies	Free-float-adjusted market cap	Composed of companies representative of the market structure of developed and emerging market countries in the Americas, Europe/Middle East, and Asia/Pacific regions. Price return and total return versions available in both USD and local currencies.
S&P Developed Ex-US BMI Energy Sector Index	Energy sector of developed global markets outside the United States	Varies	Float-adjusted market cap	Serves as a model portfolio for the SPDR® S&P Energy Sector Exchange-Traded Fund (ETF).
Bloomberg Barclays Global Aggregate Bond Index	Investment-grade bonds in the North American, European, and Asian markets	Varies	Market cap	Formerly known as Lehman Brothers Global Aggregate Bond Index.
Markit iBoxx Euro High-Yield Bond Indexes	Sub-investment-grade euro-denominated corporate bonds	Varies	Market cap and variations	Rebalanced monthly. Represents tradable part of market. Price and total return versions available with such analytical values as yield, duration, modified duration, and convexity. Provides platform for research and structured products.
FTSE EPRA/NAREIT Global Real Estate Index	Real estate securities in the North American, European, and Asian markets	Varies	Float-adjusted market cap	The stocks of REITs that constitute the index trade on public stock exchanges and may be constituents of equity market indexes.

Index	Representing	Number of Securities	Weighting Method	Comments
HFRX Global Hedge Fund Index	Overall composition of the HFR database	Varies	Asset weighted	Comprises all eligible hedge fund strategies. Examples include convertible arbitrage, distressed securities, market neutral, event driven, macro, and relative value arbitrage. Constituent strategies are asset weighted on the basis of asset distribution within the hedge fund industry.
HFRX Equal Weighted Strategies EUR Index	Overall composition of the HFR database	Varies	Equal weighted	Denominated in euros and is constructed from the same strategies as the HFRX Global Hedge Fund Index.
Morningstar Style Indexes	US stocks classified by market cap and value/growth orientation	Varies	Float-adjusted market cap	The nine indexes defined by combinations of market cap (large, mid, and small) and value/growth orientation (value, core, growth) have mutually exclusive constituents and are exhaustive with respect to the Morningstar US Market Index. Each is a model portfolio for one of the iShares Morningstar ETFs.

SUMMARY

This reading explains and illustrates the construction, management, and uses of security market indexes. It also discusses various types of indexes. Security market indexes are invaluable tools for investors, who can select from among thousands of indexes representing a variety of security markets, market segments, and asset classes. These indexes range from those representing the global market for major asset classes to those representing alternative investments in specific geographic markets. To benefit from the use of security market indexes, investors must understand their construction and determine whether the selected index is appropriate for their purposes. Frequently, an index that is well suited for one purpose may not be well suited for other purposes. Users of indexes must be familiar with how various indexes are constructed in order to select the index or indexes most appropriate for their needs.

Among the key points made in this reading are the following:

- Security market indexes are intended to measure the values of different target markets (security markets, market segments, or asset classes).
- The constituent securities selected for inclusion in the security market index are intended to represent the target market.
- A price return index reflects only the prices of the constituent securities.
- A total return index reflects not only the prices of the constituent securities but also the reinvestment of all income received since the inception of the index.
- Methods used to weight the constituents of an index range from the very simple, such as price and equal weightings, to the more complex, such as market-capitalization and fundamental weightings.

- Choices in index construction—in particular, the choice of weighting method—affect index valuation and returns.
- Index management includes 1) periodic rebalancing to ensure that the index maintains appropriate weightings and 2) reconstitution to ensure the index represents the desired target market.
- Rebalancing and reconstitution create turnover in an index. Reconstitution can dramatically affect prices of current and prospective constituents.
- Indexes serve a variety of purposes. They gauge market sentiment and serve as benchmarks for actively managed portfolios. They act as proxies for measuring systematic risk and risk-adjusted performance. They also serve as proxies for asset classes in asset allocation models and as model portfolios for investment products.
- Investors can choose from security market indexes representing various asset classes, including equity, fixed-income, commodity, real estate, and hedge fund indexes.
- Within most asset classes, index providers offer a wide variety of indexes, ranging from broad market indexes to highly specialized indexes based on the issuer's geographic region, economic development group, or economic sector or other factors.
- Proper use of security market indexes depends on understanding their construction and management.

PRACTICE PROBLEMS

1 A security market index represents the:

A risk of a security market.

B security market as a whole.

C security market, market segment, or asset class.

2 Security market indexes are:

A constructed and managed like a portfolio of securities.

B simple interchangeable tools for measuring the returns of different asset classes.

C valued on a regular basis using the actual market prices of the constituent securities.

3 When creating a security market index, an index provider must first determine the:

A target market.

B appropriate weighting method.

C number of constituent securities.

4 One month after inception, the price return version and total return version of a single index (consisting of identical securities and weights) will be equal if:

A market prices have not changed.

B capital gains are offset by capital losses.

C the securities do not pay dividends or interest.

5 The values of a price return index and a total return index consisting of identical equal-weighted dividend-paying equities will be equal:

A only at inception.

B at inception and on rebalancing dates.

C at inception and on reconstitution dates.

6 An analyst gathers the following information for an equal-weighted index comprised of assets Able, Baker, and Charlie:

Security	Beginning of Period Price (€)	End of Period Price (€)	Total Dividends (€)
Able	10.00	12.00	0.75
Baker	20.00	19.00	1.00
Charlie	30.00	30.00	2.00

The price return of the index is:

A 1.7%.

B 5.0%.

C 11.4%.

7 An analyst gathers the following information for an equal-weighted index comprised of assets Able, Baker, and Charlie:

© 2011 CFA Institute. All rights reserved.

Security	Beginning of Period Price (€)	End of Period Price (€)	Total Dividends (€)
Able	10.00	12.00	0.75
Baker	20.00	19.00	1.00
Charlie	30.00	30.00	2.00

The total return of the index is:

A 5.0%.

B 7.9%.

C 11.4%.

8 An analyst gathers the following information for a price-weighted index comprised of securities ABC, DEF, and GHI:

Security	Beginning of Period Price (£)	End of Period Price (£)	Total Dividends (£)
ABC	25.00	27.00	1.00
DEF	35.00	25.00	1.50
GHI	15.00	16.00	1.00

The price return of the index is:

A –4.6%.

B –9.3%.

C –13.9%.

9 An analyst gathers the following information for a market-capitalization-weighted index comprised of securities MNO, QRS, and XYZ:

Security	Beginning of Period Price (¥)	End of Period Price (¥)	Dividends per Share (¥)	Shares Outstanding
MNO	2,500	2,700	100	5,000
QRS	3,500	2,500	150	7,500
XYZ	1,500	1,600	100	10,000

The price return of the index is:

A –9.33%.

B –10.23%.

C –13.90%.

10 An analyst gathers the following information for a market-capitalization-weighted index comprised of securities MNO, QRS, and XYZ:

Security	Beginning of Period Price (¥)	End of Period Price (¥)	Dividends Per Share (¥)	Shares Outstanding
MNO	2,500	2,700	100	5,000
QRS	3,500	2,500	150	7,500
XYZ	1,500	1,600	100	10,000

The total return of the index is:

A 1.04%.

B –5.35%.

C –10.23%.

11 When creating a security market index, the target market:

A determines the investment universe.

B is usually a broadly defined asset class.

C determines the number of securities to be included in the index.

12 An analyst gathers the following data for a price-weighted index:

	Beginning of Period		End of Period	
Security	Price (€)	Shares Outstanding	Price (€)	Shares Outstanding
A	20.00	300	22.00	300
B	50.00	300	48.00	300
C	26.00	2,000	30.00	2,000

The price return of the index over the period is:

A 4.2%.

B 7.1%.

C 21.4%.

13 An analyst gathers the following data for a value-weighted index:

	Beginning of Period		End of Period	
Security	Price (£)	Shares Outstanding	Price (£)	Shares Outstanding
A	20.00	300	22.00	300
B	50.00	300	48.00	300
C	26.00	2,000	30.00	2,000

The return on the value-weighted index over the period is:

A 7.1%.

B 11.0%.

C 21.4%.

14 An analyst gathers the following data for an equally-weighted index:

	Beginning of Period		End of Period	
Security	Price (¥)	Shares Outstanding	Price (¥)	Shares Outstanding
A	20.00	300	22.00	300
B	50.00	300	48.00	300
C	26.00	2,000	30.00	2,000

The return on the index over the period is:

A 4.2%.

B 6.8%.

C 7.1%.

15 Which of the following index weighting methods requires an adjustment to the divisor after a stock split?

A Price weighting.

B Fundamental weighting.

C Market-capitalization weighting.

16 If the price return of an equal-weighted index exceeds that of a market-capitalization-weighted index comprised of the same securities, the *most likely* explanation is:

A stock splits.

B dividend distributions.

C outperformance of small-market-capitalization stocks.

17 A float-adjusted market-capitalization-weighted index weights each of its constituent securities by its price and:

A its trading volume.

B the number of its shares outstanding.

C the number of its shares available to the investing public.

18 Which of the following index weighting methods is most likely subject to a value tilt?

A Equal weighting.

B Fundamental weighting.

C Market-capitalization weighting.

19 Rebalancing an index is the process of periodically adjusting the constituent:

A securities' weights to optimize investment performance.

B securities to maintain consistency with the target market.

C securities' weights to maintain consistency with the index's weighting method.

20 Which of the following index weighting methods requires the most frequent rebalancing?

A Price weighting.

B Equal weighting.

C Market-capitalization weighting.

21 Reconstitution of a security market index reduces:

A portfolio turnover.

B the need for rebalancing.

C the likelihood that the index includes securities that are not representative of the target market.

22 Security market indexes are used as:

A measures of investment returns.

B proxies to measure unsystematic risk.

C proxies for specific asset classes in asset allocation models.

23 Uses of market indexes do not include serving as a:

A measure of systemic risk.

B basis for new investment products.

C benchmark for evaluating portfolio performance.

24 Which of the following statements regarding sector indexes is *most* accurate? Sector indexes:

A track different economic sectors and cannot be aggregated to represent the equivalent of a broad market index.

B provide a means to determine whether an active investment manager is more successful at stock selection or sector allocation.

C apply a universally agreed upon sector classification system to identify the constituent securities of specific economic sectors, such as consumer goods, energy, finance, health care.

25 Which of the following is an example of a style index? An index based on:

A geography.

B economic sector.

C market capitalization.

26 Which of the following statements regarding fixed-income indexes is *most* accurate?

A Liquidity issues make it difficult for investors to easily replicate fixed-income indexes.

B Rebalancing and reconstitution are the only sources of turnover in fixed-income indexes.

C Fixed-income indexes representing the same target market hold similar numbers of bonds.

27 An aggregate fixed-income index:

A comprises corporate and asset-backed securities.

B represents the market of government-issued securities.

C can be subdivided by market or economic sector to create more narrowly defined indexes.

28 Fixed-income indexes are *least likely* constructed on the basis of:

A maturity.

B type of issuer.

C coupon frequency.

29 Commodity index values are based on:

A futures contract prices.

B the market price of the specific commodity.

C the average market price of a basket of similar commodities.

30 Which of the following statements is *most* accurate?

A Commodity indexes all share similar weighting methods.

B Commodity indexes containing the same underlying commodities offer similar returns.

C The performance of commodity indexes can be quite different from that of the underlying commodities.

31 Which of the following is *not* a real estate index category?

A Appraisal index.

B Initial sales index.

C Repeat sales index.

32 A unique feature of hedge fund indexes is that they:

A are frequently equal weighted.

B are determined by the constituents of the index.

C reflect the value of private rather than public investments.

33 The returns of hedge fund indexes are *most likely*:

A biased upward.

B biased downward.

C similar across different index providers.

34 In comparison to equity indexes, the constituent securities of fixed-income indexes are:

A more liquid.

B easier to price.

C drawn from a larger investment universe.

SOLUTIONS

1. C is correct. A security market index represents the value of a given security market, market segment, or asset class.
2. A is correct. Security market indexes are constructed and managed like a portfolio of securities.
3. A is correct. The first decision is identifying the target market that the index is intended to represent because the target market determines the investment universe and the securities available for inclusion in the index.
4. C is correct. The difference between a price return index and a total return index consisting of identical securities and weights is the income generated over time by the underlying securities. If the securities in the index do not generate income, both indexes will be identical in value.
5. A is correct. At inception, the values of the price return and total return versions of an index are equal.
6. B is correct. The price return is the sum of the weighted returns of each security. The return of Able is 20 percent [(12 – 10)/10]; of Baker is –5 percent [(19 – 20)/20]; and of Charlie is 0 percent [(30 – 30)/30]. The price return index assigns a weight of 1/3 to each asset; therefore, the price return is 1/3 × [20% + (–5%) + 0%] = 5%.
7. C is correct. The total return of an index is calculated on the basis of the change in price of the underlying securities plus the sum of income received or the sum of the weighted total returns of each security. The total return of Able is 27.5 percent; of Baker is 0 percent; and of Charlie is 6.7 percent:

   Able: (12 – 10 + 0.75)/10 = 27.5%

   Baker: (19 – 20 + 1)/20 = 0%

   Charlie: (30 – 30 + 2)/30 = 6.7%

   An equal-weighted index applies the same weight (1/3) to each security's return; therefore, the total return = 1/3 × (27.5% + 0% + 6.7%) = 11.4%.
8. B is correct. The price return of the price-weighted index is the percentage change in price of the index: (68 – 75)/75 = –9.33%.

Security	**Beginning of Period Price (£)**	**End of Period Price (£)**
ABC	25.00	27.00
DEF	35.00	25.00
GHI	15.00	16.00
TOTAL	75.00	68.00

9. B is correct. The price return of the index is (48,250,000 – 53,750,000)/53,750,000 = –10.23%.

Security	**Beginning of Period Price (¥)**	**Shares Outstanding**	**Beginning of Period Value (¥)**	**End of Period Price (¥)**	**End of Period Value (¥)**
MNO	2,500	5,000	12,500,000	2,700	13,500,000
QRS	3,500	7,500	26,250,000	2,500	18,750,000

(continued)

Security	Beginning of Period Price (¥)	Shares Outstanding	Beginning of Period Value (¥)	End of Period Price (¥)	End of Period Value (¥)
XYZ	1,500	10,000	15,000,000	1,600	16,000,000
Total			53,750,000		48,250,000

10 B is correct. The total return of the market-capitalization-weighted index is calculated below:

Security	Beginning of Period Value (¥)	End of Period Value (¥)	Total Dividends (¥)	Total Return (%)
MNO	12,500,000	13,500,000	500,000	12.00
QRS	26,250,000	18,750,000	1,125,000	–24.29
XYZ	15,000,000	16,000,000	1,000,000	13.33
Total	53,750,000	48,250,000	2,625,000	–5.35

11 A is correct. The target market determines the investment universe and the securities available for inclusion in the index.

12 A is correct. The sum of prices at the beginning of the period is 96; the sum at the end of the period is 100. Regardless of the divisor, the price return is 100/96 – 1 = 0.042 or 4.2 percent.

13 B is correct. It is the percentage change in the market value over the period:

Market value at beginning of period: (20 × 300) + (50 × 300) + (26 × 2,000) = 73,000

Market value at end of period: (22 × 300) + (48 × 300) + (30 × 2,000) = 81,000

Percentage change is 81,000/73,000 – 1 = 0.1096 or 11.0 percent with rounding.

14 C is correct. With an equal-weighted index, the same amount is invested in each security. Assuming $1,000 is invested in each of the three stocks, the index value is $3,000 at the beginning of the period and the following number of shares is purchased for each stock:

Security A: 50 shares

Security B: 20 shares

Security C: 38.46 shares.

Using the prices at the beginning of the period for each security, the index value at the end of the period is $3,213.8: ($22 × 50) + ($48 × 20) + ($30 × 38.46). The price return is $3,213.8/$3,000 – 1 = 7.1%.

15 A is correct. In the price weighting method, the divisor must be adjusted so the index value immediately after the split is the same as the index value immediately prior to the split.

16 C is correct. The main source of return differences arises from outperformance of small-cap securities or underperformance of large-cap securities. In an equal-weighted index, securities that constitute the largest fraction of the market are underrepresented and securities that constitute only a small fraction of the market are overrepresented. Thus, higher equal-weighted index returns will occur if the smaller-cap equities outperform the larger-cap equities.

17 C is correct. “Float” is the number of shares available for public trading.

18 B is correct. Fundamental weighting leads to indexes that have a value tilt.

19 C is correct. Rebalancing refers to adjusting the weights of constituent securities in an index to maintain consistency with the index's weighting method.

20 B is correct. Changing market prices will cause weights that were initially equal to become unequal, thus requiring rebalancing.

21 C is correct. Reconstitution is the process by which index providers review the constituent securities, re-apply the initial criteria for inclusion in the index, and select which securities to retain, remove, or add. Constituent securities that no longer meet the criteria are replaced with securities that do. Thus, reconstitution reduces the likelihood that the index includes securities that are not representative of the target market.

22 C is correct. Security market indexes play a critical role as proxies for asset classes in asset allocation models.

23 A is correct. Security market indexes are used as proxies for measuring market or systematic risk, not as measures of systemic risk.

24 B is correct. Sector indexes provide a means to determine whether a portfolio manager is more successful at stock selection or sector allocation.

25 C is correct. Style indexes represent groups of securities classified according to market capitalization, value, growth, or a combination of these characteristics.

26 A is correct. The large number of fixed-income securities—combined with the lack of liquidity of some securities—makes it costly and difficult for investors to replicate fixed-income indexes.

27 C is correct. An aggregate fixed-income index can be subdivided by market sector (government, government agency, collateralized, corporate), style (maturity, credit quality), economic sector, or some other characteristic to create more narrowly defined indexes.

28 C is correct. Coupon frequency is not a dimension on which fixed-income indexes are based.

29 A is correct. Commodity indexes consist of futures contracts on one or more commodities.

30 C is correct. The performance of commodity indexes can be quite different from that of the underlying commodities because the indexes consist of futures contracts on the commodities rather than the actual commodities.

31 B is correct. It is not a real estate index category.

32 B is correct. Hedge funds are not required to report their performance to any party other than their investors. Therefore, each hedge fund decides to which database(s) it will report its performance. Thus, for a hedge fund index, constituents determine the index rather than index providers determining the constituents.

33 A is correct. Voluntary performance reporting may lead to survivorship bias, and poorer performing hedge funds will be less likely to report their performance.

34 C is correct. The fixed-income market has more issuers and securities than the equity market.

READING

35

Market Efficiency

by Sean Cleary, PhD, CFA, Howard J. Atkinson, CIMA, ICD.D, CFA, and Pamela Peterson Drake, PhD, CFA

Sean Cleary, PhD, CFA, is at Queen's University (Canada). Howard J. Atkinson, CIMA, ICD.D, CFA, is at Horizons ETF Management (Canada) Inc. (Canada). Pamela Peterson Drake, PhD, CFA, is at James Madison University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe market efficiency and related concepts, including their importance to investment practitioners;
☐	**b.** contrast market value and intrinsic value;
☐	**c.** explain factors that affect a market's efficiency;
☐	**d.** contrast weak-form, semi-strong-form, and strong-form market efficiency;
☐	**e.** explain the implications of each form of market efficiency for fundamental analysis, technical analysis, and the choice between active and passive portfolio management;
☐	**f.** describe market anomalies;
☐	**g.** describe behavioral finance and its potential relevance to understanding market anomalies.

INTRODUCTION

1

Market efficiency concerns the extent to which market prices incorporate available information. If market prices do not fully incorporate information, then opportunities may exist to make a profit from the gathering and processing of information. The subject of market efficiency is, therefore, of great interest to investment managers, as illustrated in Example 1.

© 2019 CFA Institute. All rights reserved.

EXAMPLE 1

Market Efficiency and Active Manager Selection

The chief investment officer (CIO) of a major university endowment fund has listed eight steps in the active manager selection process that can be applied both to traditional investments (e.g., common equity and fixed-income securities) and to alternative investments (e.g., private equity, hedge funds, and real assets). The first step specified is the evaluation of market opportunity:

> What is the opportunity and why is it there? To answer this question, we start by studying capital markets and the types of managers operating within those markets. We identify market inefficiencies and try to understand their causes, such as regulatory structures or behavioral biases. We can rule out many broad groups of managers and strategies by simply determining that the degree of market inefficiency necessary to support a strategy is implausible. Importantly, we consider the past history of active returns meaningless unless we understand why markets will allow those active returns to continue into the future.[1]

The CIO's description underscores the importance of not assuming that past active returns that might be found in a historical dataset will repeat themselves in the future. **Active returns** refer to returns earned by strategies that do *not* assume that all information is fully reflected in market prices.

Governments and market regulators also care about the extent to which market prices incorporate information. Efficient markets imply informative prices—prices that accurately reflect available information about fundamental values. In market-based economies, market prices help determine which companies (and which projects) obtain capital. If these prices do not efficiently incorporate information about a company's prospects, then it is possible that funds will be misdirected. By contrast, prices that are informative help direct scarce resources and funds available for investment to their highest-valued uses.[2] Informative prices thus promote economic growth. The efficiency of a country's capital markets (in which businesses raise financing) is an important characteristic of a well-functioning financial system.

The remainder of this reading is organized as follows. Sections 2 and 3 provide specifics on how the efficiency of an asset market is described and discuss the factors affecting (i.e., contributing to and impeding) market efficiency. Sections 4 and 5 present an influential three-way classification of the efficiency of security markets and discuss its implications for fundamental analysis, technical analysis, and portfolio management. Sections 6 and 7 present several market anomalies (apparent market inefficiencies that have received enough attention to be individually identified and named) and describe how these anomalies relate to investment strategies. Section 8 introduces behavioral finance and how that field of study relates to market efficiency. A summary concludes the reading.

---

[1] The CIO is Christopher J. Brightman, CFA, of the University of Virginia Investment Management Company, as reported in Yau, Schneeweis, Robinson, and Weiss (2007, pp. 481–482).

[2] This concept is known as allocative efficiency.

THE CONCEPT OF MARKET EFFICIENCY

2

a describe market efficiency and related concepts, including their importance to investment practitioners;

b contrast market value and intrinsic value;

2.1 The Description of Efficient Markets

An **informationally efficient market** (an **efficient market**) is a market in which asset prices reflect new information quickly and rationally. An efficient market is thus a market in which asset prices reflect all past and present information.[3]

In this section we expand on this definition by clarifying the time frame required for an asset's price to incorporate information as well as describing the elements of information releases assumed under market efficiency. We discuss the difference between market value and intrinsic value and illustrate how inefficiencies or discrepancies between these values can provide profitable opportunities for active investment. As financial markets are generally not considered being either completely efficient or inefficient, but rather falling within a range between the two extremes, we describe a number of factors that contribute to and impede the degree of efficiency of a financial market. Finally, we conclude our overview of market efficiency by illustrating how the costs incurred by traders in identifying and exploiting possible market inefficiencies affect how we interpret market efficiency.

Investment managers and analysts, as noted, are interested in market efficiency because the extent to which a market is efficient affects how many profitable trading opportunities (market inefficiencies) exist. Consistent, superior, risk-adjusted returns (net of all expenses) are not achievable in an efficient market.[4] In an efficient market, a **passive investment** strategy (i.e., buying and holding a broad market portfolio) that does not seek superior risk-adjusted returns can be preferred to an **active investment** strategy because of lower costs (for example, transaction and information-seeking costs). By contrast, in a very inefficient market, opportunities may exist for an active investment strategy to achieve superior risk-adjusted returns (net of all expenses in executing the strategy) as compared with a passive investment strategy. In inefficient markets, an active investment strategy may outperform a passive investment strategy on a risk-adjusted basis. Understanding the characteristics of an efficient market and being able to evaluate the efficiency of a particular market are important topics for investment analysts and portfolio managers.

An efficient market is a market in which asset prices reflect information quickly. But what is the time frame of "quickly"? Trades are the mechanism by which information can be incorporated into asset transaction prices. The time needed to execute trades to exploit an inefficiency may provide a baseline for judging speed of adjustment.[5] The time frame for an asset's price to incorporate information must be at least as long as the shortest time a trader needs to execute a transaction in the asset. In certain markets, such as foreign exchange and developed equity markets, market efficiency relative to certain types of information has been studied using time frames as short as one minute or less. If the time frame of price adjustment allows many traders to earn profits with little risk, then the market is relatively inefficient. These considerations lead to the observation that market efficiency can be viewed as falling on a continuum.

---

[3] This definition is convenient for making several instructional points. The definition that most simply explains the sense of the word *efficient* in this context can be found in Fama (1976): "An efficient capital market is a market that is efficient in processing information" (p. 134).

[4] The technical term for *superior* in this context is *positive abnormal* in the sense of higher than expected given the asset's risk (as measured, according to capital market theory, by the asset's contribution to the risk of a well-diversified portfolio).

[5] Although the original theory of market efficiency does not quantify this speed, the basic idea is that it is sufficiently swift to make it impossible to consistently earn abnormal profits. Chordia, Roll, and Subrahmanyam (2005) suggest that the adjustment to information on the New York Stock Exchange (NYSE) is between 5 and 60 minutes.

Finally, an important point is that in an efficient market, prices should be expected to react only to the elements of information releases that are not anticipated fully by investors—that is, to the "unexpected" or "surprise" element of such releases. Investors process the unexpected information and revise expectations (for example, about an asset's future cash flows, risk, or required rate of return) accordingly. The revised expectations enter or get incorporated in the asset price through trades in the asset. Market participants who process the news and believe that at the current market price an asset does not offer sufficient compensation for its perceived risk will tend to sell it or even sell it short. Market participants with opposite views should be buyers. In this way the market establishes the price that balances the various opinions after expectations are revised.

EXAMPLE 2

Price Reaction to the Default on a Bond Issue

Suppose that a speculative-grade bond issuer announces, just before bond markets open, that it will default on an upcoming interest payment. In the announcement, the issuer confirms various reports made in the financial media in the period leading up to the announcement. Prior to the issuer's announcement, the financial news media reported the following: 1) suppliers of the company were making deliveries only for cash payment, reducing the company's liquidity; 2) the issuer's financial condition had probably deteriorated to the point that it lacked the cash to meet an upcoming interest payment; and 3) although public capital markets were closed to the company, it was negotiating with a bank for a private loan that would permit it to meet its interest payment and continue operations for at least nine months. If the issuer defaults on the bond, the consensus opinion of analysts is that bondholders will recover approximately $0.36 to $0.38 per dollar face value.

1 If the market for the bond is efficient, the bond's market price is *most likely* to fully reflect the bond's value after default:
   - **A** in the period leading up to the announcement.
   - **B** in the first trade prices after the market opens on the announcement day.
   - **C** when the issuer actually misses the payment on the interest payment date.

2 If the market for the bond is efficient, the piece of information that bond investors *most likely* focus on in the issuer's announcement is that the issuer had:
   - **A** failed in its negotiations for a bank loan.
   - **B** lacked the cash to meet the upcoming interest payment.
   - **C** been required to make cash payments for supplier deliveries.

Solution to 1:

B is correct. The announcement removed any uncertainty about default. In the period leading up to the announcement, the bond's market price incorporated a probability of default, but the price would not have fully reflected the bond's value after default. The possibility that a bank loan might permit the company to avoid default was not eliminated until the announcement.

Solution to 2:

A is correct. The failure of the loan negotiations first becomes known in this announcement. The failure implies default.

2.2 Market Value versus Intrinsic Value

Market value is the price at which an asset can currently be bought or sold. **Intrinsic value** (sometimes called **fundamental value**) is, broadly speaking, the value that would be placed on it by investors if they had a complete understanding of the asset's investment characteristics.[6] For a bond, for example, such information would include its interest (coupon) rate, principal value, the timing of its interest and principal payments, the other terms of the bond contract (indenture), a precise understanding of its default risk, the liquidity of its market, and other issue-specific items. In addition, market variables such as the term structure of interest rates and the size of various market premiums applying to the issue (for default risk, etc.) would enter into a discounted cash flow estimate of the bond's intrinsic value (discounted cash flow models are often used for such estimates). The word *estimate* is used because in practice, intrinsic value can be estimated but is not known for certain.

If investors believe a market is highly *efficient*, they will usually accept market prices as accurately reflecting intrinsic values. Discrepancies between market price and intrinsic value are the basis for profitable active investment. Active investors seek to own assets selling below perceived intrinsic value in the marketplace and to sell or sell short assets selling above perceived intrinsic value.

If investors believe an asset market is relatively *inefficient*, they may try to develop an independent estimate of intrinsic value. The challenge for investors and analysts is estimating an asset's intrinsic value. Numerous theories and models, including the dividend discount model, can be used to estimate an asset's intrinsic value, but they all require some form of judgment regarding the size, timing, and riskiness of the future cash flows associated with the asset. The more complex an asset's future cash flows, the more difficult it is to estimate its intrinsic value. These complexities and the estimates of an asset's market value are reflected in the market through the buying and selling of assets. The market value of an asset represents the intersection of supply and demand—the point that is low enough to induce at least one investor to buy while being high enough to induce at least one investor to sell. Because information relevant to valuation flows continually to investors, estimates of intrinsic value change, and hence, market values change.

EXAMPLE 3

Intrinsic Value

1 An analyst estimates that a security's intrinsic value is lower than its market value. The security appears to be:

---

[6] Intrinsic value is often defined as the present value of all expected future cash flows of the asset.

  - **A** undervalued.
  - **B** fairly valued.
  - **C** overvalued.

2 A market in which assets' market values are, on average, equal to or nearly equal to intrinsic values is *best described* as a market that is attractive for:

  - **A** active investment.
  - **B** passive investment.
  - **C** both active and passive investment.

3 Suppose that the future cash flows of an asset are accurately estimated. The asset trades in a market that you believe is efficient based on most evidence, but your estimate of the asset's intrinsic value exceeds the asset's market value by a moderate amount. The *most likely* conclusion is that you have:

  - **A** overestimated the asset's risk.
  - **B** underestimated the asset's risk.
  - **C** identified a market inefficiency.

Solution to 1:

C is correct. The market is valuing the asset at more than its true worth.

Solution to 2:

B is correct because an active investment is not expected to earn superior risk-adjusted returns if the market is efficient. The additional costs of active investment are not justified in such a market.

Solution to 3:

B is correct. If risk is underestimated, the discount rate being applied to find the present value of the expected cash flows (estimated intrinsic value) will be too low and the intrinsic value estimate will be too high.

3

FACTORS AFFECTING MARKET EFFICIENCY INCLUDING TRADING COSTS

a describe market efficiency and related concepts, including their importance to investment practitioners;

c explain factors that affect a market's efficiency;

For markets to be efficient, prices should adjust quickly and rationally to the release of new information. In other words, prices of assets in an efficient market should "fully reflect" all information. Financial markets, however, are generally not classified at the two extremes as either completely inefficient or completely efficient but, rather, as exhibiting various degrees of efficiency. In other words, market efficiency should be viewed as falling on a continuum between extremes of completely efficient, at one end, and completely inefficient, at the other. Asset prices in a highly efficient market, by definition, reflect information more quickly and more accurately than in a less-efficient market. These degrees of efficiency also vary through time, across geographical markets, and by type of market. A number of factors contribute to and impede the degree of efficiency in a financial market.

3.1 Market Participants

One of the most critical factors contributing to the degree of efficiency in a market is the number of market participants. Consider the following example that illustrates the relationship between the number of market participants and market efficiency.

EXAMPLE 4

Illustration of Market Efficiency

Assume that the shares of a small market capitalization (cap) company trade on a public stock exchange. Because of its size, it is not considered "blue-chip" and not many professional investors follow the activities of the company.[7] A small-cap fund analyst reports that the most recent annual operating performance of the company has been surprisingly good, considering the recent slump in its industry. The company's share price, however, has been slow to react to the positive financial results because the company is not being recommended by the majority of research analysts. This mispricing implies that the market for this company's shares is less than fully efficient. The small-cap fund analyst recognizes the opportunity and immediately recommends the purchase of the company's shares. The share price gradually increases as more investors purchase the shares once the news of the mispricing spreads through the market. As a result, it takes a few days for the share price to fully reflect the information.

Six months later, the company reports another solid set of interim financial results. But because the previous mispricing and subsequent profit opportunities have become known in the market, the number of analysts following the company's shares has increased substantially. As a result, as soon as unexpected information about the positive interim results are released to the public, a large number of buy orders quickly drive up the stock price, thereby making the market for these shares more efficient than before.

A large number of investors (individual and institutional) follow the major financial markets closely on a daily basis, and if mispricings exist in these markets, as illustrated by the example, investors will act so that these mispricings disappear quickly. Besides the number of investors, the number of financial analysts who follow or analyze a security or asset should be positively related to market efficiency. The number of market participants and resulting trading activity can vary significantly through time. A lack of trading activity can cause or accentuate other market imperfections that impede market efficiency. In fact, in many of these markets, trading in many of the listed stocks is restricted for foreigners. By nature, this limitation reduces the number of market participants, restricts the potential for trading activity, and hence reduces market efficiency.

---

[7] A "blue-chip" share is one from a well-recognized company that is considered to be high quality but low risk. This term generally refers to a company that has a long history of earnings and paying dividends.

EXAMPLE 5

Factors Affecting Market Efficiency

The expected effect on market efficiency of opening a securities market to trading by foreigners would *most likely* be to:

A decrease market efficiency.

B leave market efficiency unchanged.

C increase market efficiency.

Solution:

C is correct. The opening of markets as described should increase market efficiency by increasing the number of market participants.

3.2 Information Availability and Financial Disclosure

Information availability (e.g., an active financial news media) and financial disclosure should promote market efficiency. Information regarding trading activity and traded companies in such markets as the New York Stock Exchange, the London Stock Exchange, and the Tokyo Stock Exchange is readily available. Many investors and analysts participate in these markets, and analyst coverage of listed companies is typically substantial. As a result, these markets are quite efficient. In contrast, trading activity and material information availability may be lacking in smaller securities markets, such as those operating in some emerging markets.

Similarly, significant differences may exist in the efficiency of different types of markets. For example, many securities trade primarily or exclusively in dealer or over-the-counter (OTC) markets, including bonds, money market instruments, currencies, mortgage-backed securities, swaps, and forward contracts. The information provided by the dealers that serve as market makers for these markets can vary significantly in quality and quantity, both through time and across different product markets.

Treating all market participants fairly is critical for the integrity of the market and explains why regulators place such an emphasis on "fair, orderly, and efficient markets."[8] A key element of this fairness is that all investors have access to the information necessary to value securities that trade in the market. Rules and regulations that promote fairness and efficiency in a market include those pertaining to the disclosure of information and illegal insider trading.

For example, US Securities and Exchange Commission's (SEC's) Regulation FD (Fair Disclosure) requires that if security issuers provide nonpublic information to some market professionals or investors, they must also disclose this information to the public.[9] This requirement helps provide equal and fair opportunities, which is important in encouraging participation in the market. A related issue deals with illegal insider trading. The SEC's rules, along with court cases, define illegal insider trading as trading in securities by market participants who are considered insiders "while in possession of material, nonpublic information about the security."[10] Although these rules cannot guarantee that some participants will not have an advantage over others and that insiders will not trade on the basis of inside information, the civil and criminal penalties associated with breaking these rules are intended to discourage illegal insider trading and promote fairness. In the European Union, insider trading laws are generally enshrined in legislation and enforced by regulatory and judicial authorities.[11]

---

[8] "The Investor's Advocate: How the SEC Protects Investors, Maintains Market Integrity, and Facilitates Capital Formation," US Securities and Exchange Commission (www.sec.gov/about/whatwedo.shtml).

[9] Regulation FD, "Selective Disclosure and Insider Trading," 17 CFR Parts 240, 243, and 249, effective 23 October 2000.

3.3 Limits to Trading

Arbitrage is a set of transactions that produces riskless profits. Arbitrageurs are traders who engage in such trades to benefit from pricing discrepancies (inefficiencies) in markets. Such trading activity contributes to market efficiency. For example, if an asset is traded in two markets but at different prices, the actions of buying the asset in the market in which it is underpriced and selling the asset in the market in which it is overpriced will eventually bring these two prices together. The presence of these arbitrageurs helps pricing discrepancies disappear quickly. Obviously, market efficiency is impeded by any limitation on arbitrage resulting from operating inefficiencies, such as difficulties in executing trades in a timely manner, prohibitively high trading costs, and a lack of transparency in market prices.

Some market experts argue that restrictions on short selling limit arbitrage trading, which impedes market efficiency. **Short selling** is the transaction whereby an investor sells shares that he or she does not own by borrowing them from a broker and agreeing to replace them at a future date. Short selling allows investors to sell securities they believe to be overvalued, much in the same way they can buy those they believe to be undervalued. In theory, such activities promote more efficient pricing. Regulators and others, however, have argued that short selling may exaggerate downward market movements, leading to crashes in affected securities. In contrast, some researchers report evidence indicating that when investors are unable to borrow securities, that is to short the security, or when costs to borrow shares are high, market prices may deviate from intrinsic values.[12] Furthermore, research suggests that short selling is helpful in price discovery (that is, it facilitates supply and demand in determining prices).[13]

3.4 Transaction Costs and Information-Acquisition Costs

The costs incurred by traders in identifying and exploiting possible market inefficiencies affect the interpretation of market efficiency. The two types of costs to consider are transaction costs and information-acquisition costs.

- *Transaction costs*: Practically, transaction costs are incurred in trading to exploit any perceived market inefficiency. Thus, "efficient" should be viewed as efficient within the bounds of transaction costs. For example, consider a violation of the principle that two identical assets should sell for the same price in different markets. Such a violation can be considered to be a rather simple possible exception to market efficiency because prices appear to be inconsistently processing information. To exploit the violation, a trader could arbitrage by simultaneously shorting the asset in the higher-price market and buying the asset in the lower-price market. If the price discrepancy between the two markets is smaller than the transaction costs involved in the arbitrage for the lowest cost traders, the arbitrage will not occur, and both prices are in effect efficient within the bounds of arbitrage. These bounds of arbitrage are relatively narrow in highly liquid markets, such as the market for US Treasury bills, but could be wide in illiquid markets.

---

[10] Although not the focus of this particular reading, it is important to note that a party is considered an insider not only when the individual is a corporate insider, such as an officer or director, but also when the individual is aware that the information is nonpublic information [Securities and Exchange Commission, Rules 10b5-1 ("Trading on the Basis of Material Nonpublic Information in Insider Trading Cases") and Rule 10b5-2 "Duties of Trust or Confidence in Misappropriation Insider Trading Cases")].

[11] See the European Union's Market Abuse Regulation (Regulation (EU) no. 596/2014 of the European Parliament and of the Council of 16 April 2014 on market abuse) and Directive for Criminal Sanctions for Market Abuse (Directive 2014/57/EU of the European Parliament and of the Council of 16 April 2014 on criminal sanctions for market abuse).

[12] See Deng, Mortal, and Gupta (2017) and references therein."

[13] See Bris, Goetzmann, and Zhu (2009).

- *Information-acquisition costs:* Practically, expenses are always associated with gathering and analyzing information. New information is incorporated in transaction prices by traders placing trades based on their analysis of information. Active investors who place trades based on information they have gathered and analyzed play a key role in market prices adjusting to reflect new information. The classic view of market efficiency is that active investors incur information acquisition costs but that money is wasted because prices already reflect all relevant information. This view of efficiency is very strict in the sense of viewing a market as inefficient if active investing can recapture any part of the costs, such as research costs and active asset selection. Grossman and Stiglitz (1980) argue that prices must offer a return to information acquisition; in equilibrium, if markets are efficient, returns net of such expenses are just fair returns for the risk incurred. The modern perspective views a market as inefficient if, after deducting such costs, active investing can earn superior returns. Gross of expenses, a return should accrue to information acquisition in an efficient market.

In summary, a modern perspective calls for the investor to consider transaction costs and information-acquisition costs when evaluating the efficiency of a market. A price discrepancy must be sufficiently large to leave the investor with a profit (adjusted for risk) after taking account of the transaction costs and information-acquisition costs to reach the conclusion that the discrepancy may represent a market inefficiency. Prices may somewhat less than fully reflect available information without there being a true market opportunity for active investors.

4

FORMS OF MARKET EFFICIENCY

d contrast weak-form, semi-strong-form, and strong-form market efficiency;

Eugene Fama developed a framework for describing the degree to which markets are efficient.[14] In his efficient market hypothesis, markets are efficient when prices reflect *all* relevant information at any point in time. This means that the market prices observed for securities, for example, reflect the information available at the time.

In his framework, Fama defines three forms of efficiency: weak, semi-strong, and strong. Each form is defined with respect to the available information that is reflected in prices.

---

[14] Fama (1970).

Forms of Market Efficiency	**Market Prices Reflect:** **Past Market Data**	 **Public Information**	 **Private Information**
Weak form of market efficiency	✓		
Semi-strong form of market efficiency	✓	✓	
Strong form of market efficiency	✓	✓	✓

A finding that investors can consistently earn **abnormal returns** by trading on the basis of information is evidence contrary to market efficiency. In general, abnormal returns are returns in excess of those expected given a security's risk and the market's return. In other words, abnormal return equals actual return less expected return.

4.1 Weak Form

In the **weak-form efficient market hypothesis**, security prices fully reflect *all past market data*, which refers to all historical price and trading volume information. If markets are weak-form efficient, past trading data are already reflected in current prices and investors cannot predict future price changes by extrapolating prices or patterns of prices from the past.[15]

Tests of whether securities markets are weak-form efficient require looking at patterns of prices. One approach is to see whether there is any serial correlation in security returns, which would imply a predictable pattern.[16] Although there is some weak correlation in daily security returns, there is not enough correlation to make this a profitable trading rule after considering transaction costs.

An alternative approach to test weak-form efficiency is to examine specific trading rules that attempt to exploit historical trading data. If any such trading rule consistently generates abnormal risk-adjusted returns after trading costs, this evidence will contradict weak-form efficiency. This approach is commonly associated with **technical analysis**, which involves the analysis of historical trading information (primarily pricing and volume data) in an attempt to identify recurring patterns in the trading data that can be used to guide investment decisions. Many technical analysts, also referred to as "technicians," argue that many movements in stock prices are based, in large part, on psychology. Many technicians attempt to predict how market participants will behave, based on analyses of past behavior, and then trade on those predictions. Technicians often argue that simple statistical tests of trading rules are not conclusive because they are not applied to the more sophisticated trading strategies that can be used and that the research excludes the technician's subjective judgment. Thus, it is difficult to definitively refute this assertion because there are an unlimited number of possible technical trading rules.

Can technical analysts profit from trading on past trends? Overall, the evidence indicates that investors cannot consistently earn abnormal profits using past prices or other technical analysis strategies in developed markets.[17] Some evidence suggests, however, that there are opportunities to profit on technical analysis in countries with developing markets, including Hungary, Bangladesh, and Turkey, among others.[18]

---

[15] Market efficiency should not be confused with the random walk hypothesis, in which price changes over time are independent of one another. A random walk model is one of many alternative expected return generating models. Market efficiency does not require that returns follow a random walk.

[16] Serial correlation is a statistical measure of the degree to which the returns in one period are related to the returns in another period.

[17] Bessembinder and Chan (1998) and Fifield, Power, and Sinclair (2005).

[18] Fifield, Power, and Sinclair (2005), Chen and Li (2006), and Mobarek, Mollah, and Bhuyan (2008).

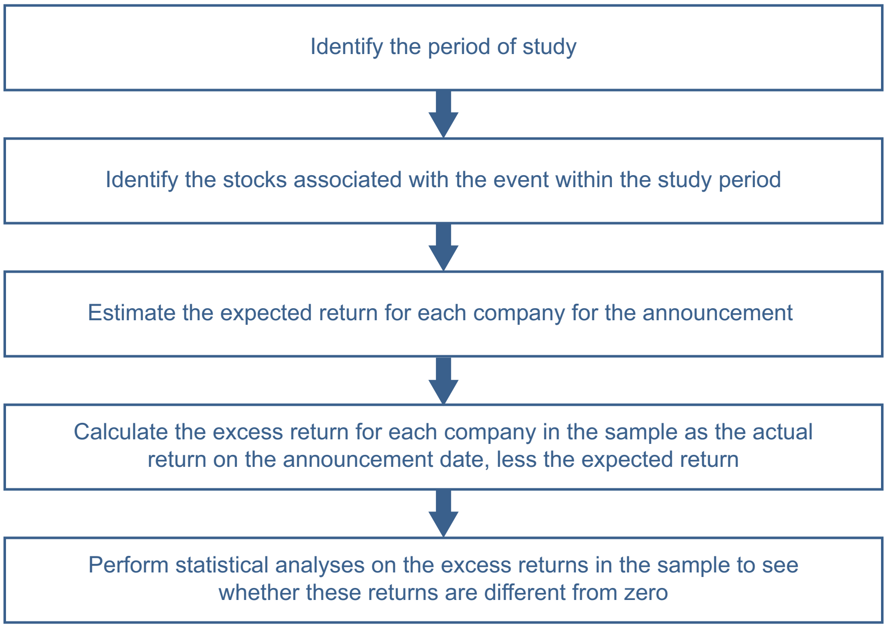
4.2 Semi-Strong Form

In a **semi-strong-form efficient market**, prices reflect all publicly known and available information. Publicly available information includes financial statement data (such as earnings, dividends, corporate investments, changes in management, etc.) and financial market data (such as closing prices, shares traded, etc.). Therefore, the semi-strong form of market efficiency encompasses the weak form. In other words, if a market is semi-strong efficient, then it must also be weak-form efficient. A market that quickly incorporates all publicly available information into its prices is semi-strong efficient.

In a semi-strong market, efforts to analyze publicly available information are futile. That is, analyzing earnings announcements of companies to identify underpriced or overpriced securities is pointless because the prices of these securities already reflect all publicly available information. If markets are semi-strong efficient, no single investor has access to information that is not already available to other market participants, and as a consequence, no single investor can gain an advantage in predicting future security prices. In a semi-strong efficient market, prices adjust quickly and accurately to new information. Suppose a company announces earnings that are higher than expected. In a semi-strong efficient market, investors would not be able to act on this announcement and earn abnormal returns.

A common empirical test of investors' reaction to information releases is the event study. Suppose a researcher wants to test whether investors react to the announcement that the company is paying a special dividend. The researcher identifies a sample period and then those companies that paid a special dividend in the period and the date of the announcement. Then, for each company's stock, the researcher calculates the expected return on the share for the event date. This expected return may be based on many different models, including the capital asset pricing model, a simple market model, or a market index return. The researcher calculates the excess return as the difference between the actual return and the expected return. Once the researcher has calculated the event's excess return for each share, statistical tests are conducted to see whether the abnormal returns are statistically different from zero. The process of an event study is outlined in Exhibit 1.

Exhibit 1 The Event Study Process


Identify the period of study
Identify the stocks associated with the event within the study period
Estimate the expected return for each company for the announcement
Calculate the excess return for each company in the sample as the actual return on the announcement date, less the expected return
Perform statistical analyses on the excess returns in the sample to see whether these returns are different from zero

How do event studies relate to efficient markets? In a semi-strong efficient market, share prices react quickly and accurately to public information. Therefore, if the information is good news, such as better-than-expected earnings, one would expect the company's shares to increase immediately at the time of the announcement; if it is bad news, one would expect a swift, negative reaction. If actual returns exceed what is expected in absence of the announcement and these returns are confined to the announcement period, then they are consistent with the idea that market prices react quickly to new information. In other words, the finding of excess returns at the time of the announcement does not necessarily indicate market inefficiency. In contrast, the finding of consistent excess returns following the announcement would suggest a trading opportunity. Trading on the basis of the announcement—that is, once the announcement is made—would not, on average, yield abnormal returns.

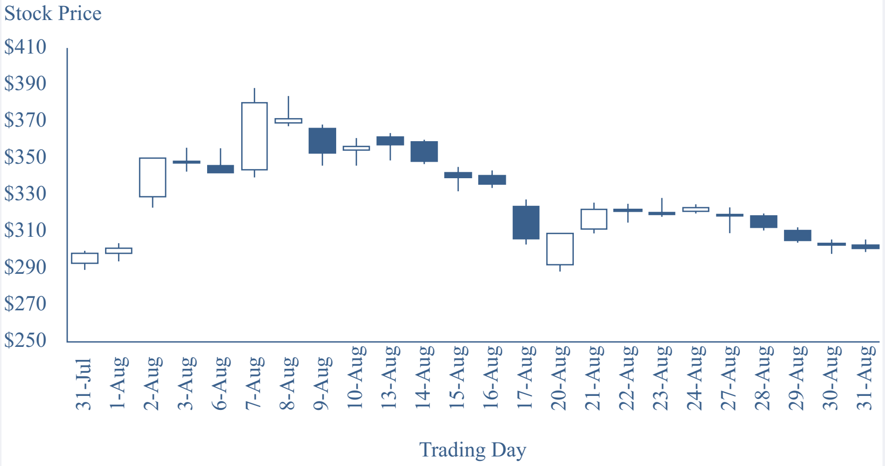
EXAMPLE 6

Information Arrival and Market Reaction

Consider an example of a news item and its effect on a share's price. The following events related to Tesla, Inc. in August of 2018:

1 August 2018	After the market closes, Tesla, Inc., publicly reports that there was a smaller-than expected cash burn for the most recent quarter.
2 August 2018	Elon Musk, Chairman and CEO of Tesla, Inc., notifies Tesla's board of directors that he wants to take the company private. This is not public information at this point.
7 August 2018	Before the market opens, the *Financial Times* reports that a Saudi fund has a $2 billion investment in Tesla. During market trading, Musk announces on Twitter "Am considering taking Tesla private at $420. Funding secured." [Twitter, Elon Musk @elonmusk, 9:48 a.m., 7 August 2018]
24 August 2018	After the market closed, Musk announces that he no longer intends on taking Tesla private.

Exhibit 2 Price of Tesla, Inc. Stock: 31 July 2018–31 August 2018



Note: Open-High-Low-Close graph of Tesla's stock price, with white rectangles indicating upward movement in the day and black rectangles indicating downward movement during the day.
Source of data: Yahoo! Finance.

Is the fact that the price of Tesla moves up immediately on the day after the Q2 earnings (the first day of trading with this information) indicative of efficiency regarding information? Most likely.

Does the fact that the price of Tesla moves up but does not reach $420 on the day the going-private Twitter announcement is made mean that investors underreacted? Not necessarily. There was confusion and uncertainty about the going-private transaction at the time, so the price did not close in on the proposed $420 per share for going private.

Does the fact that the market price of the stock declined well before the issue of going-private was laid to rest by Musk mean that the market is inefficient? Not necessarily. There were numerous analyses, discussions, and other news regarding the likelihood of the transaction, all of which was incorporated in the price of the stock before the going-private transaction was dismissed by Musk.

Researchers have examined many different company-specific information events, including stock splits, dividend changes, and merger announcements, as well as economy-wide events, such as regulation changes and tax rate changes. The results of most research are consistent with the view that developed securities markets might be semi-strong efficient. But some evidence suggests that the markets in developing countries may not be semi-strong efficient.[19]

4.3 Strong Form

In a **strong-form efficient market**, security prices fully reflect both public and private information. A market that is strong-form efficient is, by definition, also semi-strong- and weak-form efficient. In the case of a strong-form efficient market, insiders would not be able to earn abnormal returns from trading on the basis of private information. A strong-form efficient market also means that prices reflect all private information, which means that prices reflect everything that the management of a company knows about the financial condition of the company that has not been publicly released. However, this is not likely because of the strong prohibitions against insider trading that are found in most countries. If a market is strong-form efficient, those with insider information cannot earn abnormal returns.

---

[19] See Gan, Lee, Hwa, and Zhang (2005) and Raja, Sudhahar, and Selvam (2009).

Researchers test whether a market is strong-form efficient by testing whether investors can earn abnormal profits by trading on nonpublic information. The results of these tests are consistent with the view that securities markets are not strong-form efficient; many studies have found that abnormal profits can be earned when nonpublic information is used.[20]

IMPLICATIONS OF THE EFFICIENT MARKET HYPOTHESIS

5

e explain the implications of each form of market efficiency for fundamental analysis, technical analysis, and the choice between active and passive portfolio management;

The implications of efficient markets to investment managers and analysts are important because they affect the value of securities and how these securities are managed. Several implications can be drawn from the evidence on efficient markets for developed markets:

- Securities markets are weak-form efficient, and therefore, investors cannot earn abnormal returns by trading on the basis of past trends in price.
- Securities markets are semi-strong efficient, and therefore, analysts who collect and analyze information must consider whether that information is already reflected in security prices and how any new information affects a security's value.
- Securities markets are not strong-form efficient because securities laws are intended to prevent exploitation of private information.

5.1 Fundamental Analysis

Fundamental analysis is the examination of publicly available information and the formulation of forecasts to estimate the intrinsic value of assets. Fundamental analysis involves the estimation of an asset's value using company data, such as earnings and sales forecasts, and risk estimates as well as industry and economic data, such as economic growth, inflation, and interest rates. Buy and sell decisions depend on whether the current market price is less than or greater than the estimated intrinsic value.

The semi-strong form of market efficiency says that all available public information is reflected in current prices. So, what good is fundamental analysis? Fundamental analysis is necessary in a well-functioning market because this analysis helps the market participants understand the value implications of information. In other words, fundamental analysis facilitates a semi-strong efficient market by disseminating value-relevant information. And, although fundamental analysis requires costly information, this analysis can be profitable in terms of generating abnormal returns if the analyst creates a comparative advantage with respect to this information.[21]

---

[20] Evidence that finds that markets are not strong-form efficient include Jaffe (1974) and Rozeff and Zaman (1988).

5.2 Technical Analysis

Investors using **technical analysis** attempt to profit by looking at patterns of prices and trading volume. Although some price patterns persist, exploiting these patterns may be too costly and, hence, would not produce abnormal returns.

Consider a situation in which a pattern of prices exists. With so many investors examining prices, this pattern will be detected. If profitable, exploiting this pattern will eventually affect prices such that this pattern will no longer exist; it will be arbitraged away. In other words, by detecting and exploiting patterns in prices, technical analysts assist markets in maintaining weak-form efficiency. Does this mean that technical analysts cannot earn abnormal profits? Not necessarily, because there may be a possibility of earning abnormal profits from a pricing inefficiency. But would it be possible to earn abnormal returns on a consistent basis from exploiting such a pattern? No, because the actions of market participants will arbitrage this opportunity quickly, and the inefficiency will no longer exist.

5.3 Portfolio Management

If securities markets are weak-form and semi-strong-form efficient, the implication is that active trading, whether attempting to exploit price patterns or public information, is not likely to generate abnormal returns. In other words, portfolio managers cannot beat the market on a consistent basis, so therefore, passive portfolio management should outperform active portfolio management. Researchers have observed that mutual funds do not, on average, outperform the market on a risk-adjusted basis.[22] Mutual funds perform, on average, similar to the market before considering fees and expenses and perform worse than the market, on average, once fees and expenses are considered. Even if a mutual fund is not actively managed, there are costs to managing these funds, which reduces net returns.

So, what good are portfolio managers? The role of a portfolio manager is not necessarily to beat the market but, rather, to establish and manage a portfolio consistent with the portfolio's objectives, with appropriate diversification and asset allocation, while taking into consideration the risk preferences and tax situation of the investor.

6

MARKET PRICING ANOMALIES - TIME SERIES AND CROSS-SECTIONAL

f describe market anomalies;

Although considerable evidence shows that markets are efficient, researchers have identified a number of apparent market inefficiencies or anomalies. These market anomalies, if persistent, are exceptions to the notion of market efficiency. Researchers conclude that a **market anomaly** may be present if a change in the price of an asset or security cannot directly be linked to current relevant information known in the market or to the release of new information into the market.

---

[21] Brealey (1983).

[22] See Malkiel (1995). One of the challenges to evaluating mutual fund performance is that the researcher must control for survivorship bias.

The validity of any evidence supporting the potential existence of a market inefficiency or anomaly must be *consistent* over reasonably long periods. Otherwise, a detected market anomaly may largely be an artifact of the sample period chosen. In the widespread search for discovering profitable anomalies, many findings could simply be the product of a process called **data mining**, also known as **data snooping**. In generally accepted research practice, an initial hypothesis is developed which is based on economic rationale. Tests are then conducted on objectively selected data to either confirm or reject the original hypothesis. However, with data mining the process is typically reversed: data are examined with the intent to develop a hypothesis, instead of developing a hypothesis first. This is done by analyzing data in various manners, and even utilizing different empirical approaches until you find support for a desired result, in this case a profitable anomaly.

Can researchers look back on data and find a trading strategy that would have yielded abnormal returns? Absolutely. Enough data snooping often can detect a trading strategy that would have worked in the past by chance alone. But in an efficient market, such a strategy is unlikely to generate abnormal returns on a consistent basis in the future. Also, although identified anomalies may appear to produce excess returns, it is generally difficult to profitably exploit the anomalies after accounting for risk, trading costs, and so on.

Several well-known anomalies are listed in Exhibit 3. This list is by no means exhaustive, but it provides information on the breadth of the anomalies. A few of these anomalies are discussed in more detail in the following sections. The anomalies are placed into categories based on the research method that identified the anomaly. Time-series anomalies were identified using time series of data. Cross-sectional anomalies were identified based on analyzing a cross section of companies that differ on some key characteristics. Other anomalies were identified by a variety of means, including event studies.

Exhibit 3 Sampling of Observed Pricing Anomalies

Time Series	Cross-Sectional	Other
January effect	Size effect	Closed-end fund discount
Day-of-the-week effect	Value effect	Earnings surprise
Weekend effect	Book-to-market ratios	Initial public offerings
Turn-of-the-month effect	P/E ratio effect	Distressed securities effect
Holiday effect	Value Line enigma	Stock splits
Time-of-day effect		Super Bowl
Momentum		
Overreaction		

6.1 Time-Series Anomalies

Two of the major categories of time-series anomalies that have been documented are 1) calendar anomalies and 2) momentum and overreaction anomalies.

6.1.1 *Calendar Anomalies*

In the 1980s, a number of researchers reported that stock market returns in January were significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January. Since its first documentation in the 1980s, this pattern, known as the **January effect**, has been observed in most equity markets around the world. This anomaly is also known as the **turn-of-the-year effect**, or even often referred to as the "small firm in January effect" because it is most frequently observed for the returns of small market capitalization stocks.[23]

The January effect contradicts the efficient market hypothesis because excess returns in January are not attributed to any new and relevant information or news. A number of reasons have been suggested for this anomaly, including tax-loss selling. Researchers have speculated that, in order to reduce their tax liabilities, investors sell their "loser" securities in December for the purpose of creating capital losses, which can then be used to offset any capital gains. A related explanation is that these losers tend to be small-cap stocks with high volatility.[24] This increased supply of equities in December depresses their prices, and then these shares are bought in early January at relatively attractive prices. This demand then drives their prices up again. Overall, the evidence indicates that tax-loss selling may account for a portion of January abnormal returns, but it does not explain all of it.

Another possible explanation for the anomaly is so-called "window dressing", a practice in which portfolio managers sell their riskier securities prior to 31 December. The explanation is as follows: many portfolio managers prepare the annual reports of their portfolio holdings as of 31 December. Selling riskier securities is an attempt to make their portfolios appear less risky. After 31 December, a portfolio manager would then simply purchase riskier securities in an attempt to earn higher returns. However, similar to the tax-loss selling hypothesis, the research evidence in support of the window dressing hypothesis explains some, but not all, of the anomaly.

Recent evidence for both stock and bond returns suggests that the January effect is not persistent and, therefore, is not a pricing anomaly. Once an appropriate adjustment for risk is made, the January "effect" does not produce abnormal returns.[25]

Several other calendar effects, including the day-of-the-week and the weekend effects,[26] have been found. These anomalies are summarized in Exhibit 4.[27] But like the size effect, which will be described later, most of these anomalies have been eliminated over time. One view is that the anomalies have been exploited such that the effect has been arbitraged away. Another view, however, is that increasingly sophisticated statistical methodologies fail to detect pricing inefficiencies.

---

[23] There is also evidence of a January effect in bond returns that is more prevalent in high-yield corporate bonds, similar to the small-company effect for stocks.

[24] See Roll (1983).

[25] See, for example, Kim (2006).

[26] For a discussion of several of these anomalous patterns, see Jacobs and Levy (1988).

[27] The weekend effect consists of a pattern of returns around the weekend: abnormal positive returns on Fridays followed by abnormally negative returns on Mondays. This is a day-of-the-week effect that specifically links Friday and Monday returns. It is interesting to note that in 2009, the weekend effect in the United States was inverted, with 80 percent of the gains from March 2009 onward coming from the first trading day of the week.

Exhibit 4 Calendar-Based Anomalies

Anomaly	Observation
Turn-of-the-month effect	Returns tend to be higher on the last trading day of the month and the first three trading days of the next month.
Day-of-the-week effect	The average Monday return is negative and lower than the average returns for the other four days, which are all positive.
Weekend effect	Returns on weekends tend to be lower than returns on weekdays.
Holiday effect	Returns on stocks in the day prior to market holidays tend to be higher than other days.

6.1.2 *Momentum and Overreaction Anomalies*

Momentum anomalies relate to short-term share price patterns. One of the earliest studies to identify this type of anomaly was conducted by Werner DeBondt and Richard Thaler, who argued that investors overreact to the release of unexpected public information.[28] Therefore, stock prices will be inflated (depressed) for those companies releasing good (bad) information. This anomaly has become known as the overreaction effect. Using the overreaction effect, they proposed a strategy that involved buying "loser" portfolios and selling "winner" portfolios. They defined stocks as winners or losers based on their total returns over the previous three- to five-year period. They found that in a subsequent period, the loser portfolios outperformed the market, while the winner portfolios underperformed the market. Similar patterns have been documented in many, but not all, global stock markets as well as in bond markets. One criticism is that the observed anomaly may be the result of statistical problems in the analysis.

A contradiction to weak-form efficiency occurs when securities that have experienced high returns in the short term tend to continue to generate higher returns in subsequent periods.[29] Empirical support for the existence of momentum in stock returns in most stock markets around the world is well documented. If investors can trade on the basis of momentum and earn abnormal profits, then this anomaly contradicts the weak form of the efficient market hypothesis because it represents a pattern in prices that can be exploited by simply using historical price information.[30]

Researchers have argued that the existence of momentum is rational and not contrary to market efficiency because it is plausible that there are shocks to the expected growth rates of cash flows to shareholders and that these shocks induce a serial correlation that is rational and short lived.[31] In other words, having stocks with some degree of momentum in their security returns may not imply irrationality but, rather, may reflect prices adjusting to a shock in growth rates.

---

[28] DeBondt and Thaler (1985).

[29] Notice that this pattern lies in sharp contrast to DeBondt and Thaler's reversal pattern that is displayed over longer periods of time. In theory, the two patterns could be related. In other words, it is feasible that prices are bid up extremely high, perhaps too high, in the short term for companies that are doing well. In the longer term (three-to-five years), the prices of these short-term winners correct themselves and they do poorly.

[30] Jegadeesh and Titman (2001).

[31] Johnson (2002).

6.2 Cross-Sectional Anomalies

Two of the most researched cross-sectional anomalies in financial markets are the size effect and the value effect.

6.2.1 *Size Effect*

The size effect results from the observation that equities of small-cap companies tend to outperform equities of large-cap companies on a risk-adjusted basis. Many researchers documented a small-company effect soon after the initial research was published in 1981. This effect, however, was not apparent in subsequent studies.[32] Part of the reason that the size effect was not confirmed by subsequent studies may be because of the fact that if it were truly an anomaly, investors acting on this effect would reduce any potential returns. But some of the explanation may simply be that the effect as originally observed was a chance outcome and, therefore, not actually an inefficiency.

6.2.2 *Value Effect*

A number of global empirical studies have shown that value stocks, which are generally referred to as stocks that have below-average price-to-earnings (P/E) and market-to-book (M/B) ratios, and above-average dividend yields, have consistently outperformed growth stocks over long periods of time.[33] If the effect persists, the value stock anomaly contradicts semi-strong market efficiency because all the information used to categorize stocks in this manner is publicly available.

Fama and French developed a three-factor model to predict stock returns.[34] In addition to the use of market returns as specified by the capital asset pricing model (CAPM), the Fama and French model also includes the size of the company as measured by the market value of its equity and the company's book value of equity divided by its market value of equity, which is a value measure. The Fama and French model captures risk dimensions related to stock returns that the CAPM model does not consider. Fama and French find that when they apply the three-factor model instead of the CAPM, the value stock anomaly disappears.

7

OTHER ANOMALIES, IMPLICATIONS OF MARKET PRICING ANOMALIES

f describe market anomalies;

A number of additional anomalies has been documented in the financial markets, including the existence of closed-end investment fund discounts, price reactions to the release of earnings information, returns of initial public offerings, and the predictability of returns based on prior information.

---

[32] Although a large number of studies documents a small-company effect, these studies are concentrated in a period similar to that of the original research and, therefore, use a similar data set. The key to whether something is a true anomaly is persistence in out-of-sample tests. Fama and French (2008) document that the size effect is apparent only in microcap stocks but not in small- and large-cap stocks and these microcap stocks may have a significant influence in studies that document a size effect.

[33] For example, see Capaul, Rowley, and Sharpe (1993) and Fama and French (1998).

[34] Fama and French (1995).

7.1 Closed-End Investment Fund Discounts

A closed-end investment fund issues a fixed number of shares at inception and does not sell any additional shares after the initial offering. Therefore, the fund capitalization is fixed unless a secondary public offering is made. The shares of closed-end funds trade on stock markets like any other shares in the equity market (i.e., their prices are determined by supply and demand).

Theoretically, these shares should trade at a price approximately equal to their net asset value (NAV) per share, which is simply the total market value of the fund's security holdings less any liabilities divided by the number of shares outstanding. An abundance of research, however, has documented that, on average, closed-end funds trade at a discount from NAV. Most studies have documented average discounts in the 4–10 percent range, although individual funds have traded at discounts exceeding 50 percent and others have traded at large premiums.[35]

The closed-end fund discount presents a puzzle because conceptually, an investor could purchase all the shares in the fund, liquidate the fund, and end up making a profit. Some researchers have suggested that these discounts are attributed to management fees or expectations of the managers' performance, but these explanations are not supported by the evidence.[36] An alternative explanation for the discount is that tax liabilities are associated with unrealized capital gains and losses that exist prior to when the investor bought the shares, and hence, the investor does not have complete control over the timing of the realization of gains and losses.[37] Although the evidence supports this hypothesis to a certain extent, the tax effect is not large enough to explain the entire discount. Finally, it has often been argued that the discounts exist because of liquidity problems and errors in calculating NAV. The illiquidity explanation is plausible if shares are recorded at the same price as more liquid, publicly traded stocks; some evidence supports this assertion. But as with tax reasons, liquidity issues explain only a portion of the discount effect.

Can these discounts be exploited to earn abnormal returns if transaction costs are taken into account? No. First, the transaction costs involved in exploiting the discount—buying all the shares and liquidating the fund—would eliminate any profit.[38] Second, these discounts tend to revert to zero over time. Hence, a strategy to trade on the basis of these discounts would not likely be profitable.[39]

7.2 Earnings Surprise

Although most event studies have supported semi-strong market efficiency, some researchers have provided evidence that questions semi-strong market efficiency. One of these studies relates to the extensively examined adjustment of stock prices to earnings announcements.[40] The unexpected part of the earnings announcement, or **earnings surprise**, is the portion of earnings that is unanticipated by investors and, according to the efficient market hypothesis, merits a price adjustment. Positive (negative) surprises should cause appropriate and rapid price increases (decreases). Several studies have been conducted using data from numerous markets around the world. Most of the results indicate that earnings surprises are reflected quickly in stock prices, but the adjustment process is not always efficient. In particular, although a substantial adjustment occurs prior to and at the announcement date, an adjustment also occurs after the announcement.[41]

---

[35] See Dimson and Minio-Kozerski (1999) for a review of this literature.

[36] See Lee, Sheifer, and Thaler (1990).

[37] The return to owners of closed-end fund shares has three parts: 1) the price appreciation or depreciation of the shares themselves, 2) the dividends earned and distributed to owners by the fund, and 3) the capital gains and losses earned by the fund that are distributed by the fund. The explanation of the anomalous pricing has to do with the timing of the distribution of capital gains.

[38] See, for example, the study by Pontiff (1996), which shows how the cost of arbitraging these discounts eliminates the profit.

[39] See Pontiff (1995).

[40] See Jones, Rendleman, and Latané (1984).

As a result of these slow price adjustments, companies that display the largest positive earnings surprises subsequently display superior stock return performance, whereas poor subsequent performance is displayed by companies with low or negative earnings surprises.[42] This finding implies that investors could earn abnormal returns using publicly available information by buying stocks of companies that had positive earnings surprises and selling those with negative surprises.

Although there is support for abnormal returns associated with earnings surprises, and some support for such returns beyond the announcement period, there is also evidence indicating that these observed abnormal returns are an artifact of studies that do not sufficiently control for transaction costs and risk.[43]

7.3 Initial Public Offerings (IPOs)

When a company offers shares of its stock to the public for the first time, it does so through an initial public offering (or IPO). This offering involves working with an investment bank that helps price and market the newly issued shares. After the offering is complete, the new shares trade on a stock market for the first time. Given the risk that investment bankers face in trying to sell a new issue for which the true price is unknown, it is perhaps not surprising to find that, on average, the initial selling price is set too low and that the price increases dramatically on the first trading day. The percentage difference between the issue price and the closing price at the end of the first day of trading is often referred to as the degree of underpricing.

The evidence suggests that, on average, investors who are able to buy the shares of an IPO at their offering price may be able to earn abnormal profits. For example, during the internet bubble of 1995–2000, many IPOs ended their first day of trading up by more than 100 percent. Such performance, however, is not always the case. Sometimes the issues are priced too high, which means that share prices drop on their first day of trading. In addition, the evidence also suggests that investors buying after the initial offering are not able to earn abnormal profits because prices adjust quickly to the "true" values, which supports semi-strong market efficiency. In fact, the subsequent long-term performance of IPOs is generally found to be below average. Taken together, the IPO underpricing and the subsequent poor performance suggests that the markets are overly optimistic initially (i.e., investors overreact).

Some researchers have examined closely why IPOs may appear to have anomalous returns. Because of the small size of the IPO companies and the method of equally weighting the samples, what appears to be an anomaly may simply be an artifact of the methodology.[44]

---

[41] Not surprisingly, it is often argued that this slow reaction contributes to a momentum pattern.

[42] A similar pattern has been documented in the corporate bond market, where bond prices react too slowly to new company earnings announcements as well as to changes in company debt ratings.

[43] See Brown (1997) for a summary of evidence supporting the existence of this anomaly. See Zarowin (1989) for evidence regarding the role of size in explaining abnormal returns to surprises; Alexander, Goff, and Peterson (1989) for evidence regarding transaction costs and unexpected earnings strategies; and Kim and Kim (2003) for evidence indicating that the anomalous returns can be explained by risk factors.

[44] See Brav, Geczy, and Gompers (1995).

7.4 Predictability of Returns Based on Prior Information

A number of researchers have documented that equity returns are related to prior information on such factors as interest rates, inflation rates, stock volatility, and dividend yields.[45] But finding that equity returns are affected by changes in economic fundamentals is not evidence of market inefficiency and would not result in abnormal trading returns.[46]

Furthermore, the relationship between stock returns and the prior information is not consistent over time. For example, in one study, the relationship between stock prices and dividend yields changed from positive to negative in different periods.[47] Hence, a trading strategy based on dividend yields would not yield consistent abnormal returns.

7.5 Implications for Investment Strategies

Although it is interesting to consider the anomalies just described, attempting to benefit from them in practice is not easy. In fact, most researchers conclude that observed anomalies are not violations of market efficiency but, rather, are the result of statistical methodologies used to detect the anomalies. As a result, if the methodologies are corrected, most of these anomalies disappear.[48] Another point to consider is that in an efficient market, overreactions may occur, but then so do under-reactions.[49] Therefore, on average, the markets are efficient. In other words, investors face challenges when they attempt to translate statistical anomalies into economic profits. Consider the following quote regarding anomalies from the *Economist* ("Frontiers of Finance Survey," 9 October 1993):

> Many can be explained away. When transactions costs are taken into account, the fact that stock prices tend to over-react to news, falling back the day after good news and bouncing up the day after bad news, proves unexploitable: price reversals are always within the bid-ask spread. Others such as the small-firm effect, work for a few years and then fail for a few years. Others prove to be merely proxies for the reward for risk taking. Many have disappeared since (and because) attention has been drawn to them.

It is difficult to envision entrusting your retirement savings to a manager whose strategy is based on buying securities on Mondays, which tends to have negative returns on average, and selling them on Fridays. For one thing, the negative Monday returns are merely an average, so on any given week, they could be positive. In addition, such a strategy would generate large trading costs. Even more importantly, investors would likely be uncomfortable investing their funds in a strategy that has no compelling underlying economic rationale.

BEHAVIORAL FINANCE

8

g describe behavioral finance and its potential relevance to understanding market anomalies.

---

[45] See, for example, Fama and Schwert (1977) and Fama and French (1988).

[46] See Fama and French (2008).

[47] Schwert (2003, Chapter 15).

[48] Fama (1998).

[49] This point is made by Fama (1998).

Behavioral finance examines investor behavior to understand how people make decisions, individually and collectively. Behavioral finance does not assume that people consider all available information in decision-making and act rationally by maximizing utility within budget constraints and updating expectations consistent with Bayes' formula. The resulting behaviors may affect what is observed in the financial markets.

In a broader sense, behavioral finance attempts to explain why individuals make the decisions that they do, whether these decisions are rational or irrational. The focus of much of the work in this area is on the behavioral biases that affect investment decisions. The behavior of individuals, in particular their behavioral biases, has been offered as a possible explanation for a number of pricing anomalies.

Most asset-pricing models assume that markets are rational and that the intrinsic value of a security reflects this rationality. But market efficiency and asset-pricing models do not require that each individual is rational—rather, only that the market is rational. If individuals deviate from rationality, other individuals are assumed to observe this deviation and respond accordingly. These responses move the market toward efficiency. If this does not occur in practice, it may be possible to explain some market anomalies referencing observed behaviors and behavioral biases.

8.1 Loss Aversion

In most financial models, the assumption is that investors are risk averse. **Risk aversion** refers to the tendency of people to dislike risk and to require higher expected returns to compensate for exposure to additional risk. Behavioral finance allows for the possibility that the dissatisfaction resulting from a loss exceeds the satisfaction resulting from a gain of the same magnitude. **Loss aversion** refers to the tendency of people to dislike losses more than they like comparable gains. This results in a strong preference for avoiding losses as opposed to achieving gains.[50] Some argue that behavioral theories of loss aversion can explain observed overreaction in markets. If loss aversion is more important than risk aversion, researchers should observe that investors overreact.[51] Although loss aversion can explain the overreaction anomaly, evidence also suggests that under reaction is just as prevalent as overreaction, which counters these arguments.

8.2 Herding

Herding behavior has been advanced as a possible explanation of under reaction and overreaction in financial markets. **Herding** occurs when investors trade on the same side of the market in the same securities, or when investors ignore their own private information and/or analysis and act as other investors do. Herding is clustered trading that may or may not be based on information.[52] Herding may result in under- or over-reaction to information depending upon the direction of the herd.

8.3 Overconfidence

A behavioral bias offered to explain pricing anomalies is overconfidence. If investors are overconfident, they overestimate their ability to process and interpret information about a security. Overconfident investors may not process information appropriately, and if there is a sufficient number of these investors, stocks will be mispriced.[53] But most researchers argue that this mispricing is temporary, with prices correcting eventually. If it takes a sufficiently long time for prices to become correctly priced and the mispricing is predictable, it may be possible for investors to earn abnormal profits.

---

[50] See DeBondt and Thaler (1985) and Tversky and Kahneman (1981).

[51] See Fama (1998).

[52] The term used when there is herding without information is "spurious herding."

Evidence has suggested that overconfidence results in mispricing for US, UK, German, French, and Japanese markets.[54] This overconfidence, however, is predominantly in higher-growth companies, whose prices react slowly to new information.[55]

8.4 Information Cascades

An application of behavioral theories to markets and pricing focuses on the role of personal learning in markets. Personal learning is what investors learn by observing outcomes of trades and what they learn from "conversations"—ideas shared among investors about specific assets and the markets.[56] Social interaction and the resultant contagion is important in pricing and may explain such phenomena as price changes without accompanying news and mistakes in valuation.

Biases that investors possess can lead to herding behavior or information cascades. Herding and information cascades are related but not identical concepts. An **information cascade** is the transmission of information from those participants who act first and whose decisions influence the decisions of others. Those who are acting on the choices of others may be ignoring their own preferences in favor of imitating the choices of others. In particular, information cascades may occur with respect to the release of accounting information because accounting information may be difficult to interpret and may be noisy. For example, the release of earnings is difficult to interpret because it is necessary to understand how the number was arrived at and noisy because it is uncertain what the current earnings imply about future earnings.

Information cascades may result in serial correlation of stock returns, which is consistent with overreaction anomalies. Do information cascades result in correct pricing? Some argue that if a cascade is leading toward an incorrect value, this cascade is "fragile" and will be corrected because investors will ultimately give more weight to public information or the trading of a recognized informed trader.[57] Information cascades, although documented in markets, do not necessarily mean that investors can exploit knowledge of them as profitable trading opportunities.

Are information cascades rational? If the informed traders act first and uninformed traders imitate the informed traders, this behavior is consistent with rationality. The imitation trading by the uninformed traders may help the market incorporate relevant information and improve market efficiency.[58] However, the imitation trading may lead to an overreaction to information. The empirical evidence indicates that information cascades are greater for a stock when the information quality regarding the company is poor.[59] Information cascades may enhance the information available to investors.

---

[53] Another aspect to overconfidence is that investors who are overconfident in their ability to select investments and manage a portfolio tend to use less diversification, investing in what is most familiar. Therefore, investor behavior may affect investment results—returns and risk—without implications for the efficiency of markets.

[54] Scott, Stumpp, and Xu (2003) and Boujelbene Abbes, Boujelbene, and Bouri (2009).

[55] Scott, Stumpp, and Xu (2003).

[56] Hirshleifer and Teoh (2009).

[57] Avery and Zemsky (1999).

[58] Another alternative is that the uninformed traders are the majority of the market participants and the imitators are imitating not because they agree with the actions of the majority but because they are looking to act on the actions of the uninformed traders.

[59] Avery and Zemsky (1999) and Bikhchandani, Hirshleifer, and Welch (1992).

8.5 Other Behavioral Biases

Other behavioral biases that have been put forth to explain observed investor behavior include the following:

- **representativeness**—investors assess new information and probabilities of outcomes based on similarity to the current state or to a familiar classification;
- **mental accounting**—investors keep track of the gains and losses for different investments in separate mental accounts and treat those accounts differently;
- **conservatism**—investors tend to be slow to react to new information and continue to maintain their prior views or forecasts; and
- **narrow framing**—investors focus on issues in isolation and respond to the issues based on how the issues are posed.[60]

The basic idea behind behavioral finance is that investors are humans and, therefore, imperfect. These observed less than rational behaviors may help explain observed pricing anomalies. The beliefs investors have about a given asset's value may not be homogeneous. But an issue, which is controversial, is whether these insights can help someone identify and exploit any mispricing. In other words, can investors use knowledge of behavioral biases to predict how asset prices will be affected and act based on the predictions to earn abnormal profits?

8.6 Behavioral Finance and Investors

Behavior biases can affect all market participants, from the novice investor to the most experienced investment manager. An understanding of behavioral finance can help market participants recognize their own and others' behavioral biases. As a result of this recognition, they may be able to respond and make improved decisions, individually and collectively.

8.7 Behavioral Finance and Efficient Markets

The use of behavioral finance to explain observed pricing is an important part of the understanding of how markets function and how prices are determined. Whether there is a behavioral explanation for market anomalies remains a debate. Pricing anomalies are continually being uncovered, and then statistical and behavioral explanations are offered to explain these anomalies.

On the one hand, if investors must be rational for efficient markets to exist, then all the imperfections of human investors suggest that markets cannot be efficient. On the other hand, if all that is required for markets to be efficient is that investors cannot consistently beat the market on a risk-adjusted basis, then the evidence does support market efficiency.

SUMMARY

This reading has provided an overview of the theory and evidence regarding market efficiency and has discussed the different forms of market efficiency as well as the implications for fundamental analysis, technical analysis, and portfolio management.

---

[60] For a review of these behavioral issues, see Hirshleifer (2001).

The general conclusion drawn from the efficient market hypothesis is that it is not possible to beat the market on a consistent basis by generating returns in excess of those expected for the level of risk of the investment.

Additional key points include the following:

- The efficiency of a market is affected by the number of market participants and depth of analyst coverage, information availability, and limits to trading.
- There are three forms of efficient markets, each based on what is considered to be the information used in determining asset prices. In the weak form, asset prices fully reflect all market data, which refers to all past price and trading volume information. In the semi-strong form, asset prices reflect all publicly known and available information. In the strong form, asset prices fully reflect all information, which includes both public and private information.
- Intrinsic value refers to the true value of an asset, whereas market value refers to the price at which an asset can be bought or sold. When markets are efficient, the two should be the same or very close. But when markets are not efficient, the two can diverge significantly.
- Most empirical evidence supports the idea that securities markets in developed countries are semi-strong-form efficient; however, empirical evidence does not support the strong form of the efficient market hypothesis.
- A number of anomalies have been documented that contradict the notion of market efficiency, including the size anomaly, the January anomaly, and the winners–losers anomalies. In most cases, however, contradictory evidence both supports and refutes the anomaly.
- Behavioral finance uses human psychology, such as behavioral biases, in an attempt to explain investment decisions. Whereas behavioral finance is helpful in understanding observed decisions, a market can still be considered efficient even if market participants exhibit seemingly irrational behaviors, such as herding.

REFERENCES

Alexander, John C., Delbert Goff, and Pamela P. Peterson. 1989. “Profitability of a Trading Strategy Based on Unexpected Earnings.” *Financial Analysts Journal*, vol. 45, no. 4:65–71.

Avery, Christopher, and Peter Zemsky. 1998. “Multi-Dimensional Uncertainty and Herding in Financial Markets.” *American Economic Review*, vol. 88, no. 4:724–748.

Bessembinder, Hendrik, and Kalok Chan. 1998. “Market Efficiency and the Returns to Technical Analysis.” *Financial Management*, vol. 27, no. 2:5–17.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. “A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades.” *Journal of Political Economy*, vol. 100, no. 5:992–1026.

Bouljelbene Abbes, Mouna, Younes Boujelbene, and Abdelfettah Bouri. 2009. “Overconfidence Bias: Explanation of Market Anomalies French Market Case.” *Journal of Applied Economic Sciences*, vol. 4, no. 1:12–25.

Brav, Alon, Christopher Geczy, and Paul A. Gompers. 1995. “The Long-Run Underperformance of Seasoned Equity Offerings Revisited.” Working paper, Harvard University.

Brealey, Richard. 1983. “Can Professional Investors Beat the Market?” *An Introduction to Risk and Return from Common Stocks*, 2nd edition. Cambridge, MA: MIT Press.

Bris, Arturo, William N. Goetzmann, and Ning Zhu. 2009. “Efficiency and the Bear: Short Sales and Markets around the World.” *Journal of Finance*, vol. 62, no. 3:1029–1079.

Brown, Laurence D. 1997. “Earning Surprise Research: Synthesis and Perspectives.” *Financial Analysts Journal*, vol. 53, no. 2:13–19.

Capaul, Carlo, Ian Rowley, and William Sharpe. 1993. “International Value and Growth Stock Returns.” *Financial Analysts Journal*, vol. 49:27–36.

Chen, Kong-Jun, and Xiao-Ming Li. 2006. “Is Technical Analysis Useful for Stock Traders in China? Evidence from the Szse Component A-Share Index.” *Pacific Economic Review*, vol. 11, no. 4:477–488.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam. 2005. “Evidence on the Speed of Convergence to Market Efficiency.” *Journal of Financial Economics*, vol. 76, no. 2:271–292.

DeBondt, Werner, and Richard Thaler. 1985. "Does the Stock Market Overreact?" *Journal of Finance*, vol. 40, no. 3:793–808.

Deng, Xiaohu, Sandra Mortal, and Vishal Gupta. 2017. "The Real Effects of Short Selling Constraints: Cross-Country Evidence." Working paper.

Dimson, Elroy, and Carolina Minio-Kozerski. 1999. "Closed-End Funds: A Survey." *Financial Markets, Institutions & Instruments*, vol. 8, no. 2:1–41.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance*, vol. 25, no. 2:383–417.

Fama, Eugene F. 1976. *Foundations of Finance*. New York: Basic Books.

Fama, Eugene F. 1998. "Market Efficiency, Long-Term Returns, and Behavioral Finance." *Journal of Financial Economics*, vol. 50, no. 3:283–306.

Fama, Eugene F., and G. William Schwert. 1977. "Asset Returns and Inflation." *Journal of Financial Economics*, vol. 5, no. 2:115–146.

Fama, Eugene F., and Kenneth R. French. 1988. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, vol. 22, no. 1:3–25.

Fama, Eugene F., and Kenneth R. French. 1995. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Finance*, vol. 50, no. 1:131–155.

Fama, Eugene F., and Kenneth R. French. 1998. "Value versus Growth: The International Evidence." *Journal of Finance*, vol. 53:1975–1999.

Fama, Eugene F., and Kenneth R. French. 2008. "Dissecting Anomalies." *Journal of Finance*, vol. 63, no. 4:1653–1678.

Fifield, Suzanne, David Power, and C. Donald Sinclair. 2005. "An Analysis of Trading Strategies in Eleven European Stock Markets." *European Journal of Finance*, vol. 11, no. 6:531–548.

Gan, Christopher, Minsoo Lee, Au Yong Hue Hwa, and Jun Zhang. 2005. "Revisiting Share Market Efficiency: Evidence from the New Zealand, Australia, US and Japan Stock Indices." *American Journal of Applied Sciences*, vol. 2, no. 5:996–1002.

Grossman, Sanford J., and Joseph E. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review*, vol. 70, no. 3:393–408.

Hirshleifer, David. 2001. "Investor Psychology and Asset Pricing." *Journal of Finance*, vol. 56, no. 4:1533–1597.

Hirshleifer, David, and Siew Hong Teoh. 2009. "Thought and Behavior Contagion in Capital Markets." In *Handbook of Financial Markets: Dynamics and Evolution*. Edited by Klaus Reiner Schenk-Hoppe and Thorstein Hens. Amsterdam: North Holland.

Jacobs, Bruce I., and Kenneth N. Levy. 1988. "Calendar Anomalies: Abnormal Returns at Calendar Turning Points." *Financial Analysts Journal*, vol. 44, no. 6:28–39.

Jaffe, Jeffrey. 1974. "Special Information and Insider Trading." *Journal of Business*, vol. 47, no. 3:410–428.

Jegadeesh, Narayan, and Sheridan Titman. 2001. "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations." *Journal of Finance*, vol. 56:699–720.

Johnson, Timothy C. 2002. "Rational Momentum Effects." *Journal of Finance*, vol. 57, no. 2:585–608.

Jones, Charles P., Richard J. Rendleman, and Henry. A. Latané. 1984. "Stock Returns and SUEs during the 1970's." *Journal of Portfolio Management*, vol. 10:18–22.

Kim, Donchoi, and Myungsun Kim. 2003. "A Multifactor Explanation of Post-Earnings Announcement Drift." *Journal of Financial and Quantitative Analysis*, vol. 38, no. 2:383–398.

Kim, Dongcheol. 2006. "On the Information Uncertainty Risk and the January Effect." *Journal of Business*, vol. 79, no. 4:2127–2162.

Lee, Charles M.C., Andrei Sheifer, and Richard H. Thaler. 1990. "Anomalies: Closed-End Mutual Funds." *Journal of Economic Perspectives*, vol. 4, no. 4:153–164.

Malkiel, Burton G. 1995. "Returns from Investing in Equity Mutual Funds 1971 to 1991." *Journal of Finance*, vol. 50:549–572.

Mobarek, Asma, A. Sabur Mollah, and Rafiqul Bhuyan. 2008. "Market Efficiency in Emerging Stock Market." *Journal of Emerging Market Finance*, vol. 7, no. 1:17–41.

Pontiff, Jeffrey. 1995. "Closed-End Fund Premia and Returns: Implications for Financial Market Equilibrium." *Journal of Financial Economics*, vol. 37:341–370.

Pontiff, Jeffrey. 1996. "Costly Arbitrage: Evidence from Closed-End Funds." *Quarterly Journal of Economics*, vol. 111, no. 4:1135–1151.

Raja, M., J. Clement Sudhahar, and M. Selvam. 2009. "Testing the Semi-Strong Form Efficiency of Indian Stock Market with Respect to Information Content of Stock Split Announcement—A Study of IT Industry." *International Research Journal of Finance and Economics*, vol. 25:7–20.

Roll, Richard. 1983. "On Computing Mean Returns and the Small Firm Premium." *Journal of Financial Economics*, vol. 12:371–386.

Rozeff, Michael S., and Mir A. Zaman. 1988. "Market Efficiency and Insider Trading: New Evidence." *Journal of Business*, vol. 61:25–44.

Schwert, G. William. 2003. "Anomalies and Market Efficiency." *Handbook of the Economics of Finance*. Edited by George M. Constantinides, M. Harris, and Rene Stulz. Amsterdam: Elsevier Science, B. V.

Scott, James, Margaret Stumpp, and Peter Xu. 2003. "Overconfidence Bias in International Stock Prices." *Journal of Portfolio Management*, vol. 29, no. 2:80–89.

Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science*, vol. 211, no. 30:453–458.

Yau, Jot, Thomas Schneeweis, Thomas Robinson, and Lisa Weiss. 2007. "Alternative Investments Portfolio Management." *Managing Investment Portfolios: A Dynamic Process*. Hoboken, NJ: John Wiley & Sons.

Zarowin, P. 1989. "Does the Stock Market Overreact to Corporate Earnings Information?" *Journal of Finance*, vol. 44:1385–1399.

PRACTICE PROBLEMS

1 In an efficient market, the change in a company's share price is *most likely* the result of:

A insiders' private information.

B the previous day's change in stock price.

C new information coming into the market.

2 Regulation that restricts some investors from participating in a market will *most likely*:

A impede market efficiency.

B not affect market efficiency.

C contribute to market efficiency.

3 With respect to efficient market theory, when a market allows short selling, the efficiency of the market is *most likely* to:

A increase.

B decrease.

C remain the same.

4 Which of the following regulations will *most likely* contribute to market efficiency? Regulatory restrictions on:

A short selling.

B foreign traders.

C insiders trading with nonpublic information.

5 Which of the following market regulations will *most likely* impede market efficiency?

A Restricting traders' ability to short sell.

B Allowing unrestricted foreign investor trading.

C Penalizing investors who trade with nonpublic information.

6 If markets are efficient, the difference between the intrinsic value and market value of a company's security is:

A negative.

B zero.

C positive.

7 The intrinsic value of an undervalued asset is:

A less than the asset's market value.

B greater than the asset's market value.

C the value at which the asset can currently be bought or sold.

8 The market value of an undervalued asset is:

A greater than the asset's intrinsic value.

B the value at which the asset can currently be bought or sold.

C equal to the present value of all the asset's expected cash flows.

9 With respect to the efficient market hypothesis, if security prices reflect *only* past prices and trading volume information, then the market is:

A weak-form efficient.

© 2011 CFA Institute. All rights reserved.

B strong-form efficient.

C semi-strong-form efficient.

10 Which one of the following statements *best* describes the semi-strong form of market efficiency?

A Empirical tests examine the historical patterns in security prices.

B Security prices reflect all publicly known and available information.

C Semi-strong-form efficient markets are not necessarily weak-form efficient.

11 If markets are semi-strong efficient, standard fundamental analysis will yield abnormal trading profits that are:

A negative.

B equal to zero.

C positive.

12 If prices reflect all public and private information, the market is *best* described as:

A weak-form efficient.

B strong-form efficient.

C semi-strong-form efficient.

13 If markets are semi-strong-form efficient, then passive portfolio management strategies are *most likely* to:

A earn abnormal returns.

B outperform active trading strategies.

C underperform active trading strategies.

14 If a market is semi-strong-form efficient, the risk-adjusted returns of a passively managed portfolio relative to an actively managed portfolio are *most likely*:

A lower.

B higher.

C the same.

15 Technical analysts assume that markets are:

A weak-form efficient.

B weak-form inefficient.

C semi-strong-form efficient.

16 Fundamental analysts assume that markets are:

A weak-form inefficient.

B semi-strong-form efficient.

C semi-strong-form inefficient.

17 If a market is weak-form efficient but semi-strong-form inefficient, then which of the following types of portfolio management is *most likely* to produce abnormal returns?

A Passive portfolio management.

B Active portfolio management based on technical analysis.

C Active portfolio management based on fundamental analysis.

18 An increase in the time between when an order to trade a security is placed and when the order is executed *most likely* indicates that market efficiency has:

A decreased.

B remained the same.

C increased.

19 With respect to efficient markets, a company whose share price changes gradually after the public release of its annual report *most likely* indicates that the market where the company trades is:

A semi-strong-form efficient.

B subject to behavioral biases.

C receiving additional information about the company.

20 Which of the following is *least likely* to explain the January effect anomaly?

A Tax-loss selling.

B Release of new information in January.

C Window dressing of portfolio holdings.

21 If a researcher conducting empirical tests of a trading strategy using time series of returns finds statistically significant abnormal returns, then the researcher has *most likely* found:

A a market anomaly.

B evidence of market inefficiency.

C a strategy to produce future abnormal returns.

22 Which of the following market anomalies is inconsistent with weak-form market efficiency?

A Earnings surprise.

B Momentum pattern.

C Closed-end fund discount.

23 Researchers have found that value stocks have consistently outperformed growth stocks. An investor wishing to exploit the value effect should purchase the stock of companies with above-average:

A dividend yields.

B market-to-book ratios.

C price-to-earnings ratios.

24 With respect to rational and irrational investment decisions, the efficient market hypothesis requires:

A only that the market is rational.

B that all investors make rational decisions.

C that some investors make irrational decisions.

25 Observed overreactions in markets can be explained by an investor's degree of:

A risk aversion.

B loss aversion.

C confidence in the market.

26 Like traditional finance models, the behavioral theory of loss aversion assumes that investors dislike risk; however, the dislike of risk in behavioral theory is assumed to be:

A leptokurtic.

B symmetrical.

C asymmetrical.

SOLUTIONS

1 C is correct. Today's price change is independent of the one from yesterday, and in an efficient market, investors will react to new, independent information as it is made public.

2 A is correct. Reducing the number of market participants can accentuate market imperfections and impede market efficiency (e.g., restrictions on foreign investor trading).

3 A is correct. According to theory, reducing the restrictions on trading will allow for more arbitrage trading, thereby promoting more efficient pricing. Although regulators argue that short selling exaggerates downward price movements, empirical research indicates that short selling is helpful in price discovery.

4 C is correct. Regulation to restrict unfair use of nonpublic information encourages greater participation in the market, which increases market efficiency. Regulators (e.g., US SEC) discourage illegal insider trading by issuing penalties to violators of their insider trading rules.

5 A is correct. Restricting short selling will reduce arbitrage trading, which promotes market efficiency. Permitting foreign investor trading increases market participation, which makes markets more efficient. Penalizing insider trading encourages greater market participation, which increases market efficiency.

6 B is correct. A security's intrinsic value and market value should be equal when markets are efficient.

7 B is correct. The intrinsic value of an undervalued asset is greater than the market value of the asset, where the market value is the transaction price at which an asset can be currently bought or sold.

8 B is correct. The market value is the transaction price at which an asset can be currently bought or sold.

9 A is correct. The weak-form efficient market hypothesis is defined as a market where security prices fully reflect all market data, which refers to all past price and trading volume information.

10 B is correct. In semi-strong-form efficient markets, security prices reflect all publicly available information.

11 B is correct. If all public information should already be reflected in the market price, then the abnormal trading profit will be equal to zero when fundamental analysis is used.

12 B is correct. The strong-form efficient market hypothesis assumes all information, public or private, has already been reflected in the prices.

13 B is correct. Costs associated with active trading strategies would be difficult to recover; thus, such active trading strategies would have difficulty outperforming passive strategies on a consistent after-cost basis.

14 B is correct. In a semi-strong-form efficient market, passive portfolio strategies should outperform active portfolio strategies on a risk-adjusted basis.

15 B is correct. Technical analysts use past prices and volume to predict future prices, which is inconsistent with the weakest form of market efficiency (i.e., weak-form market efficiency). Weak-form market efficiency states that investors cannot earn abnormal returns by trading on the basis of past trends in price and volume.

16 C is correct. Fundamental analysts use publicly available information to estimate a security's intrinsic value to determine if the security is mispriced, which is inconsistent with the semi-strong form of market efficiency. Semi-strong-form market efficiency states that investors cannot earn abnormal returns by trading based on publicly available information.

17 C is correct. If markets are not semi-strong-form efficient, then fundamental analysts are able to use publicly available information to estimate a security's intrinsic value and identify misvalued securities. Technical analysis is not able to earn abnormal returns if markets are weak-form efficient. Passive portfolio managers outperform fundamental analysis if markets are semi-strong-form efficient.

18 A is correct. Operating inefficiencies reduce market efficiency.

19 C is correct. If markets are efficient, the information from the annual report is reflected in the stock prices; therefore, the gradual changes must be from the release of additional information.

20 B is correct. The excess returns in January are not attributed to any new information or news; however, research has found that part of the seasonal pattern can be explained by tax-loss selling and portfolio window dressing.

21 A is correct. Finding significant abnormal returns does not necessarily indicate that markets are inefficient or that abnormal returns can be realized by applying the strategy to future time periods. Abnormal returns are considered market anomalies because they may be the result of the model used to estimate the expected returns or may be the result of underestimating transaction costs or other expenses associated with implementing the strategy, rather than because of market inefficiency.

22 B is correct. Trading based on historical momentum indicates that price patterns exist and can be exploited by using historical price information. A momentum trading strategy that produces abnormal returns contradicts the weak form of the efficient market hypothesis, which states that investors cannot earn abnormal returns on the basis of past trends in prices.

23 A is correct. Higher than average dividend yield is a characteristic of a value stock, along with low price-to-earnings and low market-to-book ratios. Growth stocks are characterized by low dividend yields and high price-to-earnings and high market-to-book ratios.

24 A is correct. The efficient market hypothesis and asset-pricing models only require that the market is rational. Behavioral finance is used to explain *some* of the market anomalies as irrational decisions.

25 B is correct. Behavioral theories of loss aversion can explain observed overreaction in markets, such that investors dislike losses more than comparable gains (i.e., risk is not symmetrical).

26 C is correct. Behavioral theories of loss aversion allow for the possibility that the dislike for risk is not symmetrical, which allows for loss aversion to explain observed overreaction in markets such that investors dislike losses more than they like comparable gains.

EQUITY INVESTMENTS STUDY SESSION

12

Equity Investments (2)

This study session focuses on the characteristics, analysis, and valuation of equity securities. Various equity types including public and private equities are described. The various industry classification approaches for global equities and useful frameworks for conducting industry and individual company analysis are presented. Coverage of the three main equity valuation approaches (present value, multiplier, and asset based) conclude the session.

READING ASSIGNMENTS

Reading 36	Overview of Equity Securities by Ryan C. Fuhrmann, CFA, and Asjeet S. Lamba, PhD, CFA
Reading 37	Introduction to Industry and Company Analysis by Patrick W. Dorsey, CFA, Anthony M. Fiore, CFA, and Ian Rossa O'Reilly, CFA
Reading 38	Equity Valuation: Concepts and Basic Tools by John J. Nagorniak, CFA, and Stephen E. Wilcox, PhD, CFA

© 2021 CFA Institute. All rights reserved.

READING

36

Overview of Equity Securities

by Ryan C. Fuhrmann, CFA, and Asjeet S. Lamba, PhD, CFA

Ryan C. Fuhrmann, CFA, is at Fuhrmann Capital LLC (USA). Asjeet S. Lamba, PhD, CFA, is at the University of Melbourne (Australia).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe characteristics of types of equity securities;
☐	**b.** describe differences in voting rights and other ownership characteristics among different equity classes;
☐	**c.** compare and contrast public and private equity securities;
☐	**d.** describe methods for investing in non-domestic equity securities;
☐	**e.** compare the risk and return characteristics of different types of equity securities;
☐	**f.** explain the role of equity securities in the financing of a company's assets;
☐	**g.** contrast the market value and book value of equity securities;
☐	**h.** compare a company's cost of equity, its (accounting) return on equity, and investors' required rates of return.

IMPORTANCE OF EQUITY SECURITIES

1

Equity securities represent ownership claims on a company's net assets. As an asset class, equity plays a fundamental role in investment analysis and portfolio management because it represents a significant portion of many individual and institutional investment portfolios.

The study of equity securities is important for many reasons. First, the decision on how much of a client's portfolio to allocate to equities affects the risk and return characteristics of the entire portfolio. Second, different types of equity securities have different ownership claims on a company's net assets, which affect their risk and return characteristics in different ways. Finally, variations in the features of equity securities are reflected in their market prices, so it is important to understand the valuation implications of these features.

© 2019 CFA Institute. All rights reserved.

This reading provides an overview of equity securities and their different features and establishes the background required to analyze and value equity securities in a global context. It addresses the following questions:

- What distinguishes common shares from preference shares, and what purposes do these securities serve in financing a company's operations?
- What are convertible preference shares, and why are they often used to raise equity for unseasoned or highly risky companies?
- What are private equity securities, and how do they differ from public equity securities?
- What are depository receipts and their various types, and what is the rationale for investing in them?
- What are the risk factors involved in investing in equity securities?
- How do equity securities create company value?
- What is the relationship between a company's cost of equity, its return on equity, and investors' required rate of return?

The remainder of this reading is organized as follows. Section 1.1 provides an overview of global equity markets and their historical performance. Section 2 examines the different types and characteristics of equity securities, and Section 3 outlines the differences between public and private equity securities. Section 4 provides an overview of the various types of equity securities listed and traded in global markets. Section 5 discusses the risk and return characteristics of equity securities. Section 6 examines the role of equity securities in creating company value and the relationship between a company's cost of equity, its return on equity, and investors' required rate of return. The final section summarizes the reading.

1.1 Equity Securities in Global Financial Markets

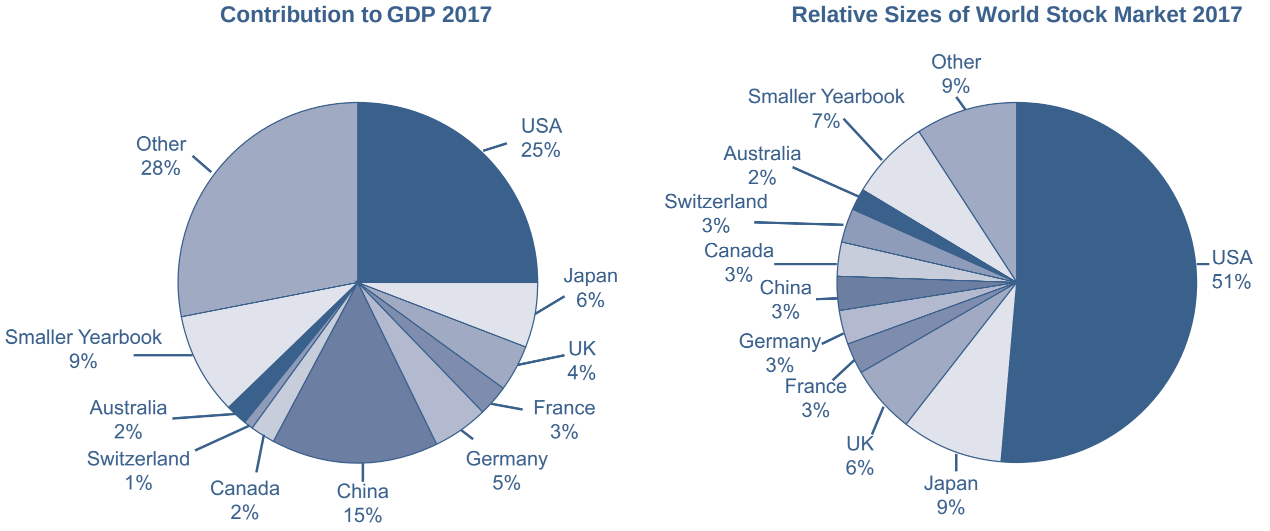
This section highlights the relative importance and performance of equity securities as an asset class. We examine the total market capitalization and trading volume of global equity markets and the prevalence of equity ownership across various geographic regions. We also examine historical returns on equities and compare them to the returns on government bonds and bills.

Exhibit 1 summarizes the contributions of selected countries and geographic regions to global gross domestic product (GDP) and global equity market capitalization. Analysts may examine the relationship between equity market capitalization and GDP as a rough indicator of whether the global equity market (or a specific country's or region's equity market) is under, over, or fairly valued, particularly compared to its long-run average.

Exhibit 1 illustrates the significant value that investors attach to publicly traded equities relative to the sum of goods and services produced globally every year. It shows the continued significance, and the potential over-representation, of US equity markets relative to their contribution to global GDP. That is, while US equity markets contribute around 51 percent to the total capitalization of global equity markets, their contribution to the global GDP is only around 25 percent. Following the stock market turmoil in 2008, however, the market capitalization to GDP ratio of the United States fell to 59 percent, which is significantly lower than its long-run average of 79 percent.

As equity markets outside the United States develop and become increasingly global, their total capitalization levels are expected to grow closer to their respective world GDP contributions. Therefore, it is important to understand and analyze equity securities from a global perspective.

Exhibit 1 Country and Regional Contributions to Global GDP and Equity Market Capitalization (2017)



Source: The WorldBank Databank *2017*, and Dimson, Marsh, and Staunton (2018).

Exhibit 2 lists the top 10 equity markets at the end of 2017 based on total market capitalization (in billions of US dollars), trading volume, and the number of listed companies.[1] Note that the rankings differ based on the criteria used. For example, the top three markets based on total market capitalization are the NYSE Euronext (US), NASDAQ OMX, and the Japan Exchange Group; however, the top three markets based on total US dollar trading volume are the Nasdaq OMX, NYSE Euronext (US), and the Shenzhen Stock Exchange, respectively.[2]

Exhibit 2 Equity Markets Ranked by Total Market Capitalization at the End of 2017 (Billions of US Dollars)

Rank	Name of Market	Total US Dollar Market Capitalization	Total US Dollar Trading Volume	Number of Listed Companies
1	NYSE Euronext (US)	\$22,081.4	\$16,140.1	2,286
2	NASDAQ OMX	\$10,039.4	\$33,407.1	2,949
3	Japan Exchange Group[a]	\$6,220.0	\$6,612.1	3,604
4	Shanghai Stock Exchange	\$5,084.4	\$7,589.3	1,396
5	Euronext[b]	\$4,393.0	\$1,981.6	1,255
6	Hong Kong Exchanges	\$4,350.5	\$1,958.8	2,118
7	Shenzhen Stock Exchanges	\$3,617.9	\$9,219.7	2,089

(continued)

---

[1] The market capitalization of an individual stock is computed as the share price multiplied by the number of shares outstanding. The total market capitalization of an equity market is the sum of the market capitalizations of each individual stock listed on that market. Similarly, the total trading volume of an equity market is computed by value weighting the total trading volume of each individual stock listed on that market. Total dollar trading volume is computed as the average share price multiplied by the number of shares traded.

[2] NASDAQ is the acronym for the National Association of Securities Dealers Automated Quotations.

Exhibit 2 (Continued)

Rank	Name of Market	Total US Dollar Market Capitalization	Total US Dollar Trading Volume	Number of Listed Companies
8	National Stock Exchange of India	\$2,351.5	\$1,013.3	1,897
9	BSE Limited[c]	\$2,331.6	\$183.0	5,616
10	Deutsche Börse	\$2,262.2	\$1,497.9	499

Notes:

[a] Japan Exchange Group is the merged entity containing the Tokyo Stock Exchange and Osaka Securities Exchange.

[b] From 2001, includes Netherlands, France, England, Belgium, and Portugal.

[c] Bombay Stock Exchange.

Source: Adapted from the *World Federation of Exchanges 2017 Report* (see http://www.world-exchanges.org). Note that market capitalization by company is calculated by multiplying its stock price by the number of shares outstanding. The market's overall capitalization is the aggregate of the market capitalizations of all companies traded on that market. The number of listed companies includes both domestic and foreign companies whose shares trade on these markets.

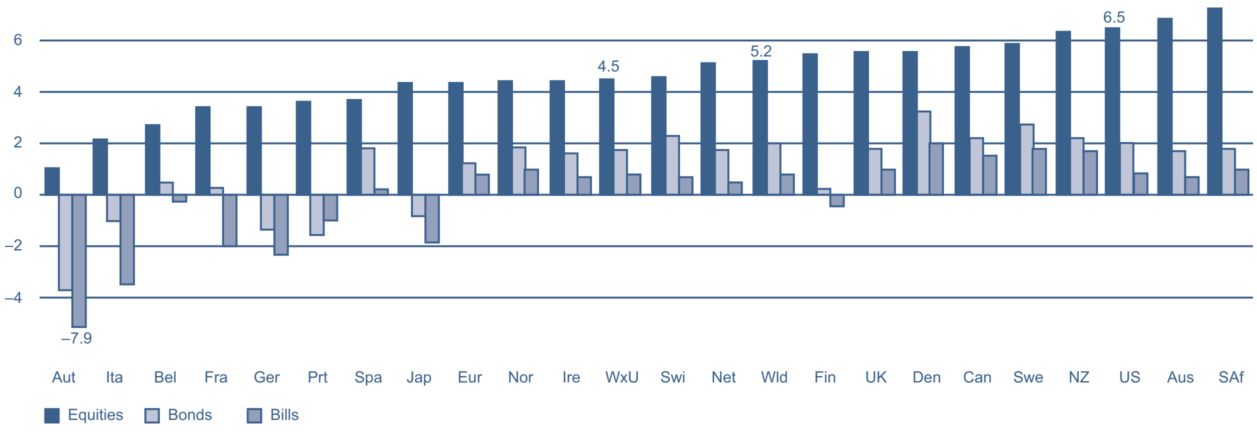
Exhibit 3 compares the *real* (or inflation-adjusted) compounded returns on government bonds, government bills, and equity securities in 21 countries plus the world index ("Wld"), the world ex-US ("WxU"), and Europe ("Eur") during the 118 years 1900–2017.[3] In real terms, government bonds and bills have essentially kept pace with the inflation rate, earning annualized real returns of less than 2 percent in most countries.[4] By comparison, real returns in equity markets have generally been around 3.5 percent per year in most markets—with a world average return of around 5.2 percent and a world average return excluding the United States just under 5 percent. During this period, South Africa and Australia were the best performing markets followed by the United States, New Zealand, and Sweden.

---

[3] The real return for a security is approximated by taking the nominal return and subtracting the observed inflation rate in that country.

[4] The exceptions are Austria, Belgium, Finland, France, Germany, Portugal, and Italy—where the average real returns on government bonds and/or bills have been negative. In general, that performance reflects the very high inflation rates in these countries during the World War years.

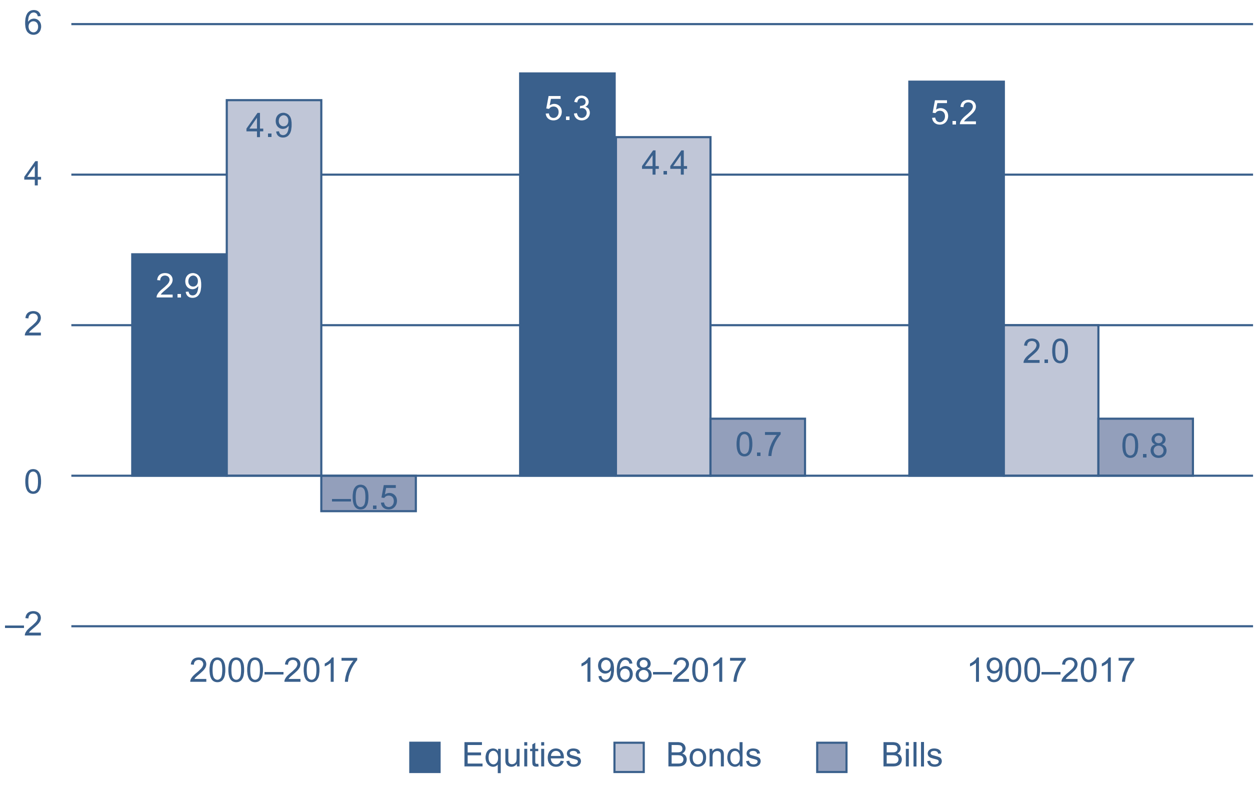
Exhibit 3 Real Returns on Global Equity Securities, Bonds, and Bills During 1900–2017



Source: Dimson, Marsh, and Staunton (2018).

Exhibit 4 shows the annualized real returns on major asset classes for the world index over 1900–2017.

Exhibit 4 Annualized Real Returns on Asset Classes for the World Index, 1900–2017

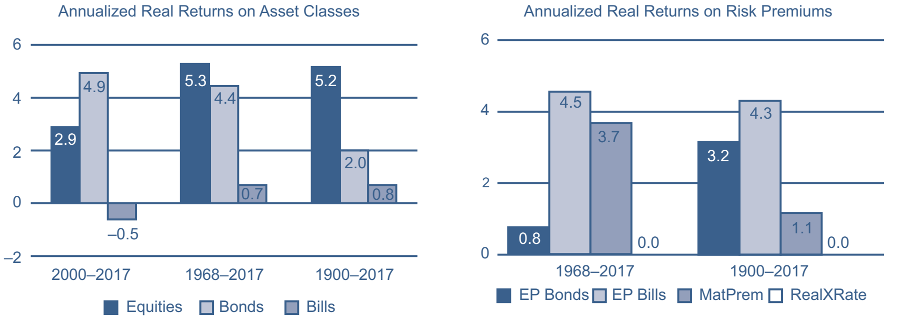


Source: Dimson, Marsh, and Staunton (2018).

The volatility in asset market returns is further highlighted in Exhibit 5, which shows the annualized risk premia for equity relative to bonds (EP Bonds), and equity relative to treasury bills (EP Bills). Maturity premium for government bond returns relative to treasury bill returns (Mat Prem) is also shown.

These observations and historical data are consistent with the concept that the return on securities is directly related to risk level. That is, equity securities have higher risk levels when compared with government bonds and bills, they earn higher rates of return to compensate investors for these higher risk levels, and they also tend to be more volatile over time.

Exhibit 5 **Annualized Real Returns on Asset Classes and Risk Premiums for the World Index since 1900–2017**



Notes: Equities are total returns, including reinvested dividend income. Bonds are total return, including reinvested coupons, on long-term government bonds. Bills denotes the total return, including any income, from Treasury bills. All returns are adjusted for inflation and are expressed as geometric mean returns. EP bonds denotes the equity risk premium relative to long-term government bonds. EP Bills denotes the equity premium relative to Treasury bills. MatPrem denotes the maturity premium for government bond returns relative to bill returns. RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.
Source: Dimson, Marsh, and Staunton (2018).

Given the high risk levels associated with equity securities, it is reasonable to expect that investors' tolerance for risk will tend to differ across equity markets. This is illustrated in Exhibit 6, which shows the results of a series of studies conducted by the Australian Securities Exchange on international differences in equity ownership. During the 2004–2014 period, equity ownership as a percentage of the population was lowest in South Korea (averaging 9.0 percent), followed by Germany (14.5 percent) and Sweden (17.7 percent). In contrast, Australia and New Zealand had the highest equity ownership as a percentage of the population (averaging more than 20 percent). In addition, there has been a relative decline in share ownership in several countries over recent years, which is not surprising given the recent overall uncertainty in global economies and the volatility in equity markets that this uncertainty has created.

Exhibit 6 International Comparisons of Stock Ownership: 2004–2014[5]

	2004	2006	2008	2010	2012	2014
Australia – Direct/ Indirect	55%	46%	41%	43%	38%	36%
South Korea – Shares	8	7	10	10	10	N/A
Germany – Shares/ Funds	16	16	14	13	15	13
Sweden – Shares	22	20	18	17	15	14
United Kingdom – Shares/Funds	22	20	18	N/A	17	N/A
New Zealand – Direct	23	26	N/A	22	23	26

Source: Adapted from the *2014 Australian Share Ownership Study* conducted by the Australian Securities Exchange (see http://www.asx.com.au). For Australia and the United States, the data pertain to direct and indirect ownership in equity markets; for other countries, the data pertain to direct ownership in shares and share funds. Data not available in specific years are shown as "N/A."

An important implication from the above discussion is that equity securities represent a key asset class for global investors because of their unique return and risk characteristics. We next examine the various types of equity securities traded on global markets and their salient characteristics.

CHARACTERISTICS OF EQUITY SECURITIES

2

a describe characteristics of types of equity securities;

b describe differences in voting rights and other ownership characteristics among different equity classes;

Companies finance their operations by issuing either debt or equity securities. A key difference between these securities is that debt is a liability of the issuing company, whereas equity is not. This means that when a company issues debt, it is contractually obligated to repay the amount it borrows (i.e., the principal or face value of the debt) at a specified future date. The cost of using these funds is called interest, which the company is contractually obligated to pay until the debt matures or is retired.

When the company issues equity securities, it is not contractually obligated to repay the amount it receives from shareholders, nor is it contractually obligated to make periodic payments to shareholders for the use of their funds. Instead, shareholders have a claim on the company's assets after all liabilities have been paid. Because of this residual claim, equity shareholders are considered to be owners of the company. Investors who purchase equity securities are seeking total return (i.e., capital or price appreciation and dividend income), whereas investors who purchase debt securities

---

[5] The percentages reported in the exhibit are based on samples of the adult population in each country who own equity securities either directly or indirectly through investment or retirement funds. For example, 36 percent of the adult population of Australia in 2014 (approximately 6.5 million people) owned equity securities either directly or indirectly. As noted in the study, it is not appropriate to make absolute comparisons across countries given the differences in methodology, sampling, timing, and definitions that have been used in different countries. However, trends across different countries can be identified.

(and hold until maturity) are seeking interest income. As a result, equity investors expect the company's management to act in their best interest by making operating decisions that will maximize the market price of their shares (i.e., shareholder wealth).

In addition to common shares (also known as ordinary shares or common stock), companies may also issue preference shares (also known as preferred stock), the other type of equity security. The following sections discuss the different types and characteristics of common and preference securities.

2.1 Common Shares

Common shares represent an ownership interest in a company and are the predominant type of equity security. As a result, investors share in the operating performance of the company, participate in the governance process through voting rights, and have a claim on the company's net assets in the case of liquidation. Companies may choose to pay out some, or all, of their net income in the form of cash dividends to common shareholders, but they are not contractually obligated to do so.[6]

Voting rights provide shareholders with the opportunity to participate in major corporate governance decisions, including the election of its board of directors, the decision to merge with or take over another company, and the selection of outside auditors. Shareholder voting generally takes place during a company's annual meeting. As a result of geographic limitations and the large number of shareholders, it is often not feasible for shareholders to attend the annual meeting in person. For this reason, shareholders may **vote by proxy**, which allows a designated party—such as another shareholder, a shareholder representative, or management—to vote on the shareholders' behalf.

Regular shareholder voting, where each share represents one vote, is referred to as **statutory voting**. Although it is the common method of voting, it is not always the most appropriate one to use to elect a board of directors. To better serve shareholders who own a small number of shares, **cumulative voting** is often used. Cumulative voting allows shareholders to direct their total voting rights to specific candidates, as opposed to having to allocate their voting rights evenly among all candidates. Total voting rights are based on the number of shares owned multiplied by the number of board directors being elected. For example, under cumulative voting, if four board directors are to be elected, a shareholder who owns 100 shares is entitled to 400 votes and can either cast all 400 votes in favor of a single candidate or spread them across the candidates in any proportion. In contrast, under statutory voting, a shareholder would be able to cast only a maximum of 100 votes for each candidate.

The key benefit to cumulative voting is that it allows shareholders with a small number of shares to apply all of their votes to one candidate, thus providing the opportunity for a higher level of representation on the board than would be allowed under statutory voting.

Exhibit 7 describes the rights of Viacom Corporation's shareholders. In this case, a dual-share arrangement allows the founding chairman and his family to control more than 70 percent of the voting rights through the ownership of Class A shares. This arrangement gives them the ability to exert control over the board of director election process, corporate decision making, and other important aspects of managing the company. A cumulative voting arrangement for any minority shareholders of Class A shares would improve their board representation.

---

[6] It is also possible for companies to pay more than the current period's net income as dividends. Such payout policies are, however, generally not sustainable in the long run.

Exhibit 7 Share Class Arrangements at Viacom Corporation[7]

Viacom has two classes of common stock: Class A, which is the voting stock, and Class B, which is the non-voting stock. There is no difference between the two classes except for voting rights; they generally trade within a close price range of each other. There are, however, far more shares of Class B outstanding, so most of the trading occurs in that class.

- **Voting Rights**—Holders of Class A common stock are entitled to one vote per share. Holders of Class B common stock do not have any voting rights, except as required by Delaware law. Generally, all matters to be voted on by Viacom stockholders must be approved by a majority of the aggregate voting power of the shares of Class A common stock present in person or represented by proxy, except as required by Delaware law.
- **Dividends**—Stockholders of Class A common stock and Class B common stock will share ratably in any cash dividend declared by the Board of Directors, subject to any preferential rights of any outstanding preferred stock. Viacom does not currently pay a cash dividend, and any decision to pay a cash dividend in the future will be at the discretion of the Board of Directors and will depend on many factors.
- **Conversion**—So long as there are 5,000 shares of Class A common stock outstanding, each share of Class A common stock will be convertible at the option of the holder of such share into one share of Class B common stock.
- **Liquidation Rights**—In the event of liquidation, dissolution, or winding-up of Viacom, all stockholders of common stock, regardless of class, will be entitled to share ratably in any assets available for distributions to stockholders of shares of Viacom common stock subject to the preferential rights of any outstanding preferred stock.
- **Split, Subdivision, or Combination**—In the event of a split, subdivision, or combination of the outstanding shares of Class A common stock or Class B common stock, the outstanding shares of the other class of common stock will be divided proportionally.
- **Preemptive Rights**—Shares of Class A common stock and Class B common stock do not entitle a stockholder to any preemptive rights enabling a stockholder to subscribe for or receive shares of stock of any class or any other securities convertible into shares of stock of any class of Viacom.

As seen in Exhibit 7, companies can issue different classes of common shares (Class A and Class B shares), with each class offering different ownership rights.[8] For example, as shown in Exhibit 8, the Ford Motor Company has Class A shares ("Common Stock"), which are owned by the investing public. It also has Class B shares, which are owned only by the Ford family. The exhibit contains an excerpt from Ford's *2017 Annual Report* (p. 144). Class A shareholders have 60 percent voting rights, whereas Class B shareholders have 40 percent. In the case of liquidation, however, Class B shareholders will not only receive the first US$0.50 per share that is available for distribution (as will Class A shareholders), but they will also receive the next US$1.00 per share that is available for distribution before Class A shareholders receive anything else. Thus, Class B shareholders have an opportunity to receive a larger proportion of distributions upon liquidation than do Class A shareholders.[9]

---

[7] This information has been adapted from Viacom's investor relations website and its 10-K filing with the US Securities and Exchange Commission; see www.viacom.com.

[8] In some countries, including the United States, companies can issue different classes of shares, with Class A shares being the most common. The role and function of different classes of shares is described in more detail in Exhibit 8.

Exhibit 8 Share Class Arrangements at Ford Motor Company[10]

NOTE 21. CAPITAL STOCK AND AMOUNTS PER SHARE

All general voting power is vested in the holders of Common Stock and Class B Stock. Holders of our Common Stock have 60% of the general voting power and holders of our Class B Stock are entitled to such number of votes per share as will give them the remaining 40%. Shares of Common Stock and Class B Stock share equally in dividends when and as paid, with stock dividends payable in shares of stock of the class held.

If liquidated, each share of Common Stock is entitled to the first $0.50 available for distribution to holders of Common Stock and Class B Stock, each share of Class B Stock is entitled to the next $1.00 so available, each share of Common Stock is entitled to the next $0.50 so available, and each share of Common and Class B Stock is entitled to an equal amount thereafter.

2.2 Preference Shares

Preference shares (or preferred stock) rank above common shares with respect to the payment of dividends and the distribution of the company's net assets upon liquidation.[11] However, preference shareholders generally do not share in the operating performance of the company and do not have any voting rights, unless explicitly allowed for at issuance. Preference shares have characteristics of both debt securities and common shares. Similar to the interest payments on debt securities, the dividends on preference shares are fixed and are generally higher than the dividends on common shares. However, unlike interest payments, preference dividends are not contractual obligations of the company. Similar to common shares, preference shares can be perpetual (i.e., no fixed maturity date), can pay dividends indefinitely, and can be callable or putable.

Exhibit 9 provides an example of callable preference shares issued by the GDL Fund to raise capital to redeem the remaining outstanding Series B Preferred shares. In this case, the purchaser of the shares will receive an ongoing dividend from the GDL Fund. If the GDL Fund chooses to buy back the shares, it must do so at the $50 a share liquidation preference price. The purchasers of the shares also have the right to put back the shares to GDL at the $50 a share price.

---

[9] For example, if US$2.00 per share is available for distribution, the Common Stock (Class A) shareholders will receive US$0.50 per share, while the Class B shareholders will receive US$1.50 per share. However, if there is US$3.50 per share available for distribution, the Common Stock shareholders will receive a total of US$1.50 per share and the Class B shareholders will receive a total of US$2.00 per share.

[10] Extracted from Ford Motor Company's *2017 Annual Report* (http://s22.q4cdn.com/857684434/files/doc_financials/2017/annual/Final-Annual-Report-2017.pdf).

[11] Preference shares have a lower priority than debt in the case of liquidation. That is, debt holders have a higher claim on a firm's assets in the event of liquidation and will receive what is owed to them first, followed by preference shareholders and then common shareholders.

Exhibit 9 Callable Stock offering by the GDL Fund[12]

RYE, NY—March 26, 2018—The GDL Fund (NYSE:GDL) (the "Fund") is pleased to announce the completion of a rights offering (the "Offering") in which the Fund issued 2,624,025 Series C Cumulative Puttable and Callable Preferred Shares (the "Series C Preferred"), totaling $131,201,250. Pursuant to the Offering, the Fund issued one non-transferable right (a "Right") for each outstanding Series B Cumulative Puttable and Callable Preferred Share (the "Series B Preferred") of the Fund to Series B Preferred shareholders of record as of February 14, 2018. Holders of Rights were entitled to purchase the Series C Preferred with any combination of cash or surrender of the Series B Preferred at liquidation preference. Therefore, one Right plus $50.00, or one Right plus one share of Series B Preferred with a liquidation value of $50.00 per share, was required to purchase each share of the Series C Preferred. The Offering expired at 5:00 PM Eastern Time on March 20, 2018.

Dividends on preference shares can be cumulative, non-cumulative, participating, non-participating, or some combination thereof (i.e., cumulative participating, cumulative non-participating, non-cumulative participating, non-cumulative non-participating).

Dividends on **cumulative preference shares** accrue so that if the company decides not to pay a dividend in one or more periods, the unpaid dividends accrue and must be paid in full before dividends on common shares can be paid. In contrast, **non-cumulative preference shares** have no such provision. This means that any dividends that are not paid in the current or subsequent periods are forfeited permanently and are not accrued over time to be paid at a later date. However, the company is still not permitted to pay any dividends to common shareholders in the current period unless preferred dividends have been paid first.

Participating preference shares entitle the shareholders to receive the standard preferred dividend plus the opportunity to receive an additional dividend if the company's profits exceed a pre-specified level. In addition, participating preference shares can also contain provisions that entitle shareholders to an additional distribution of the company's assets upon liquidation, above the par (or face) value of the preference shares. **Non-participating preference shares** do not allow shareholders to share in the profits of the company. Instead, shareholders are entitled to receive only a fixed dividend payment and the par value of the shares in the event of liquidation. The use of participating preference shares is much more common for smaller, riskier companies where the possibility of future liquidation is more of a concern to investors.

Preference shares can also be convertible. **Convertible preference shares** entitle shareholders to convert their shares into a specified number of common shares. This conversion ratio is determined at issuance. Convertible preference shares have the following advantages:

- They allow investors to earn a higher dividend than if they invested in the company's common shares.
- They allow investors the opportunity to share in the profits of the company.
- They allow investors to benefit from a rise in the price of the common shares through the conversion option.
- Their price is less volatile than the underlying common shares because the dividend payments are known and more stable.

---

[12] https://www.businesswire.com/news/home/20180326005609/en/GDL-Fund-Successfully-Completes-Offering-Issues-131

As a result, the use of convertible preference shares is a popular financing option in venture capital and private equity transactions in which the issuing companies are considered to be of higher risk and when it may be years before the issuing company "goes public" (i.e., issues common shares to the public).

Exhibit 10 provides examples of the types and characteristics of preference shares as issued by Tsakos Energy Navigation Ltd (TNP.PRE).

Exhibit 10 Examples of Preference Shares Issued by TEN Ltd[13]

Athens, Greece, June 21, 2018—TEN Ltd. ("TEN") (NYSE: TNP), a leading diversified crude, product and LNG tanker operator, today announced the pricing of its public offering of its Series F Fixed-to-Floating Rate Cumulative Redeemable Perpetual Preferred Shares, par value $1.00 per share, liquidation preference $25.00 per share ("Series F Preferred Shares"). TEN will issue 5,400,000 Series F Preferred Shares at a price to the public of $25.00 per share. Dividends will be payable on the Series F Preferred Shares to July 30, 2028 at a fixed rate equal to 9.50% per annum and from July 30, 2028, if not redeemed, at a floating rate. In connection with the offering, TEN has granted the underwriters a 30-day option to purchase 810,000 additional Series F Preferred Shares, which, if exercised in full, would result in total gross proceeds of $155,250,000. TEN intends to use the net proceeds from the offering for general corporate purposes, which may include making vessel acquisitions and/or strategic investments and preferred share redemptions. Following the offering, TEN intends to file an application to list the Series F Preferred Shares on the New York Stock Exchange. The offering is expected to close on or about June 28, 2018.

3

PRIVATE VERSUS PUBLIC EQUITY SECURITIES

c compare and contrast public and private equity securities;

Our discussion so far has focused on equity securities that are issued and traded in public markets and on exchanges. Equity securities can also be issued and traded in private equity markets. **Private equity securities** are issued primarily to institutional investors via non-public offerings, such as private placements. Because they are not listed on public exchanges, there is no active secondary market for these securities. As a result, private equity securities do not have "market determined" quoted prices, are highly illiquid, and require negotiations between investors in order to be traded. In addition, financial statements and other important information needed to determine the fair value of private equity securities may be difficult to obtain because the issuing companies are typically not required by regulatory authorities to publish this information.

There are three primary types of private equity investments: venture capital, leveraged buyouts, and private investment in public equity (or PIPE). **Venture capital** investments provide "seed" or start-up capital, early-stage financing, or mezzanine financing to companies that are in the early stages of development and require additional capital for expansion. These funds are then used to finance the company's product development and growth. Venture capitalists range from family and friends to wealthy individuals and private equity funds. Because the equity securities issued to venture capitalists are not publicly traded, they generally require a commitment of funds for a relatively long period of time; the opportunity to "exit" the investment is typically within 3 to 10 years from the initial start-up. The exit return earned by these private equity investors is based on the price that the securities can be sold for if and when the start-up company first goes public, either via an **initial public offering** (IPO) on the stock market or by being sold to other investors.

---

[13] https://www.tenn.gr/wp-content/uploads/2018/06/tenn062118.pdf

A **leveraged buyout** (LBO) occurs when a group of investors (such as the company's management or a private equity partnership) uses a large amount of debt to purchase all of the outstanding common shares of a publicly traded company. In cases where the group of investors acquiring the company is primarily comprised of the company's existing management, the transaction is referred to as a **management buyout** (MBO). After the shares are purchased, they cease to trade on an exchange and the investor group takes full control of the company. In other words, the company is taken "private" or has been privatized. Companies that are candidates for these types of transactions generally have large amounts of undervalued assets (which can be sold to reduce debt) and generate high levels of cash flows (which are used to make interest and principal payments on the debt). The ultimate objective of a buyout (LBO or MBO) is to restructure the acquired company and later take it "public" again by issuing new shares to the public in the primary market.

The third type of private investment is a **private investment in public equity**, or PIPE.[14] This type of investment is generally sought by a public company that is in need of additional capital quickly and is willing to sell a sizeable ownership position to a private investor or investor group. For example, a company may require a large investment of new equity funds in a short period of time because it has significant expansion opportunities, is facing high levels of indebtedness, or is experiencing a rapid deterioration in its operations. Depending on how urgent the need is and the size of the capital requirement, the private investor may be able to purchase shares in the company at a significant discount to the publicly-quoted market price. Exhibit 11 contains a recent PIPE transaction for the health care company TapImmune, which also included the proposed merger with Maker Therapeutics.

Exhibit 11 Example of a PIPE Transaction[15]

JACKSONVILLE, Florida, June 8, 2018—TapImmune Inc. (NASDAQ: TPIV), a clinical-stage immuno-oncology company, today announced that it has entered into security purchase agreements with certain institutional and accredited investors in connection with a private placement of its equity securities. The private placement will be led by New Enterprise Associates (NEA) with participation from Aisling Capital and Perceptive Advisors, among other new and existing investors. The private placement is expected to be completed concurrently with the closing of the proposed merger between TapImmune Inc. and Marker Therapeutics, Inc., which was previously announced on May 15, 2018.

Upon closing the private placement, TapImmune will issue 17,500,000 shares of its common stock at a price of $4.00 per share. The aggregate offering size, before deducting the placement agent fees and other offering expenses, is expected to be $70 million. Additionally, TapImmune will issue warrants to purchase 13,125,000 shares of TapImmune common stock at an exercise price of $5.00 per share that will be exercisable for a period of five years from the date of issuance. The closing of the transaction, which is subject to the closing of the

(continued)

---

[14] The term PIPE is widely used in the United States and is also used internationally, including in emerging markets.

[15] https://tapimmune.com/2018/06/tapimmune-announces-pricing-of-70-million-private-placement/

Exhibit 11 (Continued)

merger with Marker, the approval by TapImmune's stockholders as required by NASDAQ Stock Market Rules, and other customary closing conditions, is anticipated to occur by the end of the third quarter of 2018.

While the global private equity market is relatively small in comparison to the global public equity market, it has experienced considerable growth over the past three decades. According to a study of the private equity market sponsored by the *World Economic Forum* and spanning the period 1970–2007, approximately US$3.6 trillion in debt and equity were acquired in leveraged buyouts. Of this amount, approximately 75 percent or US$2.7 trillion worth of transactions occurred during 2001–2007.[16] This pace continued with a further US$2.9 trillion in transactions occurring during 2008-2017.[17] While the US and the UK markets were the focus of most private equity investments during the 1980s and 1990s, private equity investments outside of these markets have grown substantially in recent years. In addition, the number of companies operating under private equity ownership has also grown. For example, during the mid-1990s, fewer than 2,000 companies were under LBO ownership compared to more than 20,000 companies that were under LBO ownership globally at the beginning of 2017. The holding period for private equity investments has also increased during this time period from 3 to 5 years (1980s and 1990s) to approximately 10 years.[18]

The move to longer holding periods has given private equity investors the opportunity to more effectively and patiently address any underlying operational issues facing the company and to better manage it for long-term value creation. Because of the longer holding periods, more private equity firms are issuing convertible preference shares because they provide investors with greater total return potential through their dividend payments and the ability to convert their shares into common shares during an IPO.

In operating a publicly traded company, management often feels pressured to focus on short-term results[19] (e.g., meeting quarterly sales and earnings targets from analysts biased toward near-term price performance) instead of operating the company to obtain long-term sustainable revenue and earnings growth. By "going private," management can adopt a more long-term focus and can eliminate certain costs that are necessary to operate a publicly traded company—such as the cost of meeting regulatory and stock exchange filing requirements, the cost of maintaining investor relations departments to communicate with shareholders and the media, and the cost of holding quarterly analyst conference calls.

As described above, public equity markets are much larger than private equity networks and allow companies more opportunities to raise capital that is subsequently actively traded in secondary markets. By operating under public scrutiny, companies are incentivized to be more open in terms of corporate governance and executive compensation to ensure that they are acting for the benefit of shareholders. In fact, some studies have shown that private equity firms score lower in terms of corporate governance effectiveness, which may be attributed to the fact that shareholders, analysts, and other stakeholders are able to influence management when corporate governance and other policies are public.

---

[16] Stromberg (2008).

[17] https://www.statista.com/statistics/270195/global-private-equity-deal-value/

[18] See, for example, Bailey, Wirth, and Zapol (2005).

[19] See, for example, Graham, Harvey, and Rajgopal (2005).

NON-DOMESTIC EQUITY SECURITIES

d describe methods for investing in non-domestic equity securities;

Technological innovations and the growth of electronic information exchanges (electronic trading networks, the internet, etc.) have accelerated the integration and growth of global financial markets. As detailed previously, global capital markets have expanded at a much more rapid rate than global GDP in recent years; both primary and secondary international markets have benefited from the enhanced ability to rapidly and openly exchange information. Increased integration of equity markets has made it easier and less expensive for companies to raise capital and to expand their shareholder base beyond their local market. Integration has also made it easier for investors to invest in companies that are located outside of their domestic markets. This has enabled investors to further diversify and improve the risk and return characteristics of their portfolios by adding a class of assets with lower correlations to local country assets.

One barrier to investing globally is that many countries still impose "foreign restrictions" on individuals and companies from other countries that want to invest in their domestic companies. There are three primary reasons for these restrictions. The first is to limit the amount of control that foreign investors can exert on domestic companies. For example, some countries prevent foreign investors from acquiring a majority interest in domestic companies. The second is to give domestic investors the opportunity to own shares in the foreign companies that are conducting business in their country. For example, the Swedish home furnishings retailer IKEA abandoned efforts to invest in parts of the Asia/Pacific region because local governments did not want IKEA to maintain complete ownership of its stores. The third reason is to reduce the volatility of capital flows into and out of domestic equity markets. For example, one of the main consequences of the Asian Financial Crisis in 1997–98 was the large outflow of capital from such emerging market countries as Thailand, Indonesia, and South Korea. These outflows led to dramatic declines in the equity markets of these countries and significant currency devaluations and resulted in many governments placing restrictions on capital flows. Today, many of these same markets have built up currency reserves to better withstand capital outflows inherent in economic contractions and periods of financial market turmoil.

Studies have shown that reducing restrictions on foreign ownership has led to improved equity market performance over the long term.[20] Although restrictions vary widely, more countries are allowing increasing levels of foreign ownership. For example, Australia has sought tax reforms as a means to encourage international demand for its managed funds in order to increase its role as an international financial center.

Over the past two decades, three trends have emerged: a) an increasing number of companies have issued shares in markets outside of their home country; b) the number of companies whose shares are traded in markets outside of their home has increased; and c) an increasing number of companies are dual listed, which means that their shares are simultaneously issued and traded in two or more markets. Companies located in emerging markets have particularly benefited from these trends because they no longer have to be concerned with capital constraints or lack of liquidity in their domestic markets. These companies have found it easier to raise capital in the markets of developed countries because these markets generally have higher levels of liquidity and more stringent financial reporting requirements and accounting standards. Being listed on an international exchange has a number of benefits. It can increase investor awareness about the company's products and services, enhance the liquidity of the company's shares, and increase corporate transparency because of the additional market exposure and the need to meet a greater number of filing requirements.

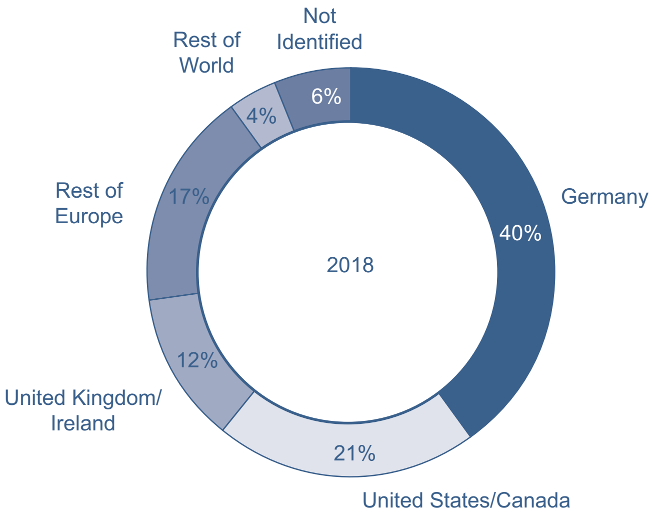
---

[20] See, for example, Henry and Chari (2004).

Technological advancements have made it easier for investors to trade shares in foreign markets. The German insurance company Allianz SE recently delisted its shares from the NYSE and certain European markets because international investors increasingly traded its shares on the Frankfurt Stock Exchange. Exhibit 12 illustrates the extent to which the institutional shareholder base at BASF, a large German chemical corporation, has become increasingly global in nature.

Exhibit 12 Example of Increased Globalization of Share Ownership[21]

BASF is one of the largest publicly owned companies with over 500,000 shareholders and a high free float. An analysis of the shareholder structure carried out in March 2018 showed that, at 21% of share capital, the United States and Canada made up the largest regional group of institutional investors. Institutional investors from Germany made up 12%. Shareholders from United Kingdom and Ireland held 12% of BASF shares, while a further 17% are held by institutional investors from the rest of Europe. Around 28% of the company's share capital is held by private investors, most of whom are resident in Germany.



4.1 Direct Investing

Investors can use a variety of methods to invest in the equity of companies outside of their local market. The most obvious is to buy and sell securities directly in foreign markets. However, this means that all transactions—including the purchase and sale of shares, dividend payments, and capital gains—are in the company's, not the investor's, domestic currency. In addition, investors must be familiar with the trading, clearing, and settlement regulations and procedures of that market. Investing directly often results in less transparency and more volatility because audited financial information may not be provided on a regular basis and the market may be less liquid. Alternatively, investors can use such securities as depository receipts and global registered shares, which represent the equity of international companies and are traded on local exchanges and in the local currencies. With these securities, investors have to worry less about currency conversions (price quotations and dividend payments are in the investor's local currency), unfamiliar market practices, and differences in accounting standards. The sections that follow discuss various securities that investors can invest in outside of their home market.

---

[21] Adapted from BASF's investor relations website (www.basf.com). **Free float** refers to the extent that shares are readily and freely tradable in the secondary market.

4.2 Depository Receipts

A **depository receipt**[22] (DR) is a security that trades like an ordinary share on a local exchange and represents an economic interest in a foreign company. It allows the publicly listed shares of a foreign company to be traded on an exchange outside its domestic market. A depository receipt is created when the equity shares of a foreign company are deposited in a bank (i.e., the depository) in the country on whose exchange the shares will trade. The depository then issues receipts that represent the shares that were deposited. The number of receipts issued and the price of each DR is based on a ratio, which specifies the number of depository receipts to the underlying shares. Consequently, a DR may represent one share of the underlying stock, many shares of the underlying stock, or a fractional share of the underlying stock. The price of each DR will be affected by factors that affect the price of the underlying shares, such as company fundamentals, market conditions, analysts' recommendations, and exchange rate movements. In addition, any short-term valuation discrepancies between shares traded on multiple exchanges represent a quick arbitrage profit opportunity for astute traders to exploit. The responsibilities of the **depository bank** that issues the receipts include acting as custodian and as a registrar. This entails handling dividend payments, other taxable events, stock splits, and serving as the transfer agent for the foreign company whose securities the DR represents. The Bank of New York Mellon is the largest depository bank; however, Deutsche Bank, JPMorgan, and Citibank also offer depository services.[23]

A DR can be **sponsored** or **unsponsored**. A sponsored DR is when the foreign company whose shares are held by the depository has a direct involvement in the issuance of the receipts. Investors in sponsored DRs have the same rights as the direct owners of the common shares (e.g., the right to vote and the right to receive dividends). In contrast, with an unsponsored DR, the underlying foreign company has no involvement with the issuance of the receipts. Instead, the depository purchases the foreign company's shares in its domestic market and then issues the receipts through brokerage firms in the depository's local market. In this case, the depository bank, not the investors in the DR, retains the voting rights. Sponsored DRs are generally subject to greater reporting requirements than unsponsored DRs. In the United States, for example, sponsored DRs must be registered (meet the reporting requirements) with the US Securities and Exchange Commission (SEC). Exhibit 13 contains an example of a sponsored DR issued by Alibaba in September 2014.

Exhibit 13 Sponsored Depository Receipts[24]

NEW YORK—(BUSINESS WIRE)—Citi today announced that Alibaba Group Holding Limited ("Alibaba Group") has appointed Citi's Issuer Services business, acting through Citibank, N.A., as the depositary bank for its American Depositary

(continued)

---

[22] Note that the spellings *depositary* and *depository* are used interchangeably in financial markets. In this reading, we use the spelling *depository* throughout.

[23] Boubakri, Cosset, and Samet (2010).

[24] https://www.businesswire.com/news/home/20140924005984/en/Citi-Appointed-Depositary-Bank-Alibaba-Group-Holding

Exhibit 13 (Continued)

Receipt ("ADR") program. Alibaba Group's ADRs, which began trading on September 19, 2014, represent the largest Depositary Receipt program in initial public offering market history.

Alibaba Group's ADR program was established through a $25.03 billion initial public offering of 368,122,000 American Depositary Shares ("ADSs"), representing ordinary shares of Alibaba Group, which was priced at $68 per ADS on September 18, 2014. The IPO ranks as the largest in history. The ADRs are listed on the New York Stock Exchange (the "NYSE") under the trading symbol BABA. Each ADS represents one ordinary share of the Company. In its role as depositary bank, Citibank will hold the underlying ordinary shares through its local custodian and issue ADSs representing such shares. Alibaba Group's ADSs trade on the NYSE in ADR form.

There are two types of depository receipts: Global depository receipts (GDRs) and American depository receipts (ADRs), which are described below.

4.2.1 *Global Depository Receipts*

A **global depository receipt** (GDR) is issued outside of the company's home country and outside of the United States. The depository bank that issues GDRs is generally located (or has branches) in the countries on whose exchanges the shares are traded. A key advantage of GDRs is that they are not subject to the foreign ownership and capital flow restrictions that may be imposed by the issuing company's home country because they are sold outside of that country. The issuing company selects the exchange where the GDR is to be traded based on such factors as investors' familiarity with the company or the existence of a large international investor base. The London and Luxembourg exchanges were the first ones to trade GDRs. Some other stock exchanges trading GDRs are the Dubai International Financial Exchange and the Singapore Stock Exchange. Currently, the London and Luxembourg exchanges are where most GDRs are traded because they can be issued in a more timely manner and at a lower cost. Regardless of the exchange they are traded on, the majority of GDRs are denominated in US dollars, although the number of GDRs denominated in pound sterling and euros is increasing. Note that although GDRs cannot be listed on US exchanges, they can be privately placed with institutional investors based in the United States.

4.2.2 *American Depository Receipts*

An **American depository receipt** (ADR) is a US dollar-denominated security that trades like a common share on US exchanges. First created in 1927, ADRs are the oldest type of depository receipts and are currently the most commonly traded depository receipts. They enable foreign companies to raise capital from US investors. Note that an ADR is one form of a GDR; however, not all GDRs are ADRs because GDRs cannot be publicly traded in the United States. The term **American depository share** (ADS) is often used in tandem with the term ADR. A depository share is a security that is actually traded in the issuing company's domestic market. That is, while American depository receipts are the certificates that are traded on US markets, American depository shares are the underlying shares on which these receipts are based.

There are four primary types of ADRs, with each type having different levels of corporate governance and filing requirements. Level I Sponsored ADRs trade in the over-the-counter (OTC) market and do not require full registration with the Securities and Exchange Commission (SEC). Level II and Level III Sponsored ADRs can trade on the New York Stock Exchange (NYSE), NASDAQ, and American Stock Exchange (AMEX). Level II and III ADRs allow companies to raise capital and make acquisitions using these securities. However, the issuing companies must fulfill all SEC requirements.

The fourth type of ADR, an SEC Rule 144A or a Regulation S depository receipt, does not require SEC registration. Instead, foreign companies are able to raise capital by privately placing these depository receipts with qualified institutional investors or to offshore non-US investors. Exhibit 14 summarizes the main features of ADRs.

Exhibit 14 Summary of the Main Features of American Depository Receipts

	Level I (Unlisted)	Level II (Listed)	Level III (Listed)	Rule 144A (Unlisted)
Objectives	Develop and broaden US investor base with existing shares	Develop and broaden US investor base with existing shares	Develop and broaden US investor base with existing/new shares	Access qualified institutional buyers (QIBs)
Raising capital on US markets?	No	No	Yes, through public offerings	Yes, through private placements to QIBs
SEC registration	Form F-6	Form F-6	Forms F-1 and F-6	None
Trading	Over the counter (OTC)	NYSE, NASDAQ, or AMEX	NYSE, NASDAQ, or AMEX	Private offerings, resales, and trading through automated linkages such as PORTAL
Listing fees	Low	High	High	Low
Size and earnings requirements	None	Yes	Yes	None

Source: Adapted from Boubakri, Cosset, and Samet (2010): Table 1.

More than 2,000 DRs, from over 80 countries, currently trade on US exchanges. Based on current statistics, the total market value of DRs issued and traded is estimated at approximately US$2 trillion, or 15 percent of the total dollar value of equities traded in US markets.[25]

4.2.3 *Global Registered Share*

A **global registered share** (GRS) is a common share that is traded on different stock exchanges around the world in different currencies. Currency conversions are not needed to purchase or sell them, because identical shares are quoted and traded in different currencies. Thus, the same share purchased on the Swiss exchange in Swiss francs can be sold on the Tokyo exchange for Japanese yen. As a result, GRSs offer more flexibility than depository receipts because the shares represent an actual ownership interest in the company that can be traded anywhere and currency conversions are not needed to purchase or sell them. GRSs were created and issued by Daimler Chrysler in 1998 and by UBS AG in 2011.

---

[25] *JPMorgan Depositary Receipt Guide* (2005):4.

4.2.4 *Basket of Listed Depository Receipts*

Another type of global security is a **basket of listed depository receipts** (BLDR), which is an exchange-traded fund (ETF) that represents a portfolio of depository receipts. An ETF is a security that tracks an index but trades like an individual share on an exchange. An equity-ETF is a security that contains a portfolio of equities that tracks an index. It trades throughout the day and can be bought, sold, or sold short, just like an individual share. Like ordinary shares, ETFs can also be purchased on margin and used in hedging or arbitrage strategies. The BLDR is a specific class of ETF security that consists of an underlying portfolio of DRs and is designed to track the price performance of an underlying DR index. For example, the Invesco BLDRS Asia 50 ADR Index Fund is a capitalization-weighted ETF designed to track the performance of 50 Asian market-based ADRs.

5

RISK AND RETURN CHARACTERISTICS

e compare the risk and return characteristics of different types of equity securities;

Different types of equity securities have different ownership claims on a company's net assets. The type of equity security and its features affect its risk and return characteristics. The following sections discuss the different return and risk characteristics of equity securities.

5.1 Return Characteristics of Equity Securities

There are two main sources of equity securities' total return: price change (or capital gain) and dividend income. The price change represents the difference between the purchase price (P_{t-1}) and the sale price (P_t) of a share at the end of time $t - 1$ and t, respectively. Cash or stock dividends (D_t) represent distributions that the company makes to its shareholders during period t. Therefore, an equity security's total return is calculated as:

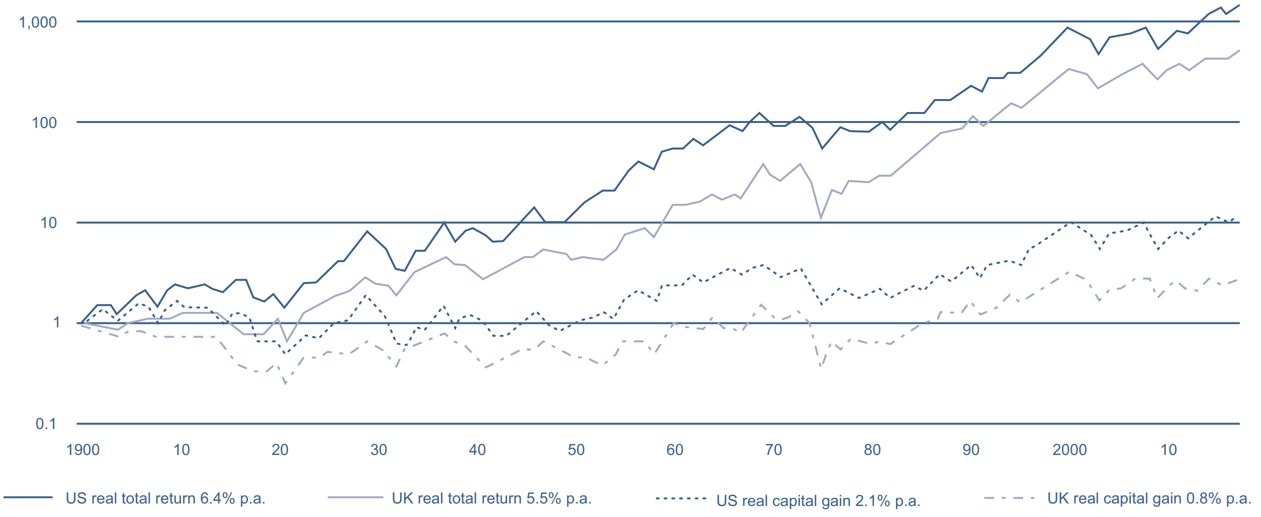
$$\text{Total return, } R_t = (P_t - P_{t-1} + D_t)/P_{t-1} \tag{1}$$

For non-dividend-paying stocks, the total return consists of price appreciation only. Companies that are in the early stages of their life cycle generally do not pay dividends because earnings and cash flows are reinvested to finance the company's growth. In contrast, companies that are in the mature phase of their life cycle may not have as many profitable growth opportunities; therefore, excess cash flows are often returned to investors via the payment of regular dividends or through share repurchases.

For investors who purchase depository receipts or foreign shares directly, there is a third source of return: **foreign exchange gains (or losses)**. Foreign exchange gains arise because of the change in the exchange rate between the investor's currency and the currency that the foreign shares are denominated in. For example, US investors who purchase the ADRs of a Japanese company will earn an additional return if the yen appreciates relative to the US dollar. Conversely, these investors will earn a lower total return if the yen depreciates relative to the US dollar. For example, if the total return for a Japanese company was 10 percent in Japan and the yen depreciated by 10 percent against the US dollar, the total return of the ADR would be (approximately) 0 percent. If the yen had instead appreciated by 10 percent against the US dollar, the total return of the ADR would be (approximately) 20 percent.

Investors that only consider price appreciation overlook an important source of return: the compounding that results from reinvested dividends. Reinvested dividends are cash dividends that the investor receives and uses to purchase additional shares. As Exhibit 15 shows, in the long run total returns on equity securities are dramatically influenced by the compounding effect of reinvested dividends. Between 1900 and 2016, US$1 invested in US equities in 1900 would have grown in *real* terms to US$1,402 with dividends reinvested, but to just US$11.9 when taking only the price appreciation or capital gain into account. This corresponds to a real compounded return of 6.4 percent per year with dividends reinvested, versus only 2.1 percent per year without dividends reinvested. The comparable ending real wealth for bonds and bills are US$9.8 and US$2.60, respectively. These ending real wealth figures correspond to annualized real compounded returns of 2.0 percent on bonds and 0.8 percent on bills.

Exhibit 15 Impact of Reinvested Dividends on Cumulative Real Returns in the US and UK Equity Market: 1900–2016



Source: Dimson, Marsh, and Staunton (2017). This chart is updated annually and can be found at http://publications.credit-suisse.com/index.cfm/publikationen-shop/research-institute/.

5.2 Risk of Equity Securities

The risk of any security is based on the uncertainty of its future cash flows. The greater the uncertainty of its future cash flows, the greater the risk and the more variable or volatile the security's price. As discussed above, an equity security's total return is determined by its price change and dividends. Therefore, the risk of an equity security can be defined as the uncertainty of its expected (or future) total return. Risk is most often measured by calculating the standard deviation of the equity's expected total return.

A variety of different methods can be used to estimate an equity's expected total return and risk. One method uses the equity's average historical return and the standard deviation of this return as proxies for its expected future return and risk. Another method involves estimating a range of future returns over a specified period of time, assigning probabilities to those returns, and then calculating an expected return and a standard deviation of return based on this information.

The type of equity security, as well as its characteristics, affects the uncertainty of its future cash flows and therefore its risk. In general, preference shares are less risky than common shares for three main reasons:

1 Dividends on preference shares are known and fixed, and they account for a large portion of the preference shares' total return. Therefore, there is less uncertainty about future cash flows.
2 Preference shareholders receive dividends and other distributions before common shareholders.
3 The amount preference shareholders will receive if the company is liquidated is known and fixed as the par (or face) value of their shares. However, there is no guarantee that investors will receive that amount if the company experiences financial difficulty.

With common shares, however, a larger portion of shareholders' total return (or all of their total return for non-dividend shares) is based on future price appreciation and future dividends are unknown. If the company is liquidated, common shareholders will receive whatever amount (if any) is remaining after the company's creditors and preference shareholders have been paid. In summary, because the uncertainty surrounding the total return of preference shares is less than common shares, preference shares have lower risk and lower expected return than common shares.

It is important to note that some preference shares and common shares can be riskier than others because of their associated characteristics. For example, from an investor's point of view, putable common or preference shares are less risky than their callable or non-callable counterparts because they give the investor the option to sell the shares to the issuer at a pre-determined price. This pre-determined price establishes a minimum price that investors will receive and reduces the uncertainty associated with the security's future cash flow. As a result, putable shares generally pay a lower dividend than non-putable shares.

Because the major source of total return for preference shares is dividend income, the primary risk affecting all preference shares is the uncertainty of future dividend payments. Regardless of the preference shares' features (callable, putable, cumulative, etc.), the greater the uncertainty surrounding the issuer's ability to pay dividends, the greater risk. Because the ability of a company to pay dividends is based on its future cash flows and net income, investors try to estimate these amounts by examining past trends or forecasting future amounts. The more earnings and the greater amount of cash flow that the company has had, or is expected to have, the lower the uncertainty and risk associated with its ability to pay future dividends.

Callable common or preference shares are riskier than their non-callable counterparts because the issuer has the option to redeem the shares at a pre-determined price. Because the call price limits investors' potential future total return, callable shares generally pay a higher dividend to compensate investors for the risk that the shares could be called in the future. Similarly, putable preference shares have lower risk than non-putable preference shares. Cumulative preference shares have lower risk than non-cumulative preference shares because the cumulative feature gives investors the right to receive any unpaid dividends before any dividends can be paid to common shareholders.

6 EQUITY AND COMPANY VALUE

f explain the role of equity securities in the financing of a company's assets;

g contrast the market value and book value of equity securities;

h compare a company's cost of equity, its (accounting) return on equity, and investors' required rates of return.

Companies issue equity securities on primary markets to raise capital and increase liquidity. This additional liquidity also provides the corporation an additional "currency" (its equity), which it can use to make acquisitions and provide stock option-based incentives to employees. The primary goal of raising capital is to finance the company's revenue-generating activities in order to increase its net income and maximize the wealth of its shareholders. In most cases, the capital that is raised is used to finance the purchase of long-lived assets, capital expansion projects, research and development, the entry into new product or geographic regions, and the acquisition of other companies. Alternatively, a company may be forced to raise capital to ensure that it continues to operate as a going concern. In these cases, capital is raised to fulfill regulatory requirements, improve capital adequacy ratios, or to ensure that debt covenants are met.

The ultimate goal of management is to increase the book value (shareholders' equity on a company's balance sheet) of the company and maximize the market value of its equity. Although management actions can directly affect the book value of the company (by increasing net income or by selling or purchasing its own shares), they can only indirectly affect the market value of its equity. The book value of a company's equity—the difference between its total assets and total liabilities—increases when the company retains its net income. The more net income that is earned and retained, the greater the company's book value of equity. Because management's decisions directly influence a company's net income, they also directly influence its book value of equity.

The market value of the company's equity, however, reflects the collective and differing expectations of investors concerning the amount, timing, and uncertainty of the company's future cash flows. Rarely will book value and market value be equal. Although management may be accomplishing its objective of increasing the company's book value, this increase may not be reflected in the market value of the company's equity because it does not affect investors' expectations about the company's future cash flows. A key measure that investors use to evaluate the effectiveness of management in increasing the company's book value is the accounting return on equity.

6.1 Accounting Return on Equity

Return on equity (ROE) is the primary measure that equity investors use to determine whether the management of a company is effectively and efficiently using the capital they have provided to generate profits. It measures the total amount of net income available to common shareholders generated by the total equity capital invested in the company. It is computed as net income available to ordinary shareholders (i.e., after preferred dividends have been deducted) divided by the average total book value of equity (BVE). That is:

$$\text{ROE}_t = \frac{\text{NI}_t}{\text{Average BVE}_t} = \frac{\text{NI}_t}{\left(\text{BVE}_t + \text{BVE}_{t-1}\right)/2} \quad (2)$$

where NI_t is the net income in year t and the average book value of equity is computed as the book values at the beginning and end of year t divided by 2. Return on equity assumes that the net income produced in the current year is generated by the equity existing at the beginning of the year and any new equity that was invested during the year. Note that some formulas only use shareholders' equity at the beginning of year t

(that is, the end of year t – 1) in the denominator. This assumes that only the equity existing at the beginning of the year was used to generate the company's net income during the year. That is:

$$\text{ROE}_t = \frac{\text{NI}_t}{\text{BVE}_{t-1}} \tag{3}$$

Both formulas are appropriate to use as long as they are applied consistently. For example, using beginning of the year book value is appropriate when book values are relatively stable over time or when computing ROE for a company annually over a period of time. Average book value is more appropriate if a company experiences more volatile year-end book values or if the industry convention is to use average book values in calculating ROE.

One caveat to be aware of when computing and analyzing ROE is that net income and the book value of equity are directly affected by management's choice of accounting methods, such as those relating to depreciation (straight line versus accelerated methods) or inventories (first in, first out versus weighted average cost). Different accounting methods can make it difficult to compare the return on equity of companies even if they operate in the same industry. It may also be difficult to compare the ROE of the same company over time if its accounting methods have changed during that time.

Exhibit 16 contains information on the net income and total book value of shareholders' equity for three **blue chip** (widely held large market capitalization companies that are considered financially sound and are leaders in their respective industry or local stock market) pharmaceutical companies: Pfizer, Novartis AG, and GlaxoSmithKline. The data are for their financial years ending December 2015 through December 2017.[26]

Exhibit 16 Net Income and Book Value of Equity for Pfizer, Novartis AG, and GlaxoSmithKline (in Thousands of US Dollars)

	Financial Year Ending		
	31 Dec 2015	31 Dec 2016	31 Dec 2017
Pfizer			
Net income	$6,960,000	$7,215,000	$21,308,000
Total stockholders' equity	$64,998,000	$59,840,000	$71,287,000
Novartis AG			
Net income	$17,783,000	$6,712,000	$7,703,000
Total stockholders' equity	$77,122,000	$74,891,000	$74,227,000
GlaxoSmithKline			
Net income	$12,420,000	$1,126,000	$2,070,700
Total stockholders' equity	$11,309,250	$6,127,800	$4,715,800

Using the average book value of equity, the return on equity for Pfizer for the years ending December 2016 and 2017 can be calculated as:

---

[26] Pfizer uses US GAAP to prepare its financial statements; Novartis and GlaxoSmithKline use International Financial Reporting Standards. Therefore, it would be inappropriate to compare the ROE of Pfizer to that of Novartis or GlaxoSmithKline.

Return on equity for the year ending December 2016

$$\text{ROE}_{2016} = \frac{\text{NI}_{2016}}{\left(\text{BVE}_{2015} + \text{BVE}_{2016}\right)/2} = \frac{7{,}215{,}000}{(64{,}998{,}000 + 59{,}840{,}000)/2} = 11.6\%$$

Return on equity for the year ending December 2017

$$\text{ROE}_{2017} = \frac{\text{NI}_{2017}}{\left(\text{BVE}_{2016} + \text{BVE}_{2017}\right)/2} = \frac{21{,}308{,}000}{(59{,}840{,}000 + 71{,}287{,}000)/2} = 32.5\%$$

Exhibit 17 summarizes the return on equity for Novartis and GlaxoSmithKline in addition to Pfizer for 2016 and 2017.

Exhibit 17 Return on Equity for Pfizer, Novartis AG, and GlaxoSmithKline

	31 Dec 2016 (%)	31 Dec 2017 (%)
Pfizer	11.6	32.5
Novartis AG	8.8	10.3
GlaxoSmithKline	12.9	38.2

In the case of Pfizer, the ROE of 32.5 percent in 2017 indicates that the company was able to generate a return (profit) of US$0.325 on every US$1.00 of capital invested by shareholders. GlaxoSmithKline almost tripled its return on equity over this period, from 12.9 percent to 38.2 percent. Novartis's ROE remained relatively unchanged.

ROE can increase if net income increases at a faster rate than shareholders' equity or if net income decreases at a slower rate than shareholders' equity. In the case of GlaxoSmithKline, ROE almost tripled between 2016 and 2017 due to its net income almost doubling during the period and due to its average shareholder's fund decreasing by almost 45 percent during the period. Stated differently, in 2017 compared to 2016, GlaxoSmithKline was significantly more effective in using its equity capital to generate profits. In the case of Pfizer, its ROE increased dramatically from 11.6 percent to 32.5 percent in 2017 versus 2016 even though its average shareholder equity increased by around 5 percent due to a nearly tripling of net income during the period.

An important question to ask is whether an increasing ROE is always good. The short answer is, "it depends." One reason ROE can increase is if net income decreases at a slower rate than shareholders' equity, which is not a positive sign. In addition, ROE can increase if the company issues debt and then uses the proceeds to repurchase some of its outstanding shares. This action will increase the company's leverage and make its equity riskier. Therefore, it is important to examine the source of changes in the company's net income *and* shareholders' equity over time. The DuPont formula, which is discussed in a separate reading, can be used to analyze the sources of changes in a company's ROE.

The book value of a company's equity reflects the historical operating and financing decisions of its management. The market value of the company's equity reflects these decisions as well as investors' collective assessment and expectations about the company's future cash flows generated by its positive net present value investment opportunities. If investors believe that the company has a large number of these future cash flow-generating investment opportunities, the market value of the company's equity will exceed its book value. Exhibit 18 shows the market price per share, the total number of shares outstanding, and the total book value of shareholders' equity for

Pfizer, Novartis AG, and GlaxoSmithKline at the end of December 2017. This exhibit also shows the total market value of equity (or market capitalization) computed as the number of shares outstanding multiplied by the market price per share.

Exhibit 18 **Market Information for Pfizer, Novartis AG, and GlaxoSmithKline (in Thousands of US Dollars except market price)**

	Pfizer	Novartis AG	GlaxoSmithKline
Market price	$35.74	$90.99	$18.39
Total shares outstanding	5,952,900	2,317,500	4,892,200
Total shareholders' equity	$71,287,000	$74,227,000	$4,715,800
Total market value of equity	$212,756,646	$210,869,325	$89,967,558

Note that in Exhibit 18, the total market value of equity for Pfizer is computed as:

Market value of equity = Market price per share × Shares outstanding

Market value of equity = US$35.74 × 5,952,900 = US$212,756,646.

The book value of equity per share for Pfizer can be computed as:

Book value of equity per share = Total shareholders' equity/Shares outstanding

Book value of equity per share = US$71,287,000/5,952,900 = US$11.98.

A useful ratio to compute is a company's price-to-book ratio, which is also referred to as the market-to-book ratio. This ratio provides an indication of investors' expectations about a company's future investment and cash flow-generating opportunities. The larger the price-to-book ratio (i.e., the greater the divergence between market value per share and book value per share), the more favorably investors will view the company's future investment opportunities. For Pfizer the price-to-book ratio is:

Price-to-book ratio = Market price per share/Book value of equity per share

Price-to-book ratio = US$35.74/US$11.98= 2.98

Exhibit 19 contains the market price per share, book value of equity per share, and price-to-book ratios for Novartis and GlaxoSmithKline in addition to Pfizer.

Exhibit 19 **Pfizer, Novartis AG, and GlaxoSmithKline**

	Pfizer	Novartis AG	GlaxoSmithKline
Market price per share	$35.74	$90.99	$18.39
Book value of equity per share	$11.98	$32.03	$0.96
Price-to-book ratio	2.98	2.84	19.16

The market price per share of all three companies exceeds their respective book values, so their price-to-book ratios are all greater than 1.00. However, there are significant differences in the sizes of their price-to-book ratios. GlaxoSmithKline has the largest price-to-book ratio, while the price-to-book ratios of Pfizer and Novartis are similar to each other. This suggests that investors believe that GlaxoSmithKline has substantially higher future growth opportunities than either Pfizer or Novartis.

It is not appropriate to compare the price-to-book ratios of companies in different industries because their price-to-book ratios also reflect investors' outlook for the industry. Companies in high growth industries, such as technology, will generally have higher price-to-book ratios than companies in slower growth (i.e., mature) industries, such as heavy equipment. Therefore, it is more appropriate to compare the price-to-book ratios of companies in the same industry. A company with relatively high growth opportunities compared to its industry peers would likely have a higher price-to-book ratio than the average price-to-book ratio of the industry.

Book value and return on equity are useful in helping analysts determine value but can be limited as a primary means to estimate a company's true or intrinsic value, which is the present value of its future projected cash flows. In Exhibit 20, Warren Buffett, one of the most successful investors in the world and CEO of Berkshire Hathaway, provides an explanation of the differences between the book value of a company and its intrinsic value in a letter to shareholders. As discussed above, market value reflects the collective and differing expectations of investors concerning the amount, timing, and uncertainty of a company's future cash flows. A company's intrinsic value can only be estimated because it is impossible to predict the amount and timing of its future cash flows. However, astute investors—such as Buffett—have been able to profit from discrepancies between their estimates of a company's intrinsic value and the market value of its equity.

Exhibit 20 Book Value versus Intrinsic Value[27]

We regularly report our per-share book value, an easily calculable number, though one of limited use. Just as regularly, we tell you that what counts is intrinsic value, a number that is impossible to pinpoint but essential to estimate.

For example, in 1964, we could state with certitude that Berkshire's per-share book value was $19.46. However, that figure considerably overstated the stock's intrinsic value since all of the company's resources were tied up in a sub-profitable textile business. Our textile assets had neither going-concern nor liquidation values equal to their carrying values. In 1964, then, anyone inquiring into the soundness of Berkshire's balance sheet might well have deserved the answer once offered up by a Hollywood mogul of dubious reputation: "Don't worry, the liabilities are solid."

Today, Berkshire's situation has reversed: Many of the businesses we control are worth far more than their carrying value. (Those we don't control, such as Coca-Cola or Gillette, are carried at current market values.) We continue to give you book value figures, however, because they serve as a rough, understated, tracking measure for Berkshire's intrinsic value.

We define intrinsic value as the discounted value of the cash that can be taken out of a business during its remaining life. Anyone calculating intrinsic value necessarily comes up with a highly subjective figure that will change both as estimates of future cash flows are revised and as interest rates move. Despite its fuzziness, however, intrinsic value is all-important and is the only logical way to evaluate the relative attractiveness of investments and businesses.

To see how historical input (book value) and future output (intrinsic value) can diverge, let's look at another form of investment, a college education. Think of the education's cost as its "book value." If it is to be accurate, the cost should include the earnings that were foregone by the student because he chose college rather than a job.

(continued)

---

[27] Extracts from Berkshire Hathaway's *2008 Annual Report* (www.berkshirehathaway.com).

Exhibit 20 (Continued)

For this exercise, we will ignore the important non-economic benefits of an education and focus strictly on its economic value. First, we must estimate the earnings that the graduate will receive over his lifetime and subtract from that figure an estimate of what he would have earned had he lacked his education. That gives us an excess earnings figure, which must then be discounted, at an appropriate interest rate, back to graduation day. The dollar result equals the intrinsic economic value of the education.

6.2 The Cost of Equity and Investors' Required Rates of Return

When companies issue debt (or borrow from a bank) or equity securities, there is a cost associated with the capital that is raised. In order to maximize profitability and shareholder wealth, companies attempt to raise capital efficiently so as to minimize these costs.

When a company issues debt, the cost it incurs for the use of these funds is called the cost of debt. The cost of debt is relatively easy to estimate because it reflects the periodic interest (or coupon) rate that the company is contractually obligated to pay to its bondholders (lenders). When a company raises capital by issuing equity, the cost it incurs is called the cost of equity. Unlike debt, however, the company is not contractually obligated to make any payments to its shareholders for the use of their funds. As a result, the cost of equity is more difficult to estimate.

Investors require a return on the funds they provide to the company. This return is called the investor's minimum required rate of return. When investors purchase the company's debt securities, their minimum required rate of return is the periodic rate of interest they charge the company for the use of their funds. Because all of the bondholders receive the same periodic rate of interest, their required rate of return is the same. Therefore, the company's cost of debt and the investors' minimum required rate of return on the debt are the same.

When investors purchase the company's equity securities, their minimum required rate of return is based on the future cash flows they expect to receive. Because these future cash flows are both uncertain and unknown, the investors' minimum required rate of return must be estimated. In addition, the minimum required return may differ across investors based on their expectations about the company's future cash flows. As a result, the company's cost of equity may be different from the investors' minimum required rate of return on equity.[28] Because companies try to raise capital at the lowest possible cost, the company's cost of equity is often used as a proxy for the investors' *minimum* required rate of return.

In other words, the cost of equity can be thought of as the minimum expected rate of return that a company must offer its investors to purchase its shares in the primary market and to maintain its share price in the secondary market. If this expected rate of return is not maintained in the secondary market, then the share price will adjust so that it meets the minimum required rate of return demanded by investors. For example, if investors require a higher rate of return on equity than the company's cost of equity, they would sell their shares and invest their funds elsewhere resulting in a decline in the company's share price. As the share price declined, the cost of equity would increase to reach the higher rate of return that investors require.

---

[28] Another important factor that can cause a firm's cost of equity to differ from investors' required rate of return on equity is the flotation cost associated with equity

Two models commonly used to estimate a company's cost of equity (or investors' minimum required rate of return) are the dividend discount model (DDM) and the capital asset pricing model (CAPM). These models are discussed in detail in other curriculum readings.

The cost of debt (after tax) and the cost of equity (i.e., the minimum required rates of return on debt and equity) are integral components of the capital budgeting process because they are used to estimate a company's weighted average cost of capital (WACC). Capital budgeting is the decision-making process that companies use to evaluate potential long-term investments. The WACC represents the minimum required rate of return that the company must earn on its long-term investments to satisfy all providers of capital. The company then chooses among those long-term investments with expected returns that are greater than its WACC.

SUMMARY

Equity securities play a fundamental role in investment analysis and portfolio management. The importance of this asset class continues to grow on a global scale because of the need for equity capital in developed and emerging markets, technological innovation, and the growing sophistication of electronic information exchange. Given their absolute return potential and ability to impact the risk and return characteristics of portfolios, equity securities are of importance to both individual and institutional investors.

This reading introduces equity securities and provides an overview of global equity markets. A detailed analysis of their historical performance shows that equity securities have offered average real annual returns superior to government bills and bonds, which have provided average real annual returns that have only kept pace with inflation. The different types and characteristics of common and preference equity securities are examined, and the primary differences between public and private equity securities are outlined. An overview of the various types of equity securities listed and traded in global markets is provided, including a discussion of their risk and return characteristics. Finally, the role of equity securities in creating company value is examined as well as the relationship between a company's cost of equity, its accounting return on equity, investors' required rate of return, and the company's intrinsic value.

We conclude with a summary of the key components of this reading:

- Common shares represent an ownership interest in a company and give investors a claim on its operating performance, the opportunity to participate in the corporate decision-making process, and a claim on the company's net assets in the case of liquidation.
- Callable common shares give the issuer the right to buy back the shares from shareholders at a price determined when the shares are originally issued.
- Putable common shares give shareholders the right to sell the shares back to the issuer at a price specified when the shares are originally issued.
- Preference shares are a form of equity in which payments made to preference shareholders take precedence over any payments made to common stockholders.
- Cumulative preference shares are preference shares on which dividend payments are accrued so that any payments omitted by the company must be paid before another dividend can be paid to common shareholders. Non-cumulative

preference shares have no such provisions, implying that the dividend payments are at the company's discretion and are thus similar to payments made to common shareholders.

- Participating preference shares allow investors to receive the standard preferred dividend plus the opportunity to receive a share of corporate profits above a pre-specified amount. Non-participating preference shares allow investors to simply receive the initial investment plus any accrued dividends in the event of liquidation.
- Callable and putable preference shares provide issuers and investors with the same rights and obligations as their common share counterparts.
- Private equity securities are issued primarily to institutional investors in private placements and do not trade in secondary equity markets. There are three types of private equity investments: venture capital, leveraged buyouts, and private investments in public equity (PIPE).
- The objective of private equity investing is to increase the ability of the company's management to focus on its operating activities for long-term value creation. The strategy is to take the "private" company "public" after certain profit and other benchmarks have been met.
- Depository receipts are securities that trade like ordinary shares on a local exchange but which represent an economic interest in a foreign company. They allow the publicly listed shares of foreign companies to be traded on an exchange outside their domestic market.
- American depository receipts are US dollar-denominated securities trading much like standard US securities on US markets. Global depository receipts are similar to ADRs but contain certain restrictions in terms of their ability to be resold among investors.
- Underlying characteristics of equity securities can greatly affect their risk and return.
- A company's accounting return on equity is the total return that it earns on shareholders' book equity.
- A company's cost of equity is the minimum rate of return that stockholders require the company to pay them for investing in its equity.

REFERENCES

Bailey, Elizabeth, Meg Wirth, and David Zapol. 2005. "Venture Capital and Global Health." *Financing Global Health Ventures*, Discussion Paper (September 2005): http://www.commonscapital.com/downloads/Venture_Capital_and_Global_Health.pdf

Boubakri, Narjess, Jean-Claude Cosset, and Anis Samet. 2010. "The Choice of ADRs." *Journal of Banking and Finance*, vol. 34, no. 9:2077–2095.

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2018. *Credit Suisse Global Investment Returns Sourcebook 2017*. Credit Suisse Research Institute.

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2018. *Credit Suisse Global Investment Returns Yearbook 2018*. Credit Suisse Research Institute.

Graham, John R., Campbell R. Harvey, and Shiva Rajgopal. 2005. "The Economic Implications of Corporate Financial Reporting." *Journal of Accounting and Economics*, vol. 40, no. 1–3:3–73.

Henry, Peter Blair, and Anusha Chari. 2004. "Risk Sharing and Asset Prices: Evidence from a Natural Experiment." *Journal of Finance*, vol. 59, no. 3:1295–1324.

Strömberg, Per. 2008. "The New Demography of Private Equity." The *Global Economic Impact of Private Equity Report 2008*, World Economic Forum.

PRACTICE PROBLEMS

1 Which of the following is *not* a characteristic of common equity?

A It represents an ownership interest in the company.

B Shareholders participate in the decision-making process.

C The company is obligated to make periodic dividend payments.

2 The type of equity voting right that grants one vote for each share of equity owned is referred to as:

A proxy voting.

B statutory voting.

C cumulative voting.

3 All of the following are characteristics of preference shares *except*:

A They are either callable or putable.

B They generally do not have voting rights.

C They do not share in the operating performance of the company.

4 Participating preference shares entitle shareholders to:

A participate in the decision-making process of the company.

B convert their shares into a specified number of common shares.

C receive an additional dividend if the company's profits exceed a pre-determined level.

5 Which of the following statements about private equity securities is *incorrect*?

A They cannot be sold on secondary markets.

B They have market-determined quoted prices.

C They are primarily issued to institutional investors.

6 Venture capital investments:

A can be publicly traded.

B do not require a long-term commitment of funds.

C provide mezzanine financing to early-stage companies.

7 Which of the following statements *most accurately* describes one difference between private and public equity firms?

A Private equity firms are focused more on short-term results than public firms.

B Private equity firms' regulatory and investor relations operations are less costly than those of public firms.

C Private equity firms are incentivized to be more open with investors about governance and compensation than public firms.

8 Emerging markets have benefited from recent trends in international markets. Which of the following has *not* been a benefit of these trends?

A Emerging market companies do not have to worry about a lack of liquidity in their home equity markets.

B Emerging market companies have found it easier to raise capital in the markets of developed countries.

© 2019 CFA Institute. All rights reserved.

C Emerging market companies have benefited from the stability of foreign exchange markets.

9 When investing in unsponsored depository receipts, the voting rights to the shares in the trust belong to:

A the depository bank.

B the investors in the depository receipts.

C the issuer of the shares held in the trust.

10 With respect to Level III sponsored ADRs, which of the following is *least likely* to be accurate? They:

A have low listing fees.

B are traded on the NYSE, NASDAQ, and AMEX.

C are used to raise equity capital in US markets.

11 A basket of listed depository receipts, or an exchange-traded fund, would *most likely* be used for:

A gaining exposure to a single equity.

B hedging exposure to a single equity.

C gaining exposure to multiple equities.

12 Calculate the total return on a share of equity using the following data:

Purchase price: $50

Sale price: $42

Dividend paid during holding period: $2

A –12.0%

B –14.3%

C –16.0%

13 If a US-based investor purchases a euro-denominated ETF and the euro subsequently depreciates in value relative to the dollar, the investor will have a total return that is:

A lower than the ETF's total return.

B higher than the ETF's total return.

C the same as the ETF's total return.

14 Which of the following is *incorrect* about the risk of an equity security? The risk of an equity security is:

A based on the uncertainty of its cash flows.

B based on the uncertainty of its future price.

C measured using the standard deviation of its dividends.

15 From an investor's point of view, which of the following equity securities is the *least* risky?

A Putable preference shares.

B Callable preference shares.

C Non-callable preference shares.

16 Which of the following is *least likely* to be a reason for a company to issue equity securities on the primary market?

A To raise capital.

B To increase liquidity.

C To increase return on equity.

17 Which of the following is *not* a primary goal of raising equity capital?

A To finance the purchase of long-lived assets.

B To finance the company's revenue-generating activities.

C To ensure that the company continues as a going concern.

18 Which of the following statements is *most accurate* in describing a company's book value?

A Book value increases when a company retains its net income.

B Book value is usually equal to the company's market value.

C The ultimate goal of management is to maximize book value.

19 Calculate the book value of a company using the following information:

Number of shares outstanding	100,000
Price per share	€52
Total assets	€12,000,000
Total liabilities	€7,500,000
Net Income	€2,000,000

A €4,500,000.

B €5,200,000.

C €6,500,000.

20 Which of the following statements is *least accurate* in describing a company's market value?

A Management's decisions do not influence the company's market value.

B Increases in book value may not be reflected in the company's market value.

C Market value reflects the collective and differing expectations of investors.

21 Calculate the return on equity (ROE) of a stable company using the following data:

Total sales	£2,500,000
Net income	£2,000,000
Beginning of year total assets	£50,000,000
Beginning of year total liabilities	£35,000,000
Number of shares outstanding at the end of the year	1,000,000
Price per share at the end of the year	£20

A 10.0%.

B 13.3%.

C 16.7%.

22 Holding all other factors constant, which of the following situations will *most likely* lead to an increase in a company's return on equity?

A The market price of the company's shares increases.

B Net income increases at a slower rate than shareholders' equity.

C The company issues debt to repurchase outstanding shares of equity.

23 Which of the following measures is the *most difficult* to estimate?

A The cost of debt.

B The cost of equity.

C Investors' required rate of return on debt.

24 A company's cost of equity is often used as a proxy for investors':

A average required rate of return.

B minimum required rate of return.

C maximum required rate of return.

SOLUTIONS

1 C is correct. The company is not obligated to make dividend payments. It is at the discretion of the company whether or not it chooses to pay dividends.

2 B is correct. Statutory voting is the type of equity voting right that grants one vote per share owned.

3 A is correct. Preference shares do not have to be either callable or putable.

4 C is correct. Participating preference shares entitle shareholders to receive an additional dividend if the company's profits exceed a pre-determined level.

5 B is correct. Private equity securities do not have market-determined quoted prices.

6 C is correct. Venture capital investments can be used to provide mezzanine financing to companies in their early stage of development.

7 B is correct. Regulatory and investor relations costs are lower for private equity firms than for public firms. There are no stock exchange, regulatory, or shareholder involvements with private equity, whereas for public firms these costs can be high.

8 C is correct. The trends in emerging markets have not led to the stability of foreign exchange markets.

9 A is correct. In an unsponsored DR, the depository bank owns the voting rights to the shares. The bank purchases the shares, places them into a trust, and then sells shares in the trust—not the underlying shares—in other markets.

10 A is correct. The listing fees on Level III sponsored ADRs are high.

11 C is correct. An ETF is used to gain exposure to a basket of securities (equity, fixed income, commodity futures, etc.).

12 A is correct. The formula states $R_t = (P_t - P_{t-1} + D_t)/P_{t-1}$. Therefore, total return = $(42 - 50 + 2)/50 = -12.0\%$.

13 A is correct. The depreciated value of the euro will create an additional loss in the form of currency return that is lower than the ETF's return.

14 C is correct. Some equity securities do not pay dividends, and therefore the standard deviation of dividends cannot be used to measure the risk of all equity securities.

15 A is correct. Putable shares, whether common or preference, give the investor the option to sell the shares back to the issuer at a pre-determined price. This pre-determined price creates a floor for the share's price that reduces the uncertainty of future cash flows for the investor (i.e., lowers risk relative to the other two types of shares listed).

16 C is correct. Issuing shares in the primary (and secondary) market *reduces* a company's return on equity because it increases the total amount of equity capital invested in the company (i.e., the denominator in the ROE formula).

17 C is correct. Capital is raised to ensure the company's existence only when it is required. It is not a typical goal of raising capital.

18 A is correct. A company's book value increases when a company retains its net income.

19 A is correct. The book value of the company is equal to total assets minus total liabilities, which is €12,000,000 – €7,500,000 = €4,500,000.

20 A is correct. A company's market value is affected by management's decisions. Management's decisions can directly affect the company's *book* value, which can then affect its market value.

21 B is correct. A company's ROE is calculated as (NI_t/BVE_{t-1}). The BVE_{t-1} is equal to the beginning total assets minus the beginning total liabilities, which equals £50,000,000 – £35,000,000 = £15,000,000. Therefore, ROE = £2,000,000/£15,000,000 = 13.3%.

22 C is correct. A company's ROE will increase if it issues debt to repurchase outstanding shares of equity.

23 B is correct. The cost of equity is not easily determined. It is dependent on investors' required rate of return on equity, which reflects the different risk levels of investors and their expectations about the company's future cash flows.

24 B is correct. Companies try to raise funds at the lowest possible cost. Therefore, cost of equity is used as a proxy for the minimum required rate of return.

READING

37

Introduction to Industry and Company Analysis

by Patrick W. Dorsey, CFA, Anthony M. Fiore, CFA, and Ian Rossa O'Reilly, CFA

Patrick W. Dorsey, CFA, is at Dorsey Asset Management (USA). Anthony M. Fiore, CFA, is at Silvercrest Asset Management (USA). Ian Rossa O'Reilly, CFA (Canada).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** explain uses of industry analysis and the relation of industry analysis to company analysis;
☐	**b.** compare methods by which companies can be grouped;
☐	**c.** explain the factors that affect the sensitivity of a company to the business cycle and the uses and limitations of industry and company descriptors such as "growth," "defensive," and "cyclical";
☐	**d.** describe current industry classification systems, and identify how a company should be classified, given a description of its activities and the classification system;
☐	**e.** explain how a company's industry classification can be used to identify a potential "peer group" for equity valuation;
☐	**f.** describe the elements that need to be covered in a thorough industry analysis;
☐	**g.** describe the principles of strategic analysis of an industry;
☐	**h.** explain the effects of barriers to entry, industry concentration, industry capacity, and market share stability on pricing power and price competition;
☐	**i.** describe industry life-cycle models, classify an industry as to life-cycle stage, and describe limitations of the life-cycle concept in forecasting industry performance;

(continued)

CFA Institute would like to thank Katy Zhao, at the Whitman Group, Morgan Stanley, (USA), for the revision of Introduction to Industry and Company Analysis reading.

© 2019 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**j.** describe macroeconomic, technological, demographic, governmental, social, and environmental influences on industry growth, profitability, and risk;
☐	**k.** compare characteristics of representative industries from the various economic sectors;
☐	**l.** describe the elements that should be covered in a thorough company analysis.

1

INTRODUCTION

Understanding how industries and companies operate, together with an analysis of financial statements, provides a basis for forecasting company performance and allows analysts to determine the value of an investment in a company or its securities. **Industry analysis** is the analysis of a specific branch of manufacturing, service, or trade. Understanding the industry in which a company operates provides an essential framework for the analysis of the individual company—that is, **company analysis**. Equity analysis and credit analysis are often conducted by analysts who concentrate on one or several industries, which results in synergies and efficiencies in gathering and interpreting information.

In this reading, we will address the following questions:

- What are the similarities and differences among industry classification systems?
- How does an analyst go about choosing a peer group of companies?
- What are the key factors to consider when analyzing an industry?
- What advantages are enjoyed by companies in strategically well-positioned industries?
- How should an analyst approach research and analysis of new industries?
- What are the factors that influence individual companies?

2

USES OF INDUSTRY ANALYSIS

a explain uses of industry analysis and the relation of industry analysis to company analysis

Industry analysis is useful in a number of investment applications that make use of fundamental analysis. Its uses include the following:

- *Understanding a company's business and business environment.* Industry analysis is often a critical early step in stock selection and valuation because it provides insights into the issuer's growth opportunities, competitive dynamics, and business risks. For a credit analyst, industry analysis provides insights into the appropriateness of a company's use of debt financing and into its ability to meet its promised payments.

- *Identifying active equity investment opportunities.* Investors taking a top-down investing approach use industry analysis to identify industries with positive, neutral, or negative outlooks for profitability and growth. Generally, investors will then position their portfolios accordingly—by overweighting, market weighting, or underweighting those industries relative to the investor's benchmark if the investor judges that the industry's perceived prospects are not fully incorporated in market prices. In fact, some investors base their investment process on attempting to outperform their benchmarks or deliver absolute returns by industry rotation—that is, timing investments in industries in relation to an analysis of industry fundamentals and/or business-cycle conditions (technical analysis may also play a role in such strategies).
- *Portfolio performance attribution.* The purpose of performance attribution is to identify and understand the sources of portfolio return. Portfolio attribution output will show any positive or negative contribution to returns from the fund manager's choice of industries and/or sectors in the portfolio. Industry classification schemes play a role in such performance attribution.

EXAMPLE 1

Which of the following information about a company would *most likely* depend on an industry analysis? The company's:

A dividend distribution policy.

B competitive environment.

C trends in personnel expenses.

Solution

B is correct. To understand the competitive environment, analysts need to understand the industry as a whole. A and C are incorrect because both are specific to the company.

We will next explore how companies may be grouped into industries.

APPROACHES TO IDENTIFYING SIMILAR COMPANIES

3

b compare methods by which companies can be grouped

c explain the factors that affect the sensitivity of a company to the business cycle and the uses and limitations of industry and company descriptors such as "growth," "defensive," and "cyclical"

Industry classification attempts to place companies into groups on the basis of commonalities. In the following sections, we discuss the three major approaches to industry classification:

- products and/or services supplied,
- business-cycle sensitivities, and
- statistical similarities.

3.1 Products and/or Services Supplied

Modern classification schemes are most commonly based on grouping companies by similar products and/or services. According to this perspective, an **industry** is defined as a group of companies offering similar products and/or services. For example, major companies in the global technology industry include Microsoft, Apple, Alphabet (owner of Google), Tencent, and Samsung Electronics. Although those companies are known to belong to the technology industry, they differ from one another in numerous ways. This is an issue that analysts should be aware of and take into consideration when making recommendations and investment decisions.

Industry classification schemes typically provide multiple levels of aggregation. The term **sector** is often used to refer to a group of related industries. The health care sector, for example, consists of a number of related industries, including the pharmaceutical, biotechnology, medical device, medical supply, hospital, and managed care industries.

These classification schemes typically place a company in an industry on the basis of a determination of its principal business activity. A company's **principal business activity** is the source from which the company derives a majority of its revenues and/or earnings. For example, companies that derive a majority of their revenues from the sale of pharmaceuticals include Novartis AG, Pfizer Inc., Roche Holding AG, GSK plc, and Sanofi S.A., all of which could be grouped together as part of the global pharmaceutical industry. Companies that engage in more than one significant business activity usually report the revenues (and, in many cases, operating profits) of the different business segments in their financial statements.

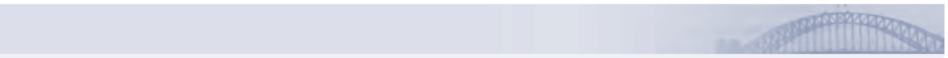
Examples of classification systems based on products and/or services include the commercial classification systems that will be discussed later—namely, the Global Industry Classification Standard (GICS) and the Industry Classification Benchmark (ICB). In addition to grouping companies by products and/or services, some of the major classification systems group consumer-related companies into cyclical and non-cyclical categories depending on a company's sensitivity to the business cycle. The next section addresses how companies can be categorized on the basis of economic sensitivity.

3.2 Business-Cycle Sensitivities

Companies are sometimes grouped on the basis of their relative sensitivity to the business cycle. This method often results in two broad groupings of companies:

- A **cyclical** company is one whose profits are strongly correlated with the strength of the overall economy. Such companies experience wider-than-average fluctuations in demand—high demand during periods of economic expansion and low demand during periods of economic contraction—and/or are subject to greater-than-average profit variability related to high operating leverage (i.e., high fixed costs). Concerning demand, cyclical products and services are often relatively expensive and represent purchases that can be delayed if necessary (e.g., because of declining disposable income). Examples of cyclical industries and broader sectors are autos, housing, basic materials, industrials, and technology.
- A **non-cyclical** company is one whose performance is largely independent of the business cycle. Non-cyclical companies produce goods or services for which demand remains relatively stable throughout the business cycle. Examples of non-cyclical industries are food and beverage, household and personal care products, health care, and utilities.

Although the classification systems we will discuss do not label their categories as cyclical or non-cyclical, certain sectors tend to experience greater economic sensitivity than others. Sectors that tend to exhibit a relatively high degree of economic sensitivity include consumer discretionary, energy, financials, industrials, technology, and materials. In contrast, sectors that exhibit relatively less economic sensitivity include consumer staples, health care, telecommunications, and utilities.



Descriptions Related to the Cyclical/Non-Cyclical Distinction

Analysts commonly encounter a number of labels related to the cyclical/non-cyclical distinction. For example, non-cyclical industries have sometimes been sorted into defensive (or stable) versus growth. Defensive industries and companies are those whose revenues and profits are least affected by fluctuations in overall economic activity. These industries/companies tend to produce staple consumer goods (e.g., bread), provide basic services (grocery stores, drug stores, fast food outlets), or have their rates and revenues determined by contracts or government regulation (e.g., cost-of-service, rate-of-return regulated public utilities). Growth industries include industries with specific demand dynamics that are so strong that they override the significance of broad economic or other external factors and generate growth regardless of overall economic conditions, although their rates of growth may slow during an economic downturn.

The usefulness of industry and company labels, such as cyclical, growth, and defensive, is limited. Cyclical industries often have growth companies within them. A cyclical industry itself, although exposed to the effects of fluctuations in overall economic activity, may grow at an above-average rate for periods spanning multiple business cycles. The label "**growth cyclical**" is sometimes used to describe companies that are growing rapidly on a long-term basis but that still experience above-average fluctuation in their revenues and profits over the course of a business cycle.

Furthermore, when fluctuations in economic activity are large, as in the deep recession of 2008–2009, few companies escape the effects of the cyclical weakness in overall economic activity.

The defensive label is also problematic. Industries may include both companies that are growth and companies that are defensive in character, making the choice between a "growth" and a "defensive" label difficult. Moreover, "defensive" cannot be understood as necessarily being descriptive of investment characteristics. Food supermarkets, for example, would typically be described as defensive but can be subject to profit-damaging price wars. So-called defensive industries/companies may sometimes face industry dynamics that make them far from defensive in the sense of preserving shareholders' capital.

One limitation of the cyclical/non-cyclical classification is that business-cycle sensitivity is a continuous spectrum rather than an "either/or" issue, so placement of companies in one of the two major groups is somewhat arbitrary. The impact of severe recessions usually reaches all parts of the economy, so non-cyclical is better understood as a relative term.

Another limitation of a business-cycle classification for global investing is that different countries and regions of the world frequently progress through the various stages of the business cycle at different times. While one region of the world may be experiencing economic expansion, other regions or countries may be in recession, which complicates the application of a business-cycle approach to industry analysis. For example, a jewelry retailer (i.e., a cyclical company) that is selling domestically in a weak economy will exhibit markedly different fundamental performance relative to a jewelry company operating in an environment where demand is robust. Comparing these two companies—that is, similar companies that are currently exposed to different demand environments—could suggest investment opportunities. Combining fundamental data from such companies, however, to establish industry benchmark values would be misleading.

3.3 Statistical Similarities

Statistical approaches to grouping companies are typically based on the correlations of past securities' returns. For example, using the technique known as cluster analysis, companies are separated (on the basis of historical correlations of stock returns) into groups *in which* correlations are relatively high but *between which* correlations are relatively low. Clustering algorithms can be based on a range of financial characteristics (e.g., total assets, total revenue, profitability, leverage) and operating characteristics (e.g., R&D intensity, employee headcount). This method of aggregation often results in non-intuitive groups of companies, and the composition of the groups may vary significantly by time period and region of the world. Moreover, statistical approaches rely on historical data, but analysts have no guarantee that past correlation values will continue in the future. In addition, such approaches carry the inherent dangers of all statistical methods—namely, (1) falsely indicating a relationship that arose by chance and is not supported by plausible economic basis or (2) falsely excluding a relationship that actually is significant.

EXAMPLE 2

1 Which of the following is the least likely to be a method for grouping companies?

- **A** Statistical approaches that assess past correlations of securities' returns
- **B** Similarity of products and services
- **C** Similarity of operating profit margins

2 Companies are classified as cyclical or non-cyclical (defensive) on the basis of the exposure of their business to:

- **A** their industry life cycle.
- **B** the business cycle of the economy.
- **C** the credit cycle, which affects their ability to borrow.

Solution to 1:

C is correct. Profit margins are not used as a method for grouping companies. Companies with similar profit margins may have little in common. A is incorrect because it describes the statistical similarities approach, which is a method for grouping companies. B is incorrect because similarity of products and services is one of the most common ways of grouping companies.

Solution to 2:

B is correct. The cyclical/non-cyclical grouping reflects companies' sensitivity to the economy's business cycle. A is incorrect because the industry life cycle describes how the industry evolves over time, not how it relates to economic fluctuations. C is incorrect because the credit cycle explains the availability and cost of credit, not the fluctuations of output of the economy to which cyclical or non-cyclical companies are exposed.

INDUSTRY CLASSIFICATION SYSTEMS

4

d describe current industry classification systems, and identify how a company should be classified, given a description of its activities and the classification system

A well-designed classification system often serves as a useful starting point for industry analysis. It allows analysts to compare industry trends and relative valuations among companies in a group. Classification systems that take a global perspective enable portfolio managers and research analysts to make global comparisons of companies in the same industry. For example, given the global nature of the automobile industry, a thorough analysis of the industry would include auto companies from many different countries and regions of the world. Classification systems are developed and used by both commercial entities and various governmental agencies. Most government and commercial classification systems are reviewed and, if necessary, updated from time to time. Generally, commercial classification systems are adjusted more frequently than government classification systems, which may be updated only every five years or so. We focus on the commercial systems because they are more commonly used in the investment industry.

4.1 Commercial Industry Classification Systems

Major index providers, including Standard & Poor's, MSCI, Russell Investments, Dow Jones, and FTSE Russell, classify companies in their equity indexes into industry groupings. Most classification schemes used by these index providers contain multiple levels of classification that start at the broadest level with a general sector grouping and then, in several further steps, subdivide or disaggregate the sectors into more granular, narrowly defined sub-industry groups.

4.1.1 *Global Industry Classification Standard*

GICS, jointly developed by Standard & Poor's and MSCI, was designed to facilitate global comparisons of industries. It classifies companies in both developed and developing economies. In June 2020, the GICS classification structure comprised four levels of detail consisting of 11 sectors, 24 industry groups, 69 industries, and 158 sub-industries. The composition of GICS historically has been adjusted over time to reflect changes in the global equity markets. One of the most significant changes happened in 2018 when GICS launched the new Communication Services sector to expand beyond simply the Telecommunication Services sector to incorporate many information technology and internet-related companies that offer communication platforms, content, and information.

4.1.2 *Industry Classification Benchmark*

The Industry Classification Benchmark (ICB), which was jointly developed by Dow Jones and FTSE, uses a four-tier structure to categorize companies globally on the basis of the source from which a company derives the majority of its revenue. In June 2020, the ICB classification system consisted of 11 industries, 20 supersectors, 45 sectors, and 173 subsectors. Although the ICB is similar to GICS in the number of tiers and the method by which companies are assigned to particular groups, the two systems use significantly different nomenclature. For example, whereas GICS uses the term "sector" to describe its broadest grouping of companies, ICB uses the term "industry" to describe the broadest category. The ICB is used by many global stock exchanges to categorize listed companies, including the London Stock Exchange, Euronext, and NASDAQ OMX. In addition, all Russell Indexes are set to adopt a newly enhanced ICB structure after the market close on 18 September 2020.

Exhibit 1 provides examples of how several companies are classified. Note how differently the systems classify PayPal, Uber, and Activision Blizzard.

Exhibit 1 Industry Classification Examples

Stock Examples	Global Industry Classification Standard	Industry Classification Benchmark
PayPal	Information Technology > Software & Services > IT Services > Data Processing & Outsourced Services	Industrials > Industrial Goods & Services > Support Services > Financial Administration
Peloton Interactive	Consumer Discretionary > Consumer Durables & Apparel > Leisure Products > Leisure Products	Consumer Services > Travel & Leisure > Travel & Leisure > Recreational Services
Uber	Industrials > Transportation > Road & Rail > Trucking	Consumer Services > Retail > General Retailers > Specialized Consumer Services
Activision Blizzard	Communication Services > Media & Entertainment > Entertainment > Interactive Home Entertainment	Consumer Goods > Personal & Household Goods > Leisure Goods > Toys

Despite the fact that some companies are assigned to different broader sectors and industries by the different systems, the two commercial classification systems use common methodologies for assigning companies to groups. Also, the broadest level of grouping for GICS and the ICB is quite similar. Specifically, GICS and the ICB each identify 11 broad groupings below which all other categories reside. Next, we describe sectors that are fairly representative of how the broadest level of classification is viewed by GICS (into "sectors") and the ICB (into "industries").

4.1.3 *Description of Representative Sectors/Industries*

Exhibit 2 provides a description of each of the 11 broad sectors.

Exhibit 2 Sector Descriptions

Materials	Companies engaged in the production of building materials, chemicals, paper and forest products, and containers and packaging, as well as metal, mineral, and mining companies.
Consumer Discretionary	Companies that derive a majority of revenue from the sale of consumer-related products or services for which demand tends to exhibit a relatively high degree of economic sensitivity. Examples of business activities that frequently fall into this category are automotive, apparel, hotel, and restaurant businesses.
Consumer Staples	Consumer-related companies whose business tends to exhibit less economic sensitivity than other companies—for example, manufacturers of food, beverages, tobacco, and personal care products.

Exhibit 2 (Continued)

Energy	Companies whose primary line of business involves exploring for, production of, or refining of natural resources used to produce energy; companies that derive a majority of revenue from the sale of equipment or through the provision of services to energy companies also fall into this category.
Financials	Companies whose primary line of business involves banking, finance, insurance, asset management, and/or brokerage services.
Health Care	Manufacturers of pharmaceutical and biotech products, medical devices, health care equipment, and medical supplies and providers of health care services.
Industrials	Manufacturers of capital goods and providers of commercial services; for example, business activities would include heavy machinery and equipment manufacture, aerospace and defense, transportation services, and commercial services and supplies.
Real Estate	Companies engaged in the development and operation of real estate. This includes companies offering real estate–related services and equity real estate investment trusts (REITs).
Information Technology	Manufacture or sale of computers, software, semiconductors, and communications equipment; other business activities that frequently fall into this category are electronic entertainment, internet services, and technology consulting and services.
Communication Services/ Telecommunications	This sector includes traditional telecommunication companies that provide fixed line and wireless communication services, with media, entertainment, and interactive media and services, such as video gaming companies.
Utilities	Electric, gas, and water utilities; telecommunication companies are sometimes included in this category.

To classify a company accurately in a particular classification scheme requires definitions of the classification categories, a statement about the criteria used in classification, and detailed information about the subject company. Example 3 introduces an exercise in such classification. In addressing the question, the reader can make use of the widely applicable sector descriptions just given and familiarity with available business products and services.

EXAMPLE 3

Classifying Companies into Industries

Exhibit 2 defines 11 representative sectors, repeated here in Exhibit 3. Assume the classification system is based on the criterion of a company's principal business activity as judged primarily by source of revenue.

Exhibit 3 List of Sectors

Sector
Materials
Consumer Discretionary
Consumer Staples
Energy
Financials
Health Care
Industrials
Real Estate
Information Technology
Communication Services/ Telecommunications
Utilities

Based on the information given, determine an appropriate sector membership for each of the following hypothetical companies:

0 An operator of shopping malls
1 A natural gas transporter and marketer
2 A manufacturer of heavy construction equipment
3 A provider of regional telephone services
4 A semiconductor company
5 A manufacturer of medical devices
6 A video conference provider
7 A manufacturer of chemicals and plastics
8 A manufacturer of automobiles
9 A cloud computing service provider
10 A food delivery company
11 A regulated supplier of electricity
12 A provider of wireless broadband services
13 A manufacturer of soaps and detergents
14 A software development company
15 An insurer
16 A regulated provider of water/wastewater services
17 A robotic-assisted surgery company
18 A manufacturer of pharmaceuticals
19 A provider of rail transportation services
20 A data center real estate investment trust
21 A developer of residential housing

Solution

Sector	Company Number
Materials	7
Consumer Discretionary	8, 10
Consumer Staples	13
Energy	1
Financials	15
Health Care	5, 17, 18
Industrials	2, 19
Real Estate	0, 20, 21
Information Technology	4, 9, 14
Communication Services/ Telecommunications	3, 6, 12
Utilities	11, 16

EXAMPLE 4

Industry Classification Schemes

1 The GICS classification system classifies companies on the basis of a company's primary business activity as measured primarily by:

A assets.

B income.

C revenue.

2 Which of the following is *least likely* to be accurately described as a cyclical company?

A An automobile manufacturer

B A producer of breakfast cereals

C An apparel company producing new, trendy clothes for teenage girls

3 Which of the following is the *most accurate* statement? A statistical approach to grouping companies into industries:

A is based on historical correlations of the securities' returns.

B frequently produces industry groups whose composition is similar worldwide.

C emphasizes the descriptive statistics of industries consisting of companies producing similar products and/or services.

Solution to 1:

C is correct.

Solution to 2:

B is correct. A producer of staple foods, such as cereals, is a classic example of a non-cyclical company. Demand for automobiles is cyclical—that is, relatively high during economic expansions and relatively low during economic contractions.

Also, demand for teenage fashions is likely to be more sensitive to the business cycle than demand for standard food items, such as breakfast cereals. When budgets have been reduced, families may try to avoid expensive clothing or may try to extend the life of existing clothing.

Solution to 3:

A is correct.

4.2 Constructing a Peer Group

e explain how a company's industry classification can be used to identify a potential "peer group" for equity valuation

A **peer group** is a group of companies engaged in similar business activities whose economics and valuation are influenced by closely related factors. Comparisons of a company in relation to a well-defined peer group can provide valuable insights into the company's performance and its relative valuation.

The construction of a peer group is a subjective process; the result often differs significantly from even the most narrowly defined categories given by the commercial classification systems. However, commercial classification systems do provide a starting point for the construction of a relevant peer group because by using such systems, an analyst can quickly discover the public companies operating in the chosen industry.

In fact, one approach to constructing a peer group is to start by identifying other companies operating in the same industry. Analysts who subscribe to one or more of the commercial classification systems can quickly generate a list of other companies in the industry in which the company operates according to that particular service provider's definition of the industry. An analyst can then investigate the business activities of these companies through a range of sources, such as companies' public disclosures or industry trade publications, and confirm that each comparable company derives a significant portion of its revenue and operating profit from a business activity similar to the primary business of the subject company.

The following questions may improve the list of peer companies:

- *What proportion of revenue and operating profit is derived from business activities similar to those of the subject company?* In general, a higher percentage results in a more meaningful comparison.
- *Does a potential peer company face a demand environment similar to that of the subject company?* For example, a comparison of growth rates, margins, and valuations may be of limited value when comparing companies that are exposed to different stages of the business cycle. (As mentioned, such differences may be the result of conducting business in geographically different markets.)
- *Does the peer company have a significant business segment(s) that is comparable to other companies even if they are not in the same commercial classification system?* For example, Amazon Web Services' business (cloud computing) is more comparable to that of such peers as Microsoft and Cisco than to that of its e-commerce peers.

Although companies with limited lines of business may be neatly categorized into a single peer group, companies with multiple divisions may be included in more than one category. For example, US-based Hewlett-Packard Company (HP), a global provider of technology and software solutions, used to be included in more than one peer group before it separated its business into HP Inc., which sells hardware, and Hewlett-Packard Enterprise (HPE), which sells software and services. Investors interested in the personal computer (PC) industry, for example, used to include HP in their peer group, but investors constructing a peer group of providers of information technology services also included HP in that group.

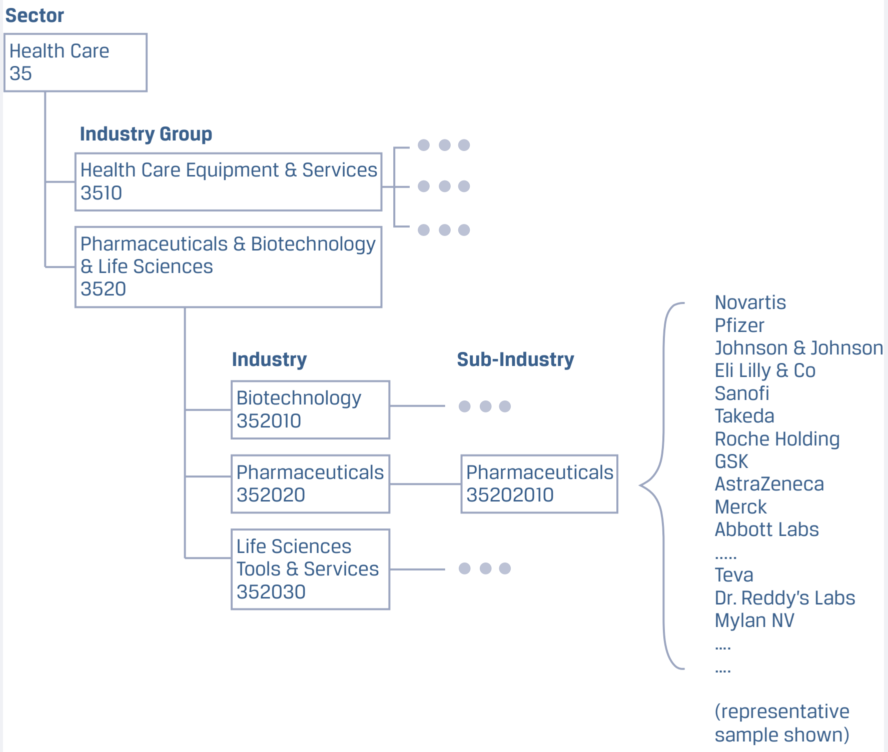
Example 5 illustrates the process of identifying a peer group of companies and shows some of the practical hurdles to determining a peer group.

EXAMPLE 5

An Analyst Researches the Peer Group of Novartis

Suppose that an analyst needs to identify the peer group of companies for Novartis for use in the valuation section of a company report. Novartis develops and produces branded pharmaceuticals and operates globally. The analyst starts by looking at Novartis's industry classification according to GICS. The most narrowly defined category that GICS uses is the sub-industry level, and in July 2020, Novartis was in the GICS sub-industry called Pharmaceuticals, together with the companies listed in Exhibit 4 (a sample of companies is shown).

Exhibit 4 GICS Industry Classification: Pharmaceuticals


Sector
Health Care
35
Industry Group
Health Care Equipment & Services
3510
Pharmaceuticals & Biotechnology
& Life Sciences
3520
Industry
Sub-Industry
Biotechnology
352010
Pharmaceuticals
352020
Life Sciences
Tools & Services
352030
Pharmaceuticals
35202010
Novartis
Pfizer
Johnson & Johnson
Eli Lilly & Co
Sanofi
Takeda
Roche Holding
GSK
AstraZeneca
Merck
Abbott Labs
.....
Teva
Dr. Reddy's Labs
Mylan NV
....
....
(representative
sample shown)

After looking over the list of companies, the analyst quickly realizes that some adjustments need to be made to the list to end up with a peer group of companies that are comparable to Novartis.

Research of the companies' disclosures, websites, and company descriptions provided by electronic data/information vendors reveals that Novartis focuses on research, development, and production of branded, patent-protected original products. Such products account for 79% of the company's revenues (source: "Novartis Annual Report 2019"). The business model is, therefore, distinct from that of companies that focus on the production of generic pharmaceuticals that do not benefit from patent protection the way branded original products do. For example, Teva and Dr. Reddy's main business is in the generic pharmaceuticals space, so those companies would not fit Novartis's peer group. In contrast, Novartis's business is subject to similar drivers and external influences that apply to a narrow range of companies, such as Roche, Pfizer, GSK, and Sanofi. The peer group that the analyst should use is, therefore, narrower than what the GICS Pharmaceuticals sub-industry would suggest.

In summary, analysts must distinguish between a company's industry—as defined by one or more of the various classification systems—and its peer group. A company's peer group should consist of companies with similar business activities whose economic activity depends on similar drivers of demand and similar factors related to cost structure and access to financial capital. In practice, these necessities frequently result in a smaller group (even a different group) of companies than the most narrowly defined categories used by the common commercial classification systems. Example 6 illustrates various aspects of developing and using peer groups.

EXAMPLE 6

The Semiconductor Industry: Business-Cycle Sensitivity and Peer-Group Determination

The GICS semiconductor and semiconductor equipment industry (453010) has two sub-industries—the semiconductor equipment sub-industry (45301010) and the semiconductors sub-industry (45301020). Members of the semiconductor equipment sub-industry include equipment suppliers, such as Lam Research Corporation and ASML Holdings NV; the semiconductors sub-industry includes integrated circuit manufacturers Intel Corporation and Taiwan Semiconductor Manufacturing Company Ltd.

Lam Research is a leading supplier of wafer fabrication equipment and services to the world's semiconductor industry. Lam also offers wafer-cleaning equipment that is used after many of the individual steps required to manufacture a finished wafer. Often, the technical advances that Lam introduces in its wafer-etching and wafer-cleaning products are also available as upgrades to its installed base. This benefit provides customers with a cost-effective way to extend the performance and capabilities of their existing wafer fabrication lines.

ASML describes itself as the world's leading provider of lithography systems (etching and printing on wafers) for the semiconductor industry. ASML manufactures complex machines that are critical to the production of integrated circuits or microchips. ASML designs, develops, integrates, markets, and services these advanced systems, which help chip makers reduce the size and increase the functionality of microchips and consumer electronic equipment. The machines are costly and thus represent a substantial capital investment for a purchaser.

Based on revenue, Intel is the world's largest semiconductor chip maker and has a dominant share of microprocessors for the personal computer market. Intel designs and produces its own proprietary semiconductors for direct sale to customers, such as personal computer makers. Intel has made significant investments in research and development (R&D) to introduce and produce new chips for new applications.

Established in 1987, Taiwan Semiconductor Manufacturing (TSMC) is the world's largest dedicated semiconductor foundry company (dedicated semiconductor foundries are semiconductor fabrication plants that execute the designs of other companies). TSMC describes itself as offering cutting-edge process technologies, pioneering design services, manufacturing efficiency, and product quality. TSMC provides design and production services to a diverse group of integrated circuit suppliers that generally do not have their own in-house manufacturing capabilities. The company's revenues represented more than 50% of the dedicated foundry segment in the semiconductor industry.

The questions that follow take the perspective of a recession similar to that of early 2009, when many economies around the world were in a recession. Based only on the information given, answer the following questions:

1 If the weak economy, similar to that of early 2009, were to recover within the next 12–18 months, which of the two sub-industries of the semiconductor and semiconductor equipment industry would most likely be the first to experience a positive improvement in business?
2 Explain whether Intel and TSMC should be considered members of the same peer group.
3 Explain whether Lam Research and ASML should be considered members of the same peer group.

Solution to 1:

In the most likely scenario, improvement in the business of the equipment makers (Lam and ASML) would lag that of semiconductor companies (Intel and TSMC). Because of the weak economy, excess manufacturing capacity should be available to meet increased demand for integrated circuits in the near term without additional equipment, which is a major capital investment. When semiconductor manufacturers believe the longer-term outlook has improved, they should begin to place orders for additional equipment.

Solution to 2:

Intel and TSMC are not likely to be considered comparable members of the same peer group because they have different sets of customers and different business models. Intel sells its proprietary semiconductors directly to customers, whereas TSMC provides design and production services to circuit suppliers that do not have their own in-house manufacturing capabilities. Standard & Poor's does not group Intel and TSMC in the same peer group; Intel was in the Semiconductors, Logic, Larger Companies group, and TSMC was in the Semiconductors, Foundry Services group.

Solution to 3:

Both Lam Research and ASML are leading companies that design and manufacture equipment to produce semiconductor chips. The companies are comparable because they both depend on the same economic factors that drive demand for their products. Their major customers are the semiconductor chip companies. Standard & Poor's grouped both companies in the same peer group—Semiconductor Equipment, Larger Front End.

DESCRIBING AND ANALYZING AN INDUSTRY AND PRINCIPLES OF STRATEGIC ANALYSIS

5

f describe the elements that need to be covered in a thorough industry analysis
g describe the principles of strategic analysis of an industry

In their work, analysts with superior knowledge about an industry's characteristics, conditions, and trends have a competitive edge in evaluating the investment merits of the companies in the industry. Analysts attempt to develop practical, reliable industry forecasts by using various approaches to forecasting. They often estimate a range of projections for a variable reflecting various possible scenarios. In order to conduct this analysis, analysts can study a wide range of factors, including but not limited to the following:

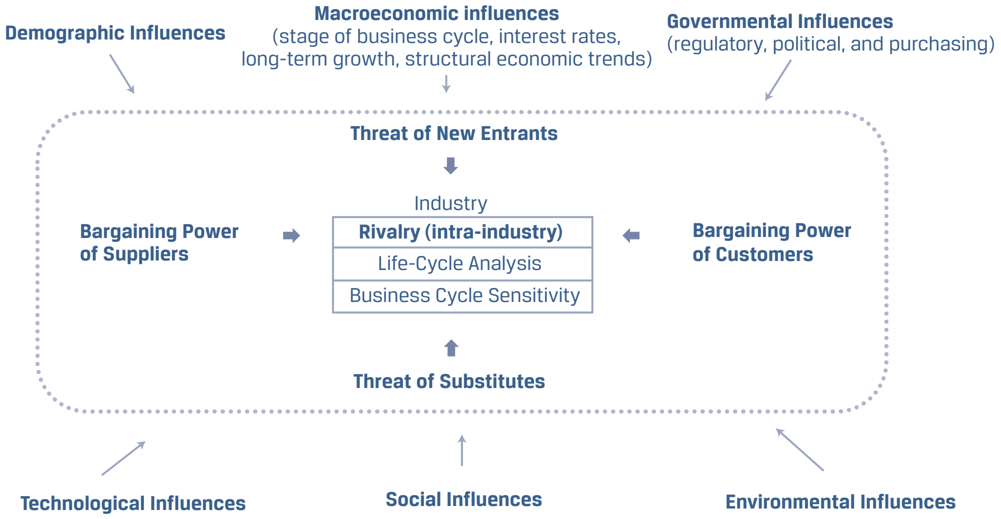
1. Statistical relationships between industry trends
2. Economic and business variables
3. Information from industry associations, from the individual subject companies they are analyzing, and from these companies' competitors, suppliers, and customers
4. Relevant industry trends and metrics to help understand and forecast trends

Investment managers and analysts also examine industry performance (1) in relation to other industries to identify industries with superior/inferior returns and (2) over time to determine the degree of consistency, stability, and risk in the returns in the industry over time. The objective of this analysis is to identify industries that offer the highest potential for investment returns on a risk-adjusted basis. The investment time horizon can be either long or short, as is the case for a rotation strategy in which portfolios are rotated into the industry groups that are expected to benefit from the next stage in the business cycle.

Analysts may also seek to compare their assumptions and projections with those made by other analysts. Doing so enables them to identify and understand any differences in methodology and any differences between their forecasts and consensus forecasts.

Exhibit 5 provides a framework designed to help analysts check that they have considered the range of forces that may affect the evolution of an industry. It shows, at the macro level, macroeconomic, demographic, environmental, governmental, social, and technological influences affecting the industry. It also depicts how an industry is affected by the forces driving industry competition (threat of new entrants, substitution threats, customer and supplier bargaining forces), the competitive forces in the industry (rivalry), life-cycle issues, and business-cycle considerations. Exhibit 5 summarizes and brings together visually the topics and concepts discussed in this section.

Exhibit 5 A Framework for Industry Analysis


Demographic Influences
Macroeconomic influences
(stage of business cycle, interest rates, long-term growth, structural economic trends)
Governmental Influences
(regulatory, political, and purchasing)
Threat of New Entrants
Bargaining Power of Suppliers
Industry
Rivalry (intra-industry)
Life-Cycle Analysis
Business Cycle Sensitivity
Bargaining Power of Customers
Threat of Substitutes
Technological Influences
Social Influences
Environmental Influences

5.1 Principles of Strategic Analysis

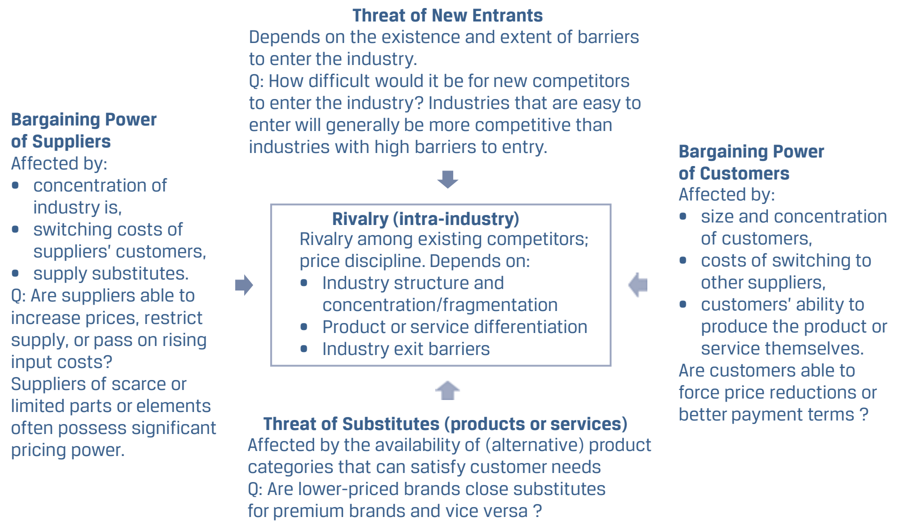
When analyzing an industry, the analyst must recognize that the economic fundamentals can vary markedly among industries. Some industries are highly competitive, with most players struggling to earn adequate returns on capital, whereas other industries have attractive characteristics that allow almost all industry participants to generate healthy profits.

Differing competitive environments are often tied to the structural attributes of an industry, which is one reason industry analysis is a vital complement to company analysis. To thoroughly analyze a company itself, the analyst needs to understand the context in which the company operates. As analysts examine the competitive structure of an industry, they should always be thinking about what attributes could change in the future. Needless to say, industry analysis must be forward looking and could ultimately help determine an overweight or underweight to sector positioning.

Analysis of the competitive environment with an emphasis on the implications of the environment for corporate strategy is known as **strategic analysis**. Michael Porter's "five forces" framework is the classic starting point for strategic analysis; although it was originally aimed more at internal managers of businesses than at external security analysts, the framework is useful to both.

Porter (2008) identified five determinants of the intensity of competition in an industry: The threat of entry, the power of suppliers, the power of buyers, the threat of substitutes, and rivalry among existing competitors. Exhibit 6 illustrates the Porter framework.

Exhibit 6 Porter's Five Forces



Threat of New Entrants
Depends on the existence and extent of barriers to enter the industry.
Q: How difficult would it be for new competitors to enter the industry? Industries that are easy to enter will generally be more competitive than industries with high barriers to entry.

Bargaining Power of Suppliers
Affected by:
- concentration of industry is,
- switching costs of suppliers' customers,
- supply substitutes.

Q: Are suppliers able to increase prices, restrict supply, or pass on rising input costs?
Suppliers of scarce or limited parts or elements often possess significant pricing power.

Rivalry (intra-industry)
Rivalry among existing competitors; price discipline. Depends on:
- Industry structure and concentration/fragmentation
- Product or service differentiation
- Industry exit barriers

Bargaining Power of Customers
Affected by:
- size and concentration of customers,
- costs of switching to other suppliers,
- customers' ability to produce the product or service themselves.

Are customers able to force price reductions or better payment terms ?

Threat of Substitutes (products or services)
Affected by the availability of (alternative) product categories that can satisfy customer needs
Q: Are lower-priced brands close substitutes for premium brands and vice versa ?

Addressing the following questions should help the analyst evaluate the threat of new entrants and the level of competition in an industry and thereby provide an effective base for describing and analyzing the industry:

- What are the barriers to entry? Is it difficult or easy for a new competitor to challenge incumbents? Relatively high (low) barriers to entry imply that the threat of new entrants is relatively low (high).
- How concentrated is the industry? Do a small number of companies control a relatively large share of the market, or does the industry have many players, each with a small market share?
- What are capacity levels? That is, based on existing investment, how much of the goods or services can be delivered in a given time frame? Does the industry suffer chronic over- or undercapacity, or do supply and demand tend to come into balance reasonably quickly in the industry?
- How stable are market shares? Do companies tend to rapidly gain or lose share, or is the industry stable?
- Where is the industry in its life cycle? Does it have meaningful growth prospects, or is demand stagnant/declining?
- How important is price to the customer's purchase decision?

The answers to these questions are elements of any thorough industry analysis. They are explored each in turn in the sections that follow.

5.2 Barriers to Entry

h explain the effects of barriers to entry, industry concentration, industry capacity, and market share stability on pricing power and price competition

When a company is earning economic profits, the chances that it will be able to sustain them over time are greater, all else being equal, if the industry has high barriers to entry. The ease with which new competitors can challenge incumbents is often an important factor in determining the competitive landscape of an industry. If new competitors can easily enter the industry, the industry is likely to be highly competitive because high returns on invested capital will quickly be competed away by new entrants eager to grab their share of economic profits. As a result, industries with low barriers to entry often have little pricing power because price increases that raise companies' returns on capital will eventually attract new competitors to the industry.

If incumbents are protected by barriers to entry, the threat of new entrants is lower and incumbents may enjoy a more benign competitive environment. Often, these barriers to entry can lead to greater pricing power, because potential competitors would find it difficult to enter the industry and undercut incumbents' prices. Of course, high barriers to entry do not guarantee pricing power, because incumbents may compete fiercely among each other.

A classic example of an industry with low barriers to entry is restaurants. Anyone with a modest amount of capital and some culinary skill can open a restaurant, and popular restaurants quickly attract competition. As a result, the industry is very competitive, and many restaurants fail in their first few years of business.

At the other end of the spectrum of barriers to entry are the global credit card networks, such as MasterCard and Visa, both of which often post operating margins greater than 30%. Such high profits should attract competition, but the barriers to entry are extremely high. Capital costs are one hurdle; also, building a massive data-processing network would not be cheap. Imagine for a moment that a venture capitalist were willing to fund the construction of a network that would replicate the physical infrastructure of the incumbents; the new card-processing company would have to convince millions of consumers to use the new card and convince thousands of merchants to accept the card. Consumers would not want to use a card that merchants did not accept, and merchants would not want to accept a card that few consumers carried. This problem would be difficult to solve, which is why the barriers to entering this industry are quite high. The barriers help preserve the profitability of the incumbent players.

One way of understanding barriers to entry is simply by thinking about what it would take for new players to compete in an industry. How much money would they need to spend? What kind of intellectual capital would they need to acquire? How easy would it be to attract enough customers to become successful?

Another way to investigate the issue is by looking at historical data. How often have new companies tried to enter the industry? Is a list of industry participants today markedly different from what it was 5 or 10 years ago? These kinds of data can be very helpful because the information is based on the real-world experience of many entrepreneurs and businesses making capital allocation decisions. If an industry has seen a flood of new entrants over the past several years, the odds that the barriers are low are good; conversely, if the same 10 companies that dominate an industry today dominated it 10 years ago, barriers to entry are probably fairly high.

Do not confuse barriers to *entry*, however, with barriers to *success*. In some industries, entering may be easy but becoming successful enough to threaten the incumbents might be quite hard.

Also, high barriers to entry do not automatically lead to good pricing power and attractive industry economics. Consider the cases of the auto manufacturing, commercial aircraft manufacturing, and oil refining industries. Starting up a new company in any of these industries is very difficult. Aside from the massive capital costs, there would be significant other barriers to entry: A new automaker would need manufacturing expertise and a dealer network, an aircraft manufacturer would need a significant amount of intellectual capital, and a refiner would need process expertise and regulatory approvals. Despite the high barriers to entry, the industries are all quite competitive with limited pricing power. Very few industry participants reliably generate returns on capital in excess of their costs of capital. Exiting the industries and re-deploying capacity to other uses when demand conditions worsen are also difficult, meaning that the industries can be prone to overcapacity.

A final consideration when analyzing barriers to entry is that they can change over time as new technologies or ways of conducting business become available. For example outsourcing of some activities means that companies can focus on product or service design, without the need to build and own production capacity.

Exhibit 7 provides an overview of barriers to entry for three representative industries.

Exhibit 7 Elements of a Strategic Analysis for Three Industries: Barriers to Entry

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Major Companies	Pfizer, Novartis, Merck, GSK plc	Schlumberger, Baker Hughes, Halliburton	Hershey, Mars/Wrigley, Nestle
Barriers to Entry	*Very High*: Substantial financial and intellectual capital required to compete effectively. A potential new entrant would need to create a sizable R&D operation, a global distribution network, and large-scale manufacturing capacity.	*High*: Technological expertise is required, but a high level of innovation allows niche companies to enter the industry and compete in specific areas.	*Medium/Low*: Low financial or technological hurdles, but new players would lack the established brands that drive consumer purchase decisions and potential supply chain operations.

5.3 Industry Concentration

Much like industries with barriers to entry, industries that are concentrated among a relatively small number of players often experience relatively less price competition. Again, there are important exceptions, so the reader should not automatically assume that concentrated industries always have pricing power or that fragmented industries do not. The degree of concentration is often measured using the Herfindahl–Hirschman Index.

An analysis of industry concentration should start with market share: What percentage of the market does each of the largest players have, and how large are those shares relative to each other and relative to the remainder of the market? Often, the *relative* market shares of competitors matter as much as their *absolute* market shares.

For example, the global market for commercial aircraft (for 100 or more passengers) is extremely concentrated; only Boeing and Airbus manufacture this type of plane on significant scale. The two companies tend to have broadly similar market shares, however, and control essentially the entire market. Because neither enjoys a scale advantage relative to its competitor and because any business gained by one is lost by the other, competition tends to be fierce.

This situation contrasts with the market for home improvement products in the United States, which is dominated by Home Depot and Lowe's. These two companies had about 11% and 7% market shares for the decade up to 2020, respectively, which does not sound very large. However, the next largest competitor had only 2% of the market, and most market participants are tiny, with miniscule market shares. Both Home Depot and Lowe's have historically posted high returns on invested capital, in part because they could profitably grow by squeezing out smaller competitors rather than engaging in fierce competition with each other.

Fragmented industries tend to be highly price competitive for several reasons. First, the large number of companies makes coordination difficult because there are too many competitors for each industry member to monitor effectively. Second, each player has such a small piece of the market that even a small gain in market share can make a meaningful difference to its fortunes, which increases the incentive of each company to undercut prices and attempt to steal market share. Finally, the large number of players encourages industry members to think of themselves individualistically rather than as members of a larger group, which can lead to fierce competitive behavior.

In concentrated industries, in contrast, each player can relatively easily keep track of what its competitors are doing, which makes tacit coordination much more feasible. Also, leading industry members are large, which means they have more to lose—and proportionately less to gain—by destructive price behavior. Large companies are also more tied to the fortunes of the industry as a whole, making them more likely to consider the long-run effects of a price war on overall industry economics.

As with barriers to entry, the level of industry concentration is a guideline rather than a hard and fast rule when thinking about the level of pricing power in an industry. For example, Exhibit 8 shows a rough classification of industries based on analysis by Morningstar after asking its equity analysts whether industries were characterized by strong or weak pricing power and whether those industries were concentrated or fragmented. Examples of companies in industries are included in parentheses. In the upper right quadrant ("concentrated with weak pricing power"), those industries that are capital intensive and/or sell commodity-like products are shown in boldface.

Exhibit 8 A Two-Factor Analysis of Industries

Concentrated with Strong Pricing Power

Orthopedic Devices (Zimmer, Smith & Nephew)
Biotech (Amgen, Genzyme)
Pharmaceuticals (Merck & Co., Novartis)
Industrial Gases (Praxair, Air Products and Chemicals)
Enterprise Networking (Cisco Systems)
US Defense (General Dynamics)
Heavy Construction Equipment (Caterpillar, Komatsu)
Credit Card Networks (MasterCard, Visa)
Investment Banking/Mergers & Acquisitions (Goldman Sachs, UBS)
Futures Exchanges (Chicago Mercantile Exchange, Intercontinental Exchange)
Tobacco (Philip Morris, British American Tobacco)
Alcoholic Beverages (Diageo, Pernod Ricard)

Concentrated with Weak Pricing Power

Commercial Aircraft (Boeing, Airbus)
Automobiles (General Motors, Toyota, Daimler)
Memory (DRAM & Flash Product, Samsung, Hynix)
Semiconductor Equipment (Applied Materials, Tokyo Electron)
Generic Drugs (Teva Pharmaceutical Industries, Sandoz)
Printers/Office Machines (HP, Lexmark)
Refiners (Valero, Marathon Oil)
Major Integrated Oil (BP, ExxonMobil)

Fragmented with Strong Pricing Power

For-Profit Education (Apollo Group, DeVry University)
Analog Chips (Texas Instruments, STMicroelectronics)
Industrial Distribution (Fastenal, W.W. Grainger)
Propane Distribution (AmeriGas, Ferrellgas)
Private Banking (Northern Trust, Credit Suisse)

Fragmented with Weak Pricing Power

Consumer Packaged Goods (Procter & Gamble, Unilever)
Retail (Walmart, Carrefour Group)
Marine Transportation (Maersk Line, Frontline)
Solar Panels
Airlines
Restaurants
Life Insurance

Source: Adapted from Morningstar Equity Research.

The industries in the top right quadrant defy the "concentration is good for pricing" guideline. When we examine these concentrated-yet-competitive industries, a clear theme emerges: Many industries in this quadrant (the boldface ones) are highly capital intensive and sell commodity-like products. Exiting the industries and re-deploying capacity to other uses when demand conditions worsen are also difficult, meaning that the industries can be prone to overcapacity. Also, if the industry sells a commodity product that is difficult—or impossible—to differentiate, the incentive to compete on price increases because a lower price frequently results in greater market share.

Generally, industry concentration is a good indicator that an industry has pricing power and rational competition, but other factors may override the importance of concentration. Industry fragmentation is a much stronger signal that the industry is competitive with limited pricing power.

The industry characteristics discussed here are guidelines meant to steer the analyst in a particular direction, not rules that should cause the analyst to ignore other relevant analytical factors.

Exhibit 9 Elements of a Strategic Analysis for Three Industries: Levels of Concentration

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Major Companies	Pfizer, Novartis, Merck, GSK plc	Schlumberger, Baker Hughes, Halliburton	Hershey, Mars/Wrigley, Nestle
Level of Concentration	*Concentrated*: A small number of companies control the bulk of the global market for branded drugs. Recent mergers have increased the level of concentration.	*Fragmented*: Although only a small number of companies provide a full range of services, many smaller players compete effectively in specific areas. Service arms of national oil companies may control significant market share in their own countries, and some product lines are concentrated in the mature US market.	*Very Concentrated*: The top four companies have a large proportion of global market share. Recent mergers have increased the level of concentration.

5.4 Industry Capacity

The effect of industry capacity (the maximum amount of a good or service that can be supplied in a given time period) on pricing is clear: Tight, or limited, capacity gives participants more pricing power because demand for the product or service exceeds supply, whereas overcapacity leads to price cutting and a very competitive environment because excess supply chases demand. An analyst should think about not only current capacity conditions but future changes in capacity levels. How quickly can companies in the industry adjust to fluctuations in demand? How flexible is the industry in bringing supply and demand into balance? What will be the effect of that process on industry pricing power or on industry margins?

5.4.1 *Capacity: Short term vs. Long Term*

Generally, capacity is fixed in the short term and variable in the long term because capacity can be increased (e.g., new factories can be built) if time is sufficient. What is considered "sufficient" time—and, therefore, the duration of the short term, in which capacity cannot be increased—may vary dramatically among industries. Sometimes, adding capacity takes years to complete, as in the case of the construction of a new manufacturing plant for pharmaceuticals, which is complex and subject to regulatory requirements. In other situations, capacity may be added or reduced relatively quickly, as in the case of service industries, such as advertising. In cyclical markets, such as commercial paper and paperboard, capacity conditions can change rapidly. Strong demand in the early stages of an economic recovery can result in the addition of supply. Given the long lead times to build manufacturing plants, new supply may reach the market just as demand slows, rapidly changing capacity conditions from tight to loose. Such considerations underscore the importance of forecasting long-term industry demand in evaluating industry investments in capacity.

5.4.2 *Physical Capacity*

Generally, if new capacity is physical—for example, an auto manufacturing plant or a massive cargo ship—it will take longer for new capacity to come on line to meet an increase in demand, resulting in a longer period of tight conditions. Unfortunately, capacity additions frequently overshoot long-run demand, and because physical capital is often hard to re-deploy, industries reliant on physical capacity may get stuck in conditions of excess capacity and diminished pricing power for an extended period.

Note that capacity need not be physical. After Hurricane Katrina caused enormous damage to the southeastern United States in 2005, reinsurance rates quickly spiked as customers sought to increase their financial protection from future hurricanes. However, these high reinsurance rates enticed a flood of fresh capital into the reinsurance market, and a number of new reinsurance companies were founded, which brought rates back down.

Financial and human capital, in contrast, can be quickly shifted to new uses. In the reinsurance example, for instance, financial capital was quick to enter the reinsurance market and take advantage of tight capacity conditions, but if too much capital had entered the market, some portion of that capital could easily have left to seek higher returns elsewhere. Money can be used for many things, but massive bulk cargo vessels are not useful for much more than transporting heavy goods across oceans. Exhibit 10 shows the elements of a strategic analysis relating to industry capacity.

Exhibit 10 Elements of a Strategic Analysis: Industry Capacity

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Major Companies	Pfizer, Novartis, Merck, GSK plc	Schlumberger, Baker Hughes, Halliburton	Hershey, Mars/Wrigley, Nestle
Impact of Industry Capacity	*Not Applicable*: Pharmaceutical pricing is primarily determined by patent protection and regulatory issues, including government approvals of drugs and of manufacturing facilities. Manufacturing capacity is of little importance.	*Medium/High*: Demand can fluctuate quickly depending on commodity prices, and industry players often find themselves with too few (or too many) employees on the payroll.	*Not Applicable*: Pricing is driven primarily by brand strength. Manufacturing capacity has little effect.

EXAMPLE 7

Which of the following companies is most likely to have the greatest ability to quickly increase capacity?

A Legal services provider

B Manufacturing company producing heavy machinery

C Company operating cargo ships

Solution

A is correct. Capacity increases in providing legal services would not require significant fixed capital investments. B and C are incorrect because the companies would require capital investments and capacity expansion would take time to implement.

5.5 Market Share Stability

Examining the stability of industry market shares over time is similar to thinking about barriers to entry and the frequency with which new players enter an industry. In fact, barriers to entry and the frequency of new product introductions, together with such factors as product differentiation, affect market shares. Stable market shares typically indicate less competitive industries; unstable market shares often indicate highly competitive industries that have limited pricing power. Exhibit 11 illustrates the development of market shares in two medical devices markets.

Exhibit 11 Market Shares in Orthopedic and Metal Mesh Devices

Over 2010–2020, the orthopedic device industry—mainly artificial hips and knees—has been a relatively stable global oligopoly led by Stryker, Zimmer Biomet, Smith & Nephew, and Johnson & Johnson, which jointly control about three quarters of the global market.

In contrast, the market for stents—small metal mesh devices used to prop open blocked arteries—while controlled by a handful of companies, has seen market shares change from being very stable to being marked by rapid change. Johnson & Johnson, which together with Boston Scientific, dominated the US stent market for many years, went from having about half the market in 2007 to having only 15% in early 2009 and exited the market in 2011; over the same period, Abbott Laboratories increased its market share from 0% to around 30%. The reason for this change was the launch of new stents by Abbott and Medtronic, which took market share from Johnson & Johnson's and Boston Scientific's established stents.

Orthopedic device companies have experienced more stability in their market shares for two reasons. First, artificial hips and knees are complicated to implant, and each manufacturer's products are slightly different. As a result, orthopedic surgeons become proficient at using one or several companies' devices and may be reluctant to incur the time and cost of learning how to implant products from a competing company. The second reason is the relatively slow pace of innovation in the orthopedic device industry, making the benefit of switching among product lines relatively low. In addition, the number of orthopedic device companies has remained fairly static over many years.

In contrast, the US stent market has experienced rapid shifts in market shares because of several factors. First, interventional cardiologists seem to be more open to implanting stents from different manufacturers; that tendency may reflect lower switching costs for stents relative to orthopedic devices and the interchangeability of stents. More importantly, however, the pace of innovation in the stent market has become quite rapid, giving cardiologists added incentive to switch to newer stents, with potentially better patient outcomes, as they became available.

Exhibit 11 (Continued)

Low switching costs plus a relatively high benefit from switching caused market shares to change quickly in the stent market. High switching costs for orthopedic devices coupled with slow innovation resulted in a lower benefit from switching, which led to greater market share stability in orthopedic devices.

Exhibit 12 briefly describes market share stability in the three representative industries.

Exhibit 12 Elements of a Strategic Analysis: Market Share

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Major Companies	Pfizer, Novartis, Merck, GSK plc	Schlumberger, Baker Hughes, Halliburton	Hershey, Mars/Wrigley, Nestle
Industry Stability	*Stable*: The branded pharmaceutical market is dominated by major companies and consolidation via mega-mergers. Market shares shift quickly, however, as new drugs are approved and gain acceptance or lose patent protection.	*Unstable*: Market shares may shift frequently depending on technology offerings and demand levels.	*Very Stable*: Market shares change glacially slowly.

5.6 Price Competition

A useful tool for analyzing an industry is attempting to think like a customer of the industry. Whatever factor most influences customer purchase decisions is likely to also be the focus of competitive rivalry in the industry. In general, industries for which price is a large factor in customer purchase decisions tend to be more competitive than industries in which customers value other attributes more highly.

Although this depiction may sound like the description of a commodity industry versus a non-commodity industry, it is, in fact, a bit more subtle. Commercial aircraft and passenger cars are certainly more differentiated than lumps of coal or gallons of gasoline, but price nonetheless weighs heavily in the purchase decisions of buyers of aircraft and cars, because fairly good substitutes are easily available. If Airbus charges too much for an aircraft, such as the A350, an airline can buy as an alternative a Boeing 777 (a small amount of "path dependence" characterizes the airline industry, in that an airline with a large fleet of a particular Airbus model will be marginally more likely to stick with that model for a new purchase than it will be to buy a Boeing). If BMW's price for a four-door premium sedan rises too high, customers can switch to an alternative premium brand with similar features. Similar switching can be expected as a result of a unilateral price increase in the case of most industries in the "Weak Pricing Power" column of the two-factor analysis in Exhibit 8.

Contrast industries in software services that can be characterized by strong pricing power. Such software companies as Adobe, for example, sell a suite of products for creative professionals. Although Adobe's individual products do face competition, its successful transition to selling most of its products via its Creative Cloud has enabled the company to enjoy gross margins above 80%. In addition, the growth of digital companies and content bodes well for Adobe's pricing positioning in the industry.

Returning to a more capital-intensive industry, consider heavy-equipment manufacturers, such as Caterpillar, JCB, and Komatsu. A large wheel loader or combine harvester requires a large capital outlay, so price certainly plays a part in the buyers' decisions. However, other factors are important enough to customers to allow these companies a small amount of pricing power. Construction equipment is typically used as a complement to other gear on a large project, which means that downtime for repairs increases costs because, for example, hourly laborers must wait for a bulldozer to be fixed. Broken equipment is also expensive for agricultural users, who may have only a few days in which to harvest a season's crop. Because of the importance to users of their products' reliability and their large service networks—which are important "differentiators," or factors bestowing a competitive advantage—Caterpillar, Komatsu, and Deere have historically been able to price their equipment at levels that have generated solid returns on invested capital.

5.7 Industry Life Cycle

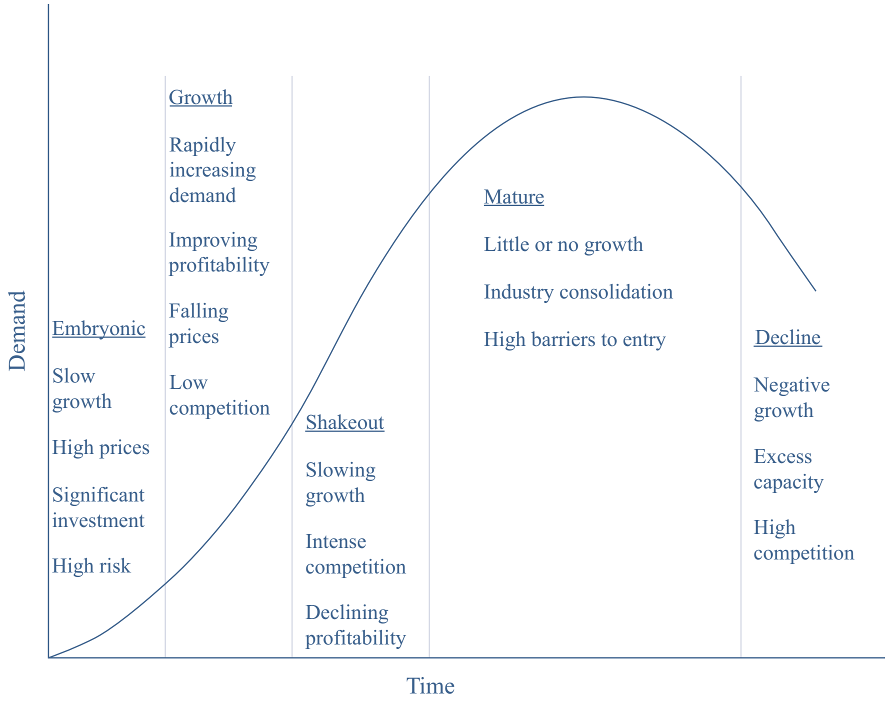
i. describe industry life-cycle models, classify an industry as to life-cycle stage, and describe limitations of the life-cycle concept in forecasting industry performance

An industry's life-cycle position often has a large impact on its competitive dynamics, making this position an important component of the strategic analysis of an industry.

5.7.1 *Description of an Industry Life-Cycle Model*

Industries, like individual companies, tend to evolve over time and usually experience significant changes in the rate of growth and levels of profitability along the way. Just as an investment in an individual company requires careful monitoring, industry analysis is a continuous process to identify changes that may be occurring or likely to occur. A useful framework for analyzing the evolution of an industry is an industry life-cycle model, which identifies the sequential stages that an industry typically goes through. The five stages of an industry life-cycle model are embryonic, growth, shakeout, mature, and decline. Each stage is characterized by different opportunities and threats. Exhibit 13 shows the model as a curve illustrating the level and growth rate of demand at each stage.

Exhibit 13 An Industry Life-Cycle Model


Demand
Embryonic
Slow growth
High prices
Significant investment
High risk
Growth
Rapidly increasing demand
Improving profitability
Falling prices
Low competition
Shakeout
Slowing growth
Intense competition
Declining profitability
Mature
Little or no growth
Industry consolidation
High barriers to entry
Decline
Negative growth
Excess capacity
High competition
Time

Source: Based on Figure 2.4 in Hill and Jones (2008).

Embryonic An embryonic industry is one that is just beginning to develop. For example, in 2014, the mobile food delivery industry was in the embryonic stage (it has grown to become a nearly US$15 billion industry in 2020), and in 1997, the global social media industry was just getting started. Characteristics of the embryonic stage include slow growth and high prices because customers tend to be unfamiliar with the industry's product and volumes are not yet sufficient to achieve meaningful economies of scale. Increasing product awareness and developing distribution channels are key strategic initiatives of companies during this stage. Substantial investment is generally required, and the risk of failure is high. A majority of startup companies do not succeed.

Growth A growth industry tends to be characterized by rapidly increasing demand, improving profitability, falling prices, and relativity low competition among companies in the industry. Demand is fueled by new customers entering the market, and prices fall as economies of scale are achieved and as distribution channels develop. The threat of new competitors entering the industry is usually highest during the growth stage, when barriers to entry are relatively low. Competition tends to be relatively limited, however, because rapidly expanding demand provides companies with an opportunity to grow without needing to capture market share from competitors. Industry profitability improves as volumes rise and economies of scale are attained.

Shakeout The shakeout stage is usually characterized by slowing growth, intense competition, and declining profitability. During the shakeout stage, demand approaches market saturation levels because few new customers are left to enter the market. Competition is intense as growth becomes increasingly dependent on market share gains. Excess industry capacity begins to develop as the rate at which companies continue to invest exceeds the overall growth of industry demand. In an effort to boost volumes to fill excess capacity, companies often cut prices, so industry profitability begins to decline. During the shakeout stage, companies increasingly focus on reducing their cost structure (restructuring) and building brand loyalty. Marginal companies may fail or merge with others.

Mature Characteristics of a mature industry include little or no growth, industry consolidation, and relatively high barriers to entry. Industry growth tends to be limited to replacement demand and population expansion because the market at this stage is completely saturated. As a result of the shakeout, mature industries often consolidate and become oligopolies. The surviving companies tend to have brand loyalty and relatively efficient cost structures, both of which are significant barriers to entry. During periods of stable demand, companies in mature industries tend to recognize their interdependence and try to avoid price wars. Periodic price wars do occur, however, most notably during periods of declining demand (such as during economic downturns). Companies with superior products or services are likely to gain market share and experience above-industry-average growth and profitability.

Decline During the decline stage, industry growth turns negative, excess capacity develops, and competition increases. Industry demand at this stage may decline for a variety of reasons, including technological substitution (for example, the newspaper industry has been declining for many years as more people turn to video platforms, the internet, and 24-hour cable news networks for information), social changes, and global competition. As demand falls, excess capacity in the industry forms and companies respond by cutting prices, which often leads to price wars. The weaker companies often exit the industry at this point, merge, or redeploy capital into different products and services.

When overall demand for an industry's products or services is declining, the opportunity for individual companies to earn above-average returns on invested capital tends to be less than when demand is stable or increasing, because of price cutting and higher per-unit costs as production is cut back. Exhibit 14 shows the life cycle stages for three representative industries.

Exhibit 14 Elements of a Strategic Analysis: Industry Life Cycle

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Major Companies	Pfizer, Novartis, Merck, GSK plc	Schlumberger, Baker Hughes, Halliburton	Hershey, Mars/Wrigley, Nestle
Life Cycle	*Mature*: Overall demand does not change greatly from year to year.	*Mature*: Demand does fluctuate with energy prices, but normalized revenue growth is only in the mid-single digits.	*Mature*: Growth is driven by population trends and pricing.

5.7.2 *Using an Industry Life-Cycle Model*

In general, new industries tend to be more competitive (with lots of players entering and exiting) than mature industries, which often have stable competitive environments and players that are more interested in protecting what they have than in gaining lots of market share. However, as industries move from maturity to decline, competitive pressures may increase again as industry participants perceive a zero-sum environment and fight over pieces of an ever-shrinking pie.

5.7.3 *Relating Management Behavior to Industry Life Cycle*

An important point for the analyst to think about is where a company is relative to where its industry sits in the life cycle. Companies in growth industries should be building customer loyalty as they introduce consumers to new products or services, building scale, and reinvesting heavily in their operations to capitalize on increasing demand. They are probably not focusing strongly on internal efficiency. Growth companies typically reinvest their cash flows in new products and product platforms rather than return cash flows to shareholders because these companies still have many opportunities to deploy their capital to make higher returns. Some may be able to sustain or even accelerate growth through innovation or by expansion into new markets.

Companies in mature industries are likely to be pursuing replacement demand or incremental demand rather than new buyers and are probably focused on extending successful product lines rather than introducing revolutionary new products. They are also probably focusing on cost rationalization and efficiency gains rather than on taking lots of market share. Importantly, these companies have fewer growth opportunities than in the previous stage and thus more limited avenues for profitably reinvesting capital, but they often have strong cash flows. Given their strong cash flows and relatively limited reinvestment opportunities, such companies should be, according to a common perspective, returning capital to shareholders via share repurchases or dividends.

5.7.4 *Spotting Warning Signs*

What can be a concern is a middle-aged company acting like a young growth company and pouring capital into projects with low return prospects in an effort to pursue size for its own sake. Many companies have a difficult time managing the transition from growth to maturity, and their returns on capital—and shareholder returns—may suffer until management decides to allocate capital in a manner more appropriate to the company's life-cycle stage.

5.7.5 *Limitations of Industry Life-Cycle Analysis*

Although models can provide a useful framework for thinking about an industry, the evolution of an industry does not always follow a predictable pattern. Various external factors that we will discuss in the subsequent sections influence life cycles. These include technological, demographic, social, environmental, or regulatory factors that may significantly affect the shape of the pattern, causing some stages to be longer or shorter than expected and, in certain cases, causing some stages to be skipped altogether. Technological changes may cause an industry to experience an abrupt shift from growth to decline, thus skipping the shakeout and mature stages. For example, the movie rental industry experienced rapid change as consumers switched from renting physical movies (on VHS tapes and DVDs) to on-demand services, such as downloading movies from the internet or through their cable providers.

Regulatory changes can also have a profound impact on the structure of an industry. A prime example is the deregulation of the telecommunications industry in the United States and Europe in the 1990s, which transformed a monopolistic industry into an intensely competitive one.

Social changes also have the ability to affect the profile of an industry. The casual dining industry has benefited over the past 30 years from the increase in the number of dual-income families in many large markets, who often have more income but less time to cook meals to eat at home. Thus, life-cycle models tend to be most useful for analyzing industries during periods of relative stability. They are less practical when the industry is experiencing rapid change over a short time period because of external or other special circumstances.

Another limiting factor of life-cycle models is that not all companies in an industry experience similar performance. The key objective for the analyst is to identify the potential winners while avoiding potential losers. Highly profitable companies can exist in competitive industries with below-average profitability—and vice versa, as shown by the cellular phone manufacturer Nokia for a number of years starting in the 1990s. It has been able to use its scale to generate levels of profitability that were well above average despite operating in a highly competitive industry. In contrast, despite the historically above-average growth and profitability of the software industry, countless examples exist of software companies that failed to ever generate a profit and eventually went out of business.

EXAMPLE 8

Industry Life Cycle

1 An industry experiencing slow growth and high prices is best characterized as being in the:

- **A** mature stage.
- **B** shakeout stage.
- **C** embryonic stage.

2 Which of the following statements about the industry life-cycle model is *least* accurate?

- **A** The model is more appropriately used during a period of rapid change than during a period of relative stability.
- **B** External factors may cause some stages of the model to be longer or shorter than expected, and in certain cases, a stage may be skipped entirely.
- **C** Not all companies in an industry will experience similar performance, and very profitable companies can exist in an industry with below-average profitability.

Solution to 1:

C is correct. Both slow growth and high prices are associated with the embryonic stage. High price is not a characteristic of the mature or shakeout stage.

Solution to 2:

A is correct. This statement is the least accurate. The model is best used during a period of relative stability rather than during a period of rapid change.

6

EXTERNAL INFLUENCES ON INDUSTRY

j describe macroeconomic, technological, demographic, governmental, social, and environmental influences on industry growth, profitability, and risk

We now turn our attention to external factors that affect industries. We will explain macroeconomic, technological, demographic, governmental, social, and environmental influences.

6.1 Macroeconomic Influences

Trends in overall economic activity generally have significant effects on the demand for an industry's products or services. These trends can be cyclical (i.e., related to the changes in economic activity caused by the business cycle) or structural (i.e., related to enduring changes in the composition or magnitude of economic activity). Among the economic variables that usually affect an industry's revenues and profits are the following:

- GDP, either in current or constant currency (inflation-adjusted) terms;
- interest rates, which represent the cost of debt to consumers and businesses and are important ingredients in financial institutions' revenues and costs;
- the availability of credit, which affects business and consumer spending and financial solvency; and
- inflation, which reflects the changes in prices of goods and services and influences costs, interest rates, and consumer and business confidence.

6.2 Technological Influences

New technologies create new or improved products that can radically change an industry and can also change how other industries that use the products conduct their operations. Exhibit 15 provides an illustration of how innovation bringing new technologies changed the compute hardware and software industry.

Exhibit 15 Technological Influences

The computer hardware industry provides one of the best examples of how technological change can affect industries. The 1958 invention of the microchip (also known as an "integrated circuit") enabled the computer hardware industry to eventually create a new market of personal computing for the general public and radically extended the use of computers in business, government, and educational institutions.

Moore's law states that the number of transistors that can be inexpensively placed on an integrated circuit doubles approximately every two years. Several other measures of digital technology have improved at exponential rates related to Moore's law, including the size, cost, density, and speed of components. As a result of these trends, the computer hardware industry came to dominate the fields of hardware for word processing and many forms of electronic communication and home entertainment. The computing industry's integrated circuit innovation increased economies of scale and erected large barriers to new entrants because the capital costs of innovation and production became very high. Intel capitalized on both factors, which allowed it to garner an industry market leadership position and to become the dominant supplier of the PC industry's highest-value component (the microprocessor). Thus, Intel became dominant because of its cost advantage, brand power, and access to capital.

Along the way, the computer hardware industry was supported and greatly assisted by the complementary industries of computer software and telecommunications (particularly in regard to development of the internet); also important were other industries—entertainment (television, movies, games), retailing, and print media. Ever more powerful integrated circuits, advances in wireless technology, and the convergence of media, which the internet and new wireless technology have facilitated, continue to reshape the uses and the roles of PC hardware in business and personal life. In the middle of the 20th century,

(continued)

Exhibit 15 (Continued)

few people in the world would have imagined they would ever have any use for a home computer. Today, an estimated 4.3 billion people, or over half the world's population, have access to connected computing. For the United States, the estimate is at least 95% of the population; it is much less in emerging and underdeveloped countries. More than 8 billion mobile cellular telephone subscriptions exist in the world today.

Exhibit 16 explains how three representative industries are affected by technological change.

Exhibit 16 Technological Influences: Representative Industries

	Branded Pharmaceuticals	**Oil Services**	**Confections/Candy**
Technological Influences	*Medium/High*: Biologic (large-molecule) drugs are pushing new therapeutic boundaries, and many large pharmaceutical companies have a relatively small presence in biotech.	*Medium/High*: Industry is reasonably innovative, and players must reinvest in R&D to remain competitive. Temporary competitive advantages are possible via commercialization of new processes or exploitation of accumulated expertise.	*Very Low*: Innovation does not play a major role in the industry.

6.3 Demographic Influences

Changes in population size, in the distributions of age and gender, and in other demographic characteristics may have significant effects on economic growth and on the amounts and types of goods and services consumed.

The effects of demographics on industries are exemplified by the impact of Japan's aging population, which has one of the highest percentages of elderly residents (26% over the age of 65) and a very low birth rate. Japan's ministry of health estimates that by 2055, the percentage of the population over 65 will rise to 40% and the total population will fall by 25%. These demographic changes are expected by some observers to have negative effects on the overall economy because, essentially, they imply a declining workforce. However, some sectors of the economy stand to benefit from these trends—for example, the health care industry. Another example of demographic influence on an industry is the experience of "baby boomers" (those born in the years after World War II), who reached adulthood in significant numbers in the early 1970s, contributing to a strong demand for new housing throughout the 1970s and during 1995–2005 when their children reached adulthood.

6.4 Governmental Influences

Governmental influence on industries' revenues and profits is pervasive and important. In setting tax rates for corporations and individuals, governments affect profits and incomes, which, in turn, affect corporate and personal spending. Governments also set rules and regulations to protect consumers, employees, and the environment, affecting how companies operate, with implications for profitability. In addition, governments are also major purchasers of goods and services from a range of industries.

Often, governments exert their influence indirectly by empowering other regulatory or self-regulatory organizations (e.g., stock exchanges, medical associations, utility rate setters, and other regulatory commissions) to govern the affairs of an industry. By setting the terms of entry into various sectors, such as financial services and health care, and the rules that companies and individuals must adhere to in these fields, governments control the supply, quality, and nature of many products and services and the public's access to them. For example, in the financial industry, the acceptance of savings deposits from and the issuance of securities to the investing public are usually tightly controlled by governments and their agencies. This control is imposed through rules designed to protect investors from fraudulent operators and to ensure that investors receive adequate disclosure about the nature and risks of their investments. Exhibit 17 illustrates government influences on three representative industries.

Exhibit 17 Government Influences: Representative Industries

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Government & Regulatory Influences	*Very High*: All drugs must be approved for sale by national safety regulators. Patent regimes may differ among countries. Also, health care is heavily regulated in most countries.	*Medium*: Regulatory frameworks can affect energy demand at the margin. Also, governments play an important role in allocating exploration opportunities to E&P companies, which can indirectly affect the amount of work flowing down to service companies.	*Low*: Industry is not regulated, but childhood obesity concerns in developed markets are a low-level potential threat. Also, high-growth emerging markets may block entry of established players into their markets, possibly limiting growth.

EXAMPLE 9

The Effects of Purchases by Government-Related Entities on the Aerospace Industry

The aerospace, construction, and firearms industries are prime examples of industries for which governments are major customers and whose revenues and profits are significantly—in some cases, predominantly—affected by their sales to governments. An example is Airbus, a major player in aerospace, defense, and related services with head offices in Paris and Ottobrunn, Germany. In 2017, Airbus generated revenues of €66.8 billion and employed an international workforce of about 130,000. Besides being a leading manufacturer of commercial aircraft, a business whose customers include government-owned airlines, Airbus also includes Airbus Military, providing tanker, transport, and mission aircraft; Airbus Helicopters SAS, the world's largest helicopter supplier; and EADS Astrium, the European leader in space programs, including Ariane and Galileo. Its Defence & Security Division is a provider of comprehensive systems solutions and makes Airbus the major partner in the Eurofighter consortium and a stakeholder in missile systems provider MBDA. These divisions are highly dependent on orders from governments.

6.5 Social Influences

Societal changes involving how people work, spend their money, enjoy their leisure time, and conduct other aspects of their lives can have significant effects on the sales of various industries.

Tobacco consumption in the United Kingdom provides a good example of the effects of social influences on an industry. Although the role of government in curbing tobacco advertising, legislating health warnings on the purchases of tobacco products, and banning smoking in public places (such as restaurants, bars, public houses, and transportation vehicles) probably has been the most powerful apparent instrument of changes in tobacco consumption, the forces underlying that change have really been social in nature. These forces are increasing consciousness on the part of the population of the damage to the health of tobacco users and those in their vicinity from smoking, the increasing cost to individuals and governments of the chronic illnesses caused by tobacco consumption, and the accompanying shift in public perception of smokers from socially correct to socially incorrect or even inconsiderate or reckless. As a result of these changes in society's views of smoking, cigarette consumption in the United Kingdom declined from 102.5 billion cigarettes in 1990 to less than 40.0 billion three decades later, placing downward pressure on tobacco companies' unit sales.

6.6 Environmental Influences

As industries continue to adapt to new technologies and strategies in order to compete and grow, the need to evaluate and mitigate environmental impact is an important influence. For example, climate change is shifting the way entire industries are perceived and evolving. With limited resources on our planet as a whole, climate change poses a real threat to many industries' growth and profitability that should be taken into consideration when evaluating external influences. These influences can be seen through three channels:

- consumer perception for certain brands, products, and services;
- increased government regulations and protections;
- potential disruptions to supply chains and the ability to operate, such as an increase in natural disasters or resource shortages in water or energy.

Take, for example, the agriculture industry. Global greenhouse gas emissions from agriculture, forestry, and other land use account for 25% of total carbon emissions globally. This is more than two times the pollution generated from all the cars on the planet. This factor and the rise of organic, natural, cruelty-free, and vegan foods will cause the agriculture industry to face strong environmental influences throughout the 21st century. The shift toward healthier and plant-based diets has become a major trend in the last 10 years. Organic food sales in the United States rose almost 6% in 2018 to reach $47.9 billion (according to the 2019 Organic Industry Survey).

Because of consumer awareness of climate change and a growing movement to protect animal rights, companies in the livestock industry could face significant reputational risk and disruption to their brands along with potential government intervention/fines. Public awareness of the environmental impact of raising livestock has been increasing—for example, that cattle are a significant source of methane (a more potent contributor to climate change than CO_2) and that cattle require very large quantities of land, fertilizer, water, and grain. Additionally, livestock treatment has been a huge concern and subject of debate among many organizations, including animal rights groups. As the trend toward ESG (environmental, social, and governance) and sustainable investing continues to grow, there is an increasing awareness among investors of the responsibility to work with food businesses to better ensure the mitigation of potential headline risk of animal cruelty and supply chain disruption from contaminated meat. Innovation in this space has evolved to include plant-based meat substitutes (e.g. Beyond Meat and Impossible Foods), which has become a multi-billion-dollar food category over the last decade.

EXAMPLE 10

1 Company A is a mining company with operations in several countries. Company B is a technology company providing video conferencing services. Company C is a consumer company, an apparel producer, with facilities in a number of Latin American and Asian countries. Which of the companies belongs to an industry most likely to be affected by environmental factors that analysts should evaluate?

- **A** Company A
- **B** Company B
- **C** Company C

2 Which of the following statements best describes social influences that should be considered in evaluating the Consumer Goods industry?

Statement 1	Supply chain management, selling practices, and product labeling
Statement 2	Access and affordability, pricing policy
Statement 3	Impact of the industry on the environment

3 Which of the following describes one of the ways governments influence large companies that produce and offer services relating to heating, air conditioning, and lighting systems?

- **A** Purchasing goods and services
- **B** Determining availability of credit
- **C** Providing raw materials

Solution to 1:

A is correct. Environmental considerations should be incorporated in the industry analysis of Company A. It operates in a sector that has high exposure to greenhouse gas emissions, as well as water management and waste and hazardous materials management, which could incur additional costs to running businesses in this sector.

Solution to 2:

Statement 1 is correct. Statements 2 and 3 describe characteristics that, although important, are not considered to be of social nature.

Solution to 3:

A is correct. Governments are a major buyer of goods and services. B is incorrect because governments do not normally provide credit to the industry. C is incorrect because the private sector provides raw materials to customers.

EXAMPLE 11

The Airline Industry: A Case Study of Many Influences

The global airline industry exemplifies many of the concepts and influences we have discussed.

Life-Cycle Stage

The industry can be described as having some mature characteristics because average annual growth in global passenger traffic has remained relatively stable, growing at a single-digit rate for the first two decades of the 21st century, with the Asian and Middle Eastern markets growing faster than the more mature markets of North America and Europe.

Sensitivity to the Business Cycle

The airline industry is a cyclical industry; global economic activity produces swings in revenues and, especially, profitability, because of the industry's high fixed costs and operating leverage. In 2009, for example, global passenger traffic declined by approximately 3.5% and airlines lost close to US$11.0 billion, which was down from a global industry profit of US$12.9 billion in 2007. The industry tends to respond early to upward and downward moves in economic cycles; depending on the region, air travel changes at 1.5 times to 2.0 times GDP growth.

Regulation

The industry is highly regulated, with governments and airport authorities playing a large role in allocating routes and airport take-off and landing slots. Government agencies and the International Airline Transport Association set rules for aircraft and flight safety.

Product Differentiation and Customer Behavior

Airline customers tend to have low brand loyalty (except at the extremes of high and low prices and service); leisure travelers focus mainly on price, and business travelers focus mostly on schedules and service. Product and service differentiation at particular price points is low because aircraft, cabin configuration, and catering tend to be quite similar in most cases. For leisure travelers, the price competition is intense and is led by low-cost discount carriers, including Southwest Airlines in the United States, Ryanair in Europe, and Air Asia in Asia. For business travelers, the major scheduled airlines and a few service-quality specialists, such as Singapore Airlines, are the main contenders.

Technology and Costs

Fuel costs (typically more than 25% of total costs and highly volatile) and labor costs (around 10% of total costs) have been the focus of management cost-reduction efforts. The airline industry is highly unionized, and labor strikes have frequently been a source of costly disruptions to the industry. Technology and innovation have always played a major role, impacting fuel efficiency and operating costs. Technology also poses a threat to the growth of business air travel in the form of improved and widely adopted telecommunications—notably, videoconferencing and webcasting.

Social and Environmental Influences

Arguably, the airline industry has been a great force in shaping demography by permitting difficult-to-access geographical areas to be settled with large populations. At the same time, large numbers of post–World War II baby boomers have been a factor in generating growth in demand for air travel in the past half century. Environmental issues now play a role in the airline industry; carbon emissions, for example, have come under scrutiny by environmentalists and governments.

6.7 Industry Comparison

k compare characteristics of representative industries from the various economic sectors

To illustrate how these elements might be applied, Exhibit 18 reproduces the factors discussed in stages and illustrated in the context of three representative industries.

Exhibit 18 Elements of a Strategic Analysis for Three Industries

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Major Companies	Pfizer, Novartis, Merck, GSK plc	Schlumberger, Baker Hughes, Halliburton	Hershey, Mars/Wrigley, Nestle
Barriers to Entry	*Very High*: Substantial financial and intellectual capital required to compete effectively. A potential new entrant would need to create a sizable R&D operation, a global distribution network, and large-scale manufacturing capacity.	*Medium*: Technological expertise is required, but a high level of innovation allows niche companies to enter the industry and compete in specific areas.	*Very High*: Low financial or technological hurdles, but new players would lack the established brands that drive consumer purchase decisions.
Level of Concentration	*Concentrated*: A small number of companies control the bulk of the global market for branded drugs. Recent mergers have increased the level of concentration.	*Fragmented*: Although only a small number of companies provide a full range of services, many smaller players compete effectively in specific areas. Service arms of national oil companies may control significant market share in their own countries, and some product lines are concentrated in the mature US market.	*Very Concentrated*: The top four companies have a large proportion of global market share. Recent mergers have increased the level of concentration.

(continued)

Exhibit 18 (Continued)

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Impact of Industry Capacity	*Not Applicable*: Pharmaceutical pricing is primarily determined by patent protection and regulatory issues, including government approvals of drugs and of manufacturing facilities. Manufacturing capacity is of little importance.	*Medium/High*: Demand can fluctuate quickly depending on commodity prices, and industry players often find themselves with too few (or too many) employees on the payroll.	*Not Applicable*: Pricing is driven primarily by brand strength. Manufacturing capacity has little effect.
Industry Stability	*Stable*: The branded pharmaceutical market is dominated by major companies and consolidation via mega-mergers. Market shares shift quickly, however, as new drugs are approved and gain acceptance or lose patent protection.	*Unstable*: Market shares may shift frequently depending on technology offerings and demand levels.	*Very Stable*: Market shares change glacially slowly.
Life Cycle	*Mature*: Overall demand does not change greatly from year to year.	*Mature*: Demand does fluctuate with energy prices, but normalized revenue growth is only in the mid-single digits.	*Mature*: Growth is driven by population trends and pricing.
Price Competition	*Low/Medium*: In the United States, price is a minimal factor because of a consumer- and provider-driven, deregulated health care system. Price is a larger part of the decision process in single-payer systems, where efficacy hurdles are higher.	*High*: Price is a major factor in purchasers' decisions. Some companies have modest pricing power because of a wide range of services or best-in-class technology, but primary customers (major oil companies) can usually substitute with in-house services if prices are too high. Also, innovation tends to diffuse quickly throughout the industry.	*Low*: A lack of private-label competition keeps pricing stable among established players, and brand familiarity plays a much larger role in consumer purchase decisions than price.
Demographic Influences	*Positive*: Populations of developed markets are aging, which slightly increases demand.	*Not Applicable*	*Not Applicable*
Government & Regulatory Influences	*Very High*: All drugs must be approved for sale by national safety regulators. Patent regimes may differ among countries. Also, health care is heavily regulated in most countries.	*Medium*: Regulatory frameworks can affect energy demand at the margin. Also, governments play an important role in allocating exploration opportunities to E&P companies, which can indirectly affect the amount of work flowing down to service companies.	*Low*: Industry is not regulated, but childhood obesity concerns in developed markets are a low-level potential threat. Also, high-growth emerging markets may block entry of established players into their markets, possibly limiting growth.

Exhibit 18 (Continued)

	Branded Pharmaceuticals	Oil Services	Confections/Candy
Technological Influences	*Medium/High*: Biologic (large-molecule) drugs are pushing new therapeutic boundaries, and many large pharmaceutical companies have a relatively small presence in biotech.	*Medium/High*: Industry is reasonably innovative, and players must reinvest in R&D to remain competitive. Temporary competitive advantages are possible via commercialization of new processes or exploitation of accumulated expertise.	*Very Low*: Innovation does not play a major role in the industry.
Growth vs. Defensive vs. Cyclical	*Defensive*: Demand for most health care services does not fluctuate with the economic cycle, but demand is not strong enough to be considered "growth."	*Cyclical*: Demand is highly variable and depends on oil prices, exploration budgets, and the economic cycle.	*Defensive*: Demand for candy and gum is extremely stable.

Example 12 reviews some of the information presented in Exhibit 18.

EXAMPLE 12

External Influences

1 Which of the following industries is *most* affected by government regulation?
   - A Oil services
   - B Pharmaceuticals
   - C Confections and candy

2 Which of the following industries is *least* affected by technological innovation?
   - A Oil services
   - B Pharmaceuticals
   - C Confections and candy

3 Which of the following statements about industry characteristics is *least* accurate?
   - A Manufacturing capacity has little effect on pricing in the confections/candy industry.
   - B The branded pharmaceutical industry is considered to be defensive rather than a growth industry.
   - C With respect to the worldwide market, the oil services industry has a high level of concentration with a limited number of service providers.

Solution to 1:

B is correct. Exhibit 18 states that the pharmaceutical industry has a high amount of government and regulatory influences.

Solution to 2:

C is correct. Exhibit 18 states that innovation does not play a large role in the candy industry.

Solution to 3:

C is correct; it is a false statement. From a worldwide perspective, the industry is considered fragmented. Although a small number of companies provide the full range of services, competition by many smaller players occurs in niche areas. In addition, national oil service companies control significant market share in their home countries.

7 COMPANY ANALYSIS

l. describe the elements that should be covered in a thorough company analysis

Company analysis includes an analysis of the company's financial position, products and/or services, and **competitive strategy** (its plans for responding to the threats and opportunities presented by the external environment). Company analysis takes place after the analyst has gained an understanding of a company's external environment—the macroeconomic, demographic, governmental, technological, environmental, and social forces influencing the industry's competitive structure. The analyst should seek to determine whether the strategy is primarily defensive or offensive in its nature and how the company intends to implement the strategy.

Porter (2008) identified two chief competitive strategies: a low-cost strategy (cost leadership) and a product/service differentiation strategy:

- *Low-cost strategy*: Companies strive to become the low-cost producers and to gain market share by offering their products and services at lower prices than their competition while still making a profit margin sufficient to generate a superior rate of return based on the higher revenues achieved. Low-cost strategies may be pursued defensively to protect market positions and returns or offensively to gain market share and increase returns. Pricing also can be defensive (when the competitive environment is one of low rivalry) or aggressive (when rivalry is intense). In the case of intense rivalry, pricing may even become predatory—that is, aimed at rapidly driving competitors out of business at the expense of near-term profitability. The hope in such a strategy is that having achieved a larger market share, the company can later increase prices to generate higher returns than before. For example, the ride-sharing industry has produced fierce competition among companies that seek to capture market share by offering incentives and discount pricing for rides. Companies seeking to follow low-cost strategies must have tight cost controls, efficient operating and reporting systems, and appropriate managerial incentives. In addition, they must commit themselves to painstaking scrutiny of production systems and their labor forces and to low-cost designs and product distribution. In some cases, they must be able to invest in productivity-improving capital equipment and to finance that investment at a low cost of capital.
- *Differentiation strategies*: Companies attempt to establish themselves as the suppliers or producers of products and services that are unique either in quality, type, or means of distribution. To be successful, their price premiums must be above their costs of differentiation and the differentiation must be appealing to customers and sustainable over time. Corporate managers who successfully

pursue differentiation strategies tend to have strong market research teams to identify and match customer needs with product development and marketing. Such a strategy puts a premium on employing creative and inventive people.

7.1 Elements That Should Be Covered in a Company Analysis

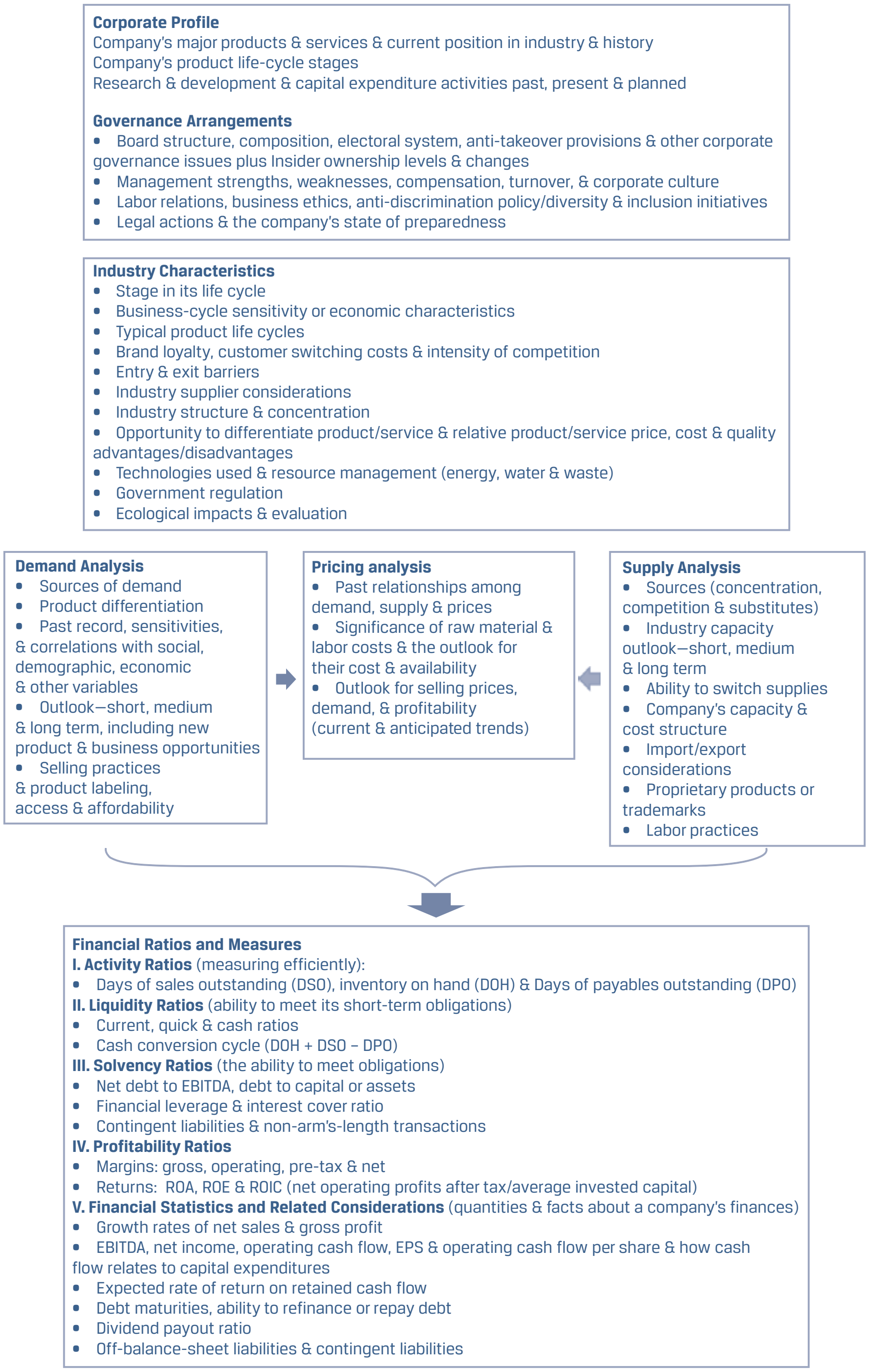
A thorough company analysis, particularly as presented in a research report, should

- provide an overview of the company (corporate profile), including a basic understanding of its businesses, investment activities, corporate governance, and perceived strengths and weaknesses;
- explain relevant industry characteristics;
- analyze the demand for the company's products and services;
- analyze the supply of products and services, which includes an analysis of costs;
- explain the company's pricing environment; and
- present and interpret relevant financial ratios, including comparisons over time and comparisons with competitors, using multi-year spreadsheets with historical and forecast data

Company analysis often includes forecasting the company's financial statements, particularly when the purpose of the analysis is to use a discounted cash flow method to value the company's common equity.

Exhibit 19 provides a checklist of points to cover in a company analysis. The list is not exhaustive and may need to be adapted to serve the needs of a particular company analysis.

Exhibit 19 A Checklist for Company Analysis



Corporate Profile
Company's major products & services & current position in industry & history
Company's product life-cycle stages
Research & development & capital expenditure activities past, present & planned

Governance Arrangements
- Board structure, composition, electoral system, anti-takeover provisions & other corporate governance issues plus Insider ownership levels & changes
- Management strengths, weaknesses, compensation, turnover, & corporate culture
- Labor relations, business ethics, anti-discrimination policy/diversity & inclusion initiatives
- Legal actions & the company's state of preparedness

Industry Characteristics
- Stage in its life cycle
- Business-cycle sensitivity or economic characteristics
- Typical product life cycles
- Brand loyalty, customer switching costs & intensity of competition
- Entry & exit barriers
- Industry supplier considerations
- Industry structure & concentration
- Opportunity to differentiate product/service & relative product/service price, cost & quality advantages/disadvantages
- Technologies used & resource management (energy, water & waste)
- Government regulation
- Ecological impacts & evaluation

Demand Analysis
- Sources of demand
- Product differentiation
- Past record, sensitivities, & correlations with social, demographic, economic & other variables
- Outlook—short, medium & long term, including new product & business opportunities
- Selling practices & product labeling, access & affordability

Pricing analysis
- Past relationships among demand, supply & prices
- Significance of raw material & labor costs & the outlook for their cost & availability
- Outlook for selling prices, demand, & profitability (current & anticipated trends)

Supply Analysis
- Sources (concentration, competition & substitutes)
- Industry capacity outlook—short, medium & long term
- Ability to switch supplies
- Company's capacity & cost structure
- Import/export considerations
- Proprietary products or trademarks
- Labor practices

Financial Ratios and Measures
I. Activity Ratios (measuring efficiently):
- Days of sales outstanding (DSO), inventory on hand (DOH) & Days of payables outstanding (DPO)

II. Liquidity Ratios (ability to meet its short-term obligations)
- Current, quick & cash ratios
- Cash conversion cycle (DOH + DSO – DPO)

III. Solvency Ratios (the ability to meet obligations)
- Net debt to EBITDA, debt to capital or assets
- Financial leverage & interest cover ratio
- Contingent liabilities & non-arm's-length transactions

IV. Profitability Ratios
- Margins: gross, operating, pre-tax & net
- Returns: ROA, ROE & ROIC (net operating profits after tax/average invested capital)

V. Financial Statistics and Related Considerations (quantities & facts about a company's finances)
- Growth rates of net sales & gross profit
- EBITDA, net income, operating cash flow, EPS & operating cash flow per share & how cash flow relates to capital expenditures
- Expected rate of return on retained cash flow
- Debt maturities, ability to refinance or repay debt
- Dividend payout ratio
- Off-balance-sheet liabilities & contingent liabilities

Evaluation of past and current company performance and forecasting future cash flows and a company's prospects is a major topic that is the subject of extensive coverage elsewhere in the CFA Program curriculum. Here, we provide only a brief introduction.

To evaluate a company's performance, the key measures presented in the "checklist" in Exhibit 19 should be compared over time and between companies (particularly peer companies). The following formula can be used to analyze how and why a company's return on equity (ROE) differs from that of other companies or its own ROE in other periods by tracing the differences to changes in its profit margin, the productivity of its assets, or its financial leverage:

ROE = (Net profit margin: Net earnings/Net sales) × (Asset turnover: Net sales/Average total assets) × (Financial leverage: Average total assets/Average common equity).

The financial statements of a company over time provide numerous insights into the effects of industry conditions on its performance and the success or failure of its strategies. They also provide a framework for forecasting the company's operating performance when given the analyst's assumptions for numerous variables in the future. The financial ratios listed in Exhibit 19 are applicable to a wide range of companies and industries, but other statistics and ratios are often also used.

7.2 Spreadsheet Modeling

Spreadsheet modeling of financial statements to analyze and forecast revenues, operating and net income, and cash flows has become one of the most widely used tools in company analysis. Although spreadsheet models are a valuable tool for understanding past financial performance and forecasting future performance, the complexity of such models can at times be a problem. Because modeling requires the analyst to predict and input numerous items in financial statements, there is a risk of errors—either in assumptions made or in formulas in the model—which can compound, leading to erroneous forecasts. Yet, those forecasts may seem precise because of the sheer complexity of the model. The result is often a false sense of understanding and security on the part of those who rely on the models. To guard against this, before or after a model is completed, a "reality check" of the model is useful.

Such testing for reasonableness can be done by, first, asking what the few most important changes in income statement items are likely to be from last year to this year and the next year and, second, attempting to quantify the effects of these significant changes or "swing factors" on the bottom line. If an analyst cannot summarize in a few points what factors are realistically expected to change income from year to year and is not convinced that these assumptions are correct, then he or she does not really understand the output of the modeling efforts. In general, financial models should be in a format that matches the company's reporting of its financial results or supplementary disclosures or that can be accurately derived from these reports. Otherwise, there will be no natural reality check when the company issues its financial results and the analyst will not be able to compare his or her estimates with actual reported results.

EXAMPLE 13

1 An analyst makes the following statement:

> "Analysis of a company's supply includes examination of labor relations; sources of and access to raw materials, including concentration of suppliers; and analysis of the company's production capacity."

The statement is:

A correct.

B incorrect because the analysis of supply should also include analysis of the stage in the product life cycle.

   C incorrect because the analysis of supply should consider the product's differentiating characteristics.

2 The challenge with using spreadsheet forecasting models stems from the fact that

   A the analyst may have a false sense of understanding the business.

   B spreadsheet models cannot be used to compare companies with their peer groups.

   C spreadsheet models require precise inputs.

Solution to 1:

A is correct. The statement accurately describes the analysis of a company's supply.

Solution to 2:

A is correct. B is incorrect because spreadsheet models are frequently used to compare companies. C is incorrect because although spreadsheet models require precise inputs, this requirement does not mean that they are not helpful.

SUMMARY

In this reading, we have provided an overview of industry analysis and illustrated approaches that are widely used by analysts to examine an industry.

- Company analysis and industry analysis are closely interrelated. Company and industry analysis together can provide insight into sources of industry revenue growth and competitors' market shares and thus the future of an individual company's top-line growth and bottom-line profitability.
- Industry analysis is useful for
  - understanding a company's business and business environment,
  - identifying active equity investment opportunities,
  - formulating an industry or sector rotation strategy, and
  - portfolio performance attribution.
- The three main approaches to classifying companies are
  - products and/or services supplied,
  - business-cycle sensitivities, and
  - statistical similarities.
- A cyclical company is one whose profits are strongly correlated with the strength of the overall economy.
- A non-cyclical company is one whose performance is largely independent of the business cycle.
- Commercial industry classification systems include
  - The Global Industry Classification Standard (GICS)
  - The Industry Classification Benchmark (ICB)

- A limitation of current classification systems is that the narrowest classification unit assigned to a company generally cannot be assumed to constitute its peer group for the purposes of detailed fundamental comparisons or valuation.
- A peer group is a group of companies engaged in similar business activities whose economics and valuation are influenced by closely related factors.
- The steps in constructing a preliminary list of peer companies are as follows:
  - Examine commercial classification systems if available. These systems often provide a useful starting point for identifying companies operating in the same industry.
  - Review the subject company's annual report for a discussion of the competitive environment. Companies frequently cite specific competitors.
  - Review competitors' annual reports to identify other potential comparables.
  - Review industry trade publications to identify additional peer companies.
  - Confirm that each comparable or peer company derives a significant portion of its revenue and operating profit from a business activity similar to that of the subject company.
- Not all industries are created equal. Some are highly competitive, with many companies struggling to earn returns in excess of their cost of capital, and other industries have attractive characteristics that enable a majority of industry participants to generate healthy profits.
- Differing competitive environments are determined by the structural attributes of the industry. For this important reason, industry analysis is a vital complement to company analysis. The analyst needs to understand the context in which a company operates to fully understand the opportunities and threats that a company faces.
- The framework for strategic analysis known as "Porter's five forces" can provide a useful starting point. Porter maintained that the profitability of companies in an industry is determined by five forces: (1) the threat of new entrants, which, in turn, is determined by economies of scale, brand loyalty, absolute cost advantages, customer switching costs, and government regulation; (2) the bargaining power of suppliers, which is a function of the feasibility of product substitution, the concentration of the buyer and supplier groups, and switching costs and entry costs in each case; (3) the bargaining power of buyers, which is a function of switching costs among customers and the ability of customers to produce their own product; (4) the threat of substitutes; and (5) the intensity of rivalry among existing competitors, which, in turn, is a function of industry competitive structure, demand conditions, cost conditions, and the height of exit barriers.
- The concept of barriers to entry refers to the ease with which new competitors can challenge incumbents and can be an important factor in determining the competitive environment of an industry. If new competitors can easily enter the industry, the industry is likely to be highly competitive because incumbents that attempt to raise prices will be undercut by newcomers. As a result, industries with low barriers to entry tend to have low pricing power. Conversely, if incumbents are protected by barriers to entry, they may enjoy a more benign competitive environment that gives them greater pricing power over their customers because they do not have to worry about being undercut by startups.
- Industry concentration is often, although not always, a sign that an industry may have pricing power and rational competition. Industry fragmentation is a much stronger signal, however, that the industry is competitive and pricing power is limited.

- The effect of industry capacity on pricing is clear: Tight capacity gives participants more pricing power because demand for products or services exceeds supply; overcapacity leads to price cutting and a highly competitive environment as excess supply chases demand. The analyst should think about not only current capacity conditions but also future changes in capacity levels—how long it takes for supply and demand to come into balance and what effect that process has on industry pricing power and returns.
- Examining the market share stability of an industry over time is similar to thinking about barriers to entry and the frequency with which new players enter an industry. Stable market shares typically indicate less competitive industries, whereas unstable market shares often indicate highly competitive industries with limited pricing power.
- An industry's position in its life cycle often has a large impact on its competitive dynamics, so it is important to keep this positioning in mind when performing strategic analysis of an industry. Industries, like individual companies, tend to evolve over time and usually experience significant changes in the rate of growth and levels of profitability along the way. Just as an investment in an individual company requires careful monitoring, industry analysis is a continuous process that must be repeated over time to identify changes that may be occurring.
- A useful framework for analyzing the evolution of an industry is an industry life-cycle model, which identifies the sequential stages that an industry typically goes through. The five stages of an industry life cycle according to the Hill and Jones model are
  - embryonic,
  - growth,
  - shakeout,
  - mature, and
  - decline.
- Price competition and thinking like a customer are important factors that are often overlooked when analyzing an industry. Whatever factors most influence customer purchasing decisions are also likely to be the focus of competitive rivalry in the industry. Broadly, industries for which price is a large factor in customer purchase decisions tend to be more competitive than industries in which customers value other attributes more highly.
- External influences on industry growth, profitability, and risk include
  - technology,
  - demographics,
  - government,
  - social factors, and
  - environmental factors.
- Company analysis takes place after the analyst has gained an understanding of the company's external environment and includes answering questions about how the company will respond to the threats and opportunities presented by the external environment. This intended response is the individual company's competitive strategy. The analyst should seek to determine whether the strategy is primarily defensive or offensive in its nature and how the company intends to implement it.
- Porter identified two chief competitive strategies:

  - A low-cost strategy (cost leadership) is one in which companies strive to become the low-cost producers and to gain market share by offering their products and services at lower prices than their competition while still making a profit margin sufficient to generate a superior rate of return based on the higher revenues achieved.
  - A product/service differentiation strategy is one in which companies attempt to establish themselves as the suppliers or producers of products and services that are unique either in quality, type, or means of distribution. To be successful, the companies' price premiums must be above their costs of differentiation and the differentiation must be appealing to customers and sustainable over time.
- A checklist for company analysis includes a thorough investigation of
  - the corporate profile,
  - industry characteristics,
  - demand for products/services,
  - supply of products/services,
  - pricing,
  - financial ratios, and
  - sustainability metrics.
- Spreadsheet modeling of financial statements to analyze and forecast revenues, operating and net income, and cash flows is one of the most widely used tools in company analysis. Spreadsheet modeling can be used to quantify the effects of the changes in certain swing factors on the various financial statements. The analyst should be aware that the output of the model will depend significantly on the assumptions that are made.

REFERENCES

Hill, Charles, and Gareth Jones. 2008. "External Analysis: The Identification of Opportunities and Threats." In Strategic Management: An Integrated Approach. Boston: Houghton Mifflin Co.

Porter, Michael E. 2008. "The Five Competitive Forces That Shape Strategy." Harvard Business Review 86 (1): 78–93.

PRACTICE PROBLEMS

1 Which of the following is *least likely* to involve industry analysis?

- **A** Sector rotation strategy
- **B** Top-down fundamental investing
- **C** Tactical asset allocation strategy

2 A sector rotation strategy involves investing in a sector by:

- **A** making regular investments in it.
- **B** investing in a pre-selected group of sectors on a rotating basis.
- **C** timing investment to take advantage of business-cycle conditions.

3 Which of the following information about a company would *most likely* depend on an industry analysis? The company's:

- **A** treatment of long-lived assets on its financial statements.
- **B** competitive environment.
- **C** trends in corporate expenses.

4 In which sector would a manufacturer of personal care products be classified?

- **A** Health care
- **B** Consumer staples
- **C** Consumer discretionary

5 An automobile manufacturer is *most likely* classified in which of the following industry sectors?

- **A** Consumer staples
- **B** Industrial durables
- **C** Consumer discretionary

6 Which of the following statements about commercial and government industry classification systems is *most* accurate?

- **A** Many commercial classification systems include private for-profit companies.
- **B** Both commercial and government classification systems exclude not-for-profit companies.
- **C** Commercial classification systems are generally updated more frequently than government classification systems.

7 Which of the following is *not* a limitation of the cyclical/non-cyclical descriptive approach to classifying companies?

- **A** A cyclical company may have a growth component in it.
- **B** Business-cycle sensitivity is a discrete phenomenon rather than a continuous spectrum.
- **C** A global company can experience economic expansion in one part of the world while experiencing recession in another part.

8 A cyclical company is *most likely* to:

- **A** have low operating leverage.
- **B** sell relatively inexpensive products.
- **C** experience wider-than-average fluctuations in demand.

© 2017 CFA Institute. All rights reserved.

9 A company that is sensitive to the business cycle would *most likely*:

A not have growth opportunities.

B experience below-average fluctuation in demand.

C sell products that customers can purchase at a later date if necessary.

10 Which of the following factors would *most likely* be a limitation of applying business-cycle analysis to global industry analysis?

A Some industries are relatively insensitive to the business cycle.

B Correlations of security returns between different world markets are relatively low.

C One region or country of the world may experience recession while another region experiences expansion.

11 Which of the following statements about peer groups is *most* accurate?

A Constructing a peer group for a company follows a standardized process.

B Commercial industry classification systems often provide a starting point for constructing a peer group.

C A peer group is generally composed of all the companies in the most narrowly defined category used by the commercial industry classification system.

12 With regard to forming a company's peer group, which of the following statements is *not* correct?

A Comments from the management of the company about competitors are generally not used when selecting the peer group.

B The higher the proportion of revenue and operating profit of the peer company derived from business activities similar to those of the subject company, the more meaningful the comparison.

C Comparing the company's performance measures with those for a potential peer-group company is of limited value when the companies are exposed to different stages of the business cycle.

13 When selecting companies for inclusion in a peer group, a company operating in three different business segments would:

A be in only one peer group.

B possibly be in more than one peer group.

C not be included in any peer group.

14 An industry that *most likely* has both high barriers to entry and high barriers to exit is the:

A restaurant industry.

B advertising industry.

C automobile industry.

15 Which factor is *most likely* associated with stable market share?

A Low switching costs

B Low barriers to entry

C Slow pace of product innovation

16 Which of the following companies *most likely* has the greatest ability to quickly increase its capacity to offer goods or services?

A A restaurant

B A steel producer

C An insurance company

17 A population that is rapidly aging would *most likely* cause the growth rate of the industry producing eye glasses and contact lenses to:

A decrease.

B increase.

C not change.

18 If over a long period of time a country's average level of educational accomplishment increases, this development would *most likely* lead to the country's amount of income spent on consumer discretionary goods to:

A decrease.

B increase.

C not change.

19 If the technology for an industry involves high fixed capital investment, then one way to seek higher profit growth is by pursuing:

A economies of scale.

B diseconomies of scale.

C removal of features that differentiate the product or service provided.

20 Which of the following life-cycle phases is typically characterized by high prices?

A Mature

B Growth

C Embryonic

21 In which of the following life-cycle phases are price wars *most likely* to be absent?

A Mature

B Decline

C Growth

22 When graphically depicting the life-cycle model for an industry as a curve, the variables on the axes are:

A price and time.

B demand and time.

C demand and stage of the life cycle.

23 Industry consolidation and high barriers to entry *most likely* characterize which life-cycle stage?

A Mature

B Growth

C Embryonic

24 Which of the following is *most likely* a characteristic of a concentrated industry?

A Infrequent, tacit coordination

B Difficulty in monitoring other industry members

C Industry members attempting to avoid competition on price

25 Which of the following industry characteristics is generally *least likely* to produce high returns on capital?

A High barriers to entry

B High degree of concentration

- **C** Short lead time to build new plants

26 An industry with high barriers to entry and weak pricing power *most likely* has:

- **A** high barriers to exit.
- **B** stable market shares.
- **C** significant numbers of issued patents.

27 Economic value is created for an industry's shareholders when the industry earns a return:

- **A** below the cost of capital.
- **B** equal to the cost of capital.
- **C** above the cost of capital.

28 Which of the following industries is *most likely* to be characterized as concentrated with strong pricing power?

- **A** Asset management
- **B** Alcoholic beverages
- **C** Household and personal products

29 With respect to competitive strategy, a company with a successful cost leadership strategy is *most likely* characterized by:

- **A** a low cost of capital.
- **B** reduced market share.
- **C** the ability to offer products at higher prices than those of its competitors.

30 When conducting a company analysis, the analysis of demand for a company's product is *least likely* to consider the:

- **A** company's cost structure.
- **B** motivations of the customer base.
- **C** product's differentiating characteristics.

31 Which of the following statements about company analysis is *most* accurate?

- **A** The complexity of spreadsheet modeling ensures precise forecasts of financial statements.
- **B** The interpretation of financial ratios should focus on comparing the company's results over time but not with competitors.
- **C** The corporate profile would include a description of the company's business, investment activities, governance, and strengths and weaknesses.

SOLUTIONS

1 C is correct. Tactical asset allocation involves timing investments in asset classes and does not make use of industry analysis.

2 C is correct. A sector rotation strategy is conducted by investors wishing to time investment in industries through an analysis of fundamentals and/or business-cycle conditions.

3 B is correct. Determination of a company's competitive environment depends on understanding its industry.

4 B is correct. Personal care products are classified as consumer staples in the "Description of Representative Sectors."

5 C is correct. Automobile manufacturers are classified as consumer discretionary. Consumer discretionary companies derive a majority of revenue from the sale of consumer-related products for which demand tends to exhibit a high degree of economic sensitivity—that is, high demand during periods of economic expansion and low demand during periods of contraction.

6 C is correct. Commercial systems are generally updated more frequently than government systems and include only publicly traded for-profit companies.

7 B is correct. Business-cycle sensitivity falls on a continuum and is not a discrete "either/or" phenomenon.

8 C is correct. Cyclical companies are sensitive to the business cycle, with low product demand during periods of economic contraction and high product demand during periods of economic expansion. They, therefore, experience wider-than-average fluctuations in product demand.

9 C is correct. Customers' flexibility as to when they purchase a product makes the product more sensitive to the business cycle.

10 C is correct. Varying conditions of recession or expansion around the world would affect the comparisons of companies with sales in different regions of the world.

11 B is correct. Constructing a peer group is a subjective process, and a logical starting point is to begin with a commercially available classification system. This system will identify a group of companies that may have properties comparable to the business activity of interest.

12 A is correct because it is a false statement. Reviewing the annual report to find management's discussion about the competitive environment and specific competitors is a suggested step in the process of constructing a peer group.

13 B is correct. The company could be in more than one peer group depending on the demand drivers for the business segments, although the multiple business segments may make it difficult to classify the company.

14 C is correct. For the automobile industry, the high capital requirements and other elements mentioned in the reading provide high barriers to entry, and recognition that auto factories are generally only of use for manufacturing cars implies a high barrier to exit.

15 C is correct. A slow pace of product innovation often means that customers prefer to stay with suppliers they know, implying stable market shares.

16 C is correct. Capacity increases in providing insurance services would not involve several factors that would be important to the other two industries, including the need for substantial fixed capital investments or, in the case of a restaurant, outfitting rental or purchased space. These requirements would tend to slow down, respectively, steel production and restaurant expansion.

17 B is correct. Vision typically deteriorates at advanced ages. An increased number of older adults implies more eyewear products will be purchased.

18 B is correct. As their educational level increases, workers are able to perform more skilled tasks, earn higher wages, and as a result, have more income left for discretionary expenditures.

19 A is correct. Seeking economies of scale would tend to reduce per-unit costs and increase profit.

20 C is correct. The embryonic stage is characterized by slow growth and high prices.

21 C is correct. The growth phase is not likely to experience price wars because expanding industry demand provides companies the opportunity to grow even without increasing market share. When industry growth is stagnant, companies may only be able to grow by increasing market share—for example, by engaging in price competition.

22 B is correct. The industry life-cycle model shows how demand evolves over time as an industry passes from the embryonic stage through the stage of decline.

23 A is correct. Industry consolidation and relatively high barriers to entry are two characteristics of a mature-stage industry.

24 C is correct. The relatively few members of the industry generally try to avoid price competition.

25 C is correct. With short lead times, industry capacity can be rapidly increased to satisfy demand, but it may also lead to overcapacity and lower profits.

26 A is correct. An industry that has high barriers to entry generally requires substantial physical capital and/or financial investment. With weak pricing power in the industry, finding a buyer for excess capacity (i.e., to exit the industry) may be difficult.

27 C is correct. Economic profit is earned and value is created for shareholders when the industry earns returns above the company's cost of capital.

28 B is correct. The alcoholic beverage industry is concentrated and possesses strong pricing power.

29 A is correct. Companies with low-cost strategies must be able to invest in productivity-improving equipment and finance that investment at a low cost of capital. Market share and pricing depend on whether the strategy is pursued defensively or offensively.

30 A is correct. The cost structure is an appropriate element when analyzing the supply of the product, but analysis of demand relies on the product's differentiating characteristics and the customers' needs and wants.

31 C is correct. The corporate profile would provide an understanding of these elements.

READING

38

Equity Valuation: Concepts and Basic Tools

by John J. Nagorniak, CFA, and Stephen E. Wilcox, PhD, CFA

John J. Nagorniak, CFA (USA). Stephen E. Wilcox, PhD, CFA, is at Minnesota State University, Mankato (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** evaluate whether a security, given its current market price and a value estimate, is overvalued, fairly valued, or undervalued by the market;
☐	**b.** describe major categories of equity valuation models;
☐	**c.** describe regular cash dividends, extra dividends, stock dividends, stock splits, reverse stock splits, and share repurchases;
☐	**d.** describe dividend payment chronology;
☐	**e.** explain the rationale for using present value models to value equity and describe the dividend discount and free-cash-flow-to-equity models;
☐	**f.** calculate the intrinsic value of a non-callable, non-convertible preferred stock;
☐	**g.** calculate and interpret the intrinsic value of an equity security based on the Gordon (constant) growth dividend discount model or a two-stage dividend discount model, as appropriate;
☐	**h.** identify characteristics of companies for which the constant growth or a multistage dividend discount model is appropriate;
☐	**i.** explain the rationale for using price multiples to value equity, how the price to earnings multiple relates to fundamentals, and the use of multiples based on comparables;
☐	**j.** calculate and interpret the following multiples: price to earnings, price to an estimate of operating cash flow, price to sales, and price to book value;

(continued)

© 2019 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**k.** describe enterprise value multiples and their use in estimating equity value;
☐	**l.** describe asset-based valuation models and their use in estimating equity value;
☐	**m.** explain advantages and disadvantages of each category of valuation model.

1 INTRODUCTION

Analysts gather and process information to make investment decisions, including buy and sell recommendations. What information is gathered and how it is processed depend on the analyst and the purpose of the analysis. Technical analysis uses such information as stock price and trading volume as the basis for investment decisions. Fundamental analysis uses information about the economy, industry, and company as the basis for investment decisions. Examples of fundamentals are unemployment rates, gross domestic product (GDP) growth, industry growth, and quality of and growth in company earnings. Whereas technical analysts use information to predict price movements and base investment decisions on the direction of predicted change in prices, fundamental analysts use information to estimate the value of a security and to compare the estimated value to the market price and then base investment decisions on that comparison.

This reading introduces equity valuation models used to estimate the **intrinsic value** (synonym: **fundamental value**) of a security; intrinsic value is based on an analysis of investment fundamentals and characteristics. The fundamentals to be considered depend on the analyst's approach to valuation. In a top-down approach, an analyst examines the economic environment, identifies sectors that are expected to prosper in that environment, and analyzes securities of companies from previously identified attractive sectors. In a bottom-up approach, an analyst typically follows an industry or industries and forecasts fundamentals for the companies in those industries in order to determine valuation. Whatever the approach, an analyst who estimates the intrinsic value of an equity security is implicitly questioning the accuracy of the market price as an estimate of value. Valuation is particularly important in active equity portfolio management, which aims to improve on the return–risk trade-off of a portfolio's benchmark by identifying mispriced securities.

This reading is organized as follows. Section 2 discusses the implications of differences between estimated value and market price. Section 3 introduces three major categories of valuation model. Sections 4–8 present an overview of present value models with a focus on the dividend discount model. Sections 9–11 describe and examine the use of multiples in valuation. Section 12 explains asset-based valuation and demonstrates how these models can be used to estimate value. Section 13 states conclusions and summarizes the reading.

ESTIMATED VALUE AND MARKET PRICE

2

a evaluate whether a security, given its current market price and a value estimate, is overvalued, fairly valued, or undervalued by the market;

By comparing estimates of value and market price, an analyst can arrive at one of three conclusions: The security is *undervalued*, *overvalued*, or *fairly valued* in the marketplace. For example, if the market price of an asset is $10 and the analyst estimates intrinsic value at $10, a logical conclusion is that the security is fairly valued. If the security is selling for $20, the security would be considered overvalued. If the security is selling for $5, the security would be considered undervalued. Basically, by estimating value, the analyst is assuming that the market price is not necessarily the best estimate of intrinsic value. If the estimated value exceeds the market price, the analyst infers the security is *undervalued*. If the estimated value equals the market price, the analyst infers the security is *fairly valued*. If the estimated value is less than the market price, the analyst infers the security is *overvalued*.

In practice, the conclusion is not so straightforward. Analysts must cope with uncertainties related to model appropriateness and the correct value of inputs. An analyst's final conclusion depends not only on the comparison of the estimated value and the market price but also on the analyst's confidence in the estimated value (i.e., in the model selected and the inputs used in it). One can envision a spectrum running from relatively high confidence in the valuation model *and* the inputs to relatively low confidence in the valuation model *and/or* the inputs. When confidence is relatively low, the analyst might demand a substantial divergence between his or her own value estimate and the market price before acting on an apparent mispricing. For instance, if the estimate of intrinsic value is $10 and the market price is $10.05, the analyst might reasonably conclude that the security is fairly valued and that the 1/2 of 1 percent market price difference from the estimated value is within the analyst's confidence interval.

Confidence in the convergence of the market price to the intrinsic value over the investment time horizon relevant to the objectives of the portfolio must also be taken into account before an analyst acts on an apparent mispricing or makes a buy, sell, or hold recommendation: The ability to benefit from identifying a mispriced security depends on the market price converging to the estimated intrinsic value.

In seeking to identify mispricing and attractive investments, analysts are treating market prices with skepticism, but they are also treating market prices with respect. For example, an analyst who finds that many securities examined appear to be overvalued will typically recheck models and inputs before acting on a conclusion of overvaluation. Analysts also often recognize and factor into recommendations that different market segments—such as securities closely followed by analysts versus securities relatively neglected by analysts—may differ in how common or persistent mispricing is. Mispricing may be more likely in securities neglected by analysts.

EXAMPLE 1

Valuation and Analyst Response

1 An analyst finds that all the securities analyzed have estimated values higher than their market prices. The securities all appear to be:

A overvalued.

B undervalued.

C fairly valued.

2 An analyst finds that nearly all companies in a market segment have common shares which are trading at market prices above the analyst's estimate of the shares' values. This market segment is widely followed by analysts. Which of the following statements describes the analyst's *most appropriate* first action?

- **A** Issue a sell recommendation for each share issue.
- **B** Issue a buy recommendation for each share issue.
- **C** Reexamine the models and inputs used for the valuations.

3 An analyst, using a number of models and a range of inputs, estimates a security's value to be between ¥250 and ¥270. The security is trading at ¥265. The security appears to be:

- **A** overvalued.
- **B** undervalued.
- **C** fairly valued.

Solution to 1:

B is correct. The estimated intrinsic value for each security is greater than the market price. The securities all appear to be undervalued in the market. Note, however, that the analyst may wish to reexamine the model and inputs to check that the conclusion is valid.

Solution to 2:

C is correct. It seems improbable that all the share issues analyzed are overvalued, as indicated by market prices in excess of estimated value—particularly because the market segment is widely followed by analysts. Thus, the analyst will not issue a sell recommendation for each issue. The analyst will *most appropriately* reexamine the models and inputs prior to issuing any recommendations. A buy recommendation is not an appropriate response to an overvalued security.

Solution to 3:

C is correct. The security's market price of ¥265 is within the range estimated by the analyst. The security appears to be fairly valued.

Analysts often use a variety of models and inputs to achieve greater confidence in their estimates of intrinsic value. The use of more than one model and a range of inputs also helps the analyst understand the sensitivity of value estimates to different models and inputs.

3

CATEGORIES OF EQUITY VALUATION MODELS

b describe major categories of equity valuation models;

Three major categories of equity valuation models are as follows:

- **Present value models** (synonym: **discounted cash flow models**). These models estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security. In present value models, benefits are often defined in terms of cash expected to be distributed to shareholders (**dividend discount models**) or in terms of cash flows available to be distributed to shareholders after meeting capital expenditure and working capital needs (**free-cash-flow-to-equity models**). Many models fall within this category, ranging from the relatively simple to the very complex. In Sections 4–8, we discuss in detail two of the simpler models, the Gordon (constant) growth model and the two-stage dividend discount models.

- **Multiplier models** (synonym: **market multiple models**). These models are based chiefly on share price multiples or enterprise value multiples. The former model estimates intrinsic value of a common share from a price multiple for some fundamental variable, such as revenues, earnings, cash flows, or book value. Examples of the multiples include price to earnings (P/E, share price divided by earnings per share) and price to sales (P/S, share price divided by sales per share). The fundamental variable may be stated on a forward basis (e.g., forecasted EPS for the next year) or a trailing basis (e.g., EPS for the past year), as long as the usage is consistent across companies being examined. Price multiples are also used to compare relative values. The use of the ratio of share price to EPS—that is, the P/E multiple—to judge relative value is an example of this approach to equity valuation.

  Enterprise value (EV) multiples have the form (Enterprise value)/(Value of a fundamental variable). Two possible choices for the denominator are earnings before interest, taxes, depreciation, and amortization (EBITDA) and total revenue. Enterprise value, the numerator, is a measure of a company's total market value from which cash and short-term investments have been subtracted (because an acquirer could use those assets to pay for acquiring the company). An estimate of common share value can be calculated indirectly from the EV multiple; the value of liabilities and preferred shares can be subtracted from the EV to arrive at the value of common equity.

- **Asset-based valuation models**. These models estimate intrinsic value of a common share from the estimated value of the assets of a corporation minus the estimated value of its liabilities and preferred shares. The estimated market value of the assets is often determined by making adjustments to the **book value** (synonym: **carrying value**) of assets and liabilities. The theory underlying the asset-based approach is that the value of a business is equal to the sum of the value of the business's assets.

As already mentioned, many analysts use more than one type of model to estimate value. Analysts recognize that each model is a simplification of the real world and that there are uncertainties related to model appropriateness and the inputs to the models. The choice of model(s) will depend on the availability of information to input into the model(s) and the analyst's confidence in the information and in the appropriateness of the model(s).

EXAMPLE 2

Categories of Equity Valuation Models

1 An analyst is estimating the intrinsic value of a new company. The analyst has one year of financial statements for the company and has calculated the average values of a variety of price multiples for the industry in which the company operates. The analyst plans to use at least one model from each of the three categories of valuation models. The analyst is *least likely* to rely on the estimate(s) from the:

  A multiplier model(s).

  B present value model(s).

  - **C** asset-based valuation model(s).

2 Based on a company's EPS of €1.35, an analyst estimates the intrinsic value of a security to be €16.60. Which type of model is the analyst *most likely* to be using to estimate intrinsic value?

  - **A** Multiplier model.
  - **B** Present value model.
  - **C** Asset-based valuation model.

Solution to 1:

B is correct. Because the company has only one year of data available, the analyst is *least likely* to be confident in the inputs for a present value model. The values on the balance sheet, even before adjustment, are likely to be close to market values because the assets are all relatively new. The multiplier models are based on average multiples from the industry.

Solution to 2:

A is correct. The analyst is using a multiplier model based on the P/E multiple. The P/E multiple used was 16.60/1.35 = 12.3.

As you begin the study of specific equity valuation models in the next section, you must bear in mind that any model of value is, by necessity, a simplification of the real world. Never forget this simple fact! You may encounter models much more complicated than the ones discussed here, but even those models will be simplifications of reality.

4

BACKGROUND FOR THE DIVIDEND DISCOUNT MODEL

c describe regular cash dividends, extra dividends, stock dividends, stock splits, reverse stock splits, and share repurchases;

d describe dividend payment chronology;

Present value models follow a fundamental tenet of economics which states that individuals defer consumption—that is, they invest—for the future benefits expected. Individuals and companies make an investment because they expect thereby to earn a rate of return over the investment period. Logically, the value of an investment should be equal to the present value of the expected future benefits. For common shares, an analyst can equate benefits to the cash flows to be generated by the investment. The simplest present value model of equity valuation is the dividend discount model (DDM), which specifies cash flows from a common stock investment to be dividends.

The next section describes aspects of dividends that users of dividend discount models should understand.

4.1 Dividends: Background for the Dividend Discount Model

Generally, there are two sources of return from investing in equities: (1) cash dividends received by an investor over his or her holding period and (2) the change in the market price of equities over that holding period.

A **dividend** is a distribution paid to shareholders based on the number of shares owned, and a cash dividend is a cash distribution made to a company's shareholders. Cash dividends are typically paid out regularly at known intervals; such dividends are known as regular cash dividends. By contrast, an **extra dividend** or **special dividend** is a dividend paid by a company that does not pay dividends on a regular schedule or a dividend that supplements regular cash dividends with an extra payment. Companies in cyclical industries and companies undergoing corporate and/or financial restructuring are among those observed to use extra dividends.[1]

The payment of dividends is not a legal obligation: dividends must be declared (i.e., authorized) by a company's board of directors; in some jurisdictions, they must also be approved by shareholders. Regular cash dividends are customarily declared and paid out quarterly in the United States and Canada; semiannually in Europe and Japan; and annually in some other countries, including China.

Dividend discount models address discounting expected cash dividends. A **stock dividend** (also known as a **bonus issue of shares**) is a type of dividend in which a company distributes additional shares of its common stock (typically, 2%–10% of the shares then outstanding) to shareholders instead of cash. A stock dividend divides the "pie" (the market value of shareholders' equity) into smaller pieces without affecting the value of the pie or any shareholder's proportional ownership in the company. Thus, stock dividends are not relevant for valuation. Stock splits and reverse stock splits are similar to stock dividends in that they have no economic effect on the company or shareholders. A **stock split** involves an increase in the number of shares outstanding with a consequent decrease in share price. An example of a stock split is a two-for-one stock split in which each shareholder is issued an additional share for each share currently owned. A **reverse stock split** involves a reduction in the number of shares outstanding with a corresponding increase in share price. In a one-for-two reverse stock split, each shareholder would receive one new share for every two old shares held, thereby reducing the number of shares outstanding by half.

In contrast to stock dividends and stock splits, share repurchases are an alternative to cash dividend payments. A **share repurchase** (or **buyback**) is a transaction in which a company uses cash to buy back its own shares. Shares that have been repurchased are not considered for dividends, voting, or computing earnings per share. A share repurchase is viewed as equivalent to the payment of cash dividends of equal value in terms of the effect on shareholders' wealth, all other things being equal. Company managements have expressed several key reasons for engaging in share repurchases—namely, (1) signaling a belief that their shares are undervalued (or, more generally, to support share prices), (2) flexibility in the amount and timing of distributing cash to shareholders, (3) tax efficiency in markets where tax rates on dividends exceed tax rates on capital gains, and (4) the ability to absorb increases in outstanding shares because of the exercise of employee stock options.

The payout of regular cash dividends to common shareholders follows a fairly standard chronology that is set in motion once the company's board of directors votes to pay the dividend. First is the **declaration date**, the day that the company issues a statement declaring a specific dividend. Next comes the **ex-dividend date** (or ex-date), the first date that a share trades without (i.e., "ex") the dividend. This is followed closely (one or two business days later) by the **holder-of-record date** (also called the owner-of-record date, shareholder-of-record date, record date, date of record, or date of book closure), the date that a shareholder listed on the company's books will be deemed to have ownership of the shares for purposes of receiving the upcoming dividend; the amount of time between the ex-date and the holder-of-record date is linked to the trade settlement cycle in force. The final milestone is the **payment date** (or **payable date**), which is the day that the company actually mails out (or electronically transfers) a dividend payment to shareholders.

---

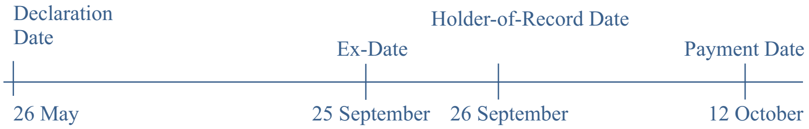
[1] Another type of dividend is a liquidating dividend, which is a return of capital rather than a distribution from earnings or retained earnings. Liquidating dividends are used when a company goes out of business and distributes its net assets, sells a portion of its business for cash and distributes the sale's proceeds, or pays a dividend that exceeds its accumulated retained earnings.

EXAMPLE 3

Total S.A. Dividend Payment Time Line

On 26 May 2017, Total S.A., one of the world's largest integrated energy companies, declared an annual dividend of €2.48 per share, payable on a quarterly basis. The first quarterly dividend of €2.48/4 = €0.62 was payable on 12 October 2017. The holder-of-record date was 26 September, and the ex-dividend date was 25 September. A timeline for the upcoming Total S.A. quarterly dividend is shown in Exhibit 1.

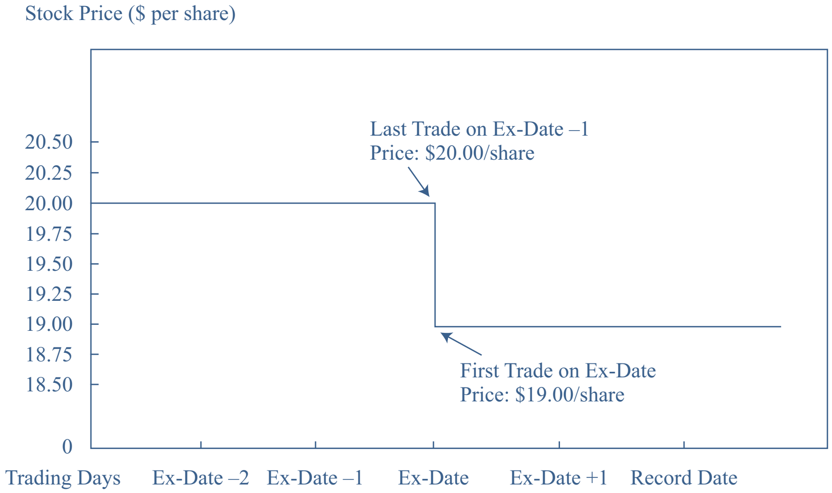
Exhibit 1 Timeline for Total S.A. Quarterly Dividend



Source: Total S.A. website: www.total.com.

Because buyers of a company's shares on the ex-dividend date are no longer eligible to receive the upcoming dividend, all else being equal, on that day the company's share price immediately decreases by the amount of the foregone dividend. Exhibit 2 illustrates the decrease in share price that occurs for a hypothetical company that has declared a $1.00 per share dividend as trading begins on its ex-dividend date.

Exhibit 2 Stock Price Change for Hypothetical Company on Ex-Dividend Date



Note: Assumes dividend declared is $1 per share and convention for stock trade settlement is $T + 3$.

DIVIDEND DISCOUNT MODEL (DDM) AND FREE-CASH-FLOW-TO-EQUITY MODEL (FCFE)

5

e explain the rationale for using present value models to value equity and describe the dividend discount and free-cash-flow-to-equity models;

m explain advantages and disadvantages of each category of valuation model.

If the issuing company is assumed to be a going concern, the intrinsic value of a share is the present value of expected future dividends. If a constant required rate of return is also assumed, then the DDM expression for the intrinsic value of a share is Equation 1:

$$V_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1+r)^t} \tag{1}$$

where

V_0 = value of a share of stock today, at $t = 0$

D_t = expected dividend in year t, assumed to be paid at the end of the year

r = required rate of return on the stock

At the shareholder level, cash received from a common stock investment includes any dividends received and the proceeds when shares are sold. If an investor intends to buy and hold a share for one year, the value of the share today is the present value of two cash flows—namely, the expected dividend *plus* the expected selling price in one year:

$$V_0 = \frac{D_1 + P_1}{(1+r)^1} = \frac{D_1}{(1+r)^1} + \frac{P_1}{(1+r)^1} \tag{2}$$

where P_1 = the expected price per share at $t = 1$.

To estimate the expected selling price, P_1, the analyst could estimate the price another investor with a one-year holding period would pay for the share in one year. If V_0 is based on D_1 and P_1, it follows that P_1 could be estimated from D_2 and P_2:

$$P_1 = \frac{D_2 + P_2}{(1 + r)}$$

Substituting the right side of this equation for P_1 in Equation 2 results in V_0 estimated as

$$V_0 = \frac{D_1}{(1 + r)} + \frac{D_2 + P_2}{(1 + r)^2} = \frac{D_1}{(1 + r)} + \frac{D_2}{(1 + r)^2} + \frac{P_2}{(1 + r)^2}$$

Repeating this process, we find the value for n holding periods is the present value of the expected dividends for the n periods plus the present value of the expected price in n periods:

$$V_0 = \frac{D_1}{(1 + r)^1} + \cdots + \frac{D_n}{(1 + r)^n} + \frac{P_n}{(1 + r)^n}$$

Using summation notation to represent the present value of the n expected dividends, we arrive at the general expression for an n-period holding period or investment horizon:

$$V_0 = \sum_{t=1}^{n} \frac{D_t}{(1 + r)^t} + \frac{P_n}{(1 + r)^n} \tag{3}$$

The expected value of a share at the end of the investment horizon—in effect, the expected selling price—is often referred to as the **terminal stock value** (or **terminal value**).

EXAMPLE 4

Estimating Share Value for a Three-Year Investment Horizon

For the next three years, the annual dividends of a stock are expected to be €2.00, €2.10, and €2.20. The stock price is expected to be €20.00 at the end of three years. If the required rate of return on the shares is 10 percent, what is the estimated value of a share?

Solution:

The present values of the expected future cash flows can be written as follows:

$$V_0 = \frac{2.00}{(1.10)^1} + \frac{2.10}{(1.10)^2} + \frac{2.20}{(1.10)^3} + \frac{20.00}{(1.10)^3}$$

Calculating and summing these present values gives an estimated share value of V_0 = 1.818 + 1.736 + 1.653 + 15.026 = €20.23.

The three dividends have a total present value of €5.207, and the terminal stock value has a present value of €15.026, for a total estimated value of €20.23.

Extending the holding period into the indefinite future, we can say that a stock's estimated value is the present value of all expected future dividends as shown in Equation 1.

Consideration of an indefinite future is valid because businesses established as corporations are generally set up to operate indefinitely. This general form of the DDM applies even in the case in which the investor has a finite investment horizon. For that investor, stock value today depends *directly* on the dividends the investor expects to receive before the stock is sold and depends *indirectly* on the expected dividends for periods subsequent to that sale, because those expected future dividends determine the expected selling price. Thus, the general expression given by Equation 1 holds irrespective of the investor's holding period.

In practice, many analysts prefer to use a free-cash-flow-to-equity (FCFE) valuation model. These analysts assume that dividend-paying *capacity* should be reflected in the cash flow estimates rather than *expected dividends*. FCFE is a measure of dividend-paying capacity. Analysts may also use FCFE valuation models for a non-dividend-paying stock. To use a DDM, the analyst needs to predict the timing and amount of the first dividend and all the dividends or dividend growth thereafter. Making these predictions for non-dividend-paying stock accurately is typically difficult, so in such cases, analysts often resort to FCFE models.

The calculation of FCFE starts with the calculation of cash flow from operations (CFO). CFO is simply defined as net income plus non-cash expenses minus investment in working capital. FCFE is a measure of cash flow generated in a period that is available for distribution to common shareholders. What does "available for distribution" mean? The entire CFO is *not* available for distribution; the portion of the CFO needed for fixed capital investment (FCInv) during the period to maintain the value of the company as a going concern is *not* viewed as available for distribution to common shareholders. Net amounts borrowed (borrowings minus repayments) are considered to be available for distribution to common shareholders. Thus, FCFE can be expressed as

$$\text{FCFE} = \text{CFO} - \text{FCInv} + \text{Net borrowing} \tag{4}$$

The information needed to calculate historical FCFE is available from a company's statement of cash flows and financial disclosures. Frequently, under the assumption that management is acting in the interest of maintaining the value of the company as a going concern, reported capital expenditure is taken to represent FCInv. Analysts must make projections of financials to forecast future FCFE. Valuation obtained by using FCFE involves discounting expected future FCFE by the required rate of return on equity; the expression parallels Equation 1:

$$V_0 = \sum_{t=1}^{\infty}\frac{\text{FCFE}_t}{(1+r)^t}$$

EXAMPLE 5

Present Value Models

1 An investor expects a share to pay dividends of \$3.00 and \$3.15 at the end of Years 1 and 2, respectively. At the end of the second year, the investor expects the shares to trade at \$40.00. The required rate of return on the shares is 8 percent. If the investor's forecasts are accurate and the market price of the shares is currently \$30, the *most likely* conclusion is that the shares are:

- **A** overvalued.
- **B** undervalued.
- **C** fairly valued.

2 Two investors with different holding periods but the same expectations and required rate of return for a company are estimating the intrinsic value of a common share of the company. The investor with the shorter holding period will *most likely* estimate a:

A lower intrinsic value.

B higher intrinsic value.

C similar intrinsic value.

3 An equity valuation model that focuses on expected dividends rather than the capacity to pay dividends is the:

A dividend discount model.

B free cash flow to equity model.

C cash flow return on investment model.

Solution to 1:

B is correct.

$$V_0 = \frac{3.00}{(1.08)^1} + \frac{3.15}{(1.08)^2} + \frac{40.00}{(1.08)^2} = 39.77$$

The value estimate of $39.77 exceeds the market price of $30, so the conclusion is that the shares are undervalued.

Solution to 2:

C is correct. The intrinsic value of a security is independent of the investor's holding period.

Solution to 3:

A is correct. Dividend discount models focus on expected dividends.

How is the required rate of return for use in present value models estimated? To estimate the required rate of return on a share, analysts frequently use the capital asset pricing model (CAPM):

$$\begin{aligned}\text{Required rate of return on share } i = {} & \text{Current expected risk} \\ & \text{–free rate of return} \\ & + \text{Beta}_i\left[\text{Market (equity) risk premium}\right]\end{aligned} \tag{5}$$

Equation 5 states that the required rate of return on a share is the sum of the current expected risk-free rate plus a risk premium that equals the product of the stock's beta (a measure of non-diversifiable risk) and the market risk premium (the expected return of the market in excess of the risk-free return, where in practice, the "market" is often represented by a broad stock market index). However, even if analysts agree that the CAPM is an appropriate model, their inputs into the CAPM may differ. Thus, there is no uniquely correct answer to the question: What is the required rate of return?

Other common methods for estimating the required rate of return for the stock of a company include adding a risk premium that is based on economic judgments, rather than the CAPM, to an appropriate risk-free rate (usually a government bond) and adding a risk premium to the yield on the company's bonds. Good business and economic judgment is paramount in estimating the required rate of return. In many investment firms, required rates of return are determined by firm policy.

PREFERRED STOCK VALUATION

6

f calculate the intrinsic value of a non-callable, non-convertible preferred stock;

m explain advantages and disadvantages of each category of valuation model.

General dividend discount models are relatively easy to apply to preferred shares. In its simplest form, **preferred stock** is a form of equity (generally, non-voting) that has priority over common stock in the receipt of dividends and on the issuer's assets in the event of a company's liquidation. It may have a stated maturity date at which time payment of the stock's par (face) value is made or it may be perpetual with no maturity date; additionally, it may be callable or convertible.

For a non-callable, non-convertible perpetual preferred share paying a level dividend D and assuming a constant required rate of return over time, Equation 1 reduces to the formula for the present value of a perpetuity. Its value is:

$$V_0 = \frac{D_0}{r} \tag{6}$$

For example, a \$100 par value non-callable perpetual preferred stock offers an annual dividend of \$5.50. If its required rate of return is 6 percent, the value estimate would be \$5.50/0.06 = \$91.67.

For a non-callable, non-convertible preferred stock with maturity at time n, the estimated intrinsic value can be estimated by using Equation 3 but using the preferred stock's par value, F, instead of P_n:

$$V_0 = \sum_{t=1}^{n} \frac{D_t}{(1+r)^t} + \frac{F}{(1+r)^n} \tag{7}$$

When Equation 7 is used, the most precise approach is to use values for n, r, and D that reflect the payment schedule of the dividends. This method is similar to the practice of fixed-income analysts in valuing a bond. For example, a non-convertible preferred stock with a par value of £20.00, maturity in six years, a nominal required rate of return of 8.20 percent, and semiannual dividends of £2.00 would be valued by using an n of 12, an r of 4.10 percent, a D of £2.00, and an F of £20.00. The result would be an estimated value of £31.01. Assuming payments are annual rather than semiannual (i.e., assuming that $n = 6$, $r = 8.20$ percent, and $D = £4.00$) would result in an estimated value of £30.84.

Preferred stock issues are frequently callable (redeemable) by the issuer at some point prior to maturity, often at par value or at prices in excess of par value that decline to par value as the maturity date approaches. Such call options tend to reduce the value of a preferred issue to an investor because the option to redeem will be exercised by the issuer when it is in the issuer's favor and ignored when it is not. For example, if an issuer can redeem shares at par value that would otherwise trade (on the basis of dividends, maturity, and required rate of return) above par value, the issuer has motivation to redeem the shares.

Preferred stock issues can also include a retraction option that enables the holder of the preferred stock to sell the shares back to the issuer prior to maturity on prespecified terms. Essentially, the holder of the shares has a put option. Such put options tend to increase the value of a preferred issue to an investor because the option to retract will be exercised by the investor when it is in the investor's favor and ignored when it is not. Although the precise valuation of issues with such embedded options is beyond the scope of this reading, Example 6 includes a case in which Equation 7 can be used to approximate the value of a callable, retractable preferred share.

EXAMPLE 6

Preferred Share Valuation: Two Cases

Case 1: Non-callable, Non-convertible, Perpetual Preferred Shares

The following facts concerning the Union Electric Company 4.75 percent perpetual preferred shares are as follows:

- Issuer: Union Electric Co. (owned by Ameren)
- Par value: US$100
- Dividend: US$4.75 per year
- Maturity: perpetual
- Embedded options: none
- Credit rating: Moody's Investors Service/Standard & Poor's Ba1/BB
- Required rate of return on Ba1/BB rated preferred shares as of valuation date: 7.5 percent.

A Estimate the intrinsic value of this preferred share.

B Explain whether the intrinsic value of this issue would be higher or lower if the issue were callable (with all other facts remaining unchanged).

Solution to 1A:

Basing the discount rate on the required rate of return on Ba1/BB rated preferred shares of 7.5 percent gives an intrinsic value estimate of US$4.75/0.075 = US$63.33.

Solution to 1B:

The intrinsic value would be lower if the issue were callable. The option to redeem or call the issue is valuable to the issuer because the call will be exercised when doing so is in the issuer's interest. The intrinsic value of the shares to the investor will typically be lower if the issue is callable. In this case, because the intrinsic value without the call is much less than the par value, the issuer would be unlikely to redeem the issue if it were callable; thus, callability would reduce intrinsic value, but only slightly.

Case 2: Retractable Term Preferred Shares

Retractable term preferred shares are a type of preferred share that has been previously issued by Canadian companies, and have now began to be offered by companies in other jurisdictions, including Japan. This type of issue specifies a "retraction date" when the preferred shareholders have the option to sell back their shares to the issuer at par value (i.e., the shares are "retractable" or "putable" at that date).[2] At predetermined dates prior to the retraction date, the issuer has the option to redeem the preferred issue at predetermined prices (which are always at or above par value).

An example of a retractable term preferred share currently outstanding is TMC (Toyota Motor Corporation), First Series Model AA class shares, with a 0.5 percent dividend rate, increasing by 0.5 percent every year until 2020 and thereafter, becoming fixed at 2.5 percent. TMC is leading global automobile manufacturer, with headquarters in Japan and global operations. The issue is in Japanese Yen. The shares have a ¥10,598 par value and pay a semiannual dividend of ¥26.5 [= (0.5 percent × ¥10,598)/2] on 31 March 2016. The semiannual dividend is expected to increase to ¥132.5 [= (2.5 percent × ¥10,598)/2] on 31 March 2020 and beyond. As of 31 December 2017 the company carried ratings from Moody's and Standard & Poor's of Aa3 and AA–, respectively. Thus, the shares are viewed by Moody's and Standard & Poor's as having "adequate" credit quality, qualified by "Aa3 and AA–," which means relatively high quality within that group. Beginning from 2 April 2021, the shares are redeemable at the option of TMC at ¥10,598 (par value). The retraction date is the last day of March, June, September, and December of each year, starting from 1 September 2020, with the shares retractable at par value. The Series AA shares have voting rights and may exercise their voting rights and other rights held by holders of common shares of TMC in the same manner. The Series AA shares were issued at a 20% premium to the common shares price in 2015, and since then the share price has decreased to ¥7,243 as at 31 December 2017, and with a current required rate of 3.05 per year (1.525 percent semiannual). Because the issue's market price is so far below the prices at which TMC could redeem or call the issue, redemption is considered to be unlikely by TMC, whereas the retraction option for the Series AA holders appears to have a significant value since they will potentially be able to put back the shares to TMC at approximately 45 percent over the current market value (¥10,598 compared to ¥7,243).

---

[2] "Retraction" refers to this option, which is a put option. The terminology is not completely settled: The type of share being called "retractable term preferred" is also known as "hard retractable preferred," with "hard" referring to payment in cash rather than common shares at the retraction date.

A Assume that the issue will be retracted in December 2020; the holders of the shares will put the shares to the company in December 2020. Based on the information given, estimate the intrinsic value of a share. Assume it is December 2017.

Solution to 2A:

An intrinsic value estimate of a share of this preferred issue is ¥10,279.

Expected semiannual dividends:

Year ended March 31, 2018: ¥79.5 [= (1.5 percent × ¥10,598)/2]

Year ended March 31, 2019: ¥106 [= (2.0 percent × ¥10,598)/2]

Year ended March 31, 2020: ¥132.5 [= (2.5 percent × ¥10,598)/2]

$$V_0 = \left[\frac{¥79.5}{1.01525} + \frac{¥79.5}{1.01525^2} + \frac{¥106}{1.01525^3} + \frac{¥106}{1.01525^4} + \frac{¥132.5}{1.01525^5} + \frac{¥132.5}{1.01525^6} + \frac{¥10{,}598}{1.01525^{6.5}}\right]$$
$$\approx ¥10{,}205$$

The difference between the current market price of ¥7,243 and the intrinsic value of ¥10,205 is the implied value of retractable option given to the holders of the Series AA shares.

THE GORDON GROWTH MODEL

7

g calculate and interpret the intrinsic value of an equity security based on the Gordon (constant) growth dividend discount model or a two-stage dividend discount model, as appropriate;

h identify characteristics of companies for which the constant growth or a multistage dividend discount model is appropriate;

m explain advantages and disadvantages of each category of valuation model.

A rather obvious problem when one is trying to implement Equation 1 for common equity is that it requires the analyst to estimate an infinite series of expected dividends. To simplify this process, analysts frequently make assumptions about how dividends will grow or change over time. The Gordon (constant) growth model (Gordon, 1962) is a simple and well-recognized DDM. The model assumes dividends grow indefinitely at a constant rate.

Because of its assumption of a constant growth rate, the Gordon growth model is particularly appropriate for valuing the equity of dividend-paying companies that are relatively insensitive to the business cycle and in a mature growth phase. Examples might include an electric utility serving a slowly growing area or a producer of a staple food product (e.g., bread). A history of increasing the dividend at a stable growth rate is another practical criterion if the analyst believes that pattern will hold in the future.

With a constant growth assumption, Equation 1 can be written as Equation 8, where g is the constant growth rate:

$$V_0 = \sum_{t=1}^{\infty} \frac{D_0 (1+g)^t}{(1+r)^t} = D_0 \left[\frac{(1+g)}{(1+r)} + \frac{(1+g)^2}{(1+r)^2} + \ldots + \frac{(1+g)^{\infty}}{(1+r)^{\infty}} \right] \tag{8}$$

If required return r is assumed to be strictly greater than growth rate g, then the square-bracketed term in Equation 8 is an infinite geometric series and sums to $[(1 + g)/(r - g)]$. Substituting into Equation 8 produces the Gordon growth model as presented in Equation 9:

$$V_0 = \frac{D_0 (1+g)}{r-g} = \frac{D_1}{r-g} \tag{9}$$

For an illustration of the expression, suppose the current (most recent) annual dividend on a share is €5.00 and dividends are expected to grow at 4 percent per year. The required rate of return on equity is 8 percent. The Gordon growth model estimate of intrinsic value is, therefore, €5.00(1.04)/(0.08 – 0.04) = €5.20/0.04 = €130 per share. Note that the numerator is D_1 not D_0. (Using the wrong numerator is a common error.)

The Gordon growth model estimates intrinsic value as the present value of a growing perpetuity. If the growth rate, g, is assumed to be zero, Equation 8 reduces to the expression for the present value of a perpetuity, given earlier as Equation 6.

In estimating a long-term growth rate, analysts use a variety of methods, including assessing the growth in dividends or earnings over time, using the industry median growth rate, and using the relationship shown in Equation 10 to estimate the sustainable growth rate:

$$g = b \times \text{ROE} \tag{10}$$

where

g = dividend growth rate
b = earnings retention rate = (1 – Dividend payout ratio)
ROE = return on equity

Example 7 illustrates the application of the Gordon growth model to the shares of a large industrial manufacturing company. The analyst believes it will continue to grow at a rate that it achieved in the previous three years and remain stable in the future. The example asks how much the dividend growth assumption adds to the intrinsic value estimate. The question is relevant to valuation because if the amount is high on a percentage basis, a large part of the value of the share depends on the realization of the growth estimate. One can answer the question by subtracting from the intrinsic value estimate determined by Equation 9 the value determined by Equation 6, which assumes no dividend growth.[3]

EXAMPLE 7

Applying the Gordon Growth Model

Siemens AG operates in the capital goods and technology space. It is involved in the engineering, manufacturing, automation, power, and transportation sectors. It operates globally and is one of the largest companies in the sectors in which it operates. It is a substantial employer in both its original, domestic German market, as well as dozens of countries around the world. Selected financial information for Siemens appears in Exhibit 3.

Exhibit 3 Selected Financial Information for Siemens AG

Year	2017	2016	2015	2014	2013
EPS	€7.45	€6.74	€8.85	€6.37	€5.08
DPS	€3.7	€3.6	€3.5	€3.3	€3.0
Payout ratio	50%	53%	40%	52%	59%
ROE	15.6%	15.9%	22.3%	18.2%	14.6%
Share price (XETRA - Frankfurt)	€119.2	€104.2	€79.94	€94.37	€89.06

Note: DPS stands for "dividends per share."
Source: Morningstar, www.siemens.com.

The analyst estimates the growth rate to be approximately 5.4 percent based on the dividend growth rate over the period 2013 to 2017 [$3(1 + g)^4 = 3.7$, so $g = 5.4\%$]. To verify that the estimated growth rate of 5.4 percent is feasible in the future, the analyst also uses the average of Siemens's retention rate and ROE for the previous five years ($g \approx 0.49 \times 17.3\% \approx 8.5\%$) to estimate the sustainable growth rate.

Using a number of approaches, including adding a risk premium to a long-term German government bond and using the CAPM, the analyst estimates a required return of 7.5 percent. The most recent dividend of €3.70 is used for D_0.

1 Use the Gordon growth model to estimate Siemens's intrinsic value.
2 How much does the dividend growth assumption add to the intrinsic value estimate?
3 Based on the estimated intrinsic value, is a share of Siemens undervalued, overvalued, or fairly valued?

---

[3] A related concept, the present value of growth opportunities (PVGO), is discussed in more advanced readings.

4 What is the intrinsic value if the growth rate estimate is lowered to 4.4 percent?

5 What is the intrinsic value if the growth rate estimate is lowered to 4.4 percent and the required rate of return estimate is increased to 8.5 percent?

Solution to 1:

$$V_0 = \frac{€3.70(1 + 0.054)}{0.075 - 0.054} = €185.70$$

Solution to 2:

$$€184.20 - \frac{€3.70}{0.075} = €136.37$$

Solution to 3:

A share of Siemens appears to be undervalued. The analyst, before making a recommendation, might consider how realistic the estimated inputs are and check the sensitivity of the estimated value to changes in the inputs.

Solution to 4:

$$V_0 = \frac{€3.70(1 + 0.044)}{0.075 - 0.044} = €124.61$$

Solution to 5:

$$V_0 = \frac{€3.70(1 + 0.044)}{0.085 - 0.044} = €94.21$$

The Gordon growth model estimate of intrinsic value is extremely sensitive to the choice of required rate of return r and growth rate g. It is possible that the growth rate assumption and the required return assumption used initially were too high. Worldwide economic growth is typically in the low single digits, which may mean that a large company such as Siemens may struggle to grow dividends at 5.4 percent into perpetuity. Exhibit 4 presents a further sensitivity analysis of Siemens's intrinsic value to the required return and growth estimates.

Exhibit 4 Sensitivity Analysis of the Intrinsic-Value Estimate for Siemens AG

	g = 2.5%	g = 3.5%	g = 4.5%	g = 5.5%	g = 6.5%
r = 6%	€108.4	€153.2	€257.8	€780.7	—
r = 7%	€84.3	€109.4	€154.7	€260.2	€788.1
r = 8%	€69.0	€85.1	€110.5	€156.1	€262.7
r = 9%	€58.3	€69.6	€85.9	€111.5	€157.6
r = 10%	€50.6	€58.9	€70.3	€86.7	€112.6

Note that no value is shown when the growth rate exceeds the required rate of return. The Gordon growth model assumes that the growth rate cannot be greater than the required rate of return.

The assumptions of the Gordon model are as follows:

- Dividends are the correct metric to use for valuation purposes.
- The dividend growth rate is forever: It is perpetual and never changes.
- The required rate of return is also constant over time.
- The dividend growth rate is strictly less than the required rate of return.

An analyst might be dissatisfied with these assumptions for many reasons. The equities being examined might not currently pay a dividend. The Gordon assumptions might be too simplistic to reflect the characteristics of the companies being evaluated. Some alternatives to using the Gordon model are as follows:

- Use a more robust DDM that allows for varying patterns of growth.
- Use a cash flow measure other than dividends for valuation purposes.
- Use some other approach (such as a multiplier method) to valuation.

Applying a DDM is difficult if the company being analyzed is not currently paying a dividend. A company may not be paying a dividend if 1) the investment opportunities the company has are all so attractive that the retention and reinvestment of funds is preferable, from a return perspective, to the distribution of a dividend to shareholders or 2) the company is in such shaky financial condition that it cannot afford to pay a dividend. An analyst might still use a DDM to value such companies by assuming that dividends will begin at some future point in time. The analyst might further assume that constant growth occurs after that date and use the Gordon growth model for valuation. Extrapolating from no current dividend, however, generally yields highly uncertain forecasts. Analysts typically choose to use one or more of the alternatives instead of or as a supplement to the Gordon growth model.

EXAMPLE 8

Gordon Growth Model in the Case of No Current Dividend

A company does not currently pay a dividend but is expected to begin to do so in five years (at $t = 5$). The first dividend is expected to be \$4.00 and to be received five years from today. That dividend is expected to grow at 6 percent into perpetuity. The required return is 10 percent. What is the estimated current intrinsic value?

Solution:

The analyst can value the share in two pieces:

1. The analyst uses the Gordon growth model to estimate the value at $t = 5$; in the model, the year-ahead dividend is \$4(1.06). Then the analyst finds the present value of this value as of $t = 0$.
2. The analyst finds the present value of the \$4 dividend not “counted” in the estimate in Piece 1 (which values dividends from $t = 6$ onward). Note that the statement of the problem implies that D_0, D_1, D_2, D_3, and D_4 are zero.

Piece 1: The value of this piece is $65.818:

$$V_n = \frac{D_n(1+g)}{r-g} = \frac{D_{n+1}}{r-g}$$

$$V_5 = \frac{\$4(1+0.06)}{0.10-0.06} = \frac{\$4.24}{0.04} = \$106$$

$$V_0 = \frac{\$106}{(1+0.10)^5} = \$65.818$$

Piece 2: The value of this piece is $2.484:

$$V_0 = \frac{\$4}{(1+0.10)^5} = \$2.484$$

The sum of the two pieces is $65.818 + $2.484 = $68.30.

Alternatively, the analyst could value the share at t = 4, the point at which dividends are expected to be paid in the following year and from which point they are expected to grow at a constant rate.

$$V_4 = \frac{\$4.00}{0.10-0.06} = \frac{\$4.00}{0.04} = \$100$$

$$V_0 = \frac{\$100}{(1+0.10)^4} = \$68.30$$

The next section addresses the application of the DDM with more flexible assumptions as to the dividend growth rate.

8

MULTISTAGE DIVIDEND DISCOUNT MODELS

g calculate and interpret the intrinsic value of an equity security based on the Gordon (constant) growth dividend discount model or a two-stage dividend discount model, as appropriate;

h identify characteristics of companies for which the constant growth or a multistage dividend discount model is appropriate;

m explain advantages and disadvantages of each category of valuation model.

Multistage growth models are often used to model rapidly growing companies. The *two-stage DDM* assumes that at some point the company will begin to pay dividends that grow at a constant rate, but prior to that time the company will pay dividends that are growing at a higher rate than can be sustained in the long run. That is, the company is assumed to experience an initial, finite period of high growth, perhaps prior to the entry of competitors, followed by an infinite period of sustainable growth. The two-stage DDM thus makes use of two growth rates: a high growth rate for an initial, finite period followed by a lower, sustainable growth rate into perpetuity. The Gordon growth model is used to estimate a terminal value at time n that reflects the present value at time n of the dividends received during the sustainable growth period.

Equation 11 will be used here as the starting point for a two-stage valuation model. The two-stage valuation model is similar to Example 8 except that instead of assuming zero dividends for the initial period, the analyst assumes that dividends will exhibit a high rate of growth during the initial period. Equation 11 values the dividends over the short-term period of high growth and the terminal value at the end of the period of high growth. The short-term growth rate, g_S, lasts for n years. The intrinsic value per share in year n, V_n, represents the year n value of the dividends received during the sustainable growth period or the terminal value at time n. V_n can be estimated by using the Gordon growth model as shown in Equation 12, where g_L is the long-term or sustainable growth rate. The dividend in year $n + 1$, D_{n+1}, can be determined by using Equation 13:

$$V_0 = \sum_{t=1}^{n} \frac{D_0(1 + g_S)^t}{(1 + r)^t} + \frac{V_n}{(1 + r)^n} \tag{11}$$

$$V_n = \frac{D_{n+1}}{r - g_L} \tag{12}$$

$$D_{n+1} = D_0(1 + g_S)^n(1 + g_L) \tag{13}$$

EXAMPLE 9

Applying the Two-Stage Dividend Discount Model

The current dividend, D_0, is \$5.00. Growth is expected to be 10 percent a year for three years and then 5 percent thereafter. The required rate of return is 15 percent. Estimate the intrinsic value.

Solution:

$$D_1 = \$5.00(1 + 0.10) = \$5.50$$

$$D_2 = \$5.00(1 + 0.10)^2 = \$6.05$$

$$D_3 = \$5.00(1 + 0.10)^3 = \$6.655$$

$$D_4 = \$5.00(1 + 0.10)^3(1 + 0.05) = \$6.98775$$

$$V_3 = \frac{\$6.98775}{0.15 - 0.05} = \$69.8775$$

$$V_0 = \frac{\$5.50}{(1 + 0.15)} + \frac{\$6.05}{(1 + 0.15)^2} + \frac{\$6.655}{(1 + 0.15)^3} + \frac{\$69.8775}{(1 + 0.15)^3} \approx \$59.68$$

The DDM can be extended to as many stages as deemed appropriate. For most publicly traded companies (that is, companies beyond the start-up stage), practitioners assume growth will ultimately fall into three stages: 1) growth, 2) transition, and 3) maturity. This assumption supports the use of a *three-stage DDM*, which makes use of three growth rates: a high growth rate for an initial finite period, followed by a lower growth rate for a finite second period, followed by a lower, sustainable growth rate into perpetuity.

One can make the case that a three-stage DDM would be most appropriate for a fairly young company, one that is just entering the growth phase. The two-stage DDM would be appropriate to estimate the value of an older company that has already moved through its growth phase and is currently in the transition phase (a period with a higher growth rate than the sustainable growth rate) prior to moving to the maturity phase (the period with a lower, sustainable growth rate).

However, the choice of a two-stage DDM need not rely solely on the age of a company. Long-established companies sometimes manage to restart above-average growth through, for example, innovation, expansion to new markets, or acquisitions. Or a company's long-run growth rate may be interrupted by a period of subnormal performance. If growth is expected to moderate (in the first case) or improve (in the second case) toward some long-term growth rate, a two-stage DDM may be appropriate. Thus, we chose a two-stage DDM to value International Business Machines Corporation in Example 10.

EXAMPLE 10

Two-Stage Dividend Discount Model: International Business Machines Corporation

International Business Machines Corporation (IBM) is a US based leading technology company. IBM was founded in 1911, initially as a company that manufactured machinery for sale and lease, ranging from commercial scales and industrial time recorders, meat and cheese slicers, to punched cards. IBM introduced the personal computer in 1981; however, by the 1990s, it began to suffer losses in its core computer manufacturing business and by the 2000s, it had begun to diversify into business consulting, which was finalized in 2005 when it sold its personal computer business to Chinese company Lenovo. IBM now operates through five segments: Cognitive Solutions, Global Business Services (GBS), Technology Services & Cloud Platforms, Systems, and Global Financing. The Cognitive Solutions segment delivers a spectrum of capabilities from descriptive, predictive, and prescriptive analytics to cognitive systems. Cognitive Solutions includes Watson, a cognitive computing platform that has the ability to interact in natural language, process big data, and learn from interactions with people and computers. The GBS segment provides clients with consulting, application management services, and global process services. The Technology Services & Cloud Platforms segment provides information technology infrastructure services. The Systems segment provides clients with infrastructure technologies. The Global Financing segment includes client financing, commercial financing, and remanufacturing and remarketing.

The 30 July 2018 *Value Line* report on IBM appears in Exhibit 5. IBM has increased its dividends every year over the past 15 years. Information from *Value Line* shows that the dividend growth is around 17.0 percent for the past 10 years, 13.5 percent for the past 5 years, and estimated to be 4.5 percent for 2015 to 2023. After a period of growth through acquisition and merger, the pattern suggests that IBM may be transitioning to a mature growth phase.

The two-stage DDM is arguably a good choice for valuing IBM because the company appears to be transitioning from a high-growth phase (note the 13.5 percent dividend growth for the past 5 years) to a lower-growth phase (note the forecast of 4.5 percent dividend growth to 2015–2023).

The CAPM can be used to estimate the required return, r, for IBM. The *Value Line* report (in the top left corner) estimates beta to be 0.95. Using the yield of about 2.0 percent on 10-year US Treasury notes as a proxy for the risk-free rate and assuming an equity risk premium of 5.0 percent, we find the estimate for r would be 6.75 percent [2.0% + 0.95(5.0%)].

To estimate the intrinsic value at the end of 2018, we use the 2018 dividend of US$6.21 from the *Value Line* report. The dividend is assumed to grow at a rate of 5.0 percent for one year and then 2.33 percent thereafter. The growth rate assumption for the first stage is consistent with the *Value Line* forecast for the

2018 and 2019 dividends. Our assumption of a 2.33 percent perpetual growth rate produces an 8-year growth rate assumption near 4.5%,[4] which is consistent with the *Value Line* forecast of 4.5 percent growth from 2015–2022. Thus:

$$D_{2019} = \text{US\$}6.21(1 + 0.05) = \text{US\$}6.5205$$

$$D_{2020} = \text{US\$}6.21(1 + 0.05)(1 + 0.0233) = \text{US\$}6.6724$$

$$D_{2021} = \text{US\$}6.21(1 + 0.05)(1 + 0.0233)^2 = \text{US\$}6.8279$$

$$V_{2020} = \frac{\text{US\$}6.8279}{0.0675 - 0.0233} = \text{US\$}154.4774$$

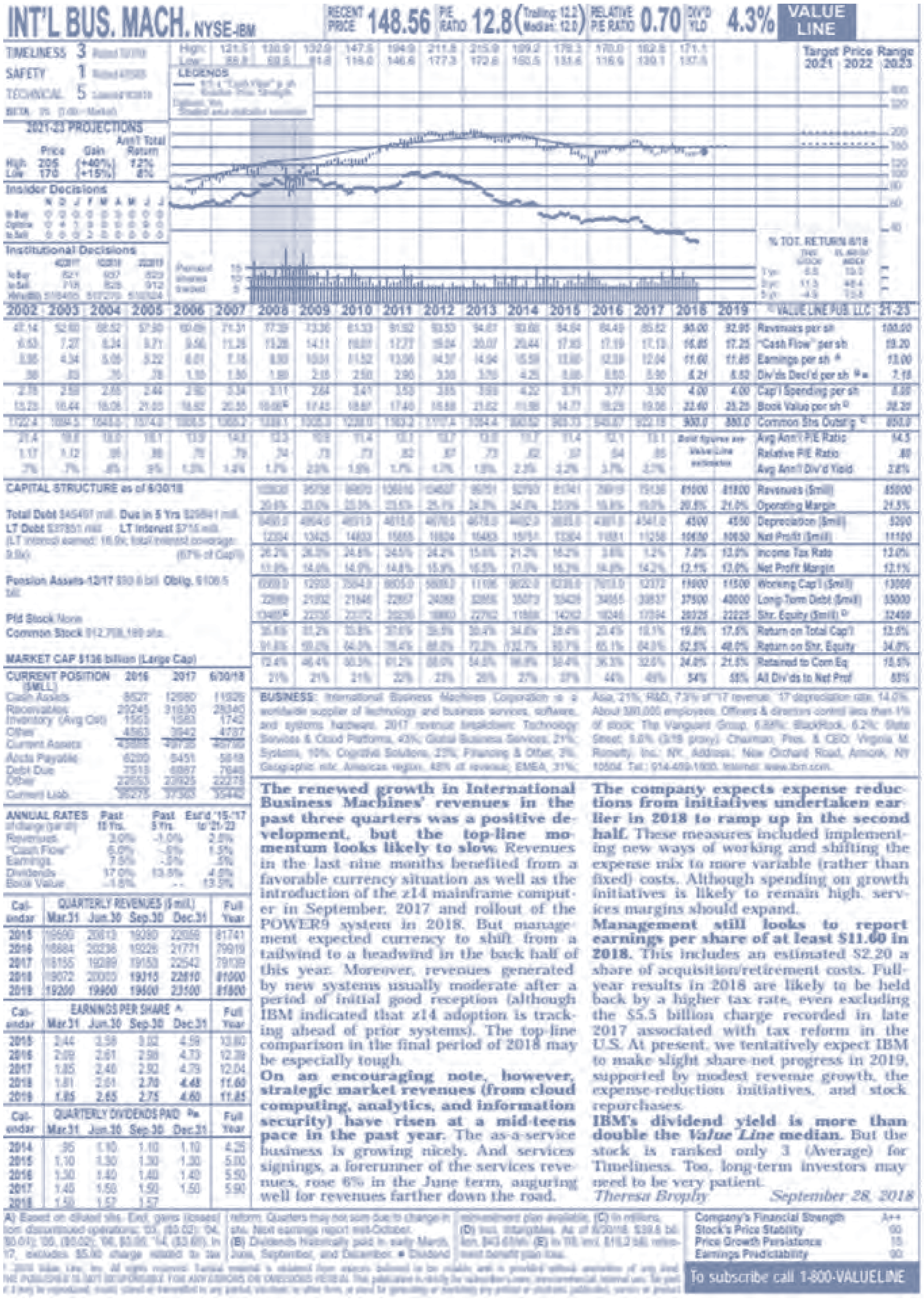
$$V_{2018} = \frac{\text{US\$}6.5205}{(1 + 0.0675)} + \frac{\text{US\$}6.6724}{(1 + 0.0675)^2} + \frac{\text{US\$}154.4774}{(1 + 0.0675)^2} \approx \text{US\$}147.523$$

Given a recent price of US$148.56, as noted at the top left corner of the *Value Line* report, the intrinsic-value estimate of US$147.523 suggests that IBM is approximately fairly valued.

---

[4] The exact geometric average annual growth rate for 2015–2023 can be determined as $[(1 + 0.10)(1 + 0.0727)(1 + 0.0525)(1 + 0.499)(1 + 0.022)(1 + 0.022)(1 + 0.022)(1 + 0.022)]^{1/8} - 1 = 4.5\%$.

Exhibit 5 *Value Line* Report on IBM



INT'L BUS. MACH. NYSE-IBM

RECENT PRICE 148.56 | P/E RATIO 12.8 (Trailing: 12.2 / Median: 12.0) | RELATIVE P/E RATIO 0.70 | DIV'D YLD 4.3% | VALUE LINE

TIMELINESS 3
SAFETY 1
TECHNICAL 5

2021-23 PROJECTIONS

	Price	Gain	Ann'l Total Return
High	205	(+40%)	12%
Low	170	(+15%)	8%

Insider Decisions

Institutional Decisions

% TOT. RETURN 8/18

CAPITAL STRUCTURE as of 6/30/18

Total Debt $45497 mill. Due in 5 Yrs $29841 mill.
LT Debt $37851 mill. LT Interest $715 mill.

Pension Assets-12/17 $90.6 bill. Oblig. $106.5 bill.

Pfd Stock None

MARKET CAP $136 billion (Large Cap)

CURRENT POSITION

ANNUAL RATES

QUARTERLY REVENUES ($ mill.)

EARNINGS PER SHARE

QUARTERLY DIVIDENDS PAID

BUSINESS: International Business Machines Corporation is a worldwide supplier of technology and business services, software, and systems hardware. 2017 revenue breakdown: Technology Services & Cloud Platforms, 43%; Global Business Services, 21%; Systems, 10%; Cognitive Solutions, 23%; Financing & Other, 3%. Geographic mix: Americas region, 48% of revenue; EMEA, 31%; Asia, 21%. R&D, 7.3% of '17 revenue. '17 depreciation rate: 14.0%. Officers & directors control less than 1% of stock. The Vanguard Group, 6.8%; BlackRock, 6.2%; State Street, 5.6% (3/18 proxy). Chairman, Pres. & CEO: Virginia M. Rometty. Inc.: NY. Address: New Orchard Road, Armonk, NY 10504. Tel.: 914-499-1900. Internet: www.ibm.com.

The renewed growth in International Business Machines' revenues in the past three quarters was a positive development, but the top-line momentum looks likely to slow. Revenues in the last nine months benefited from a favorable currency situation as well as the introduction of the z14 mainframe computer in September, 2017 and rollout of the POWER9 system in 2018. But management expected currency to shift from a tailwind to a headwind in the back half of this year. Moreover, revenues generated by new systems usually moderate after a period of initial good reception (although IBM indicated that z14 adoption is tracking ahead of prior systems). The top-line comparison in the final period of 2018 may be especially tough.

On an encouraging note, however, strategic market revenues (from cloud computing, analytics, and information security) have risen at a mid-teens pace in the past year. The as-a-service business is growing nicely. And services signings, a forerunner of the services revenues, rose 6% in the June term, auguring well for revenues farther down the road.

The company expects expense reductions from initiatives undertaken earlier in 2018 to ramp up in the second half. These measures included implementing new ways of working and shifting the expense mix to more variable (rather than fixed) costs. Although spending on growth initiatives is likely to remain high, services margins should expand.

Management still looks to report earnings per share of at least $11.60 in 2018. This includes an estimated $2.20 a share of acquisition/retirement costs. Full-year results in 2018 are likely to be held back by a higher tax rate, even excluding the $5.5 billion charge recorded in late 2017 associated with tax reform in the U.S. At present, we tentatively expect IBM to make slight share-net progress in 2019, supported by modest revenue growth, the expense-reduction initiatives, and stock repurchases.

IBM's dividend yield is more than double the *Value Line* median. But the stock is ranked only 3 (Average) for Timeliness. Too, long-term investors may need to be very patient.

Theresa Brophy *September 28, 2018*

(B) Dividends historically paid in early March, June, September, and December.

Company's Financial Strength	A++
Stock's Price Stability	90
Price Growth Persistence	15
Earnings Predictability	90

To subscribe call 1-800-VALUELINE

Source: © 2018 Value Line, Inc. All Rights Reserved Worldwide. "Value Line" is a registered trademark of Value Line, Inc. Value Line Geometric and Arithmetic Indices calculated by Thomson Reuters. Information supplied by Thomson Reuters.

MULTIPLER MODELS AND RELATIONSHIP AMONG PRICE MULTIPLES, PRESENT VALUE MODELS, AND FUNDAMENTALS

9

i. explain the rationale for using price multiples to value equity, how the price to earnings multiple relates to fundamentals, and the use of multiples based on comparables;

j. calculate and interpret the following multiples: price to earnings, price to an estimate of operating cash flow, price to sales, and price to book value;

m. explain advantages and disadvantages of each category of valuation model.

The term **price multiple** refers to a ratio that compares the share price with some sort of monetary flow or value to allow evaluation of the relative worth of a company's stock. Some practitioners use price ratios as a screening mechanism. If the ratio falls below a specified value, the shares are identified as candidates for purchase, and if the ratio exceeds a specified value, the shares are identified as candidates for sale. Many practitioners use ratios when examining a group or sector of stocks and consider the shares for which the ratio is relatively low to be attractively valued securities.

Price multiples that are used by security analysts include the following:

- Price-to-earnings ratio (P/E). This measure is the ratio of the stock price to earnings per share. P/E is arguably the price multiple most frequently cited by the media and used by analysts and investors (Block 1999). The seminal works of McWilliams (1966), Miller and Widmann (1966), Nicholson (1968), Dreman (1977), and Basu (1977) presented evidence of a return advantage to low-P/E stocks.
- Price-to-book ratio (P/B). The ratio of the stock price to book value per share. Considerable evidence suggests that P/B multiples are inversely related to future rates of return (Fama and French 1995).
- Price-to-sales ratio (P/S). This measure is the ratio of stock price to sales per share. O'Shaughnessy (2005) provided evidence that a low P/S multiple is the most useful multiple for predicting future returns.
- Price-to-cash-flow ratio (P/CF). This measure is the ratio of stock price to some per-share measure of cash flow. The measures of cash flow include free cash flow (FCF) and operating cash flow (OCF).

A common criticism of all of these multiples is that they do not consider the future. This criticism is true if the multiple is calculated from trailing or current values of the divisor. Practitioners seek to counter this criticism by a variety of techniques, including forecasting fundamental values (the divisors) one or more years into the future. The resulting forward (leading or prospective) price multiples may differ markedly from the trailing price multiples. In the absence of an explicit forecast of fundamental values, the analyst is making an implicit forecast of the future when implementing such models. The choice of price multiple—trailing or forward—should be used consistently for companies being compared.

Besides the traditional price multiples used in valuation, just presented, analysts need to know how to calculate and interpret other ratios. Such ratios include those used to analyze business performance and financial condition based on data reported in financial statements. In addition, many industries have specialized measures of business performance that analysts covering those industries should be familiar with. In analyzing cable television companies, for example, the ratio of total market value of the company to the total number of subscribers is commonly used. Another common measure is revenue per subscriber. In the oil industry, a commonly cited ratio is proved reserves per common share. Industry-specific or sector-specific ratios such as these can be used to understand the key business variables in an industry or sector as well as to highlight attractively valued securities.

9.1 Relationships among Price Multiples, Present Value Models, and Fundamentals

Price multiples are frequently used independently of present value models. One price multiple valuation approach, the method of comparables, does not involve cash flow forecasts or discounting to present value. A price multiple is often related to fundamentals through a discounted cash flow model, however, such as the Gordon growth model. Understanding such connections can deepen the analyst's appreciation of the factors that affect the value of a multiple and often can help explain reasons for differences in multiples that do not involve mispricing. The expressions that are developed can be interpreted as the *justified value* of a multiple—that is, the value justified by (based on) fundamentals or a set of cash flow predictions. These expressions are an alternative way of presenting intrinsic-value estimates.

As an example, using the Gordon growth model identified previously in Equation 9 and assuming that price equals intrinsic value ($P_0 = V_0$), we can restate Equation 9 as follows:

$$P_0 = \frac{D_1}{r - g} \tag{9.1}$$

To arrive at the model for the justified forward P/E given in Equation 14, we divide both sides of Equation 9.1 by a forecast for next year's earnings, E_1. In Equation 14, the dividend payout ratio, p, is the ratio of dividends to earnings:

$$\frac{P_0}{E_1} = \frac{D_1/E_1}{r - g} = \frac{p}{r - g} \tag{14}$$

Equation 14 indicates that the P/E is inversely related to the required rate of return and positively related to the growth rate; that is, as the required rate of return increases, the P/E declines, and as the growth rate increases, the P/E increases. The P/E and the payout ratio appear to be positively related. This relationship may not be true, however, because a higher payout ratio may imply a slower growth rate as a result of the company retaining a lower proportion of earnings for reinvestment. This phenomenon is referred to as the dividend displacement of earnings.

EXAMPLE 11

Value Estimate Based on Fundamentals

Petroleo Brasileiro SA, commonly known as Petrobras, was once labeled "the most expensive oil company" by Bloomberg.com. Data for Petrobras and the oil industry, including the trailing twelve-month (TTM) P/E and payout ratios, appear below.

	Petrobras	Industry
P/E ratio (TTM)	39.61	13.0
Return on assets (TTM) (%)	3. 0	3.2
EPS 3-year growth rate (%)	NM	66.00
EPS (MRQ) vs. Qtr. 1 yr. ago (% change)	138.96	–12.0

Note: NM stands for non-quantifiable. Petrobras EPS has decreased from a loss of BRL 1.14 per share to a profit of BRL 0.16 per share in the most recent period. MRQ stands for "most recent quarter."
Source: Reuters.

Explain how the information shown supports a higher P/E for Petrobras than for the industry.

Solution:

The data support a higher P/E for Petrobras because its (MRQ) EPS growth rate exceed those of the industry. Equation 14 implies a positive relationship between the payout ratio and the P/E multiple. Petrobras has had a negative EPS for the period 2014 to 2017, and has paid no dividend during that period. A higher payout ratio supports a higher P/E. Furthermore, to the extent that higher EPS growth implies a high growth rate in dividends, the high EPS growth rate supports a high P/E. The higher P/E ratio is due to an improvement in the underlying financial performance of the company and the expected higher growth potential of the stock compared to the median firms in the industry.

EXAMPLE 12

Determining Justified Forward P/E

Heinrich Gladisch, CFA, is estimating the justified forward P/E for Nestlé, one of the world's leading nutrition and health companies. Gladisch notes that sales for 2017 were SFr89.78 billion (US$90.3 billion) and that net income was SFr7.18 billion (US$7.25 billion). He organizes the data for EPS, dividends per share, and the dividend payout ratio for the years 2013–2017 in the following table:

	2013	2014	2015	2016	2017
Earnings per share	SFr3.24	SFr4.54	SFr2.90	SFr2.76	SFr2.32
Year over year % change		44.6%	–36.1%	–4.8%	–15.9%
Dividend per share	SFr2.15	SFr2.2	SFr2.25	SFr2.3	SFr2.35
Year over year % change		2.3%	2.3%	2.2%	2.2%
Dividend payout ratio	68.5%	48.5%	77.6%	83.3%	99.2%

Gladisch calculates that ROE averaged 15.5 percent in the period 2013–2017 but was below that level at 11.7 percent in 2017. In that year, however, Nestlé's reported net income included a large nonrecurring component. The company reported 2017 "underlying earnings," which it defined as net income "from continuing operations before impairments, restructuring costs, results on disposals and significant one-off items," to be SFr2.93, giving an adjusted 14.8% ROE. Predicting increasing improvement in Nestlé's profit margins from growth in its product markets, Gladisch estimates a long-run ROE of 21.5 percent.

Gladisch decides that the dividend payout ratios of the 2013–2016 period—averaging 67.7 percent—are more representative of Nestlé's future payout ratio than is the high 2017 dividend payout ratio (when based on reported earnings). The dividend payout ratio in 2017 was higher because management apparently based the 2017 dividend on the components of net income that were expected to continue into the future. But basing a dividend on net income including non-recurring items creates the potential need to increase dividends in the future. Rounding up the 2013–2017 average, Gladisch settles on an estimate of 68 percent for the dividend payout ratio for use in calculating a justified forward P/E using Equation 14.

Gladisch's firm estimates that the required rate of return for Nestlé's shares is 9 percent per year. Gladisch also finds the following data at the opposite ends of the spectrum of external research analyst forecasts:

	2018E	2019E
Most optimistic analyst forecast:		
EPS	SFr3.99	SFr4.33
Year over year % change	71.9%	8.5%
P/E (based on a target price of SFr105)	26.3	24.2
Least optimistic analyst forecast:		
EPS	SFr3.52	SFr3.59
Year over year % change	51.7%	2.0%
P/E (based on a target price of SFr68)	19.3	18.9

1. Based only on information and estimates developed by Gladisch and his firm, estimate Nestlé's justified forward P/E.
2. Compare and contrast the justified forward P/E estimate from Question 1 to the estimates from each end of the spectrum of external research analysts forecasts.

Solution to 1:

The estimate of the justified forward P/E is 32.38. The dividend growth rate can be estimated by using Equation 10 as (1 – Dividend payout ratio) × ROE = (1 – 0.68) × 0.21.5 = 0.069, or 6.9 percent. Therefore,

$$\frac{P_0}{E_1} = \frac{\text{Payout}}{r - g} = \frac{0.68}{0.09 - 0.069} = 32.38$$

Solution to 2:

The estimated justified forward P/E of 32.38 is higher than the justified 2018 P/E estimates of 26.3 and 19.3 of the two analysts. Using a required rate of return of 9.5 percent rather than 9 percent results in a justified forward P/E estimate of 26.2 = 0.68/(0.095 – 0.069). Using an ROE of 16.5 percent (the average ROE of the 2013–2016 period) rather than 21.5 percent results in a justified forward P/E estimate of 18.4 = 0.68/[0.09 – (0.32)(0.165)] = 0.68/(0.09 – 0.053). The justified forward P/E is very sensitive to changes in the inputs.

Justified forward P/E estimates can be sensitive to small changes in assumptions. Therefore, analysts can benefit from carrying out a sensitivity analysis, as shown in Exhibit 6, which is based on Example 12. Exhibit 6 shows how the justified forward

P/E varies with changes in the estimates for the dividend payout ratio (columns) and return on equity. The dividend growth rate (rows) changes because of changes in the retention rate (1 – Payout rate) and ROE. Recall g = ROE times retention rate.

Exhibit 6 Estimates for Nestlé's Justified Forward P/E (Required Rate of Return = 9 Percent)

Constant Dividend Growth Rate (%)	Dividend Payout Ratio				
	55%	**60%**	**65%**	**70%**	**75%**
4.0	11.0	12.0	13.0	14.0	15.0
4.5	12.2	13.3	14.4	15.6	16.7
5.0	13.8	15.0	16.3	17.5	18.8
5.5	15.7	17.1	18.6	20.0	21.4
6.0	18.3	20.0	21.7	23.3	25.0
6.5	22.0	24.0	26.0	28.0	30.0
7.0	27.5	30.0	32.5	35.0	37.5
7.5	36.7	40.0	43.3	46.7	50.0

METHOD OF COMPARABLES AND VALUATION BASED ON PRICE MULTIPLES

10

i. explain the rationale for using price multiples to value equity, how the price to earnings multiple relates to fundamentals, and the use of multiples based on comparables;

j. calculate and interpret the following multiples: price to earnings, price to an estimate of operating cash flow, price to sales, and price to book value;

m. explain advantages and disadvantages of each category of valuation model.

The method of comparables is the most widely used approach for analysts reporting valuation judgments on the basis of price multiples. This method essentially compares relative values estimated using multiples or the relative values of multiples. The economic rationale underlying the method of comparables is the **law of one price**: Identical assets should sell for the same price. The methodology involves using a price multiple to evaluate whether an asset is fairly valued, undervalued, or overvalued in relation to a benchmark value of the multiple. Choices for the benchmark multiple include the multiple of a closely matched individual stock or the average or median value of the multiple for the stock's industry. Some analysts perform trend or time-series analyses and use past or average values of a price multiple as a benchmark.

Identifying individual companies or even an industry as the "comparable" may present a challenge. Many large corporations operate in several lines of business, so the scale and scope of their operations can vary significantly. When identifying comparables (sometimes referred to as "comps"), the analyst should be careful to identify companies that are most similar according to a number of dimensions. These dimensions include (but are not limited to) overall size, product lines, and growth rate. The type of analysis shown in Section 9 relating multiples to fundamentals is a productive way to identify the fundamental variables that should be taken into account in identifying comparables.

EXAMPLE 13

Method of Comparables (1)

As noted previously, P/E is a price multiple frequently used by analysts. Using P/E in the method of comparables can be problematic, however, as a result of business cycle effects on EPS. An alternative valuation tool that is useful during periods of economic slowdown or extraordinary growth is the P/S multiple. Although sales will decline during a recession and increase during a period of economic growth, the change in sales will be less than the change in earnings in percentage terms because earnings are heavily influenced by fixed operating and financing costs (operating and financial leverage).

The following data provide the P/S for most of the major automobile manufacturers as at December 2017:

Company	P/S
Peugeot	0.28
Ford Motor	0.33
General Motors	0.36
Nissan Motor	0.38
Honda Motor	0.46
Tata Motors	0.49
Daimler	0.55
BMW	0.57
Toyota Motor	0.80

Sources: Morningstar and company websites.

Based on the data presented, which stock appears to be undervalued when compared with the others?

Solution:

The P/S analysis suggests that Peugeot shares offer the best value. An analyst must be alert for a range of potential explanations of apparently low or high multiples when performing comparables analysis, rather than just assuming a relative mispricing.

EXAMPLE 14

Method of Comparables (2)

Incorporated in the Netherlands, Airbus is active in the aerospace and defense industry. It is a dominant aerospace company in Europe. Its largest business, Airbus Commercial Aircraft, is a manufacturing company with bases in several European countries and accounts for the majority of Airbus SE profits. Airbus and its primary competitor, Boeing, control most of the global commercial airplane industry.

Comparisons are frequently made between Airbus and Boeing. As noted in Exhibit 7, the companies are broadly similar in size as measured by total revenues. Converting total forecast revenues from euros to US dollars using the average exchange rate for 2017 of US\$1.13/€ results in a value of \$75.5 billion for Airbus's total revenues. Thus, total revenues for Boeing are expected to be about a fifth higher than those for Boeing.

The companies do differ, however, in several important areas. Airbus derives a greater share of its revenue from commercial aircraft production than does Boeing, and the order backlog for Airbus is much higher than that for Boeing. Converting the Airbus order backlog from euros to US dollars using the quarter-end rate for September 2017 of \$1.1813/€ results in a value of \$1.12 billion for Airbus's order backlog. Thus, the order backlog for Airbus is more than twice as high as the backlog for Boeing.[5]

Exhibit 7 Data for EADS and Boeing

	Airbus	Boeing
Total revenues (billions, 2017)	€66.8	\$92.2
Annual revenue growth (2015–2017 average)	1.8%	–2.1%
Percent of revenues from commercial aircraft	75%	69%
Order backlog (billions)	€945	\$474
Share price, 12/Dec/17	€86.96	\$283.73
EPS (basic)	€3.33	\$10.18
DPS	€1.48	\$5.7
Dividend payout ratio	44%	56%
P/E ratio	26.1	27.9

Note: 2017 forecast data and YTD average exchange rate as of 12 December 2017. Order backlog as of 30 September 2017.

Sources: Company websites: www.airbus.com and www.boeing.com, *Financial Times*.

What data shown in Exhibit 7 support a higher P/E for Boeing than for Airbus?

Solution:

Recall from Equation 14 and the discussion that followed it that P/E is directly related to the payout ratio and the dividend growth rate. The P/E is inversely related to the required rate of return. The only data presented in Exhibit 7 that support a higher P/E for Boeing is the company's higher dividend payout ratio (expected at 56 percent versus 44 percent for Airbus).

The following implicitly supports a higher P/E for Airbus: Airbus has higher revenue growth (as reported for 2016 and expected for 2017) and a higher backlog of orders, suggesting that it may have a higher future growth rate.

---

[5] Exchange rate data are available from FRED (Federal Reserve Economic Data) at http://research.stlouisfed.org/fred2/. Each company uses slightly different methodology for calculating order backlog.

EXAMPLE 15

Method of Comparables (3)

Canon Inc. is a leading worldwide manufacturer of business machines, cameras, and optical products. Canon was founded in 1937 as a camera manufacturer and is incorporated in Tokyo. The corporate philosophy of Canon is *kyosei* or "living and working together for the common good." The following data can be used to determine a P/E for Canon over the time period 2013–2017. Analyze the P/E of Canon over time and discuss the valuation of Canon.

Year	Price (a)	EPS (b)	P/E (a) ÷ (b)
2013	¥3,330	¥200.8	16.6
2014	¥3,840.5	¥229.0	16.8
2015	¥3,675	¥201.7	18.2
2016	¥3,295	¥138	23.9
2017	¥4,200	¥222.88	18.8

Sources: EPS, year-end prices, and P/E data are from Capital IQ and the *Financial Times.*

Solution:

Trend analysis of Canon's P/E reveals a peak of 23.9 at the end of 2016. The 2013 P/E of 16.6 is the lowest of the five years reported. This finding suggests that Canon's share price may be fairly price as of year-end 2017. A bearish case for Canon's stock can be made if an analyst believes that P/E will return to its historical low (16.6 over this five-year period) or be lower. Such a bearish prediction requires that a decrease in P/E not be offset by an increase in EPS. A bullish case can be made if the analyst believes the stock deserves re-rating and an even higher than trend P/E.

10.1 Illustration of a Valuation Based on Price Multiples

Telefónica S.A., a world leader in the telecommunication sector, provides communication, information, and entertainment products and services in Europe, Africa, and Latin America. It has operated in its home country of Spain since 1924, but as of 2017, more than 75 percent of its business was outside its home market.

Deutsche Telekom AG provides network access, communication services, and value-added services via fixed and mobile networks. It generates more than half of its revenues outside its home country, Germany.

Exhibit 8 provides comparable data for these two communication giants for 2015–2017.

Exhibit 8 Data for Telefónica and Deutsche Telekom

	Telefónica			Deutsche Telekom		
	2017	2016	2015	2017	2016	2015
(1) Total assets (€ billions)	115.0	123.6	120.3	141.3	148.5	143.9
Asset growth	−6.9%	2.7%	—	−4.9%	3.2%	—
(2) Net revenues (€ billions)	52.0	52.0	54.9	77.3	75.2	71.3

Exhibit 8 (Continued)

	Telefónica			Deutsche Telekom		
	2017	2016	2015	2017	2016	2015
Revenue growth	0%	−5.2%	—	2.8%	5.5%	—
(3) Net cash flow from operating activities (€ billions)	13.8	13.3	13.6	17.2	15.5	15.0
Cash flow growth	3.4%	−2.0%	—	11.0%	3.3%	—
(4) Book value of common shareholders' equity (€ billions)	16.9	18.2	15.8	42.5	38.8	38.2
Debt ratio: *1 − [(4) ÷ (1)]*	85.3%	85.3%	86.9%	70.0%	73.9%	73.5%
(5) Net profit (€ billions)	2.9	2.1	0.4	3.5	2.7	3.3
Earnings growth	38.1%	425.0%	—	29.6%	−18.2%	—
(6) Weighted average number of shares outstanding (millions)	5,122.9	4,896.6	4,833.6	4,740.2	4,654.9	4,584.8
(7) Price per share (€)	7.93	9.52	9.13	13.42	16.21	15.44
Price-to-revenue ratio (P/R): *(7) ÷ [(2) ÷ (6)]*	0.8	0.9	0.8	0.8	1.0	1.0
P/CF: *(7) ÷ [(3) ÷ (6)]*	2.9	3.5	3.2	3.7	4.9	4.7
P/B: *(7) ÷ [(4) ÷ (6)]*	2.4	2.6	2.8	1.5	1.9	1.9
P/E: *(7) ÷ [(5) ÷ (6)]*	14.0	22.2	110.3	18.2	27.9	21.5

Sources: Company websites: www.telefonica.es and www.deutschetelekom.com.

Time-series analysis of all price multiples in Exhibit 8 suggests that both companies are currently attractively valued. For example, the 2017 price-to-revenue ratio (P/R) of 0.78 for Telefónica is below the 2015–2017 average for this ratio of approximately 0.83. The 2017 P/CF of 3.7 for Deutsche Telekom is below the 2015–2017 average for this ratio of approximately 4.4.

A comparative analysis produces somewhat mixed results. The 2017 values for Deutsche Telekom for the P/R, P/CF, P/E multiples are higher than those for Telefónica. This result suggests that Telefónica is attractively valued when compared with Deutsche Telekom. The 2017 P/B for Telefónica, however, is higher than for Deutsche Telekom.

An analyst investigating these contradictory results would look for information not reported in Exhibit 8. For example, the earnings before interest, taxes, depreciation, and amortization (EBITDA) for Telefónica was €16.4 billion in 2017. The EBITDA value for Deutsche Telekom was €20.7 billion in 2017. The 2017 price-to-EBITDA ratio for Telefónica is [(7.93 × 5,123)/16,400] or [7.92/(16,400/5,123] = 2.5, whereas the 2017 price-to-EBITDA ratio for Deutsche Telekom is 3.1. Thus, the higher P/E for Deutsche Telekom can-not be explained by higher depreciation charges, higher interest costs, and/or a greater tax burden, but appears to be due to a better quality of earnings

In summary, the major advantage of using price multiples is that they allow for relative comparisons, both cross-sectional (versus the market or another comparable) and in time series. The approach can be especially beneficial for analysts who are assigned to a particular industry or sector and need to identify the expected best performing stocks within that sector. Price multiples are popular with investors because the multiples can be calculated easily and many multiples are readily available from financial websites and newspapers.

Caution is necessary. A stock may be relatively undervalued when compared with its benchmarks but overvalued when compared with an estimate of intrinsic value as determined by one of the discounted cash flow methodologies. Furthermore, differences in reporting rules among different markets and in chosen accounting methods can result in revenues, earnings, book values, and cash flows that are not easily comparable. These differences can, in turn, result in multiples that are not easily comparable. Finally, the multiples for cyclical companies may be highly influenced by current economic conditions.

11 ENTERPRISE VALUE

k describe enterprise value multiples and their use in estimating equity value;

m explain advantages and disadvantages of each category of valuation model.

An alternative to estimating the value of equity is to estimate the value of the enterprise. Enterprise value is most frequently determined as market capitalization plus market value of preferred stock plus market value of debt minus cash and investments (cash equivalents and short-term investments). Enterprise value is often viewed as the cost of a takeover: In the event of a buyout, the acquiring company assumes the acquired company's debt but also receives its cash. Enterprise value is most useful when comparing companies with significant differences in capital structure.

Enterprise value (EV) multiples are widely used in Europe, with EV/EBITDA arguably the most common. EBITDA is a proxy for operating cash flow because it excludes depreciation and amortization. EBITDA may include other non-cash expenses, however, and non-cash revenues. EBITDA can be viewed as a source of funds to pay interest, dividends, and taxes. Because EBITDA is calculated prior to payment to any of the company's financial stakeholders, using it to estimate enterprise value is logically appropriate.

Using enterprise value instead of market capitalization to determine a multiple can be useful to analysts. Even where the P/E is problematic because of negative earnings, the EV/EBITDA multiple can generally be computed because EBITDA is usually positive. An alternative to using EBITDA in EV multiples is to use operating income.

In practice, analysts may have difficulty accurately assessing enterprise value if they do not have access to market quotations for the company's debt. When current market quotations are not available, bond values may be estimated from current quotations for bonds with similar maturity, sector, and credit characteristics. Substituting the book value of debt for the market value of debt provides only a rough estimate of the debt's market value. This is because market interest rates change and investors' perception of the issuer's credit risk may have changed since the debt was issued.

EXAMPLE 16

Estimating the Market Value of Debt and Enterprise Value

Cameco Corporation is one of the world's largest uranium producers; it accounts for 16 percent of world production from its mines in Canada and the United States. Cameco estimates it has about 458 million kilograms of proven and probable reserves and holds premier land positions in the world's most promising areas for new uranium discoveries in Canada and Australia. Cameco is also a leading provider of processing services required to produce fuel for nuclear power plants. It generates 1,000 megawatts of electricity through a partnership in North America's largest nuclear generating station located in Ontario, Canada.

For simplicity of exposition in this example, we will present share counts in thousands and all dollar amounts in thousands of Canadian dollars. In 2017, Cameco had 395,793 shares outstanding. Its 2017 year-end share price was $14.11. Therefore, Cameco's 2017 year-end market capitalization was $5,584,640.

In its 2017 Annual Report (available at www.cameco.com), Cameco reported total debt and other liabilities of $2,919,100. The company presented the following schedule for long-term debt payments:

Year	Payment
2018	$69,000
2019 and 2020	610,000
2021 and 2022	482,000
Thereafter	744,000
Total	$1,905,000

Cameco's longest maturity debt matures in 2042. We will assume that the amounts paid in 2019 and 2020, and in 2021 and 2022, will be paid equally during the two years. The "thereafter" period includes two debenture tranches, the first one maturing in 2024 for a total value of $620,000 and the second tranche maturing in 2042 for the remaining $124,000. A yield curve for zero-coupon Canadian government securities was available from the Bank of Canada. The yield-curve data and assumed risk premiums in Exhibit 9 were used to estimate the market value of Cameco's long-term debt:

Exhibit 9 Estimated Market Value

Year	Yield on Zero-Coupon Government Security (%)	Assumed Risk Premium (%)	Discount Rate (%)	Book Value	Market Value
2018	0.89	0.50	1.39	$69,000	$68,054
2019	1.11	1.00	2.11	$305,000	$292,525
2020	1.39	1.50	2.89	$305,000	$280,014
2021	1.65	2.00	3.65	$241,000	$208,804
2022	1.88	2.50	4.38	$241,000	$194,505
2023	2.10	3.00	5.10	$0	$0
2024	2.30	3.50	5.80	$620,000	$417,823
2025	2.50	4.00	6.50	$0	$0
…	…	…	…	$…	$…
2042	2.92	5.00	7.92	$124,000	$18,445
				$1,905,000	$1,480,170

Note from Exhibit 9 that the book value of long-term debt is \$1,905,000 and its estimated market value is \$1,480,170. The book value of total debt and liabilities of \$2,919,100 minus the book value of long-term debt of \$1,905,000 is \$1,014,100. If we assume that the market value of that remaining debt is equal to its book value of \$1,014,100 an estimate of the market value of total debt and liabilities is that amount plus the estimated market value of long-term debt of \$1,480,170 or \$2,494,270.

At the end of 2017, Cameco had cash and equivalents of \$591,600. Enterprise value can be estimated as the \$5,584,640 market value of stock plus the \$2,494,270 market value of debt minus the \$591,600 cash and equivalents, or \$7,487,310. Cameco's 2017 EBITDA was \$606,000; an estimate of EV/EBITDA is, therefore, \$7,487,310 divided by \$606,000, or 12.4.

EXAMPLE 17

EV/Operating Income

Exhibit 10 presents data for twelve major mining companies. Based only on the information in Exhibit 10, which two mining companies seem to be the *most* undervalued?

Exhibit 10 Data for Twelve Major Mining Companies

Company	EV (US\$ millions)	Operating Income (OI) (US\$ millions)	EV/OI
BHP Billiton	119,712.3	11,753	10.19
Rio Tinto	93,856.1	6,471	14.5
Vale	82,051.2	6,366	12.89
Glencore	80,772.0	–549	–147.13
Southern Copper	37,817.0	1,564	24.18
Freeport-McMoRan	33,452.0	–2,766	–12.09
Anglo American	32,870.3	2,562	12.83
Norilsk Nickel	22,483.0	3,377	6.66
Coal India	21,652.1	1,382	15.67
Barrick Gold	21,549.8	2,424	8.89
Newmont Mining	20,683.0	–65	–318.20
Goldcorp	12,986.7	369	35.19

Source: www.miningfeeds.com, Morningstar.

Solution:

Norilsk Nickel and Barrick Gold have the lowest EV/OI and thus appear to be the *most* undervalued or favorably priced on the basis of the EV/OI. Note the negative ratio for Glencore, Freeport-McMoRan, and Newmont Mining. Negative ratios are difficult to interpret, so other means are used to evaluate companies with negative ratios.

ASSET-BASED VALUATION

12

l. describe asset-based valuation models and their use in estimating equity value;

m. explain advantages and disadvantages of each category of valuation model.

An asset-based valuation of a company uses estimates of the market or fair value of the company's assets and liabilities. Thus, asset-based valuations work well for companies that do not have a high proportion of intangible or "off the books" assets and that do have a high proportion of current assets and current liabilities. The analyst may be able to value these companies' assets and liabilities in a reasonable fashion by starting with balance sheet items. For most companies, however, balance sheet values are different from market (fair) values, and the market (fair) values can be difficult to determine.

Asset-based valuation models are frequently used together with multiplier models to value private companies. As public companies increase reporting or disclosure of fair values, asset-based valuation may be increasingly used to supplement present value and multiplier models of valuation. Important facts that the practitioner should realize are as follows:

- Companies with assets that do not have easily determinable market (fair) values—such as those with significant property, plant, and equipment—are very difficult to analyze using asset valuation methods.
- Asset and liability fair values can be very different from the values at which they are carried on the balance sheet of a company.
- Some assets that are "intangible" are shown on the books of the company. Other intangible assets, such as the value from synergies or the value of a good business reputation, may not be shown on the books. Because asset-based valuation may not consider some intangibles, it can give a "floor" value for a situation involving a significant amount of intangibles. When a company has significant intangibles, the analyst should prefer a forward-looking cash flow valuation.
- Asset values may be more difficult to estimate in a hyper-inflationary environment.

We begin by discussing asset-based valuation for hypothetical nonpublic companies and then move on to a public company example. Analysts should consider the difficulties and rewards of using asset-based valuation for companies that are suited to this measure. Owners of small privately held businesses are familiar with valuations arrived at by valuing the assets of the company and then subtracting any relevant liabilities.

EXAMPLE 18

An Asset-Based Valuation of a Family-Owned Laundry

A family owns a laundry and the real estate on which the laundry stands. The real estate is collateral for an outstanding loan of $100,000. How can asset-based valuation be used to value this business?

Solution:

The analyst should get at least two market appraisals for the real estate (building and land) and estimate the cost to extinguish the $100,000 loan. This information would provide estimated values for everything except the laundry as a going concern. That is, the analyst has market values for the building and land and the loan but needs to value the laundry business. The analyst can value the assets of the laundry: the equipment and inventory. The equipment can be valued at depreciated value, inflation-adjusted depreciated value, or replacement cost. Replacement cost in this case means the amount that would have to be spent to buy equivalent used machines. This amount is the market value of the used machines. The analyst will recognize that any intangible value of the laundry (prime location, clever marketing, etc.) is being excluded, which will result in an inaccurate asset-based valuation.

Example 18 shows some of the subtleties present in applying asset-based valuation to determine company value. It also shows how asset-based valuation does not deal with intangibles. Example 19 emphasizes this point.

EXAMPLE 19

An Asset-Based Valuation of a Restaurant

The business being valued is a restaurant that serves breakfast and lunch. The owner/proprietor wants to sell the business and retire. The restaurant space is rented, not owned. This particular restaurant is hugely popular because of the proprietor's cooking skills and secret recipes. How can the analyst value this business?

Solution:

Because of the intangibles, setting a value on this business is challenging. A multiple of income or revenue might be considered. But even those approaches overlook the fact that the proprietor may not be selling his secret recipes and, furthermore, does not intend to continue cooking. Some (or all) of the intangible assets may vanish when the business is sold. Asset-based valuation for this restaurant would begin with estimating the value of the restaurant equipment and inventory and subtracting the value of any liabilities. This approach will provide only a good baseline, however, for a minimum valuation.

For public companies, the assets will typically be so extensive that a piece-by-piece analysis will be impossible, and the transition from book value to market value is a nontrivial task. The asset-based valuation approach is most applicable when the market value of the corporate assets is readily determinable and the intangible assets, which are typically difficult to value, are a relatively small proportion of corporate assets. Asset-based valuation has also been applied to financial companies, natural resource companies, and formerly going-concerns that are being liquidated. Even for other types of companies, however, asset-based valuation of tangible assets may provide a baseline for a minimal valuation.

EXAMPLE 20

An Asset-Based Valuation of an Airline

Consider the value of an airline company that has few routes, high labor and other operating costs, has stopped paying dividends, and is losing millions of dollars each year. Using most valuation approaches, the company will have a negative value. Why might an asset-based valuation approach be appropriate for use by one of the company's competitors that is considering acquisition of this airline?

Solution:

The airline's routes, landing rights, leases of airport facilities, and ground equipment and airplanes may have substantial value to a competitor. An asset-based approach to valuing this company would value the company's assets separately and aside from the money-losing business in which they are presently being utilized.

Analysts recognizing the uncertainties related to model appropriateness and the inputs to the models frequently use more than one model or type of model in valuation to increase their confidence in their estimates of intrinsic value. The choice of models will depend on the availability of information to put into the models. Example 21 illustrates the use of three valuation methods.

EXAMPLE 21

A Simple Example of the Use of Three Major Equity Valuation Models

Company data for dividend per share (DPS), earnings per share (EPS), share price, and price-to-earnings ratio (P/E) for the most recent five years are presented in Exhibit 11. In addition, estimates (indicated by an "E" after the amount) of DPS and EPS for the next five years are shown. The valuation date is at the end of Year 5. The company has 1,000 shares outstanding.

Exhibit 11 Company DPS, EPS, Share Price, and P/E Data

Year	DPS	EPS	Share Price	TTM P/E
10	$3.10E	$5.20E	—	—
9	$2.91E	$4.85E	—	—
8	$2.79E	$4.65E	—	—
7	$2.65E	$4.37E	—	—
6	$2.55E	$4.30E	—	—
5	$2.43	$4.00	$50.80	12.7
4	$2.32	$3.90	$51.48	13.2
3	$2.19	$3.65	$59.86	16.4
2	$2.14	$3.60	$54.72	15.2
1	$2.00	$3.30	$46.20	14.0

The company's balance sheet at the end of Year 5 is given in Exhibit 12.

Exhibit 12 Balance Sheet as of End of Year 5

Cash	$ 5,000
Accounts receivable	15,000
Inventories	30,000
Net fixed assets	50,000
Total assets	$100,000

(continued)

Exhibit 12 (Continued)

Accounts payable	$ 3,000
Notes payable	17,000
Term loans	25,000
Common shareholders' equity	55,000
Total liabilities and equity	$100,000

1 Using a Gordon growth model, estimate intrinsic value. Use a discount rate of 10 percent and an estimate of growth based on growth in dividends over the next five years.

2 Using a multiplier approach, estimate intrinsic value. Assume that a reasonable estimate of P/E is the average trailing twelve-month (TTM) P/E ratio over Years 1 through 4.

3 Using an asset-based valuation approach, estimate value per share from adjusted book values. Assume that the market values of accounts receivable and inventories are as reported, the market value of net fixed assets is 110 percent of reported book value, and the reported book values of liabilities reflect their market values.

Solution to 1:

$D_5 (1 + g)^5 = D_{10}2.43(1 + g)^5 = 3.10$

$g \approx 5.0\%$

Estimate of value = $V_5 = 2.55/(0.10 - 0.05) = \$51.00$

Solution to 2:

Average P/E = (14.0 + 15.2 + 16.4 + 13.2)/4 = 14.7

Estimate of value = $4.00 × 14.7 = $58.80

Solution to 3:

Market value of assets = 5,000 + 15,000 + 30,000 + 1.1(50,000) = $105,000

Market value of liabilities = $3,000 + 17,000 + 25,000 = $45,000

Adjusted book value = $105,000 – 45,000 = $60,000

Estimated value (adjusted book value per share) = $60,000 ÷ 1,000 shares = $60.00

Given the current share price of $50.80, the multiplier and the asset-based valuation approaches indicate that the stock is undervalued. Given the intrinsic value estimated using the Gordon growth model, the analyst is likely to conclude that the stock is fairly priced. The analyst might examine the assumptions in the multiplier and the asset-based valuation approaches to determine why their estimated values differ from the estimated value provided by the Gordon growth model and the market price.

SUMMARY

The equity valuation models used to estimate intrinsic value—present value models, multiplier models, and asset-based valuation—are widely used and serve an important purpose. The valuation models presented here are a foundation on which to base analysis and research but must be applied wisely. Valuation is not simply a numerical analysis. The choice of model and the derivation of inputs require skill and judgment.

When valuing a company or group of companies, the analyst wants to choose a valuation model that is appropriate for the information available to be used as inputs. The available data will, in most instances, restrict the choice of model and influence the way it is used. Complex models exist that may improve on the simple valuation models described in this reading; but before using those models and assuming that complexity increases accuracy, the analyst would do well to consider the "law of parsimony:" A model should be kept as simple as possible in light of the available inputs. Valuation is a fallible discipline, and any method will result in an inaccurate forecast at some time. The goal is to minimize the inaccuracy of the forecast.

Among the points made in this reading are the following:

- An analyst estimating intrinsic value is implicitly questioning the market's estimate of value.
- If the estimated value exceeds the market price, the analyst infers the security is *undervalued*. If the estimated value equals the market price, the analyst infers the security is *fairly valued*. If the estimated value is less than the market price, the analyst infers the security is *overvalued*. Because of the uncertainties involved in valuation, an analyst may require that value estimates differ markedly from market price before concluding that a misvaluation exists.
- Analysts often use more than one valuation model because of concerns about the applicability of any particular model and the variability in estimates that result from changes in inputs.
- Three major categories of equity valuation models are present value, multiplier, and asset-based valuation models.
- Present value models estimate value as the present value of expected future benefits.
- Multiplier models estimate intrinsic value based on a multiple of some fundamental variable.
- Asset-based valuation models estimate value based on the estimated value of assets and liabilities.
- The choice of model will depend upon the availability of information to input into the model and the analyst's confidence in both the information and the appropriateness of the model.
- Companies distribute cash to shareholders using dividend payments and share repurchases.
- Regular cash dividends are a key input to dividend valuation models.
- Key dates in dividend chronology are the declaration date, ex-dividend date, holder-of-record date, and payment date.
- In the dividend discount model, value is estimated as the present value of expected future dividends.
- In the free cash flow to equity model, value is estimated as the present value of expected future free cash flow to equity.
- The Gordon growth model, a simple DDM, estimates value as $D_1/(r - g)$.

- The two stage dividend discount model estimates value as the sum of the present values of dividends over a short-term period of high growth and the present value of the terminal value at the end of the period of high growth. The terminal value is estimated using the Gordon growth model.
- The choice of dividend model is based upon the patterns assumed with respect to future dividends.
- Multiplier models typically use multiples of the form: P/ measure of fundamental variable or EV/ measure of fundamental variable.
- Multiples can be based upon fundamentals or comparables.
- Asset-based valuations models estimate value of equity as the value of the assets less the value of liabilities.

REFERENCES

Basu, S. 1977. "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis." *Journal of Finance*, vol. 32, no. 3:663–682.

Block, S. 1999. "A Study of Financial Analysts: Practice and Theory." *Financial Analysts Journal*, vol. 55, no. 4:86–95.

Dreman, D. 1977. *Psychology of the Stock Market*. New York: AMACOM.

Fama, E., and K. French. 1995. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Finance*, vol. 50, no. 1:131–155.

McWilliams, J. 1966. "Prices, Earnings and P-E Ratios." *Financial Analysts Journal*, vol. 22, no. 3:137.

Miller, P., and E. Widmann. 1966. "Price Performance Outlook for High & Low P/E Stocks." *1966 Stock & Bond Issue, Commercial & Financial Chronicle*: 26–28.

Nicholson, S. 1968. "Price Ratios in Relation to Investment Results." *Financial Analysts Journal*, vol. 24, no. 1:105–109.

O'Shaughnessy, J. 2005. *What Works on Wall Street*. New York: McGraw-Hill.

PRACTICE PROBLEMS

1 An analyst estimates the intrinsic value of a stock to be in the range of €17.85 to €21.45. The current market price of the stock is €24.35. This stock is *most likely*:

A overvalued.

B undervalued.

C fairly valued.

2 An analyst determines the intrinsic value of an equity security to be equal to $55. If the current price is $47, the equity is *most likely*:

A undervalued.

B fairly valued.

C overvalued.

3 In asset-based valuation models, the intrinsic value of a common share of stock is based on the:

A estimated market value of the company's assets.

B estimated market value of the company's assets plus liabilities.

C estimated market value of the company's assets minus liabilities.

4 Which of the following is *most likely* used in a present value model?

A Enterprise value.

B Price to free cash flow.

C Free cash flow to equity.

5 Book value is *least likely* to be considered when using:

A a multiplier model.

B an asset-based valuation model.

C a present value model.

6 An analyst is attempting to calculate the intrinsic value of a company and has gathered the following company data: EBITDA, total market value, and market value of cash and short-term investments, liabilities, and preferred shares. The analyst is *least likely* to use:

A a multiplier model.

B a discounted cash flow model.

C an asset-based valuation model.

7 An analyst who bases the calculation of intrinsic value on dividend-paying capacity rather than expected dividends will *most likely* use the:

A dividend discount model.

B free cash flow to equity model.

C cash flow from operations model.

8 An investor expects to purchase shares of common stock today and sell them after two years. The investor has estimated dividends for the next two years, D_1 and D_2, and the selling price of the stock two years from now, P_2. According to the dividend discount model, the intrinsic value of the stock today is the present value of:

A next year's dividend, D_1.

© 2019 CFA Institute. All rights reserved.

- **B** future expected dividends, D_1 and D_2.
- **C** future expected dividends and price—D_1, D_2 and P_2.

9 In the free cash flow to equity (FCFE) model, the intrinsic value of a share of stock is calculated as:

- **A** the present value of future expected FCFE.
- **B** the present value of future expected FCFE plus net borrowing.
- **C** the present value of future expected FCFE minus fixed capital investment.

10 With respect to present value models, which of the following statements is *most accurate*?

- **A** Present value models can be used only if a stock pays a dividend.
- **B** Present value models can be used only if a stock pays a dividend or is expected to pay a dividend.
- **C** Present value models can be used for stocks that currently pay a dividend, are expected to pay a dividend, or are not expected to pay a dividend.

11 A Canadian life insurance company has an issue of 4.80 percent, $25 par value, perpetual, non-convertible, non-callable preferred shares outstanding. The required rate of return on similar issues is 4.49 percent. The intrinsic value of a preferred share is *closest to*:

- **A** $25.00.
- **B** $26.75.
- **C** $28.50.

12 Two analysts estimating the value of a non-convertible, non-callable, perpetual preferred stock with a constant dividend arrive at different estimated values. The *most likely* reason for the difference is that the analysts used different:

- **A** time horizons.
- **B** required rates of return.
- **C** estimated dividend growth rates.

13 The Beasley Corporation has just paid a dividend of $1.75 per share. If the required rate of return is 12.3 percent per year and dividends are expected to grow indefinitely at a constant rate of 9.2 percent per year, the intrinsic value of Beasley Corporation stock is *closest* to:

- **A** $15.54.
- **B** $56.45.
- **C** $61.65.

14 An investor is considering the purchase of a common stock with a $2.00 annual dividend. The dividend is expected to grow at a rate of 4 percent annually. If the investor's required rate of return is 7 percent, the intrinsic value of the stock is *closest* to:

- **A** $50.00.
- **B** $66.67.
- **C** $69.33.

15 An analyst gathers or estimates the following information about a stock:

Current price per share	€22.56
Current annual dividend per share	€1.60
Annual dividend growth rate for Years 1–4	9.00%
Annual dividend growth rate for Years 5+	4.00%
Required rate of return	12%

Based on a dividend discount model, the stock is *most likely*:

A undervalued.

B fairly valued.

C overvalued.

16 An analyst is attempting to value shares of the Dominion Company. The company has just paid a dividend of \$0.58 per share. Dividends are expected to grow by 20 percent next year and 15 percent the year after that. From the third year onward, dividends are expected to grow at 5.6 percent per year indefinitely. If the required rate of return is 8.3 percent, the intrinsic value of the stock is *closest* to:

A \$26.00.

B \$27.00.

C \$28.00.

17 Hideki Corporation has just paid a dividend of ¥450 per share. Annual dividends are expected to grow at the rate of 4 percent per year over the next four years. At the end of four years, shares of Hideki Corporation are expected to sell for ¥9000. If the required rate of return is 12 percent, the intrinsic value of a share of Hideki Corporation is *closest* to:

A ¥5,850.

B ¥7,220.

C ¥7,670.

18 The Gordon growth model can be used to value dividend-paying companies that are:

A expected to grow very fast.

B in a mature phase of growth.

C very sensitive to the business cycle.

19 The best model to use when valuing a young dividend-paying company that is just entering the growth phase is *most likely* the:

A Gordon growth model.

B two-stage dividend discount model.

C three-stage dividend discount model.

20 An equity analyst has been asked to estimate the intrinsic value of the common stock of Omega Corporation, a leading manufacturer of automobile seats. Omega is in a mature industry, and both its earnings and dividends are expected to grow at a rate of 3 percent annually. Which of the following is *most likely* to be the best model for determining the intrinsic value of an Omega share?

A Gordon growth model.

B Free cash flow to equity model.

C Multistage dividend discount model.

21 A price earnings ratio that is derived from the Gordon growth model is inversely related to the:

A growth rate.

B dividend payout ratio.

C required rate of return.

22 The primary difference between P/E multiples based on comparables and P/E multiples based on fundamentals is that fundamentals-based P/Es take into account:

A future expectations.

B the law of one price.

C historical information.

23 An analyst makes the following statement: "Use of P/E and other multiples for analysis is not effective because the multiples are based on historical data and because not all companies have positive accounting earnings." The analyst's statement is *most likely*:

A inaccurate with respect to both historical data and earnings.

B accurate with respect to historical data and inaccurate with respect to earnings.

C inaccurate with respect to historical data and accurate with respect to earnings.

24 An analyst has prepared a table of the average trailing twelve-month price-to-earning (P/E), price-to-cash flow (P/CF), and price-to-sales (P/S) for the Tanaka Corporation for the years 2014 to 2017.

Year	P/E	P/CF	P/S
2014	4.9	5.4	1.2
2015	6.1	8.6	1.5
2016	8.3	7.3	1.9
2017	9.2	7.9	2.3

As of the date of the valuation in 2018, the trailing twelve-month P/E, P/CF, and P/S are, respectively, 9.2, 8.0, and 2.5. Based on the information provided, the analyst may reasonably conclude that Tanaka shares are *most likely*:

A overvalued.

B undervalued.

C fairly valued.

25 An analyst has gathered the following information for the Oudin Corporation:

Expected earnings per share = €5.70
Expected dividends per share = €2.70
Dividends are expected to grow at 2.75 percent per year indefinitely
The required rate of return is 8.35 percent

Based on the information provided, the price/earnings multiple for Oudin is *closest* to:

A 5.7.

B 8.5.

C 9.4.

26 An analyst gathers the following information about two companies:

	Alpha Corp.	Delta Co.
Current price per share	\$57.32	\$18.93
Last year's EPS	\$3.82	\$1.35
Current year's estimated EPS	\$4.75	\$1.40

Which of the following statements is *most accurate*?

A Delta has the higher trailing P/E multiple and lower current estimated P/E multiple.

B Alpha has the higher trailing P/E multiple and lower current estimated P/E multiple.

C Alpha has the higher trailing P/E multiple and higher current estimated P/E multiple.

27 An analyst gathers the following information about similar companies in the banking sector:

	First Bank	Prime Bank	Pioneer Trust
P/B	1.10	0.60	0.60
P/E	8.40	11.10	8.30

Which of the companies is *most likely* to be undervalued?

A First Bank.

B Prime Bank.

C Pioneer Trust.

28 The market value of equity for a company can be calculated as enterprise value:

A minus market value of debt, preferred stock, and short-term investments.

B plus market value of debt and preferred stock minus short-term investments.

C minus market value of debt and preferred stock plus short-term investments.

29 Which of the following statements regarding the calculation of the enterprise value multiple is *most likely* correct?

A Operating income may be used instead of EBITDA.

B EBITDA may not be used if company earnings are negative.

C Book value of debt may be used instead of market value of debt.

30 An analyst has determined that the appropriate EV/EBITDA for Rainbow Company is 10.2. The analyst has also collected the following forecasted information for Rainbow Company:

EBITDA = \$22,000,000

Market value of debt = \$56,000,000

Cash = \$1,500,000

The value of equity for Rainbow Company is *closest* to:

A \$169 million.

B \$224 million.

C \$281 million.

31 Enterprise value is most often determined as market capitalization of common equity and preferred stock minus the value of cash equivalents plus the:

A book value of debt.

B market value of debt.

C market value of long-term debt.

32 Asset-based valuation models are best suited to companies where the capital structure does not have a high proportion of:

A debt.

B intangible assets.

C current assets and liabilities.

33 Which of the following is *most likely* a reason for using asset-based valuation?

A The analyst is valuing a privately held company.

B The company has a relatively high level of intangible assets.

C The market values of assets and liabilities are different from the balance sheet values.

34 A disadvantage of the EV method for valuing equity is that the following information may be difficult to obtain:

A Operating income.

B Market value of debt.

C Market value of equity.

35 Which type of equity valuation model is *most likely* to be preferable when one is comparing similar companies?

A A multiplier model.

B A present value model.

C An asset-based valuation model.

36 Which of the following is *most likely* considered a weakness of present value models?

A Present value models cannot be used for companies that do not pay dividends.

B Small changes in model assumptions and inputs can result in large changes in the computed intrinsic value of the security.

C The value of the security depends on the investor's holding period; thus, comparing valuations of different companies for different investors is difficult.

SOLUTIONS

1 A is correct. The current market price of the stock exceeds the upper bound of the analyst's estimate of the intrinsic value of the stock.

2 A is correct. The market price is less than the estimated intrinsic, or fundamental, value.

3 C is correct. Asset-based valuation models calculate the intrinsic value of equity by subtracting liabilities from the market value of assets.

4 C is correct. FCFE can be used in a form of present value, or discounted cash flow, model. Both EV and price to free cash flow are forms of multiplier models.

5 C is correct. Multiplier valuation models (in the form of P/B) and asset-based valuation models (in the form of adjustments to book value) use book value, whereas present value models typically discount future expected cash flows.

6 B is correct. To use a discounted cash flow model, the analyst will require FCFE or dividend data. In addition, the analyst will need data to calculate an appropriate discount rate.

7 B is correct. The FCFE model assumes that dividend-paying capacity is reflected in FCFE.

8 C is correct. According to the dividend discount model, the intrinsic value of a stock today is the present value of all future dividends. In this case, the intrinsic value is the present value of D_1, D_2, and P_2. Note that P_2 is the present value at Period 2 of all future dividends from Period 3 to infinity.

9 A is correct. In the FCFE model, the intrinsic value of stock is calculated by discounting expected future FCFE to present value. No further adjustments are required.

10 C is correct. Dividend discount models can be used for a stock that pays a current dividend or a stock that is expected to pay a dividend. FCFE can be used for both of those stocks and for stocks that do not, or are not expected to, pay dividends in the near future. Both of these models are forms of present value models.

11 B is correct. The expected annual dividend is 4.80% × \$25 = \$1.20. The value of a preferred share is \$1.20/0.0449 = \$26.73.

12 B is correct. The required rate of return, r, can vary widely depending on the inputs and is not unique. A preferred stock with a constant dividend would not have a growth rate to estimate, and the investor's time horizon would have no effect on the calculation of intrinsic value.

13 C is correct. $P_0 = D_1/(r - g) = 1.75(1.092)/(0.123 - 0.092) = \$61.65$.

14 C is correct. According to the Gordon growth model, $V_0 = D_1/(r - g)$. In this case, $D_1 = \$2.00 \times 1.04 = \$2.08$, so $V_0 = \$2.08/(0.07 - 0.04) = \$69.3333 = \$69.33$.

15 A is correct. The current price of €22.56 is less than the intrinsic value (V_0) of €24.64; therefore, the stock appears to be currently undervalued. According to the two-stage dividend discount model:

$$V_0 = \sum_{t=1}^{n} \frac{D_0\left(1 + g_S\right)^t}{\left(1 + r\right)^t} + \frac{V_n}{\left(1 + r\right)^n} \text{ and } V_n = \frac{D_{n+1}}{r - g_L}$$

$$D_{n+1} = D_0(1 + g_S)^n(1 + g_L)$$

$$D_1 = €1.60 \times 1.09 = €1.744$$

D_2 = €1.60 × $(1.09)^2$ = €1.901
D_3 = €1.60 × $(1.09)^3$ = €2.072
D_4 = €1.60 × $(1.09)^4$ = €2.259
D_5 = [€1.60 × $(1.09)^4$](1.04) = €2.349
V_4 = €2.349/(0.12 – 0.04) = €29.363

$$V_0 = \frac{1.744}{(1.12)^1} + \frac{1.901}{(1.12)^2} + \frac{2.072}{(1.12)^3} + \frac{2.259}{(1.12)^4} + \frac{29.363}{(1.12)^4}$$

= 1.557 + 1.515 + 1.475 + 1.436 + 18.661
= €24.64 (which is greater than the current price of €22.56)

16 C is correct.

$$\begin{aligned}V_0 &= \frac{D_1}{(1+r)} + \frac{D_2}{(1+r)^2} + \frac{P_2}{(1+r)^2}\\ &= \frac{0.70}{(1.083)} + \frac{0.80}{(1.083)^2} + \frac{31.29}{(1.083)^2}\\ &= \$28.01\end{aligned}$$

Note that D_1 = 0.58(1.20) = 0.70, D_2 = 0.58(1.20)(1.15) = 0.80, and $P_2 = D_3/(k - g)$ = 0.80(1.056)/(0.083 – 0.056) = 31.29

17 B is correct.

$$\begin{aligned}V_0 &= \frac{D_1}{(1+r)} + \frac{D_2}{(1+r)^2} + \frac{D_3}{(1+r)^3} + \frac{D_4}{(1+r)^4} + \frac{P_4}{(1+r)^4}\\ &= \frac{468}{(1.12)} + \frac{486.72}{(1.12)^2} + \frac{506.19}{(1.12)^3} + \frac{526.44}{(1.12)^4} + \frac{9000}{(1.12)^4}\\ &= ¥7{,}220\end{aligned}$$

18 B is correct. The Gordon growth model (also known as the constant growth model) can be used to value dividend-paying companies in a mature phase of growth. A stable dividend growth rate is often a plausible assumption for such companies.

19 C is correct. The Gordon growth model is best suited to valuing mature companies. The two-stage model is best for companies that are transitioning from a growth stage to a mature stage. The three-stage model is appropriate for young companies just entering the growth phase.

20 A is correct. The company is a mature company with a steadily growing dividend rate. The two-stage (or multistage) model is unnecessary because the dividend growth rate is expected to remain stable. Although an FCFE model could be used, that model is more often chosen for companies that currently pay no dividends.

21 C is correct. The justified forward P/E is calculated as follows:

$$\frac{P_0}{E_1} = \frac{\dfrac{D_1}{E_1}}{r - g}$$

P/E is inversely related to the required rate of return, *r*, and directly related to the growth rate, *g*, and the dividend payout ratio, *D*/*E*.

22 A is correct. Multiples based on comparables are grounded in the law of one price and take into account historical multiple values. In contrast, P/E multiples based on fundamentals can be based on the Gordon growth model, which takes into account future expected dividends.

23 A is correct. The statement is inaccurate in both respects. Although multiples can be calculated from historical data, forecasted values can be used as well. For companies without accounting earnings, several other multiples can be used. These multiples are often specific to a company's industry or sector and include price-to-sales and price-to-cash flow.

24 A is correct. Tanaka shares are most likely overvalued. As the table below shows, all the 2018 multiples are currently above their 2014–2017 averages.

Year	P/E	P/CF	P/R
2014	4.9	5.4	1.2
2015	6.1	8.6	1.5
2016	8.3	7.3	1.9
2017	9.2	7.9	2.3
Average	7.1	7.3	1.7

25 B is correct.

$$\frac{P_0}{E_1} = \frac{\frac{D_1}{E_1}}{r - g} = \frac{\frac{2.7}{5.7}}{0.0835 - 0.0275} = 8.5$$

26 B is correct. P/E = Current price/EPS, and Estimated P/E = Current price/ Estimated EPS.

Alpha P/E = \$57.32/\$3.82 = 15.01

Alpha estimated P/E = \$57.32/4.75 = 12.07

Delta P/E = \$18.93/\$1.35 = 14.02

Delta estimated P/E = \$18.93/\$1.40 = 13.52

27 C is correct. Relative to the others, Pioneer Trust has the lowest P/E multiple and the P/B multiple is tied for the lowest with Prime Bank. Given the law of one price, similar companies should trade at similar P/B and P/E levels. Thus, based on the information presented, Pioneer is most likely to be undervalued.

28 C is correct. Enterprise value is calculated as the market value of equity plus the market value of debt and preferred stock minus short-term investments. Therefore, the market value of equity is enterprise value minus the market value of debt and preferred stock plus short-term investments.

29 A is correct. Operating income may be used in place of EBITDA when calculating the enterprise value multiple. EBITDA may be used when company earnings are negative because EBITDA is usually positive. The book value of debt cannot be used in place of market value of debt.

30 A is correct.

$$\text{EV} = 10.2 \times 22{,}000{,}000 = \$224{,}400{,}000$$

$$\begin{aligned}\text{Equity value} &= \text{EV} - \text{Debt} + \text{Cash}\\ &= 224{,}400{,}000 - 56{,}000{,}000 + 1{,}500{,}000\\ &= \$169{,}900{,}000\end{aligned}$$

31 B is correct. The market value of debt must be calculated and taken out of the enterprise value. Enterprise value, sometimes known as the cost of a takeover, is the cost of the purchase of the company, which would include the assumption of the company's debts at market value.

32 B is correct. Intangible assets are hard to value. Therefore, asset-based valuation models work best for companies that do not have a high proportion of intangible assets.

33 A is correct. Asset-based valuations are most often used when an analyst is valuing private enterprises. Both B and C are considerations in asset-based valuations but are more likely to be reasons to avoid that valuation model rather than reasons to use it.

34 B is correct. According to the reading, analysts may have not have access to market quotations for company debt.

35 A is correct. Although all models can be used to compare various companies, multiplier models have the advantage of reducing varying fundamental data points into a format that allows direct comparisons. As long as the analyst applies the data in a consistent manner for all the companies, this approach provides useful comparative data.

36 B is correct. Very small changes in inputs, such as required rate of return or dividend growth rate, can result in large changes to the valuation model output. Some present value models, such as FCFE models, can be used to value companies without dividends. Also, the intrinsic value of a security is independent of the investor's holding period.

Fixed Income

STUDY SESSIONS

Study Session 13	Fixed Income (1)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to describe fixed-income securities and their markets, yield measures, risk factors, and valuation measures and drivers. The candidate should also be able to calculate yields and values of fixed-income securities.

Fixed-income securities continue to represent the largest capital market segment in the financial ecosystem and the primary means in which institutions, governments, and other issuers raise capital globally. Institutions and individuals use fixed-income investments in a wide range of applications including asset liability management, income generation, and principal preservation. Since the global financial crisis of 2008, evaluating risk—in particular, credit risk—for fixed-income securities has become an increasingly important aspect for this asset class.

© 2021 CFA Institute. All rights reserved.

FIXED INCOME
STUDY SESSION
13
Fixed Income (1)

This study session introduces the unique attributes that define fixed-income securities, then follows with an overview of global debt markets. Primary issuers, sectors, and bond types are explained. Key concepts for the calculation and interpretation of bond prices, yields, and spreads and coverage of interest rate risk and key related risk measures are presented. Securitization—the creation of fixed-income securities backed by certain (typically less liquid) assets—including the various types, characteristics, and risks of these investments end the session.

READING ASSIGNMENTS

Reading 39	Fixed-Income Securities: Defining Elements by Moorad Choudhry, PhD, FRM, FCSI, and Stephen E. Wilcox, PhD, CFA
Reading 40	Fixed-Income Markets: Issuance, Trading, and Funding by Moorad Choudhry, PhD, FRM, FCSI, Steven V. Mann, PhD, and Lavone F. Whitmer, CFA
Reading 41	Introduction to Fixed-Income Valuation by James F. Adams, PhD, CFA, and Donald J. Smith, PhD
Reading 42	Introduction to Asset-Backed Securities by Frank J. Fabozzi, PhD, CPA, CFA

© 2021 CFA Institute . All rights reserved.

READING

39

Fixed-Income Securities: Defining Elements

by Moorad Choudhry, PhD, FRM, FCSI, and Stephen E. Wilcox, PhD, CFA

Moorad Choudhry PhD, FRM, FCSI is at Recognise Bank (England). Stephen E. Wilcox, PhD, CFA, is at Minnesota State University, Mankato (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe basic features of a fixed-income security;
☐	**b.** describe content of a bond indenture;
☐	**c.** compare affirmative and negative covenants and identify examples of each;
☐	**d.** describe how legal, regulatory, and tax considerations affect the issuance and trading of fixed-income securities;
☐	**e.** describe how cash flows of fixed-income securities are structured;
☐	**f.** describe contingency provisions affecting the timing and/or nature of cash flows of fixed-income securities and whether such provisions benefit the borrower or the lender.

INTRODUCTION AND OVERVIEW OF A FIXED-INCOME SECURITY

1

a describe basic features of a fixed-income security

Fixed-income securities constitute the most prevalent means of raising capital globally based on total market value. These instruments allow governments, companies, and other issuers to borrow from investors, promising future interest payments and the return of principal, which are contractual (legal) obligations of the issuer. Fixed-income securities are the largest source of capital for government, not-for-profit, and other entities that do not issue equity. For private companies, fixed-income investors differ from shareholders in not having ownership rights. Payments of interest and repayment of principal (amount borrowed) are a higher priority claim on the company's

© 2021 CFA Institute. All rights reserved.

earnings and assets compared with the claim of common shareholders. Since fixed-income claims rank above shareholder claims in the capital structure, a company's fixed-income securities have, in theory, lower risk than their common shares.

Financial analysts who master these and other fixed-income concepts have a distinct edge over their peers for several reasons. First, given the nature of fixed-income cash flows and their preponderance across issuers and regions, these instruments form the basis for risk versus return comparisons both across and within specific jurisdictions. For example, as bonds issued by the US Treasury and other developed market central governments are viewed as having little to no default risk, they serve as building blocks in determining the time value of money for less certain cash flows. Fixed-income securities also fulfill an important role in portfolio management as a prime means by which individual and institutional investors can fund known future obligations, such as tuition payments or retirement obligations. Finally, while the correlation of fixed-income returns with common share returns varies, adding fixed-income securities to portfolios that include common shares can be an effective way of obtaining diversification benefits.

Among the questions to be addressed are the following:

- Which features define a fixed-income security, and how do they determine the scheduled cash flows?
- What are the legal, regulatory, and tax considerations associated with a fixed-income security, and why are they important for investors?
- What are the common interest and principal payment structures?
- What types of provisions may affect the disposal or redemption of fixed-income securities?

Note that the terms "fixed-income securities," "debt securities," and "bonds" are often used interchangeably by experts and non-experts alike. We will also follow this convention, and where any nuance of meaning is intended, it will be made clear. Moreover, the term "fixed income" is not to be understood literally: Some fixed-income securities have interest payments that change over time.

1.1 Overview of a Fixed-Income Security

A **bond** is a contractual agreement between the issuer and the bondholders. Three important elements that an investor needs to know about when considering a fixed-income security are:

- The bond's features, including the issuer, maturity, par value, coupon rate and frequency, and currency denomination. These features determine the bond's scheduled cash flows and, therefore, are key determinants of the investor's expected and actual return.
- The legal, regulatory, and tax considerations that apply to the contractual agreement between the issuer and the bondholders.
- The contingency provisions that may affect the bond's scheduled cash flows. These contingency provisions are options providing either issuers or bondholders certain rights affecting the bond's disposal or redemption.

This section describes a bond's basic features and introduces yield measures. The legal, regulatory, and tax considerations and contingency provisions are discussed in subsequent sections.

1.1.1 *Basic Features of a Bond*

All bonds, regardless of issuer, are characterized by the same basic features, which include maturity, par or principal amount, coupon size, frequency, and currency.

1.1.1.1 Issuer Many entities issue bonds: private individuals, such as the musician David Bowie; national governments, such as Singapore or Italy; and companies, such as BP, General Electric, or Tata Group.

Bond issuers are classified into categories based on the similarities of these issuers and their characteristics. Major types of issuers include the following:

- Supranational organizations, such as the World Bank or the European Investment Bank;
- Sovereign (national) governments, such as the United States or Japan. As sovereign bonds are backed by the full faith and credit of each respective government, they usually represent the lowest risk and most secure bonds in each market;
- Non-sovereign (local) governments, such as the State of Minnesota in the United States, the Catalonia region in Spain, or the City of Edmonton in Alberta, Canada;
- Quasi-government entities (i.e., agencies that are owned or sponsored by governments), such as postal services in many countries—for example, Correios in Brazil, La Poste in France, or Pos in Indonesia;
- Companies (i.e., corporate issuers). A distinction is often made between financial issuers (e.g., banks and insurance companies) and non-financial issuers; and
- Special legal entities (i.e., special purpose entities) that use specific assets, such as auto loans and credit card debt obligations, to guarantee (or secure) a bond issue known as an asset-backed security that is then sold to investors.

Market participants often classify fixed-income markets by the type of issuer, which leads to the identification of three bond market sectors: the government and government-related sector (i.e., the first four types of issuers just listed), the corporate sector (the fifth type listed), and the structured finance sector (the last type listed).

While several major local fixed-income markets, such as China and Japan, are dominated by government issuers, a significant portion of the US bond market consists of corporate issuance and asset-backed securities in addition to US treasury securities issued by the federal government.

Asset-backed securities (ABS) are created from a process called securitization, which involves moving assets from the owner of the assets into a special legal entity. This special legal entity then uses the securitized assets as guarantees to back (secure) a bond issue, leading to the creation of ABS. Assets that are typically used to create ABS include residential and commercial mortgage loans (mortgages), automobile (auto) loans, student loans, bank loans, and credit card debt. Many elements discussed in this reading apply to both traditional bonds and ABS. Considerations specific to ABS are discussed in a separate reading on asset-backed securities.

Bondholders are exposed to credit risk—that is, the risk of loss resulting from the issuer failing to make full and timely payments of interest and/or repayments of principal. Credit risk is inherent to all debt investments. Bond markets are sometimes classified into sectors based on the issuer's creditworthiness as judged by credit rating agencies. The three largest credit rating agencies are Moody's Investors Service, Standard & Poor's, and Fitch Ratings. One major distinction is between investment-grade and non-investment-grade bonds, also called high-yield or speculative bonds. For example, bonds rated Baa3 or higher by Moody's and BBB– or higher by Standard & Poor's and Fitch are considered investment grade. Although a variety of considerations enter into distinguishing the two sectors, the promised payments of investment-grade bonds are perceived as less risky than those of non-investment-grade bonds because of profitability and liquidity considerations. Some regulated financial intermediaries, such as banks and life insurance companies, may face explicit or implicit limitations of holdings of non-investment-grade bonds. The investment policy statements of some investors may also include constraints or limits on such holdings. From the issuer's perspective, an investment-grade credit rating generally allows easier access to bond markets and at lower interest rates than does a non-investment-grade credit rating.

1.1.1.2 Maturity The maturity date of a bond refers to the date when the issuer is obligated to redeem the bond by paying the outstanding principal amount. The **tenor** is the time remaining until the bond's maturity date. Tenor is an important consideration in analyzing a bond's risk and return, as it indicates the period over which the bondholder can expect to receive interest payments and the length of time until the principal is repaid in full.

Maturities typically range from overnight to 30 years or longer. Fixed-income securities with maturities at issuance (original maturity) of one year or less are known as **money market securities**. Issuers of money market securities include governments and companies. Commercial paper and certificates of deposit are examples of money market securities. Fixed-income securities with original maturities that are longer than one year are called **capital market securities**. Although very rare, **perpetual bonds**, such as the consols issued by the sovereign government in the United Kingdom, have no stated maturity date.

1.1.1.3 Par Value The **principal amount**, **principal value**, or simply **principal** of a bond is the amount that the issuer agrees to repay the bondholders on the maturity date. This amount is also referred to as the par value, or simply par, face value, nominal value, redemption value, or maturity value. Bonds can have any par value.

In practice, bond prices are quoted as a percentage of their par value. For example, assume that a bond's par value is $1,000. A quote of 95 means that the bond price is $950 (95% × $1,000). When the bond is priced at 100% of par, the bond is said to be trading at par. If the bond's price is below 100% of par, such as in the previous example, the bond is trading at a discount. Alternatively, if the bond's price is above 100% of par, the bond is trading at a premium.

1.1.1.4 Coupon Rate and Frequency The coupon rate or nominal rate of a bond is the interest rate that the issuer agrees to pay each year until the maturity date. The annual amount of interest payments made is called the coupon. A bond's coupon is determined by multiplying its coupon rate by its par value. For example, a bond with a coupon rate of 6% and a par value of $1,000 will pay annual interest of $60 (6% × $1,000).

Coupon payments may be made annually, such as those for German government bonds or Bunds. Many bonds, such as government and corporate bonds issued in the United States or government gilts issued in the United Kingdom, pay interest semi-annually. Some bonds make quarterly or monthly interest payments. The acronyms QUIBS (quarterly interest bonds) and QUIDS (quarterly income debt securities) are used by Morgan Stanley and Goldman Sachs, respectively, for bonds that make quarterly interest payments. Many **mortgage-backed securities** (MBS), which are ABS backed by residential or commercial mortgages, pay interest monthly to match the cash flows of the underlying assets. If a bond has a coupon rate of 6% and a par value of $1,000, the periodic interest payments will be $60 if coupon payments are made annually, $30 if they are made semi-annually, $15 if they are made quarterly, and $5 if they are made monthly.

A **plain vanilla bond** or **conventional bond** pays a fixed rate of interest. In this case, the coupon payment does not change during the bond's life. However, there are bonds that pay a floating rate of interest; such bonds are called **floating-rate notes** (FRNs) or **floaters**. The coupon rate of an FRN includes two components: a **market**

reference rate (MRR) plus a spread. The spread, also called margin, is typically constant and is expressed in basis points (bps). A **basis point** is equal to 0.01% (i.e., 100 basis points equals 1%). The spread is set when the bond is issued based on the issuer's creditworthiness at issuance: The higher the issuer's credit quality, the lower the spread. The MRR, however, resets periodically. Thus, as the MRR changes, the coupon rate and coupon payment change accordingly.

The market reference rate is a collective name for a set of rates covering different currencies for different maturities, ranging from overnight to one year. These rates have historically included the London Interbank Offered Rate (Libor), the Euro Interbank Offered Rate (Euribor), the Hong Kong Interbank Offered Rate (Hibor), or the Singapore Interbank Offered Rate (Sibor) for issues denominated in US dollars, euros, Hong Kong dollars, and Singapore dollars, respectively. The process of phasing out Libor and moving to new market reference rates is discussed in a subsequent CFA Level I reading.

For example, assume that the coupon rate of an FRN that makes semi-annual interest payments in June and December is expressed as the six-month MRR + 150 bps. Suppose that in December 20X0, the six-month MRR is 3.25%. The interest rate that will apply to the payment due in June 20X1 will be 4.75% (3.25% + 1.50%). Now suppose that in June 20X1, the six-month MRR has decreased to 3.15%. The interest rate that will apply to the payment due in December 20X1 will decrease to 4.65% (3.15% + 1.50%).

All bonds, whether they pay a fixed or floating rate of interest, make periodic coupon payments except for **zero-coupon bonds**. Zero-coupon, or **pure discount bonds**, are issued at a discount to par value and are redeemed at par. The interest earned on a zero-coupon bond is implied and equal to the difference between the par value and the purchase price. For example, if the par value is $1,000 and the purchase price is $950, the implied interest is $50.

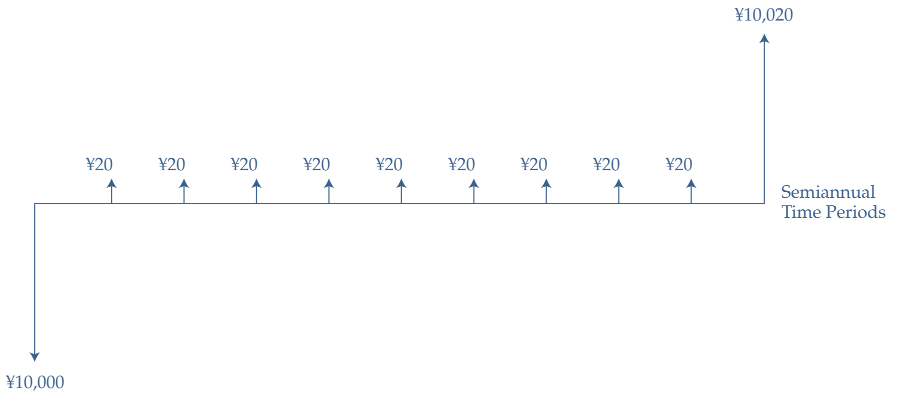
1.1.1.5 Currency Denomination Bonds can be issued in any currency, although a large number of bond issues are made in either euros or US dollars. The currency of issue may affect a bond's attractiveness. Borrowers in emerging markets often elect to issue bonds in euros or US dollars because doing so makes the bonds more attractive to international investors than bonds in a domestic currency that may be illiquid or not freely traded. Issuers may also choose to borrow in a foreign currency if they expect cash flows in that currency to offset interest payments and principal repayments. If a bond is aimed solely at a country's domestic investors, it is more likely that the borrower will issue in the local currency.

Dual-currency bonds make coupon payments in one currency and pay the par value at maturity in another currency. For example, assume that a Japanese company needs to finance a long-term project in the United States that will take several years to become profitable. The Japanese company could issue a yen/US dollar dual-currency bond. The coupon payments in yen can be made from the cash flows generated in Japan, and the principal can be repaid in dollars once the project becomes profitable.

Currency option bonds can be viewed as a combination of a single-currency bond plus a foreign currency option. They give bondholders the right to choose the currency in which they want to receive interest payments and principal repayments. Bondholders can select one of two currencies for each payment.

Exhibit 1 brings all the basic features of a bond together and illustrates how these features determine the cash flow pattern for a plain vanilla bond. The bond is a five-year Japanese government bond (JGB) with a coupon rate of 0.4% and a par value of ¥10,000. Interest payments are made semi-annually. The bond is priced at par when it is issued and is redeemed at par.

Exhibit 1 Cash Flows for a Plain Vanilla Bond



The downward-pointing arrow in Exhibit 1 represents the cash flow paid by the bond investor (received by the issuer) on the day of the bond issue—that is, ¥10,000. The upward-pointing arrows are the cash flows received by the bondholder (paid by the issuer) during the bond's life. As interest is paid semi-annually, the coupon payment is ¥20 [(0.004 × ¥10,000) ÷ 2] every six months for five years—that is, 10 coupon payments of ¥20. The last payment is equal to ¥10,020 because it includes both the last coupon payment and the payment of the par value.

EXAMPLE 1

1 An example of a sovereign bond is a bond issued by:
   - **A** the World Bank.
   - **B** the city of New York.
   - **C** the federal German government.

2 The risk of loss resulting from the issuer failing to make full and timely payment of interest is called:
   - **A** credit risk.
   - **B** systemic risk.
   - **C** interest rate risk.

3 A money market security *most likely* matures in:
   - **A** one year or less.
   - **B** between 1 and 10 years.
   - **C** over 10 years.

4 If the bond's price is higher than its par value, the bond is trading at:
   - **A** par.
   - **B** a discount.
   - **C** a premium.

5 A bond has a par value of £100 and a coupon rate of 5%. Coupon payments are made semi-annually. The periodic interest payment is:
   - **A** £2.50, paid twice a year.
   - **B** £5.00, paid once a year.
   - **C** £5.00, paid twice a year.

6 The coupon rate of a floating-rate note that makes payments in June and December is expressed as six-month MRR + 25 bps. Assuming that the six-month MRR is 3.00% at the end of June 20XX and 3.50% at the end of December 20XX, the interest rate that applies to the payment due in December 20XX is:

A 3.25%.

B 3.50%.

C 3.75%.

7 The type of bond that allows bondholders to choose the currency in which they receive each interest payment and principal repayment is a:

A pure discount bond.

B dual-currency bond.

C currency option bond.

Solution to 1:

C is correct. A sovereign bond is a bond issued by a national government, such as the federal German government. A is incorrect because a bond issued by the World Bank is a supranational bond. B is incorrect because a bond issued by a local government, such as the city of New York, is a non-sovereign bond.

Solution to 2:

A is correct. Credit risk is the risk of loss resulting from the issuer failing to make full and timely payments of interest and/or repayments of principal. B is incorrect because systemic risk is the risk of failure of the financial system. C is incorrect because interest rate risk is the risk that a change in market interest rate affects a bond's value.

Solution to 3:

A is correct. The primary difference between a money market security and a capital market security is the maturity at issuance. Money market securities mature in one year or less, whereas capital market securities mature in more than one year.

Solution to 4:

C is correct. If a bond's price is higher than its par value, the bond is trading at a premium. A is incorrect because a bond is trading at par if its price is equal to its par value. B is incorrect because a bond is trading at a discount if its price is lower than its par value.

Solution to 5:

A is correct. The annual coupon payment is 5% × £100 = £5.00. The coupon payments are made semi-annually, so £2.50 paid twice a year.

Solution to 6:

A is correct. The interest rate that applies to the payment due in December 20XX is the six-month MRR at the end of June 20XX plus 25 bps. Thus, it is 3.25% (3.00% + 0.25%).

Solution to 7:

C is correct. A currency option bond gives bondholders the right to choose the currency in which they want to receive each interest payment and principal repayment. A is incorrect because a pure discount bond is issued at a discount to par value and redeemed at par. B is incorrect because a dual-currency bond makes coupon payments in one currency and pays the par value at maturity in another currency.

1.1.2 *Yield Measures*

Several yield measures are commonly used by market participants. The **current yield** or **running yield** is equal to the bond's annual coupon divided by the bond's price, expressed as a percentage. For example, if a bond has a coupon rate of 6%, a par value of $1,000, and a price of $1,010, the current yield is 5.94% ($60 ÷ $1,010). The current yield is a measure of income that is analogous to the dividend yield for a common share.

The most common yield measure is known as the **yield-to-maturity**, also called the **yield-to-redemption** or **redemption yield**. The yield-to-maturity is the internal rate of return on a bond's expected cash flows—that is, the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. The yield-to-maturity can be considered an estimate of the bond's expected return; it reflects the annual return earned by an investor who purchases the bond today and holds it until maturity, provided they receive all promised cash flows and are able to reinvest all coupons at this same yield. There is an inverse relationship between the bond's price and its yield-to-maturity, all else being equal. That is, the higher the bond's yield-to-maturity, the lower its price. Alternatively, the higher the bond's price, the lower its yield-to-maturity. Thus, investors expecting interest rates to fall (and investors to demand a lower yield-to-maturity) anticipate a positive return from price appreciation. The topic of risk and return of fixed-income securities is covered in more detail in a subsequent reading.

2

BOND INDENTURE

b describe content of a bond indenture

c compare affirmative and negative covenants and identify examples of each

As contractual agreements between the issuer and the bondholders, bonds are subject to legal, regulatory, and tax considerations that are important to fixed-income investors.

2.1 Bond Indenture

The **trust deed** is the legal contract that describes the form of the bond, the obligations of the issuer, and the rights of the bondholders. This legal contract is often referred to as the bond **indenture**, particularly in the United States and Canada. The indenture references both the issuer and the features of the bond issue, such as the principal value, the interest or coupon rate, the interest payment dates, the maturity date, and any contingency provisions. The indenture also identifies the funding sources for interest and principal payments and any collateral, credit enhancements, or covenants. **Collateral** is a specific asset or assets or financial guarantees securing the debt obligation above and beyond the issuer's promise to pay. **Credit enhancements** are provisions used to reduce the credit risk of the bond issue. **Covenants** are clauses that specify the rights of the bondholders and any actions that the issuer is obligated to perform or prohibited from performing.

Because it would be impractical for the issuer to enter into a direct agreement with each bondholder, the indenture is usually held by a trustee. The trustee is typically a financial institution acting as a **fiduciary** (or representative protecting investor interests) to ensure the issuer complies with the obligations specified in the indenture and to take action on behalf of the bondholders, when necessary. The trustee's duties are administrative and include the maintenance of required documentation and records; holding beneficial title to, safeguarding, and appraising collateral (if any); invoicing the issuer for interest payments and principal repayments; and holding funds until they are paid, although the cash flow movements from issuers to the trustee are typically handled by the principal paying agent. In the event of default, the discretionary powers of the trustee increase considerably. For example, the trustee is responsible for calling bondholder meetings to discuss actions as well as bringing legal action against an issuer on behalf of bondholders, if necessary.

For a plain vanilla bond, the indenture is often a standard template that is updated for a bond's specific terms and conditions. For exotic bonds, the document is tailored and can often be several hundred pages.

The indenture identifies basic bond features (e.g., investor rights in the event of default) that impact the bond's risk–reward profile, such as:

- the legal identity of the bond issuer and its legal form;
- the source of repayment proceeds;
- the asset or collateral backing (if any);
- the credit enhancements (if any); and
- the covenants (if any).

We consider each of these areas in the following sections.

2.1.1 *Legal Identity of the Bond Issuer and Its Legal Form*

The bond issuer, identified in the indenture by its legal name, must make all contractual payments assigned to the bond issuer. For a sovereign bond, the legal issuer is usually the office responsible for managing the national budget, such as HM Treasury (Her Majesty's Treasury) in the United Kingdom. The legal issuer may be different from the body that administers the bond issue process. Using the UK example, the legal obligation to repay gilts lies with HM Treasury, but the bonds are issued by the UK Debt Management Office, an executive agency of HM Treasury.

For corporate bonds, the issuer is usually the corporate legal entity—for example, Wal-Mart Stores Inc., Samsung Electronics Co. Ltd., or Volkswagen AG. However, bonds are sometimes issued by a subsidiary of a parent legal entity. In this case, investors should consider the credit quality of the subsidiary, unless the indenture specifies that the bond liabilities are guaranteed by the parent. When rated, subsidiaries often receive a credit rating below that of the parent.

Bonds are sometimes issued by a holding company, or parent legal entity for a group of companies, rather than by one of the operating companies in the group. A holding company may be rated differently from its operating companies, and investors may lack recourse to assets held by those companies. If the bonds are issued by a holding company with few assets to call on in the event of default, investors face more credit risk than if the bonds were issued by an operating company in the group.

For ABS, the legal obligation to repay the bondholders often lies with the special legal entity created by the sponsor or originator, a financial institution in charge of the securitization process. The special legal entity is most frequently referred to as a special purpose entity (SPE) in the United States and a special purpose vehicle (SPV) in Europe, also called a special purpose company (SPC). The legal form for the special legal entity may be a limited partnership, a limited liability company, or a trust. Typically, special legal entities are thinly capitalized, have no independent management or employees, and have no purpose other than the transactions for which they were created.

The sponsor transfers the assets to the special legal entity to carry out a securitization or series of securitizations. A key reason for forming a special legal entity is *bankruptcy remoteness*. The transfer of assets by the sponsor is considered a legal sale; once the assets have been securitized, the sponsor no longer has ownership rights. Any party making claims following the bankruptcy of the sponsor would be unable to recover the assets or their proceeds. As a result, the special legal entity's ability to pay interest and repay the principal should remain intact even if the sponsor were to fail—hence the reason why the special legal entity is also called a bankruptcy-remote vehicle.

2.1.2 *Source of Repayment Proceeds*

The indenture usually describes how the issuer intends to service the debt (make interest payments) and repay the principal. Generally, the source of repayment for bonds issued by supranational organizations is either the repayment of previous loans made by the organization or the paid-in capital from its members. National governments may also act as guarantors for certain bond issues. If additional sources of repayment are needed, the supranational organization may call on its members to provide funds.

Sovereign bonds are backed by the "full faith and credit" of the national government. As national governments have unique powers to ensure the ability to repay debt—such as the authority to tax economic activity, print money, and control foreign currency reserves—sovereign bonds denominated in local currency are generally considered the safest of all investments. It is therefore highly probable that interest and principal will be paid fully and on time, resulting in lower yields on sovereign bonds than for similar bonds from other local issuers. Government yields are often used as a "risk-free rate" for time value of money calculations as well as a benchmark reference rate for pricing fixed-income securities from other issuers of the same tenor.

There are three major sources for repayment of non-sovereign government debt issues, and bonds are usually classified according to these sources. The first source is through the general taxing authority of the issuer. The second source is from the cash flows of the project the bond issue is financing. The third source is from special taxes or fees established specifically for the purpose of funding interest payments and principal repayments.

The source of payment for corporate bonds is the issuer's ability to generate cash flows, primarily through its operations. These cash flows depend on the issuer's financial strength and integrity. Because corporate bonds carry a higher level of credit risk than otherwise similar sovereign and non-sovereign government bonds, they typically offer a higher yield.

In contrast to corporate bonds, the source of payment for ABS depends on the cash flows generated by one or more of the underlying financial assets, such as mortgages or auto loans, rather than an operating entity. Thus, investors in ABS must pay special attention to the quality of these assets.

2.1.3 *Asset or Collateral Backing*

Collateral backing is a way to alleviate credit risk. Investors should review where they rank compared with other creditors in the event of default and analyze the quality of the collateral backing the bond issue.

2.1.3.1 Seniority Ranking **Secured bonds** are backed by assets or financial guarantees pledged to ensure debt repayment in the case of default. In contrast, unsecured bonds have no collateral; bondholders have only a general claim on the issuer's assets and cash flows. Thus, unsecured bonds are paid after secured bonds in the event of default. By lowering credit risk, collateral backing increases the bond issue's credit quality and decreases its yield.

A bond's collateral backing might not specify an identifiable asset but instead may be described as the "general plant and infrastructure" of the issuer. In such cases, investors rely on seniority ranking—that is, the systematic way in which lenders are repaid in case of bankruptcy or liquidation. What matters to investors is where they rank compared with other creditors rather than whether an asset of sufficient quality and value is in place to cover their claims. Senior debt has a priority claim over subordinated debt or junior debt. Financial institutions issue a large volume of both senior unsecured and subordinated bonds globally; it is not uncommon to see large and small banks issue such bonds. For example, banks as diverse as Royal Bank of Scotland in the United Kingdom and Prime Bank in Bangladesh issue senior unsecured bonds to institutional investors.

Debentures are a type of bond that can be secured or unsecured. In many jurisdictions, debentures are unsecured bonds, with no collateral backing assigned to the bondholders. In contrast, in the United Kingdom and in other Commonwealth countries, such as India, debentures are usually backed by an asset or pool of assets assigned as collateral support for the bond obligations and segregated from other creditor claims. Thus, it is important for investors to review the indenture to determine whether a debenture is secured or unsecured. If the debenture is secured, debenture holders rank above unsecured creditors of the company; they have a specific asset or pool of assets that the trustee can call on to realize the debt in the event of default.

2.1.3.2 Types of Collateral Backing Many types of bonds are secured by a form of collateral. Some companies issue collateral trust bonds and equipment trust certificates. **Collateral trust bonds** are backed by securities, such as common shares, other bonds, or other financial assets. These securities are pledged by the issuer and typically held by the trustee. **Equipment trust certificates** are bonds secured by specific types of equipment or physical assets, such as aircraft, railroad cars, shipping containers, or oil rigs. They are often issued to take advantage of the tax benefits of leasing. For example, suppose an airline finances the purchase of new aircraft with equipment trust certificates. The legal title to the aircraft is held by the trustee, which issues equipment trust certificates to investors in the amount of the aircraft purchase price. The trustee leases the aircraft to the airline and collects lease payments from the airline to pay the interest on the certificates. When the certificates mature, the trustee sells the aircraft to the airline, uses the proceeds to retire the principal, and cancels the lease.

Mortgaged property is one of the most common forms of ABS collateral. MBS are debt obligations that represent claims to the cash flows from pools of mortgage loans, most commonly on residential property. Mortgage loans are purchased from banks, mortgage companies, and other originators and then assembled into pools by a governmental, quasi-governmental, or private entity.

Financial institutions, particularly in Europe, issue covered bonds. **Covered bonds** are debt obligations backed by a segregated pool of assets. They are similar to ABS but offer bondholders recourse against both the financial institution and the underlying asset pool. Covered bonds will be addressed in detail in a later reading.

2.1.4 *Credit Enhancements*

Credit enhancements refer to a variety of provisions that can be used to reduce the credit risk of a bond issue. Thus, they increase the issue's credit quality and decrease the bond's yield. Credit enhancements are very often used when creating ABS.

The two primary types of credit enhancements are internal and external. Internal credit enhancement relies on structural features regarding the bond issue. External credit enhancement refers to financial guarantees received from a third party, often called a financial guarantor. We describe each type in the following sections.

2.1.4.1 Internal Credit Enhancement The most common forms of internal credit enhancement are subordination, overcollateralization, and reserve accounts.

Subordination, also known as **credit tranching**, is the most popular internal credit enhancement technique. It relies on creating more than one bond class or tranche and ordering the claim priorities for ownership or interest in an asset between the tranches. The cash flows generated by the assets are allocated with different priority to tranches of different seniority. The ordering of the claim priorities is called a senior/subordinated structure, where the tranches of highest seniority are senior followed by subordinated or junior tranches. The subordinated tranches function as credit protection for the more senior tranches, in the sense that the most senior tranche has the first claim on available cash flows. This type of protection is also commonly referred to as a waterfall structure because in the event of default, the proceeds from liquidating assets will first be used to repay the most senior creditors. Thus, if the issuer defaults, losses are allocated from the bottom up—that is, from the most junior to the most senior tranche. The most senior tranche is typically unaffected unless losses exceed the amount of the subordinated tranches, which is why the most senior tranche is usually rated Aaa/AAA.

Overcollateralization refers to the process of posting more collateral than is needed to obtain or secure financing. It represents a form of internal credit enhancement because the additional collateral can be used to absorb losses. For example, a bond issue of $100 million with collateral value of $110 million has excess collateral of $10 million. Over time, the amount of overcollateralization changes. This can happen because of amortization, prepayments, or defaults as well as changes in collateral values. For example, one of the most significant contributors to the 2007–2009 global financial crisis was a valuation problem with the residential housing assets backing MBS. While many properties were originally valued above the value of securities outstanding, as property prices fell and the number of defaults rose, the credit quality of these MBS declined sharply. This resulted in a rapid rise in yields and losses among investors in these securities.

Reserve accounts or **reserve funds** are another internal credit enhancement in the form of either a cash reserve fund or an excess spread account. A cash reserve fund is a cash deposit used to absorb losses. An excess spread account is an allocation of any asset cash flows remaining after paying interest to bondholders. The excess spread (or excess interest cash flow) can be retained and deposited into a reserve account as a first line of protection against losses. In a process called "turboing," the excess spread can be used to retire the principal, with the most senior tranche having the first claim on these funds.

2.1.4.2 External Credit Enhancement The most common forms of external credit enhancement are bank guarantees and surety bonds, letters of credit, and cash collateral accounts.

Bank guarantees and **surety bonds** are very similar in nature because they both reimburse bondholders for any losses incurred if the issuer defaults. However, there is usually a maximum amount that is guaranteed, called the penal sum. The major difference between a bank guarantee and a surety bond is that the former is issued by a bank, whereas the latter is issued by a rated and regulated insurance company. Insurance companies that specialize in providing financial guarantees are typically called monoline insurance companies or monoline insurers. Monoline insurers played an important role in securitization until the 2007–2009 global financial crisis, but they are now a less common form of credit enhancement due to the credit downgrades of these insurers during the crisis.

A **letter of credit** from a financial institution is another form of external credit enhancement for a bond issue. The financial institution provides the issuer with a credit line to reimburse any cash flow shortfalls from the assets backing the issue. Letters of credit have also become a less common form of credit enhancement since the credit downgrades of financial institutions during the financial crisis.

Bank guarantees, surety bonds, and letters of credit expose the investor to third-party (or counterparty) risk—that is, the possibility that a guarantor cannot meet its obligations. A **cash collateral account** mitigates this concern because the issuer immediately borrows the credit enhancement amount and usually invests it in highly rated short-term commercial paper. Because a cash collateral account is an actual deposit rather than a pledge of cash, a cash collateral account provider downgrade will not necessarily result in a downgrade of the bond issue backed by that provider.

2.1.5 *Covenants*

Bond covenants are legally enforceable rules that borrowers and lenders agree on at the time of a new bond issue. An indenture will frequently include affirmative (or positive) and negative covenants. Affirmative covenants enumerate what issuers are required to do, whereas negative covenants specify what issuers are prohibited from doing.

Affirmative covenants are typically administrative in nature. For example, frequently used affirmative covenants include what the issuer will do with the proceeds from the bond issue and the promise of making the contractual payments. The issuer may also promise to comply with all laws and regulations, maintain its current lines of business, insure and maintain its assets, and pay taxes as they come due. Other examples include a **pari passu** (or "equal footing") clause, which ensures that a debt obligation is treated the same as the borrower's other senior debt instruments, or a **cross-default** clause, which specifies that a borrower is considered in default if they default on another debt obligation. These types of covenants typically do not impose additional costs to the issuer and do not materially constrain the issuer's discretion regarding how to operate its business.

In contrast, negative covenants are frequently costly and materially constrain the issuer's potential business decisions. They protect bondholders from the dilution of their claims, asset withdrawals or substitutions, and suboptimal investments by the issuer. Examples of negative covenants include the following:

- *Restrictions on debt* regulate the issue of additional debt. Maximum acceptable debt usage ratios (sometimes called leverage ratios or gearing ratios) and minimum acceptable interest coverage ratios are frequently specified, permitting new debt to be issued only when justified by the issuer's financial condition.
- *Negative pledges* prevent the issuance of debt that would be senior to or rank in priority ahead of the existing bondholders' debt.
- *Restrictions on prior claims* protect unsecured bondholders by preventing the issuer from using assets that are not collateralized (called unencumbered assets) to become collateralized.
- *Restrictions on distributions to shareholders* restrict dividends and other payments to shareholders, such as share buybacks (repurchases). The restriction typically operates by reference to the borrower's profitability; that is, the covenant sets a base date, usually at or near the time of the issue, limiting dividends and share buybacks to a percentage of earnings or cumulative earnings after that date.
- *Restrictions on asset disposals* limit the amount of assets that can be disposed by the issuer during the bond's life. This limit on cumulative disposals as a percentage of a company's gross assets is intended to protect bondholder claims by preventing a break-up of the company.

- *Restrictions on investments* constrain risky investments by blocking speculative investments and ensuring an issuer devotes its capital to its going-concern business. A companion covenant may require the issuer to stay in its present line of business.
- *Restrictions on mergers and acquisitions* prevent these actions unless the company is the surviving company or the acquirer delivers a supplemental indenture to the trustee expressly assuming the old bonds and terms of the old indenture. These requirements effectively prevent a company from avoiding its obligations to bondholders by selling to another company.

These are only a few examples of negative covenants. The common characteristic of all negative covenants is ensuring that the issuer will not take any actions that would significantly reduce its ability to make interest payments and repay the principal. Bondholders, however, rarely wish to be too specific about how an issuer should run its business because doing so would imply a degree of control that bondholders legally want to avoid. In addition, very restrictive covenants may not be in the bondholders' best interest if they force the issuer to default when default is avoidable. For example, strict restrictions on debt may prevent the issuer from raising new funds that are necessary to meet its contractual obligations; strict restrictions on asset disposals may prohibit the issuer from selling assets or business units and obtaining the necessary liquidity to make interest payments or principal repayments; and strict restrictions on mergers and acquisitions may prevent the issuer from being taken over by a stronger company that would be able to honor the issuer's contractual obligations.

EXAMPLE 2

1 The term *most likely* used to refer to the legal contract under which a bond is issued is:

   A indenture.

   B debenture.

   C letter of credit.

2 The individual or entity that *most likely* assumes the role of trustee for a bond issue is:

   A a financial institution appointed by the issuer.

   B the treasurer or chief financial officer of the issuer.

   C a financial institution appointed by a regulatory authority.

3 The individual or entity *most likely* responsible for the timely payment of interest and repayment of principal to bondholders is the:

   A trustee.

   B primary or lead bank of the issuer.

   C treasurer or chief financial officer of the issuer.

4 The major advantage of issuing bonds through a special legal entity is:

   A bankruptcy remoteness.

   B beneficial tax treatments.

   C greater liquidity and lower issuing costs.

5 The category of bond *most likely* repaid from the repayment of previous loans made by the issuer is:

   A sovereign bonds.

   B supranational bonds.

C non-sovereign bonds.

6 The type of collateral used to secure collateral trust bonds is *most likely*:

A securities.

B mortgages.

C physical assets.

7 The external credit enhancement that has the *least* amount of third-party risk is a:

A surety bond.

B letter of credit.

C cash collateral account.

8 An example of an affirmative covenant is the requirement:

A that dividends will not exceed 60% of earnings.

B to insure and perform periodic maintenance on financed assets.

C that the debt-to-equity ratio will not exceed 0.4 and times interest earned will not fall below 8.0.

9 An example of a covenant that protects bondholders against the dilution of their claims is a restriction on:

A debt.

B investments.

C mergers and acquisitions.

Solution to 1:

A is correct. The contract between a bond issuer and the bondholders is very often called an indenture or deed trust. The indenture documents the terms of the issue, including the principal amount, the coupon rate, and the payments schedule. It also provides information about the funding sources for the contractual payments and specifies whether there are any collateral, credit enhancement, or covenants. B is incorrect because a debenture is a type of bond. C is incorrect because a letter of credit is an external credit enhancement.

Solution to 2:

A is correct. The issuer chooses a financial institution with trust powers, such as the trust department of a bank or a trust company, to act as a trustee for the bond issue.

Solution to 3:

A is correct. Although the issuer is ultimately the source of the contractual payments, it is the trustee that ensures timely payments. Doing so is accomplished by invoicing the issuer for interest payments and principal repayments and holding the funds until they are paid.

Solution to 4:

A is correct. A special legal entity is a bankruptcy-remote vehicle. Bankruptcy remoteness is achieved by transferring the assets from the sponsor to the special legal entity. Once this transfer is completed, the sponsor no longer has ownership rights. If the sponsor defaults, no claims can be made to recover the assets that were transferred or the proceeds from the transfer to the special legal entity.

Solution to 5:

B is correct. The source of payment for bonds issued by supranational organizations is either the repayment of previous loans made by the organization or the paid-in capital of its member states. A is incorrect because national governments rely on their taxing authority and money creation to repay their debt. C is incorrect because non-sovereign bonds are typically repaid from the issuer's taxing authority or the cash flows of the project being financed.

Solution to 6:

A is correct. Collateral trust bonds are secured by securities, such as common shares, other bonds, or other financial assets. B is incorrect because MBS are secured by mortgages. C is incorrect because equipment trust certificates are backed by physical assets, such as aircraft, railroad cars, shipping containers, or oil rigs.

Solution to 7:

C is correct. The third-party (or counterparty) risk for a surety bond and a letter of credit arises from both being future promises to pay. In contrast, a cash collateral account allows the issuer to immediately borrow the credit-enhancement amount and then invest it.

Solution to 8:

B is correct. Affirmative covenants indicate what the issuer "must do" and are administrative in nature. A covenant requiring the issuer to insure and perform periodic maintenance on financed assets is an example of an affirmative covenant. A and C are incorrect because they are negative covenants; they indicate what the issuer cannot do.

Solution to 9:

A is correct. A restriction on debt typically takes the form of a maximum acceptable debt usage ratio or a minimum acceptable interest coverage ratio. Thus, it limits the issuer's ability to issue new debt that would dilute the bondholders' claims. B and C are incorrect because they are covenants that restrict the issuer's business activities by preventing the company from making investments or being taken over, respectively.

3

LEGAL, REGULATORY, AND TAX CONSIDERATIONS

d describe how legal, regulatory, and tax considerations affect the issuance and trading of fixed-income securities

Fixed-income securities are subject to different legal and regulatory requirements across jurisdictions, depending on where they are issued and traded as well as who holds them.

An important consideration for investors is where the bonds are issued and traded because it affects the laws and regulations that apply. The *global bond* markets consist of national bond markets and the Eurobond market. A *national bond* market includes all the bonds that are issued and traded in a specific country and are denominated in the currency of that country. Bonds issued by entities that are incorporated in that country are called *domestic bonds*, whereas bonds issued by entities that are incorporated in another country are called *foreign bonds*. If Ford Motor Company issues bonds denominated in US dollars in the United States, for example, these bonds will be classified as domestic. If Volkswagen Group and Toyota Motor Corporation (or their German or Japanese subsidiaries) issue bonds denominated in US dollars in the United States, these bonds will be classified as foreign. Foreign bonds very often receive nicknames. For example, foreign bonds are called "kangaroo bonds" in Australia, "maple bonds" in Canada, "panda bonds" in China, "samurai bonds" in Japan, "kimchi bonds" in South Korea, "matryoshka bonds" in Russia, "matador bonds" in Spain, "bulldog bonds" in the United Kingdom, and "Yankee bonds" in the United States. National regulators may make distinctions both between and among resident and non-resident issuers, and they may have different requirements regarding the issuance process, the level of disclosures, or the restrictions imposed on the bond issuer and/or the investors as to who can purchase the bonds.

Governments and companies have issued foreign bonds in London since the 19th century, and foreign bond issues expanded in such countries as the United States, Japan, and Switzerland during the 1980s. But the 1960s saw the emergence of another bond market: the Eurobond market. The Eurobond market was created primarily to bypass the legal, regulatory, and tax constraints imposed on bond issuers and investors, particularly in the United States. Bonds issued and traded on the Eurobond market are called **Eurobonds**, and they are named after the currency in which they are denominated. For example, Eurodollar and Euroyen bonds are denominated in US dollars and Japanese yens, respectively. Bonds that are denominated in euros are called euro-denominated Eurobonds.

Eurobonds are issued outside the jurisdiction of any single country, are usually unsecured, and may be denominated in any currency, including the issuer's domestic currency. They are underwritten by an international syndicate—that is, a group of financial institutions from different jurisdictions—and mostly sold to investors in Europe, the Middle East, and Asia. Eurobonds denominated in US dollars cannot be sold to US investors at the time of issue because they are not registered with the US Securities and Exchange Commission (SEC). In the past, Eurobonds typically were **bearer bonds**, meaning that the trustee did not keep records of who owned the bonds; only the clearing system knew who the bond owners were. Eurobonds, domestic, and foreign bonds are now **registered bonds** for which ownership is recorded by either name or serial number.

A global bond is one issued simultaneously in the Eurobond market and in at least one domestic bond market, ensuring sufficient demand for large bond issues and access to all fixed-income investors regardless of location. For example, the World Bank is a regular issuer of global bonds. Many market participants refer to foreign bonds, Eurobonds, and global bonds as international bonds as opposed to domestic bonds.

Domestic bonds, foreign bonds, Eurobonds, and global bonds are subject to different legal, regulatory, and tax requirements. They also have different interest payment frequencies and interest payment calculation methods that affect bond cash flows and prices. Note, however, that the currency in which a bond is denominated has a stronger effect on its price than where the bond is issued or traded. This is because market interest rates have a strong influence on a bond's price, and the market interest rates that affect a bond are those associated with the currency in which the bond is denominated.

As the emergence and growth of the Eurobond market illustrates, legal and regulatory considerations affect the dynamics of the global fixed-income markets. Exhibit 2 compares the amounts of total and international debt outstanding for the 15 countries that were the largest debt issuers at the end of December 2019. The reported amounts are based on the residence of the issuer.

Exhibit 2 Total and International Debt Securities by Residence of Issuer at the End of December 2019

Issuers	Total Debt Securities (US$ billions)	International Debt Securities (US$ billions)
United States	41,232	2,356
China	14,726	217
Japan	12,825	480
United Kingdom	6,288	3,288
France	4,670	1,433
Germany	3,548	1,290
Canada	3,361	971
Italy	3,194	840
Netherlands	2,193	2,109
Spain	1,972	532
Australia	1,959	595
Luxembourg	1,084	936
Ireland	824	893
Denmark	793	124
Belgium	729	167

Source: Based on data from the Bank for International Settlements, Table C1, available at www.bis.org/statistics/secstats.htm (updated 3 June 2020).

EXAMPLE 3

1 An example of a domestic bond is a bond issued by:

- **A** LG Group from South Korea, denominated in British pounds, and sold in the United Kingdom.
- **B** the UK Debt Management Office, denominated in British pounds, and sold in the United Kingdom.
- **C** Wal-Mart from the United States, denominated in US dollars, and sold in various countries in North America, Europe, the Middle East, and Asia Pacific.

2 A bond issued by Sony in Japan, denominated in US dollars but not registered with the SEC, and sold to an institutional investor in the Middle East, is *most likely* an example of a:

- **A** Eurobond.
- **B** global bond.
- **C** foreign bond.

Solution to 1:

B is correct. A domestic bond is issued by a local issuer, denominated in local currency, and sold in the domestic market. Gilts are British pound-denominated bonds issued by the UK Debt Management Office in the United Kingdom. Thus, they are UK domestic bonds. A is incorrect because a bond issued by LG Group from South Korea, denominated in British pounds, and sold in the United

Kingdom is an example of a foreign bond. C is incorrect because a bond issued by Wal-Mart from the United States, denominated in US dollars, and sold in various countries in North America, Europe, the Middle East, and Asia Pacific is most likely an example of a global bond, particularly if it is also sold in the Eurobond market.

Solution to 2:

A is correct. A Eurobond is a bond that is issued internationally, outside the jurisdiction of any single country. Thus, a bond issued by Sony from Japan, denominated in US dollars but not registered with the SEC, is an example of a Eurobond. B is incorrect because global bonds are bonds that are issued simultaneously in the Eurobond market and in at least one domestic bond market. C is incorrect because if Sony's bond issue were a foreign bond, it would be registered with the SEC.

3.1 Tax Considerations

Bond interest is usually taxed at the ordinary income tax rate, which is typically the same tax rate that an individual would pay on wage or salary income. Tax-exempt securities are the exception to this rule. For example, interest income received by holders of local government bonds, called municipal bonds in the United States, is often exempt from federal income tax and from the income tax of the state in which the bonds are issued. The tax status of bond income may also depend on where the bond is issued and traded. For example, some domestic bonds pay their interest net of income tax. Other bonds, including some Eurobonds, make gross interest payments.

In addition to earnings from interest, a bond investment will generate a capital gain or loss if sold prior to maturity at a price different from the purchase price. This change will generate a capital gain if the bond price has increased or a capital loss if the bond price has decreased. Capital gains or losses usually face different tax treatment from taxable income, which often varies for long-term and short-term capital gains. For example, capital gains recognized over a year after the original bond purchase may be taxed at a lower long-term capital gains tax rate, whereas capital gains recognized within a year of bond purchase may be taxed as a short-term capital gain that equals the ordinary income tax rate. Exceptions exist, and not all countries have a separate capital gains tax. Differences in national and local tax legislation can add further complexity to aggregate country capital gains tax treatment.

For bonds issued at a discount, the tax status of the original issue discount is an additional tax consideration. The original issue discount is the difference between the par value and the original issue price. For example, the United States includes a prorated portion of the discount in interest income every tax year, while Japan does not. Original issue discount tax treatment varies by country, and Exhibit 3 illustrates the potential importance of this tax consideration.

Exhibit 3

Original Issue Discount Tax Provision

Assume a hypothetical country, Zinland, where the local currency is the zini (Z). The market interest rate in Zinland is 10%, and both interest income and capital gains are taxed. Companies A and B issue 20-year bonds with a par value of Z1,000. Company A issues a coupon bond with an annual coupon rate of 10%. Investors buy Company A's bonds for Z1,000. Every year, they receive and pay tax on their Z100 annual interest payments. When Company A's bonds mature,

(continued)

Exhibit 3 (Continued)

bondholders receive the par value of Z1,000. Company B issues a zero-coupon bond at a discount. Investors buy Company B's bonds for Z148.64. They do not receive any cash flows until Company B pays the par value of Z1,000 when the bonds mature.

Company A's bonds and Company B's bonds are economically identical in the sense that they have the same maturity (20 years) and the same yield to maturity (10%). Company A's bonds make periodic payments, however, whereas Company B's bonds defer payment until maturity. Investors in Company A's bonds must include the annual interest payments in taxable income. When they receive their original Z1,000 investment back at maturity, they face no capital gain or loss. Without an original issue discount tax provision, investors in Company B's bonds do not have any taxable income until the bonds mature. When they receive the par value at maturity, they face a capital gain on the original issue discount—that is, on Z851.36 (Z1,000 – Z148.64). The purpose of an original issue discount tax provision is to tax investors in Company B's bonds the same way as investors in Company A's bonds. Thus, a prorated portion of the Z851.36 original issue discount is included in taxable income every tax year until maturity. This allows investors in Company B's bonds to increase their cost basis in the bonds so that at maturity, they face no capital gain or loss.

Some jurisdictions also have tax provisions for bonds bought at a premium. They may allow investors to deduct a prorated portion of the amount paid in excess of the bond's par value from their taxable income every tax year until maturity. For example, if an investor pays $1,005 for a bond that has a par value of $1,000 and matures five years later, she can deduct $1 from her taxable income every tax year for five years. This deduction may be a choice rather than a requirement, with an investor able to decide whether to deduct a prorated portion of the premium each year or to deduct nothing and declare a capital loss when the bond is redeemed at maturity.

EXAMPLE 4

1 The coupon payment is *most likely* to be taxed as:

- **A** ordinary income.
- **B** short-term capital gain.
- **C** long-term capital gain.

2 Assume that a company issues bonds in the hypothetical country of Zinland, where the local currency is the zini (Z). There is an original issue discount tax provision in Zinland's tax code. The company issues a 10-year zero-coupon bond with a par value of Z1,000 and sells it for Z800. An investor who buys the zero-coupon bond at issuance and holds it until maturity *most likely*:

- **A** has to include Z20 in his taxable income every tax year for 10 years and has to declare a capital gain of Z200 at maturity.
- **B** has to include Z20 in his taxable income every tax year for 10 years and does not have to declare a capital gain at maturity.
- **C** does not have to include anything in his taxable income every tax year for 10 years but has to declare a capital gain of Z200 at maturity.

Solution to 1:

A is correct. Interest income is typically taxed at the ordinary income tax rate, which may be the same tax rate that individuals pay on wage and salary income.

Solution to 2:

B is correct. The original issue discount tax provision requires the investor to include a prorated portion of the original issue discount in his taxable income every tax year until maturity. The original issue discount is the difference between the par value and the original issue price—that is, Z1,000 − Z800 = Z200. The bond's maturity is 10 years. Thus, the prorated portion that must be included each year is Z200 ÷ 10 = Z20. The original issue discount tax provision allows the investor to increase his cost basis in the bond so that when the bond matures, the investor faces no capital gain or loss.

PRINCIPAL REPAYMENT STRUCTURES

4

e describe how cash flows of fixed-income securities are structured

The most common payment structure by far is that of a plain vanilla bond, as depicted in Exhibit 1. These bonds make periodic, fixed coupon payments and a lump-sum payment of principal at maturity. But there are other structures regarding both the principal repayment and the interest payments. This section discusses the major schedules observed in the global fixed-income markets. Schedules for principal repayments and interest payments are typically similar for a particular type of bond, such as 30-year US Treasury bonds. However, payment schedules vary considerably between types of bonds, such as government bonds versus corporate bonds.

4.1 Principal Repayment Structures

How the amount borrowed is repaid is an important consideration for investors as it affects credit risk. Any provision that periodically retires some of the principal amount outstanding is a way to reduce credit risk.

4.1.1 *Bullet, Fully Amortized, and Partially Amortized Bonds*

The payment structure of a plain vanilla bond has been used for nearly every government bond ever issued as well as for most corporate bonds. Such a bond is also known as a **bullet bond** because the entire payment of principal occurs at maturity.

In contrast, an **amortizing bond** has a payment schedule that calls for periodic payments of interest and repayments of principal. A fully amortized bond is characterized by a fixed periodic payment schedule that reduces the bond's outstanding principal amount to zero by the maturity date. A partially amortized bond also makes fixed periodic payments until maturity, but only a portion of the principal is repaid by the maturity date. Thus, a **balloon payment** is required at maturity to retire the bond's outstanding principal amount.

Exhibit 4 illustrates the differences in the payment schedules for a bullet bond, a fully amortized bond, and a partially amortized bond. For the three bonds, the principal amount is $1,000, the maturity is five years, the coupon rate is 6%, and interest payments are made annually. The market interest rate used to discount the bonds' expected cash flows until maturity is assumed to be constant at 6%. The bonds are issued and redeemed at par. For the partially amortized bond, the balloon payment is $200 at maturity.

Exhibit 4 Example of Payment Schedules for Bullet, Fully Amortized, and Partially Amortized Bonds

Bullet Bond

Year	Investor Cash Flows	Interest Payment	Principal Repayment	Outstanding Principal at the End of the Year
0	–$1,000.00			$1,000.00
1	60.00	$60.00	$0.00	1,000.00
2	60.00	60.00	0.00	1,000.00
3	60.00	60.00	0.00	1,000.00
4	60.00	60.00	0.00	1,000.00
5	1,060.00	60.00	1,000.00	0.00

Fully Amortized Bond

Year	Investor Cash Flows	Interest Payment	Principal Repayment	Outstanding Principal at the End of the Year
0	–$1,000.00			
1	237.40	$60.00	$177.40	$822.60
2	237.40	49.36	188.04	634.56
3	237.40	38.07	199.32	435.24
4	237.40	26.11	211.28	223.96
5	237.40	13.44	223.96	0.00

Partially Amortized Bond

Year	Investor Cash Flows	Interest Payment	Principal Repayment	Outstanding Principal at the End of the Year
0	–$1,000.00			
1	201.92	$60.00	$141.92	$858.08
2	201.92	51.48	150.43	707.65
3	201.92	42.46	159.46	548.19
4	201.92	32.89	169.03	379.17
5	401.92	22.75	379.17	0.00

Investors pay $1,000 now to purchase any of the three bonds. For the bullet bond, they receive the coupon payment of $60 (6% × $1,000) every year for five years. The last payment is $1,060 because it includes both the last coupon payment and the principal amount.

For the fully amortized bond, the annual payment, which includes both the coupon payment and the principal repayment, is constant. Thus, this annual payment can be viewed as an annuity. This annuity lasts for five years; its present value, discounted at the market interest rate of 6%, is equal to the bond price of $1,000. Therefore, the annual payment is $237.40. The first year, the interest part of the payment is $60 (6% × $1,000), which implies that the principal repayment part is $177.40 ($237.40 – $60). This repayment leaves an outstanding principal amount, which becomes the basis for the calculation of the interest the following year, of $822.60 ($1,000 – $177.40). The second year, the interest part of the payment is $49.36 (6% × $822.60), the principal repayment part is $188.04 ($237.40 – $49.36), and the outstanding principal amount is

$634.56 ($822.60 – $188.04). The fifth year, the outstanding principal amount is fully repaid. Note that the annual payment is constant, but over time the interest payment decreases while the principal repayment increases.

The partially amortized bond can be viewed as the combination of two elements: a five-year annuity plus the balloon payment at maturity. The sum of the present values of these two elements is equal to the bond price of $1,000. As for the fully amortized bond, the discount rate is the market interest rate of 6%, making the constant amount for the annuity $201.92. This amount represents the annual payment for the first four years. For Years 1 through 4, the split between interest and principal is done the same way as for the fully amortized bond. The interest part of the payment is equal to 6% multiplied by the outstanding principal at the end of the previous year; the principal repayment part is equal to $201.92 minus the interest part of the payment for the year; and the outstanding principal amount at the end of the year is equal to the outstanding principal amount at the end of the previous year minus the principal repayment for the year. In Year 5, investors receive $401.92; this amount is calculated either as the sum of the interest payment ($22.75) and the outstanding principal amount ($379.17) or as the constant amount of the annuity ($201.92) plus the balloon payment ($200). As for the fully amortized bond, the interest payment decreases and the principal repayment increases over time. Because the principal amount is not fully amortized, interest payments are higher for the partially amortized bond than for the fully amortized bond, except the first year when they are equal.

Exhibit 4 does not address the complexity of the repayment structure for some bonds, such as many ABS. For example, MBS face prepayment risk, which is the possible early repayment of mortgage principal. Borrowers usually have the right to prepay mortgages, which typically occurs when a current homeowner purchases a new home or when homeowners refinance their mortgages because market interest rates have fallen.

EXAMPLE 5

1 The structure that requires the largest repayment of principal at maturity is that of a:

A bullet bond.

B fully amortized bond.

C partially amortized bond.

2 A plain vanilla bond has a maturity of 10 years, a par value of £100, and a coupon rate of 9%. Interest payments are made annually. The market interest rate is assumed to be constant at 9%. The bond is issued and redeemed at par. The principal repayment the first year is *closest* to:

A £0.00.

B £6.58.

C £10.00.

3 Relative to a fully amortized bond, the coupon payments of an otherwise similar partially amortized bond are:

A lower or equal.

B equal.

C higher or equal.

Solution to 1:

A is correct. The entire repayment of principal occurs at maturity for a bullet (or plain vanilla) bond, whereas it occurs over time for fully and partially amortized bonds. Thus, the largest repayment of principal at maturity is that of a bullet bond.

Solution to 2:

A is correct. A plain vanilla (or bullet) bond does not make any principal repayment until the maturity date. B is incorrect because £6.58 would be the principal repayment for a fully amortized bond.

Solution to 3:

C is correct. Except at maturity, the principal repayments are lower for a partially amortized bond than for an otherwise similar fully amortized bond. Consequently, the principal amounts outstanding—and, therefore, the amounts of interest payments—are higher for a partially amortized bond than for a fully amortized bond, all else equal. The only exception is the first interest payment, which is the same for both repayment structures. This is because no principal repayment has been made by the time the first coupon is paid.

4.1.2 *Sinking Fund Arrangements*

A **sinking fund arrangement** is another approach that can be used to achieve the same goal of periodically retiring the bond's principal outstanding. The term sinking fund refers to an issuer's plans to set aside funds over time to retire the bond. Originally, a sinking fund was a specified cash reserve that was segregated from the rest of the issuer's business for the purpose of repaying the principal. More generally today, a sinking fund arrangement specifies the portion of the bond's principal outstanding, perhaps 5%, that must be repaid each year throughout the bond's life or after a specified date. This repayment occurs whether or not an actual segregated cash reserve has been created.

Typically, the issuer will forward repayment proceeds to the bond's trustee. The trustee will then either redeem bonds to this value or select by lottery the serial numbers of bonds to be paid off. The bonds for repayment may be listed in business newspapers, such as the *Wall Street Journal* or the *Financial Times*.

Another type of sinking fund arrangement operates by redeeming a steadily increasing amount of the bond's notional principal (total amount) each year. Any remaining principal is then redeemed at maturity. It is common to find utility and energy companies in the United States, the United Kingdom, and the Commonwealth countries that issue bonds with sinking fund arrangements that incorporate such a provision.

Another common variation is for the bond issue to include a call provision, which gives the issuer the option to repurchase the bonds before maturity. The issuer can usually repurchase the bonds at the market price, at par, or at a specified sinking fund price, whichever is the lowest. To allocate the burden of the call provision fairly among bondholders, the bonds to be retired are selected at random based on serial number. Usually, the issuer can repurchase only a small portion of the bond issue. Some indentures, however, allow issuers to use a doubling option to repurchase double the required number of bonds.

Sinking fund arrangements have several benefits and drawbacks. The benefits include a structure that ensures a formal plan exists for debt retirement and also reduces credit risk due to a reduction of default risk at maturity. Sinking fund drawbacks include investor reinvestment risk, which is the risk associated with having to reinvest cash flows at an interest rate below the current yield-to-maturity. If the serial number of an investor's bonds is selected, the bonds will be repaid and the investor will have to reinvest the proceeds. If market interest rates have fallen since the investor purchased the bonds, the investor probably will not be able to purchase a bond offering the same return. Another potential disadvantage for investors occurs if the issuer has the option to repurchase bonds at below market prices. For example, an issuer could exercise a call option to buy back bonds at par on bonds priced above par. In this case, investors would suffer a loss.

Exhibit 5 illustrates an example of a sinking fund arrangement.

Exhibit 5

Example of a Sinking Fund Arrangement

The notional principal of the bond issue is £200 million. The sinking fund arrangement calls for 5% of the outstanding principal amount to be retired in Years 10 through 19, with the outstanding balance paid off at maturity in 20 years.

Year	Outstanding Principal at the Beginning of the Year (£ millions)	Sinking Fund Payment (£ millions)	Outstanding Principal at the End of the Year (£ millions)	Final Principal Repayment (£ millions)
0			200.00	
1 to 9	200.00	0.00	200.00	
10	200.00	10.00	190.00	
11	190.00	9.50	180.50	
12	180.50	9.03	171.48	
13	171.48	8.57	162.90	
14	162.90	8.15	154.76	
15	154.76	7.74	147.02	
16	147.02	7.35	139.67	
17	139.67	6.98	132.68	
18	132.68	6.63	126.05	
19	126.05	6.30	119.75	
20	119.75			119.75

There is no repayment of the principal during the first nine years. Starting the 10th year, the sinking fund arrangement calls for 5% of the outstanding principal amount to be retired each year. In Year 10, £10 million (5% × £200 million) are paid off, which leaves an outstanding principal balance of £190 million. In Year 11, the principal amount repaid is £9.50 million (5% × £190 million). The final repayment of the remaining balance (£119.75 million) is a balloon payment at maturity.

COUPON PAYMENT STRUCTURES

5

e describe how cash flows of fixed-income securities are structured

A coupon is the interest payment that the bond issuer makes to the bondholder. A conventional bond pays a fixed periodic coupon over a specified time to maturity. Most frequently, the coupon is paid semi-annually for sovereign and corporate bonds; this is the case in the United States, the United Kingdom, and such Commonwealth countries as Bangladesh, India, and New Zealand. Eurobonds usually pay an annual coupon, although some Eurobonds make quarterly coupon payments. Most bonds issued in the eurozone have an annual coupon.

Fixed-rate coupons are not the only coupon payment structure, however. A wide range of coupon types is offered in the global fixed-income markets to meet the differing needs of both issuers and investors.

5.1 Floating-Rate Notes

Floating-rate notes do not have a fixed coupon; instead, their coupon rate is linked to an external reference rate, such as Euribor. Thus, an FRN's interest rate will fluctuate periodically during the bond's life, following the changes in the reference rate. Therefore, the FRN's cash flows are not known with certainty. Large issuers of FRNs include government-sponsored enterprises (GSEs), such as the Federal Home Loan Banks (FHLB), the Federal National Mortgage Association ("Fannie Mae"), and the Federal Home Loan Mortgage Corporation ("Freddie Mac") in the United States, as well as banks and financial institutions in Europe and Asia Pacific. It is rare for national governments to issue FRNs because investors in sovereign bonds generally prefer fixed coupon bonds.

Almost all FRNs have quarterly coupons, although counterexamples do exist. FRNs usually pay a fixed spread over the specified reference rate. A typical coupon rate may be the three-month US dollar MRR + 20 bps (i.e., MRR + 0.20%) for a US dollar-denominated bond or the three-month Euribor + 20 bps for a euro-denominated FRN.

Contrary to plain vanilla, which are fixed-rate securities that decline in value in a rising interest rate environment, FRNs are less affected when interest rates increase because their coupon rates vary with market interest rates and are reset at regular, short-term intervals. Thus, FRNs have little interest rate risk—that is, the risk that a change in market interest rate affects a bond's value. FRNs are frequently favored by investors who expect that interest rates will rise. That said, investors still face credit risk when investing in FRNs. If an issuer's credit risk does not change from one coupon reset date to the next, the FRN's price generally will stay close to the par value. However, if there is a change in the issuer's credit quality that affects the perceived credit risk associated with the bond, the price of the FRN will deviate from its par value. A higher level of credit risk will lead to a lower price.

Additional features observed in FRNs may include a floor or a cap. A floor (floored FRN) prevents the coupon from falling below a specified minimum rate. This feature benefits the bondholders, who are guaranteed that the interest rate will not fall below the specified rate during a time of falling interest rates. In contrast, a cap (capped FRN) prevents the coupon from rising above a specified maximum rate. This feature benefits the issuer because it sets a limit to the interest rate paid on the debt during a time of rising interest rates. It is also possible to have a collared FRN, which includes both a cap and a floor.

An inverse or reverse FRN, or simply an inverse floater, is a bond whose coupon rate has an inverse relationship to the reference rate. The basic structure is the same as an ordinary FRN except for the direction in which the coupon rate is adjusted. When interest rates fall, the coupon rate on an ordinary FRN decreases; in contrast, the coupon rate on a reverse FRN increases. Thus, inverse FRNs are typically favored by investors who expect interest rates to decline.

5.2 Step-Up Coupon Bonds

The coupon of a **step-up coupon bond**, which may be fixed or floating, increases by specified margins at specified dates. An example of a bond with a step-up coupon is a ten-year callable bond issued by the Federal Home Loan Bank on 3 August 2016.

The initial coupon rate was 1.25% and steps up to 1.50% on 3 August 2018, to 2.00% on 3 August 2020, to 2.50% on 3 August 2022, to 3.00% on 3 August 2023, to 4.00% on 3 August 2024, and finally to 6.00% on 3 August 2025 for the final year. The bond was first callable at par on 3 August 2018, at the time of the first step up.

Bonds with step-up coupons offer bondholders some protection against rising interest rates, and they may be an important feature for callable bonds. When interest rates increase, there is a higher likelihood that the issuer will not call the bonds, particularly if the bonds have a fixed rate of interest. The step-up coupon allows bondholders to receive a higher coupon, in line with the higher market interest rates. Alternatively, when interest rates decrease or remain stable, the step-up feature acts as an incentive for the issuer to call the bond before the coupon rate increases and the interest expense rises.

Redeeming the bond when the coupon increases is not automatic, however; the issuer may choose to keep the bond despite its increasing cost. This may happen if refinancing the bond is necessary and alternatives are less advantageous for this issuer. For example, a financial crisis may make it difficult for the issuer to refinance. Alternatively, the issuer's credit quality may have deteriorated, which would lead to a higher yield, potentially making the coupon rate on the new bond more expensive than that on the existing bond despite the stepped-up coupon. Although the issuer does not have to call the bond, there is an implicit expectation from investors that it will do so if the market price for the bond is above the call price. Failure to do so may be viewed negatively by market participants and reduce investors' appetite for that issuer's bonds in the future.

5.3 Credit-Linked Coupon Bonds

A **credit-linked coupon bond** has a coupon that changes when the bond's credit rating changes. An example of a bond with a credit-linked coupon is one of British Telecom's bonds maturing in 2030. It was issued with a coupon rate of 8.625%, with a 25 bp increase in coupon for each credit rating downgrade below the bond's credit rating at the time of issuance and a 25 bp decrease for every credit rating upgrade above the bond's credit rating at the time of issuance.

Bonds with credit-linked coupons are attractive to investors who are concerned about the future creditworthiness of the issuer. They may also provide some general protection against a poor economy because credit ratings tend to decline the most during recessions. A potential problem associated with these bonds is that increases in the coupon payments resulting from a downgrade may ultimately result in further deteriorations of the credit rating or even contribute to the issuer's default.

5.4 Payment-in-Kind Coupon Bonds

A payment-in-kind (PIK) coupon bond typically allows the issuer to pay interest in the form of additional amounts of the bond issue rather than as a cash payment. Such bonds are favored by issuers who are concerned that the issuer may face potential cash flow problems in the future. They are used, for example, in financing companies that have a high debt burden, such as companies going through a leveraged buyout (a form of acquisition in which the financing consists primarily of debt). Investors usually demand a higher yield for holding bonds with PIK coupons in exchange for assuming their additional credit risk.

Other forms of PIK arrangements can also be found, such as paying the bondholders with common shares worth the amount of coupon due. With a PIK "toggle" note, the borrower has the option, for each interest period, to pay interest in cash, to make the interest payment in kind, or some mix of the two. Cash payments or payments in kind are frequently at the discretion of the borrower, but whether the payment is made in cash or in kind can be determined by an earnings or cash flow trigger identified in the indenture.

5.5 Deferred Coupon Bonds

A **deferred coupon bond**, sometimes called a **split coupon bond**, pays no coupons for its first few years but then pays a higher coupon than it otherwise normally would for the remainder of its life. Issuers of deferred coupon bonds are usually seeking ways to conserve cash in the years immediately following the bond issue, which may indicate poorer credit quality. Deferred coupon bonds are also common in project financing when the assets being developed do not generate any income during the development phase. A deferred coupon bond allows the issuer to delay interest payments until the project is completed, and the cash flows generated by the assets being financed can be used to service the debt.

One of the main advantages of investing in a deferred coupon bond is that these bonds are typically priced at significant discounts to par. Investors may also find the deferred coupon structure to be very helpful in managing taxes. If taxes due on the interest income can be delayed, investors may be able to minimize taxes. This tax advantage, however, depends on the jurisdiction concerned and how its tax rules apply to deferred coupon payments.

A zero-coupon bond can be thought of as an extreme form of deferred coupon bond. These securities pay no interest to the investor and thus are issued at a deep discount to par value. At maturity, the bondholder receives the par value of the bond as payment. Effectively, a zero-coupon bond defers all interest payments until maturity.

5.6 Index-Linked Bonds

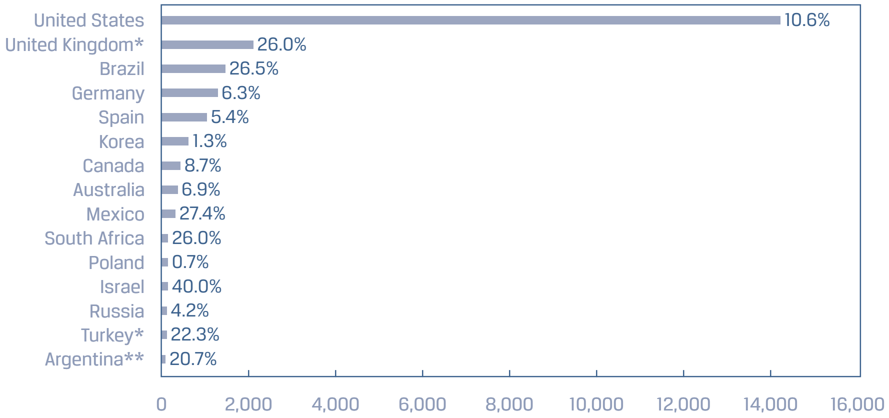
An **index-linked bond** has its coupon payments and/or principal repayment linked to a specified index. In theory, a bond can be indexed to any published variable, including an index reflecting prices, earnings, economic output, commodities, or foreign currencies. **Inflation-linked bonds** are an example of index-linked bonds. They offer investors protection against inflation by linking a bond's coupon payments and/or the principal repayment to an index of consumer prices, such as the UK Retail Price Index (RPI) or the US Consumer Price Index (CPI). The advantage of using the RPI or CPI is that these indexes are well-known, transparent, and published regularly.

Governments are large issuers of inflation-linked bonds, also called **linkers**. The United Kingdom was one of the first developed countries to issue inflation-linked bonds in 1981, offering gilts linked to the UK RPI, its main measure of the rate of inflation. In 1997, the US Treasury introduced Treasury Inflation-Protected Securities (TIPS) linked to the US Consumer Price Index (CPI). Inflation-linked bonds are now more frequently being offered by corporate issuers, including both financial and non-financial companies.

A bond's stated coupon rate represents the nominal interest rate received by the bondholders. But inflation reduces the actual value of the interest received. The interest rate that bondholders actually receive, net of inflation, is the real interest rate; it is approximately equal to the nominal interest rate minus the rate of inflation. By increasing the coupon payments and/or the principal repayment in line with increases in the price index, inflation-linked bonds reduce inflation risk. An example of an inflation-linked bond is the 1% US TIPS that matures on 15 February 2048. The coupon rate remains fixed at 1%, but the principal is adjusted every six months for changes in the CPI.

Exhibit 6 shows some of the national governments that have issued large amounts of inflation-linked bonds. It displays the amount of the linkers and the percentage of the total debt for each country. Sovereign issuers can be grouped into three categories. The first category includes such countries as Brazil and Argentina, who have issued inflation-linked bonds because they were experiencing extremely high rates of inflation when borrowing. Thus, offering inflation-linked bonds was their only available alternative to raise funds. The second category includes the United Kingdom and Australia. These countries have issued inflation-linked bonds to not only add credibility to the government's commitment to disinflationary policies but also to capitalize on investor demand. The third category, which includes the United States, Canada, and Germany, consists of national governments that are most concerned about the social welfare benefits associated with inflation-linked securities. Theoretically, inflation-linked bonds provide investors the benefit of a long-term asset with a fixed real return that is free from inflation risk.

Exhibit 6 Government Inflation-Linked Debt in Billions of USD as of Q4 2019 and the Percentage of Total Debt



** Denotes latest data from 2017.*
*** Denotes latest data from 2018.*
Source: Based on data from the Bank for International Settlements, Table C2, available at www.bis.org/statistics/secstats.htm (updated 3 June 2020).

Different methods have been used for linking the cash flows of an index-linked bond to a specified index; the link can be made via the interest payments, the principal repayment, or both. The following examples describe how the link between the cash flows and the index is established, using inflation-linked bonds as an illustration.

- *Zero-coupon-indexed bonds* pay no coupon, so the inflation adjustment is made via the principal repayment only: The principal amount to be repaid at maturity increases in line with increases in the price index during the bond's life. This type of bond has been issued in Sweden.
- *Interest-indexed bonds* pay a fixed nominal principal amount at maturity but an index-linked coupon during the bond's life. Thus, the inflation adjustment applies to the interest payments only. This is essentially a floating-rate note in which the reference rate is the inflation rate instead of an MRR, such as Euribor. These have been issued by insurance companies and major commercial banks but not typically by governments.

- *Capital-indexed bonds* pay a fixed coupon rate, but it is applied to a principal amount that increases in line with increases in the index during the bond's life. Thus, both the interest payments and the principal repayment are adjusted for inflation. Such bonds have been issued by governments in Australia, Canada, New Zealand, the United Kingdom, and the United States.
- *Indexed-annuity bonds* are fully amortized bonds, in contrast to interest-indexed and capital-indexed bonds that are non-amortizing coupon bonds. The annuity payment, which includes both payment of interest and repayment of the principal, increases in line with inflation during the bond's life. Indexed-annuity bonds linked to a price index have been issued by local governments in Australia but not by the national government.

Exhibit 7 illustrates the different methods used for inflation-linked bonds.

Exhibit 7

Examples of Inflation-Linked Bonds

Assume a hypothetical country, Lemuria, where the currency is the lemming (L). The country issued 20-year bonds linked to the domestic Consumer Price Index (CPI). The bonds have a par value of L1,000. Lemuria's economy has been free of inflation until the most recent six months, when the CPI increased by 5%.

Suppose that the bonds are zero-coupon-indexed bonds. There will never be any coupon payments. Following the 5% increase in the CPI, the principal amount to be repaid increases to L1,050 [L1,000 × (1 + 0.05)] and will continue increasing in line with inflation until maturity.

Now, suppose that the bonds are coupon bonds that make semi-annual interest payments based on an annual coupon rate of 4%. If the bonds are interest-indexed bonds, the principal amount at maturity will remain L1,000 regardless of the CPI level during the bond's life and at maturity. The coupon payments, however, will be adjusted for inflation. Prior to the increase in inflation, the semi-annual coupon payment was L20 [(0.04 × L1,000) ÷ 2]. Following the 5% increase in the CPI, the semi-annual coupon payment increases to L21 [L20 × (1 + 0.05)]. Future coupon payments will also be adjusted for inflation.

If the bonds are capital-indexed bonds, the annual coupon rate remains 4%, but the principal amount is adjusted for inflation and the coupon payment is based on the inflation-adjusted principal amount. Following the 5% increase in the CPI, the inflation-adjusted principal amount increases to L1,050 [L1,000 × (1 + 0.05)], and the new semi-annual coupon payment is L21 [(0.04 × L1,050) ÷ 2]. The principal amount will continue increasing in line with increases in the CPI until maturity, and so will the coupon payments.

If the bonds are indexed-annuity bonds, they are fully amortized. Prior to the increase in inflation, the semi-annual payment was L36.56—the annuity payment based on a principal amount of L1,000 paid back in 40 semi-annual payments with an annual discount rate of 4%. Following the 5% increase in the CPI, the annuity payment increases to L38.38 [L36.56 × (1 + 0.05)]. Future annuity payments will also be adjusted for inflation in a similar manner.

EXAMPLE 6

1 Floating-rate notes *most likely* pay:

- **A** annual coupons.
- **B** quarterly coupons.
- **C** semi-annual coupons.

2 A zero-coupon bond can *best* be considered a:

- **A** step-up bond.
- **B** credit-linked bond.
- **C** deferred coupon bond.

3 The bonds that do **not** offer protection to the investor against increases in market interest rates are:

- **A** step-up bonds.
- **B** floating-rate notes.
- **C** inverse floating-rate notes.

4 The US Treasury offers Treasury Inflation-Protected Securities (TIPS). The principal of TIPS increases with inflation and decreases with deflation based on changes in the US Consumer Price Index. When TIPS mature, an investor is paid the original principal or inflation-adjusted principal, whichever is greater. TIPS pay interest twice a year based on a fixed real coupon rate that is applied to the inflation-adjusted principal. TIPS are *most likely*:

- **A** capital-indexed bonds.
- **B** interest-indexed bonds.
- **C** indexed-annuity bonds.

5 Assume a hypothetical country, Lemuria, where the national government has issued 20-year capital-indexed bonds linked to the domestic Consumer Price Index (CPI). Lemuria's economy has been free of inflation until the most recent six months, when the CPI increased. Following the increase in inflation:

- **A** the principal amount remains unchanged but the coupon rate increases.
- **B** the coupon rate remains unchanged, but the principal amount increases.
- **C** the coupon payment remains unchanged, but the principal amount increases.

Solution to 1:

B is correct. Most FRNs pay interest quarterly and are tied to a three-month MRR.

Solution to 2:

C is correct. Because interest is effectively deferred until maturity, a zero-coupon bond can be thought of as a deferred coupon bond. A and B are incorrect because both step-up bonds and credit-linked bonds pay regular coupons. For a step-up bond, the coupon increases by specified margins at specified dates. For a credit-linked bond, the coupon changes when the bond's credit rating changes.

Solution to 3:

C is correct. The coupon rate on an inverse FRN has an inverse relationship to the market reference rate. Thus, an inverse FRN does not offer protection to the investor when market interest rates increase but rather when they decrease. A and B are incorrect because step-up bonds and FRNs both offer protection against increases in market interest rates.

Solution to 4:

A is correct. TIPS have a fixed coupon rate, and the principal is adjusted based on changes in the CPI. Thus, TIPS are an example of capital-indexed bonds. B is incorrect because with an interest-indexed bond, it is the principal repayment at maturity that is fixed and the coupon that is linked to an index. C is incorrect because indexed-annuity bonds are fully amortized bonds, not bullet bonds. The annuity payment (interest payment and principal repayment) is adjusted based on changes in an index.

Solution to 5:

B is correct. Following an increase in inflation, the coupon rate of a capital-indexed bond remains unchanged, but the principal amount is adjusted upward for inflation. Thus, the coupon payment, which is equal to the fixed coupon rate multiplied by the inflation-adjusted principal amount, increases.

6

CALLABLE AND PUTABLE BONDS

f describe contingency provisions affecting the timing and/or nature of cash flows of fixed-income securities and whether such provisions benefit the borrower or the lender

A contingency refers to some future event or circumstance that is possible but not certain. A **contingency provision** is a clause in a legal document that allows for some action if the event or circumstance does occur. For bonds, the term **embedded option** refers to various contingency provisions found in the indenture. These contingency provisions provide the issuer or the bondholders the right, but not the obligation, to take some action. These rights are called "options." These options are not independent of the bond and cannot be traded separately—hence the term "embedded." Some common types of bonds with embedded options include callable bonds, putable bonds, and convertible bonds. The options embedded in these bonds grant either the issuer or the bondholders certain rights affecting the disposal or redemption of the bond.

6.1 Callable Bonds

The most widely used embedded option is the call provision. A **callable bond** gives the issuer the right to redeem all or part of the bond before the specified maturity date. The primary reason why issuers choose to issue callable bonds rather than non-callable bonds is to protect themselves against a decline in interest rates. This decline can come either from market interest rates falling or from the issuer's credit quality improving. If market interest rates fall or credit quality improves, the issuer of a callable bond has the right to replace an old, expensive bond issue with a new, cheaper bond issue. In other words, the issuer can benefit from a decline in interest rates by being able to refinance its debt at a lower interest rate. For example, assume that the market interest rate was 6% at the time of issuance and that a company issued a bond with a coupon rate of 7%—the market interest rate plus a spread of 100 bps. Now assume that the market interest rate has fallen to 4% and that the company's creditworthiness has not changed; it can still issue at the market interest rate plus 100 bps. If the original bond is callable, the company can redeem it and replace it with a new bond paying 5% annually. If the original bond is non-callable, the company must carry on paying 7% annually and cannot benefit from the decline in market interest rates.

As illustrated in this example, callable bonds are advantageous to the issuer of the security. Put another way, the call option has value to the *issuer*. Callable bonds present investors with a higher level of reinvestment risk than non-callable bonds; that is, if the bonds are called, bondholders must reinvest funds in a lower interest rate environment. For this reason, callable bonds have to offer a higher yield and sell at a lower price than otherwise similar non-callable bonds. The higher yield and lower price compensate the bondholders for the value of the call option to the issuer.

Callable bonds have a long tradition and are commonly issued by corporate issuers. Although first issued in the US market, they are now frequently issued in every major bond market and in a variety of forms.

The details about the call provision are specified in the indenture. These details include the call price, which represents the price paid to bondholders when the bond is called. The call premium is the amount over par paid by the issuer if the bond is called. There may be restrictions on when the bond can be called, or the bond may have different call prices depending on when it is called. The call schedule specifies the dates and prices at which a bond may be called. Some callable bonds are issued with a *call protection period*, also called lockout period, cushion, or deferment period. The call protection period prohibits the issuer from calling a bond early in its life and is often added as an incentive for investors to buy the bond. The earliest time that a bond might be called is known as the call date.

Make-whole calls first appeared in the US corporate bond market in the mid-1990s and have become more commonplace ever since. A typical make-whole call requires the issuer to make a lump-sum payment to the bondholders based on the present value of the future coupon payments and outstanding principal due to early bond redemption. The discount rate used is usually a pre-determined spread over the yield-to-maturity of an appropriate sovereign bond. The typical result is a redemption value that is significantly greater than the bond's current market price. A make-whole call provision is less detrimental to bondholders than a regular call provision because it allows them to be compensated if the issuer calls the bond. Issuers, however, rarely invoke this provision because redeeming a bond that includes a make-whole provision before the maturity date is costly. Issuers tend to include a make-whole provision as a "sweetener" to make the bond issue more attractive to potential buyers and allow them to pay a lower coupon rate.

Available exercise styles on callable bonds include the following:

- *American-style* call, sometimes referred to as continuously callable, for which the issuer has the right to call a bond at any time starting on the first call date.
- *European-style* call, for which the issuer has the right to call a bond only once on the call date.
- *Bermuda-style* call, for which the issuer has the right to call bonds on specified dates following the call protection period. These dates frequently correspond to coupon payment dates.

EXAMPLE 7

Assume a hypothetical 30-year bond is issued on 15 August 2019 at a price of 98.195 (as a percentage of par). Each bond has a par value of $1,000. The bond is callable in whole or in part every 15 August from 2029 at the option of the issuer. The call prices are shown below.

Year	Call Price	Year	Call Price
2029	103.870	2035	101.548
2030	103.485	2036	101.161

(continued)

Year	Call Price	Year	Call Price
2031	103.000	2037	100.774
2032	102.709	2038	100.387
2033	102.322	2039 and thereafter	100.000
2034	101.955		

1 The call protection period is:

A 10 years.

B 11 years.

C 20 years.

2 The call premium (per $1,000 in par value) in 2033 is *closest* to:

A $2.32.

B $23.22.

C $45.14.

3 The call provision is *most likely*:

A a Bermuda call.

B a European call.

C an American call.

Solution to 1:

A is correct. The bonds were issued in 2019 and are first callable in 2029. The call protection period is 2029 – 2019 = 10 years.

Solution to 2:

B is correct. The call prices are stated as a percentage of par. The call price in 2033 is $1,023.22 (102.322% × $1,000). The call premium is the amount paid above par by the issuer. The call premium in 2033 is $23.22 ($1,023.22 – $1,000).

Solution to 3:

A is correct. The bond is callable every 15 August from 2029—that is, on specified dates following the call protection period. Thus, the embedded option is a Bermuda call.

6.2 Putable Bonds

A put provision gives the bondholders the right to sell the bond back to the issuer at a pre-determined price on specified dates. **Putable bonds** are beneficial for the bondholder by guaranteeing a pre-specified selling price at the redemption dates. If interest rates rise after issuance and bond prices fall, the bondholders can put the bond back to the issuer and reinvest the proceeds in bonds that offer higher yields, in line with higher market interest rates.

Because a put provision has value to the *bondholders*, the price of a putable bond will be higher than the price of an otherwise similar bond issued without the put provision. Similarly, the yield on a bond with a put provision will be lower than the yield on an otherwise similar non-putable bond. The lower yield compensates the issuer for the value of the put option to the investor.

The indenture lists the redemption dates and the prices applicable to the sale of the bond back to the issuer. The selling price is usually the par value of the bond. Depending on the terms set out in the indenture, putable bonds may allow buyers to force a sellback only once or multiple times during the bond's life. Putable bonds that incorporate a single sellback opportunity include a European-style put and are often referred to as one-time put bonds. Putable bonds that allow multiple sellback opportunities include a Bermuda-style put and are known as multiple put bonds. Multiple put bonds offer more flexibility for investors, so they are generally more expensive than one-time put bonds.

Typically, putable bonds incorporate one- to five-year put provisions. Their increasing popularity has often been motivated by investors wanting to protect themselves against major declines in bond prices. One benefit of this rising popularity has been an improvement in liquidity in some markets, because the put protection attracts more conservative classes of investors. The global financial crisis that started in 2008 showed that these securities can often exacerbate liquidity problems, however, because they provide a first claim on the issuer's assets. The put provision gives bondholders the opportunity to convert their claim into cash before other creditors.

CONVERTIBLE BONDS

7

f describe contingency provisions affecting the timing and/or nature of cash flows of fixed-income securities and whether such provisions benefit the borrower or the lender

A **convertible bond** is a hybrid security with both debt and equity features. It gives the bondholder the right to exchange the bond for a specified number of common shares in the issuing company. Thus, a convertible bond can be viewed as the combination of a straight bond (option-free bond) plus an embedded equity call option. Convertible bonds can also include additional provisions, the most common being a call provision.

From the investor's perspective, a convertible bond offers several advantages relative to a non-convertible bond. First, it gives the bondholder the ability to convert into equity in case of share price appreciation, and thus participate in the equity upside. At the same time, the bondholder receives downside protection; if the share price does not appreciate, the convertible bond offers the comfort of regular coupon payments and the promise of principal repayment at maturity. Even if the share price and thus the value of the equity call option decline, the price of a convertible bond cannot fall below the price of the straight bond. Consequently, the value of the straight bond acts as a floor for the price of the convertible bond.

Because the conversion provision is valuable to *bondholders*, the price of a convertible bond is higher than the price of an otherwise similar bond without the conversion provision. Similarly, the yield on a convertible bond is lower than the yield on an otherwise similar non-convertible bond. However, most convertible bonds offer investors a yield advantage; the coupon rate on the convertible bond is typically higher than the dividend yield on the underlying common share.

From the issuer's perspective, convertible bonds offer two main advantages. The first is reduced interest expense. Issuers are usually able to offer below-market coupon rates due to the conversion feature's value. The second advantage is the elimination of debt if the conversion option is exercised, but this is dilutive to existing shareholders.

Key terms regarding the conversion provision include the following:

- The **conversion price** is the price per share at which the convertible bond can be converted into shares.
- The **conversion ratio** is the number of common shares that each bond can be converted into. The indenture sometimes does not stipulate the conversion ratio but only mentions the conversion price. The conversion ratio is equal to the par

value divided by the conversion price. For example, if the par value is €1,000 and the conversion price is €20, the conversion ratio is €1,000 ÷ €20 = 50:1, or 50 common shares per bond.

- The **conversion value**, sometimes called the parity value, is the current share price multiplied by the conversion ratio. For example, if the current share price is €33 and the conversion ratio is 30:1, the conversion value is €33 × 30 = €990.
- The **conversion premium** is the difference between the convertible bond's price and its conversion value. For example, if the convertible bond's price is €1,020 and the conversion value is €990, the conversion premium is €1,020 – €990 = €30.
- Conversion parity occurs if the conversion value is equal to the convertible bond's price. Using the previous two examples, if the current share price is €34 instead of €33, then both the convertible bond's price and the conversion value are equal to €1,020 (i.e., a conversion premium equal to 0). This condition is referred to as parity. If the common share is selling for less than €34, the condition is below parity. In contrast, if the common share is selling for more than €34, the condition is above parity.

Generally, convertible bonds have maturities of five to ten years. First-time or newer issuers are usually able to issue convertible bonds of up to three years in maturity only. Although it is common for convertible bonds to reach conversion parity before maturity, bondholders rarely exercise the conversion option before that time. Early conversion would eliminate the yield advantage of continuing to hold the convertible bond; investors would typically receive in dividends less than they would receive in coupon payments. For this reason, it is common to find convertible bonds that are also callable by the issuer on a set of specified dates. If the convertible bond includes a call provision and the conversion value is above the convertible bond price, the issuer may force the bondholders to convert their bonds into common shares before maturity. For this reason, callable convertible bonds have to offer a higher yield and sell at a lower price than otherwise similar non-callable convertible bonds. Some indentures specify that the bonds can be called only if the share price exceeds a specified price, giving investors more predictability about the share price at which the issuer may force conversion.

A **warrant** is an "attached" rather than embedded option entitling the holder to buy the underlying stock of the issuing company at a fixed exercise price until the expiration date. Warrants are considered yield enhancements; they are frequently attached to bond issues as a "sweetener." Warrants are actively traded in some financial markets, such as the Deutsche Börse and the Hong Kong Stock Exchange.

Several European banks have issued a type of convertible bond called contingent convertible bonds. **Contingent convertible bonds**, nicknamed "CoCos," are bonds with contingent write-down provisions. Two main features distinguish bonds with contingent write-down provisions from the traditional convertible bonds just described. A traditional convertible bond is convertible at the option of the bondholder, and conversion occurs on the upside—that is, if the issuer's share price increases. In contrast, bonds with contingent write-down provisions are convertible on the downside. In the case of CoCos, conversion is automatic if a specified event occurs—for example, if the bank's core Tier 1 capital ratio (a measure of the bank's proportion of core equity capital available to absorb losses) falls below the minimum requirement set by the regulators. Thus, if the bank experiences losses that reduce its equity capital below the minimum requirement, CoCos are a way to reduce the bank's likelihood of default and, therefore, systemic risk—that is, the risk of failure of the financial system. When the bank's core Tier 1 capital falls below the minimum requirement, the CoCos immediately convert into equity, automatically recapitalizing the bank, lightening the debt burden, and reducing the risk of default. Because the conversion is not at the option of the bondholders but automatic, CoCos force bondholders to take losses. For this reason, CoCos must offer a higher yield than otherwise similar bonds.

EXAMPLE 8

1 Which of the following is **not** an example of an embedded option?

- **A** Warrant
- **B** Call provision
- **C** Conversion provision

2 The type of bond with an embedded option that would *most likely* sell at a lower price than an otherwise similar bond without the embedded option is a:

- **A** putable bond.
- **B** callable bond.
- **C** convertible bond.

3 The additional risk inherent to a callable bond is *best* described as:

- **A** credit risk.
- **B** interest rate risk.
- **C** reinvestment risk.

4 The put provision of a putable bond:

- **A** limits the risk to the issuer.
- **B** limits the risk to the bondholder.
- **C** does not materially affect the risk of either the issuer or the bondholder.

5 Assume that a convertible bond issued in South Korea has a par value of ₩1,000,000 and is currently priced at ₩1,100,000. The underlying share price is ₩40,000, and the conversion ratio is 25:1. The conversion condition for this bond is:

- **A** parity.
- **B** above parity.
- **C** below parity.

Solution to 1:

A is correct. A warrant is a separate, tradable security that entitles the holder to buy the underlying common share of the issuing company. B and C are incorrect because the call provision and the conversion provision are embedded options.

Solution to 2:

B is correct. The call provision is an option that benefits the issuer. Because of this, callable bonds sell at lower prices and higher yields relative to otherwise similar non-callable bonds. A and C are incorrect because the put provision and the conversion provision are options that benefit the investor. Thus, putable bonds and convertible bonds sell at higher prices and lower yields relative to otherwise similar bonds that lack those provisions.

Solution to 3:

C is correct. Reinvestment risk refers to the effect that lower interest rates have on available rates of return when reinvesting the cash flows received from an earlier investment. Because bonds are typically called following a decline in market interest rates, reinvestment risk is particularly relevant for the holder of a callable bond. A is incorrect because credit risk refers to the risk of loss resulting from the issuer failing to make full and timely payments of interest and/or repayments of principal. B is incorrect because interest rate risk is the risk that a change in market interest rate affects a bond's value.

Solution to 4:

B is correct. A putable bond limits the risk to the bondholder by guaranteeing a pre-specified selling price at the redemption dates.

Solution to 5:

C is correct. The conversion value of the bond is ₩40,000 × 25 = ₩1,000,000. The price of the convertible bond is ₩1,100,000. Thus, the conversion value of the bond is less than the bond's price, and this condition is referred to as below parity.

SUMMARY

This reading introduces the salient features of fixed-income securities while noting how these features vary among different types of securities. Important points include the following:

- The three important elements that an investor needs to know when investing in a fixed-income security are: (1) the bond's features, which determine its scheduled cash flows and thus the bondholder's expected and actual return; (2) the legal, regulatory, and tax considerations that apply to the contractual agreement between the issuer and the bondholders; and (3) the contingency provisions that may affect the bond's scheduled cash flows.
- The basic features of a bond include the issuer, maturity, par value (or principal), coupon rate and frequency, and currency denomination.
- Issuers of bonds include supranational organizations, sovereign governments, non-sovereign governments, quasi-government entities, and corporate issuers.
- Bondholders are exposed to credit risk and may use bond credit ratings to assess the credit quality of a bond.
- A bond's principal is the amount the issuer agrees to pay the bondholder when the bond matures.
- The coupon rate is the interest rate that the issuer agrees to pay to the bondholder each year. The coupon rate can be a fixed rate or a floating rate. Bonds may offer annual, semi-annual, quarterly, or monthly coupon payments depending on the type of bond and where the bond is issued.
- Bonds can be issued in any currency. Such bonds as dual-currency bonds and currency option bonds are connected to two currencies.

- The yield-to-maturity is the discount rate that equates the present value of the bond's future cash flows until maturity to its price. Yield-to-maturity can be considered an estimate of the market's expectation for the bond's return.
- A plain vanilla bond has a known cash flow pattern. It has a fixed maturity date and pays a fixed rate of interest over the bond's life.
- The bond indenture or trust deed is the legal contract that describes the form of the bond, the issuer's obligations, and the investor's rights. The indenture is usually held by a financial institution called a trustee, which performs various duties specified in the indenture.
- The issuer is identified in the indenture by its legal name and is obligated to make timely payments of interest and repayment of principal.
- For asset-backed securities, the legal obligation to repay bondholders often lies with a separate legal entity—that is, a bankruptcy-remote vehicle that uses the assets as guarantees to back a bond issue.
- How the issuer intends to service the debt and repay the principal should be described in the indenture. The source of repayment proceeds varies depending on the type of bond.
- Collateral backing is a way to alleviate credit risk. Secured bonds are backed by assets or financial guarantees pledged to ensure debt payment. Examples of collateral-backed bonds include collateral trust bonds, equipment trust certificates, mortgage-backed securities, and covered bonds.
- Credit enhancement can be internal or external. Examples of internal credit enhancement include subordination, overcollateralization, and reserve accounts. A bank guarantee, a surety bond, a letter of credit, and a cash collateral account are examples of external credit enhancement.
- Bond covenants are legally enforceable rules that borrowers and lenders agree on at the time of a new bond issue. Affirmative covenants enumerate what issuers are required to do, whereas negative covenants enumerate what issuers are prohibited from doing.
- An important consideration for investors is where the bonds are issued and traded, because it affects the laws, regulation, and tax status that apply. Bonds issued in a country in local currency are domestic bonds if they are issued by entities incorporated in the country and foreign bonds if they are issued by entities incorporated in another country. Eurobonds are issued internationally, outside the jurisdiction of any single country and are subject to a lower level of listing, disclosure, and regulatory requirements than domestic or foreign bonds. Global bonds are issued in the Eurobond market and at least one domestic market at the same time.
- Although some bonds may offer special tax advantages, as a general rule, interest is taxed at the ordinary income tax rate. Some countries also implement a capital gains tax. There may be specific tax provisions for bonds issued at a discount or bought at a premium.
- An amortizing bond is a bond whose payment schedule requires periodic payment of interest and repayment of principal. This differs from a bullet bond, whose entire payment of principal occurs at maturity. The amortizing bond's outstanding principal amount is reduced to zero by the maturity date for a fully amortized bond, but a balloon payment is required at maturity to retire the bond's outstanding principal amount for a partially amortized bond.
- Sinking fund agreements provide another approach to the periodic retirement of principal, in which an amount of the bond's principal outstanding amount is usually repaid each year throughout the bond's life or after a specified date.

- A floating-rate note, or floater, is a bond whose coupon is set based on a market reference rate (MRR) plus a spread. FRNs can be floored, capped, or collared. An inverse FRN is a bond whose coupon has an inverse relationship to the reference rate.
- Other coupon payment structures include bonds with step-up coupons, which pay coupons that increase by specified amounts on specified dates; bonds with credit-linked coupons, which change when the issuer's credit rating changes; bonds with payment-in-kind coupons, which allow the issuer to pay coupons with additional amounts of the bond issue rather than in cash; and bonds with deferred coupons, which pay no coupons in the early years following the issue but higher coupons thereafter.
- The payment structures for index-linked bonds vary considerably among countries. A common index-linked bond is an inflation-linked bond, or linker, whose coupon payments and/or principal repayments are linked to a price index. Index-linked payment structures include zero-coupon-indexed bonds, interest-indexed bonds, capital-indexed bonds, and indexed-annuity bonds.
- Common types of bonds with embedded options include callable bonds, putable bonds, and convertible bonds. These options are "embedded" in the sense that there are provisions provided in the indenture that grant either the issuer or the bondholder certain rights affecting the disposal or redemption of the bond. They are not separately traded securities.
- Callable bonds give the issuer the right to buy bonds back prior to maturity, thereby raising the reinvestment risk for the bondholder. For this reason, callable bonds have to offer a higher yield and sell at a lower price than otherwise similar non-callable bonds to compensate the bondholders for the value of the call option to the issuer.
- Putable bonds give the bondholder the right to sell bonds back to the issuer prior to maturity. Putable bonds offer a lower yield and sell at a higher price than otherwise similar non-putable bonds to compensate the issuer for the value of the put option to the bondholders.
- A convertible bond gives the bondholder the right to convert the bond into common shares of the issuing company. Because this option favors the bondholder, convertible bonds offer a lower yield and sell at a higher price than otherwise similar non-convertible bonds.

PRACTICE PROBLEMS

1 A 10-year bond was issued four years ago. The bond is denominated in US dollars, offers a coupon rate of 10% with interest paid semi-annually, and is currently priced at 102% of par. The bond's:

A tenor is six years.

B nominal rate is 5%.

C redemption value is 102% of the par value.

2 A sovereign bond has a maturity of 15 years. The bond is *best* described as a:

A perpetual bond.

B pure discount bond.

C capital market security.

3 A company has issued a floating-rate note with a coupon rate equal to the three-month MRR + 65 bps. Interest payments are made quarterly on 31 March, 30 June, 30 September, and 31 December. On 31 March and 30 June, the three-month MRR is 1.55% and 1.35%, respectively. The coupon rate for the interest payment made on 30 June is:

A 2.00%.

B 2.10%.

C 2.20%.

4 The legal contract that describes the form of the bond, the obligations of the issuer, and the rights of the bondholders can be *best* described as a bond's:

A covenant.

B indenture.

C debenture.

5 Which of the following is a type of external credit enhancement?

A Covenants

B A surety bond

C Overcollateralization

6 An affirmative covenant is *most likely* to stipulate:

A limits on the issuer's leverage ratio.

B how the proceeds of the bond issue will be used.

C the maximum percentage of the issuer's gross assets that can be sold.

7 Which of the following *best* describes a negative bond covenant? The issuer is:

A required to pay taxes as they come due.

B prohibited from investing in risky projects.

C required to maintain its current lines of business.

8 A South African company issues bonds denominated in pound sterling that are sold to investors in the United Kingdom. These bonds can be *best* described as:

A Eurobonds.

B global bonds.

C foreign bonds.

9 Relative to domestic and foreign bonds, Eurobonds are *most likely* to be:

© 2015 CFA Institute. All rights reserved.

A bearer bonds.

B registered bonds.

C subject to greater regulation.

10 An investor in a country with an original issue discount tax provision purchases a 20-year zero-coupon bond at a deep discount to par value. The investor plans to hold the bond until the maturity date. The investor will *most likely* report:

A a capital gain at maturity.

B a tax deduction in the year the bond is purchased.

C taxable income from the bond every year until maturity.

11 A bond that is characterized by a fixed periodic payment schedule that reduces the bond's outstanding principal amount to zero by the maturity date is *best* described as a:

A bullet bond.

B plain vanilla bond.

C fully amortized bond.

12 If interest rates are expected to increase, the coupon payment structure *most likely* to benefit the issuer is a:

A step-up coupon.

B inflation-linked coupon.

C cap in a floating-rate note.

13 Investors who believe that interest rates will rise *most likely* prefer to invest in:

A inverse floaters.

B fixed-rate bonds.

C floating-rate notes.

14 A 10-year, capital-indexed bond linked to the Consumer Price Index (CPI) is issued with a coupon rate of 6% and a par value of 1,000. The bond pays interest semi-annually. During the first six months after the bond's issuance, the CPI increases by 2%. On the first coupon payment date, the bond's:

A coupon rate increases to 8%.

B coupon payment is equal to 40.

C principal amount increases to 1,020.

15 The provision that provides bondholders the right to sell the bond back to the issuer at a predetermined price prior to the bond's maturity date is referred to as:

A a put provision.

B a make-whole call provision.

C an original issue discount provision.

16 Which of the following provisions is a benefit to the issuer?

A Put provision

B Call provision

C Conversion provision

17 Relative to an otherwise similar option-free bond, a:

A putable bond will trade at a higher price.

B callable bond will trade at a higher price.

C convertible bond will trade at a lower price.

18 Which type of bond *most likely* earns interest on an implied basis?

A Floater

B Conventional bond

C Pure discount bond

19 Clauses that specify the rights of the bondholders and any actions that the issuer is obligated to perform or is prohibited from performing are:

A covenants.

B collaterals.

C credit enhancements.

20 Which of the following type of debt obligation *most likely* protects bondholders when the assets serving as collateral are non-performing?

A Covered bonds

B Collateral trust bonds

C Mortgage-backed securities

21 Which of the following *best* describes a negative bond covenant? The requirement to:

A insure and maintain assets.

B comply with all laws and regulations.

C maintain a minimum interest coverage ratio.

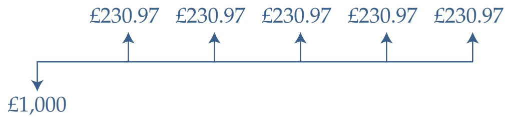
22 Contrary to positive bond covenants, negative covenants are *most likely*:

A costlier.

B legally enforceable.

C enacted at time of issue.

23 A five-year bond has the following cash flows:



The bond can *best* be described as a:

A bullet bond.

B fully amortized bond.

C partially amortized bond.

24 Investors seeking some general protection against a poor economy are *most likely* to select a:

A deferred coupon bond.

B credit-linked coupon bond.

C payment-in-kind coupon bond.

25 The benefit to the issuer of a deferred coupon bond is *most likely* related to:

A tax management.

B cash flow management.

C original issue discount price.

26 Which of the following bond types provides the *most* benefit to a bondholder when bond prices are declining?

A Callable

B Plain vanilla

C Multiple put

27 Which type of call bond option offers the *greatest* flexibility as to when the issuer can exercise the option?

A Bermuda call

B European call

C American call

28 Which of the following *best* describes a convertible bond's conversion premium?

A Bond price minus conversion value

B Par value divided by conversion price

C Current share price multiplied by conversion ratio

SOLUTIONS

1 A is correct. The tenor of the bond is the time remaining until the bond's maturity date. Although the bond had a maturity of ten years at issuance (original maturity), it was issued four years ago. Thus, there are six years remaining until the maturity date.

B is incorrect because the nominal rate is the coupon rate (i.e., the interest rate that the issuer agrees to pay each year until the maturity date). Although interest is paid semi-annually, the nominal rate is 10%, not 5%. C is incorrect because it is the bond's price, not its redemption value (also called principal amount, principal value, par value, face value, nominal value, or maturity value), that is equal to 102% of the par value.

2 C is correct. A capital market security has an original maturity longer than one year.

A is incorrect because a perpetual bond does not have a stated maturity date. Thus, the sovereign bond, which has a maturity of 15 years, cannot be a perpetual bond. B is incorrect because a pure discount bond is a bond issued at a discount to par value and redeemed at par. Some sovereign bonds (e.g., Treasury bills) are pure discount bonds, but others are not.

3 C is correct. The coupon rate that applies to the interest payment due on 30 June is based on the three-month MRR rate prevailing on 31 March. Thus, the coupon rate is 1.55% + 0.65% = 2.20%.

4 B is correct. The indenture, also referred to as trust deed, is the legal contract that describes the form of the bond, the obligations of the issuer, and the rights of the bondholders.

A is incorrect because covenants are only one element of a bond's indenture. Covenants are clauses that specify the rights of the bondholders and any actions that the issuer is obligated to perform or prohibited from performing. C is incorrect because a debenture is a type of bond.

5 B is correct. A surety bond is an external credit enhancement (i.e., a guarantee received from a third party). If the issuer defaults, the guarantor who provided the surety bond will reimburse investors for any losses, usually up to a maximum amount called the penal sum.

A is incorrect because covenants are legally enforceable rules that borrowers and lenders agree upon when the bond is issued. C is incorrect because overcollateralization is an internal, not external, credit enhancement. Collateral is a guarantee underlying the debt above and beyond the issuer's promise to pay, and overcollateralization refers to the process of posting more collateral than is needed to obtain or secure financing. Collateral, such as assets or securities pledged to ensure debt payments, is not provided by a third party. Thus, overcollateralization is not an external credit enhancement.

6 B is correct. Affirmative (or positive) covenants enumerate what issuers are required to do and are typically administrative in nature. A common affirmative covenant describes what the issuer intends to do with the proceeds from the bond issue.

A and C are incorrect because imposing a limit on the issuer's leverage ratio or on the percentage of the issuer's gross assets that can be sold are negative covenants. Negative covenants prevent the issuer from taking actions that could reduce its ability to make interest payments and repay the principal.

7 B is correct. Prohibiting the issuer from investing in risky projects restricts the issuer's potential business decisions. These restrictions are referred to as negative bond covenants.

A and C are incorrect because paying taxes as they come due and maintaining the current lines of business are positive covenants.

8 C is correct. Bonds sold in a country and denominated in that country's currency by an entity from another country are referred to as foreign bonds.

A is incorrect because Eurobonds are bonds issued outside the jurisdiction of any single country. B is incorrect because global bonds are bonds issued in the Eurobond market and at least one domestic country simultaneously.

9 A is correct. Eurobonds are typically issued as bearer bonds (i.e., bonds for which the trustee does not keep records of ownership). In contrast, domestic and foreign bonds are typically registered bonds for which ownership is recorded by either name or serial number.

B is incorrect because Eurobonds are typically issued as bearer bonds, not registered bonds. C is incorrect because Eurobonds are typically subject to lower, not greater, regulation than domestic and foreign bonds.

10 C is correct. The original issue discount tax provision requires the investor to include a prorated portion of the original issue discount in his taxable income every tax year until maturity. The original issue discount is equal to the difference between the bond's par value and its original issue price.

A is incorrect because the original issue discount tax provision allows the investor to increase his cost basis in the bond so that when the bond matures, he faces no capital gain or loss. B is incorrect because the original issue discount tax provision does not require any tax deduction in the year the bond is purchased or afterwards.

11 C is correct. A fully amortized bond calls for equal cash payments by the bond's issuer prior to maturity. Each fixed payment includes both an interest payment component and a principal repayment component such that the bond's outstanding principal amount is reduced to zero by the maturity date.

A and B are incorrect because a bullet bond or plain vanilla bond only make interest payments prior to maturity. The entire principal repayment occurs at maturity.

12 C is correct. A cap in a floating-rate note (capped FRN) prevents the coupon rate from increasing above a specified maximum rate. This feature benefits the issuer in a rising interest rate environment because it sets a limit to the interest rate paid on the debt.

A is incorrect because a bond with a step-up coupon is one in which the coupon, which may be fixed or floating, increases by specified margins at specified dates. This feature benefits the bondholders, not the issuer, in a rising interest rate environment because it allows bondholders to receive a higher coupon in line with the higher market interest rates. B is incorrect because inflation-linked bonds have their coupon payments and/or principal repayment linked to an index of consumer prices. If interest rates increase as a result of inflation, this feature is a benefit for the bondholders, not the issuer.

13 C is correct. In contrast to fixed-rate bonds that decline in value in a rising interest rate environment, floating-rate notes (FRNs) are less affected when interest rates increase because their coupon rates vary with market interest rates and are reset at regular, short-term intervals. Consequently, FRNs are favored by investors who believe that interest rates will rise.

A is incorrect because an inverse floater is a bond whose coupon rate has an inverse relationship to the reference rate, so when interest rates rise, the coupon rate on an inverse floater decreases. Thus, inverse floaters are favored by investors who believe that interest rates will decline, not rise. B is incorrect because fixed rate-bonds decline in value in a rising interest rate environment. Consequently, investors who expect interest rates to rise will likely avoid investing in fixed-rate bonds.

14 C is correct. Capital-indexed bonds pay a fixed coupon rate that is applied to a principal amount that increases in line with increases in the index during the bond's life. If the consumer price index increases by 2%, the coupon rate remains unchanged at 6%, but the principal amount increases by 2% and the coupon payment is based on the inflation-adjusted principal amount. On the first coupon payment date, the inflation-adjusted principal amount is 1,000 × (1 + 0.02) = 1,020 and the semi-annual coupon payment is equal to (0.06 × 1,020) ÷ 2 = 30.60.

15 A is correct. A put provision provides bondholders the right to sell the bond back to the issuer at a predetermined price prior to the bond's maturity date.

B is incorrect because a make-whole call provision is a form of call provision (i.e., a provision that provides the issuer the right to redeem all or part of the bond before its maturity date). A make-whole call provision requires the issuer to make a lump sum payment to the bondholders based on the present value of the future coupon payments and principal repayments not paid because of the bond being redeemed early by the issuer. C is incorrect because an original issue discount provision is a tax provision relating to bonds issued at a discount to par value. The original issue discount tax provision typically requires the bondholders to include a prorated portion of the original issue discount (i.e., the difference between the par value and the original issue price) in their taxable income every tax year until the bond's maturity date.

16 B is correct. A call provision (callable bond) gives the issuer the right to redeem all or part of the bond before the specified maturity date. If market interest rates decline or the issuer's credit quality improves, the issuer of a callable bond can redeem it and replace it by a cheaper bond. Thus, the call provision is beneficial to the issuer.

A is incorrect because a put provision (putable bond) is beneficial to the bondholders. If interest rates rise, thus lowering the bond's price, the bondholders have the right to sell the bond back to the issuer at a predetermined price on specified dates. C is incorrect because a conversion provision (convertible bond) is beneficial to the bondholders. If the issuing company's share price increases, the bondholders have the right to exchange the bond for a specified number of common shares in the issuing company.

17 A is correct. A put feature is beneficial to the bondholders. Thus, the price of a putable bond will typically be higher than the price of an otherwise similar non-putable bond.

B is incorrect because a call feature is beneficial to the issuer. Thus, the price of a callable bond will typically be lower, not higher, than the price of an otherwise similar non-callable bond. C is incorrect because a conversion feature is beneficial to the bondholders. Thus, the price of a convertible bond will typically be higher, not lower, than the price of an otherwise similar non-convertible bond.

18 C is correct. A zero-coupon, or pure discount, bond pays no interest; instead, it is issued at a discount to par value and redeemed at par. As a result, the interest earned is implied and equal to the difference between the par value and the purchase price.

19 A is correct. Covenants specify the rights of the bondholders and any actions that the issuer is obligated to perform or is prohibited from performing.

20 A is correct. A covered bond is a debt obligation backed by a segregated pool of assets called a "cover pool." When the assets that are included in the cover pool become non-performing (i.e., the assets are not generating the promised cash flows), the issuer must replace them with performing assets.

21 C is correct. Negative covenants enumerate what issuers are prohibited from doing. Restrictions on debt, including maintaining a minimum interest coverage ratio or a maximum debt usage ratio, are typical examples of negative covenants.

22 A is correct. Affirmative covenants typically do not impose additional costs to the issuer, while negative covenants are frequently costly. B is incorrect because all bond covenants are legally enforceable rules, so there is no difference in this regard between positive and negative bond covenants. C is incorrect because borrowers and lenders agree on all bond covenants at the time of a new bond issue, so there is no difference in this regard between positive and negative bond covenants.

23 B is correct. A bond that is fully amortized is characterized by a fixed periodic payment schedule that reduces the bond's outstanding principal amount to zero by the maturity date. The stream of £230.97 payments reflects the cash flows of a fully amortized bond with a coupon rate of 5% and annual interest payments.

24 B is correct. A credit-linked coupon bond has a coupon that changes when the bond's credit rating changes. Because credit ratings tend to decline the most during recessions, credit-linked coupon bonds may thus provide some general protection against a poor economy by offering increased coupon payments when credit ratings decline.

25 B is correct. Deferred coupon bonds pay no coupon for their first few years but then pay higher coupons than they otherwise normally would for the remainder of their life. Deferred coupon bonds are common in project financing when the assets being developed may not generate any income during the development phase, thus not providing cash flows to make interest payments. A deferred coupon bond allows the issuer to delay interest payments until the project is completed and the cash flows generated by the assets can be used to service the debt.

26 C is correct. A putable bond is beneficial for the bondholder by guaranteeing a prespecified selling price at the redemption date, thus offering protection when interest rates rise and bond prices decline. Relative to a one-time put bond that incorporates a single sellback opportunity, a multiple put bond offers more frequent sellback opportunities, thus providing the most benefit to bondholders.

27 C is correct. An American call option gives the issuer the right to call the bond at any time starting on the first call date.

28 A is correct. The conversion premium is the difference between the convertible bond's price and its conversion value.

READING

40

Fixed-Income Markets: Issuance, Trading, and Funding

by Moorad Choudhry, PhD, FRM, FCSI, Steven V. Mann, PhD, and Lavone F. Whitmer, CFA

Moorad Choudhry PhD, FRM, FCSI is at Recognise Bank (England). Steven V. Mann, PhD, is at the University of South Carolina (USA). Lavone F. Whitmer, CFA, is at Federal Home Loan Bank of Indianapolis (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe classifications of global fixed-income markets;
☐	**b.** describe the use of interbank offered rates as reference rates in floating-rate debt;
☐	**c.** describe mechanisms available for issuing bonds in primary markets;
☐	**d.** describe secondary markets for bonds;
☐	**e.** describe securities issued by sovereign governments;
☐	**f.** describe securities issued by non-sovereign governments, quasi-government entities, and supranational agencies;
☐	**g.** describe types of debt issued by corporations;
☐	**h.** describe structured financial instruments;
☐	**i.** describe short-term funding alternatives available to banks;
☐	**j.** describe repurchase agreements (repos) and the risks associated with them.

INTRODUCTION

1

Global fixed-income markets represent the largest subset of financial markets in terms of number of issuances and market capitalization, they bring borrowers and lenders together to allocate capital globally to its most efficient uses. Fixed-income markets include not only publicly traded securities, such as commercial paper, notes, and bonds, but also non-publicly traded loans. Although they usually attract less

© 2021 CFA Institute. All rights reserved.

attention than equity markets, fixed-income markets are more than three times the size of global equity markets. The Institute of International Finance reports that the size of the global debt market surpassed USD 253 trillion in the third quarter of 2019, representing a 322% global debt-to-GDP ratio (1 Institute of International Finance Global Debt Monitor, January 13, 2020).

Debt investors and issuers need to understand how fixed-income markets are structured and how they operate. Debt issuers have financing needs that must be met. For example, a government may need to finance an infrastructure project, a new hospital, or a new school. A start-up technology company may require funds beyond its initial seed funding from investors to expand its business and achieve scale. Financial institutions also have funding needs, and they are among the largest issuers of fixed-income securities.

Among the questions this reading addresses are the following:

- What are the key bond market sectors?
- How are bonds sold in primary markets and traded in secondary markets?
- What types of bonds are issued by governments, government-related entities, financial companies, and non-financial companies?
- What additional sources of funds are available to banks?

We first present an overview of global fixed-income markets and how these markets are classified. We identify the major issuers of and investors in fixed-income securities, present fixed-income indexes, discuss how fixed-income securities are issued in primary markets, and explore how these securities are then traded in secondary markets. We also examine different bond market sectors and discuss additional short-term funding alternatives available to banks.

2 CLASSIFICATION OF FIXED-INCOME MARKETS

a describe classifications of global fixed-income markets

b describe the use of interbank offered rates as reference rates in floating-rate debt

Although there is no standard classification of fixed-income markets, many investors and market participants use criteria to structure fixed-income markets and identify bond market sectors. This section starts by describing the most widely used ways of classifying fixed-income markets.

2.1 Classification of Fixed-Income Markets

Common criteria used to classify fixed-income markets include the type of issuer; the bonds' credit quality, maturity, currency denomination, and type of coupon; and where the bonds are issued and traded.

2.1.1 *Classification by Type of Issuer*

One way of classifying fixed-income markets is by type of issuer, which leads to the identification of four bond market sectors: households, non-financial corporates, government, and financial institutions. Exhibit 1 presents data on global debt markets at the end of the third quarter of 2019 by these sectors. Each sector is broken down into mature markets (United States, Euro area, Japan, and United Kingdom) and emerging markets. Although the sectors are of roughly comparable sizes overall, in emerging markets the debt of non-financial corporates is proportionately higher and the financial sector lower than in the mature markets.

Exhibit 1 Global Debt by Sector at End of Q3 2019 in USD

USD Trillion (% of total)	Households	Non-financial Corporates	Government	Financial Sector	Total
Mature markets	34.3	43.1	52.5	50.2	180.1
	(19.0%)	(23.9%)	(29.2%)	(27.9%)	(100.0%)
Emerging markets	13.2	31.3	16.7	11.3	72.5
	(18.2%)	(43.2%)	(23.0%)	(15.6%)	(100.0%)
Global debt	47.5	74.4	69.2	61.5	252.6
	(18.8%)	(29.5%)	(27.4%)	(24.3%)	(100.0%)

Source: Institute of International Finance Global Debt Monitor, January 2020.

Exhibit 2 displays global debt by sector as a percentage of GDP for various mature and emerging market countries at the end of the third quarter of 2019. In general, debt in proportion to GDP is higher in developed countries. However, debt takes different forms in different countries.

Exhibit 2 Global Debt by Sector as a Percentage of GDP, at End of Q3 2019

	Households	Non-financial Corporates	Government	Financial Sector	Total
United States	74.2%	74.2%	101.8%	77.1%	327.3%
Japan	55.3	101.9	226.3	157.0	540.5
United Kingdom	83.8	81.5	110.3	178.0	453.6
China	55.4	156.7	53.6	42.8	308.5
South Korea	95.1	101.6	40.2	88.8	325.7
Brazil	28.7	42.9	87.9	40.4	199.9
Mexico	16.7	26.4	35.3	16.6	95.0
Israel	41.6	69.6	60.4	10.6	182.2
Nigeria	15.6	8.3	29.2	4.3	57.4

Source: Institute of International Finance Global Debt Monitor, 13 January 2020.

2.1.2 *Classification by Credit Quality*

Investors who hold bonds are exposed to credit risk, which is the risk of loss resulting from the issuer failing to make full and timely payments of interest and/or principal. Bond markets can be classified based on the issuer's creditworthiness as judged by credit rating agencies. Ratings of Baa3 or above by Moody's Investors Service or BBB– or above by Standard & Poor's (S&P) and Fitch Ratings are considered investment grade. In contrast, ratings below these levels are referred to as non-investment grade, high yield, speculative, or "junk." An important point to understand is that credit ratings assess the issuer's creditworthiness at a certain point in time; they are not a recommendation to buy or sell the issuer's securities. Credit ratings are not static, however; they will change if a credit rating agency perceives that the probability of default for an issuer has changed.

One of the reasons why the distinction between investment-grade and high-yield bond markets matters is because institutional investors may be prohibited from investing in, or restricted in their exposure to, lower-quality or lower-rated securities. Prohibition or restriction in high-yield bond holdings generally arise because of a more restrictive risk–reward profile that forms part of the investor's investment objectives and constraints. For example, regulated banks, life insurance companies, and pension schemes are usually limited to investing in very highly rated securities. In contrast, hedge funds or some sovereign wealth funds (Qatar and Kuwait) have no formal restrictions on what type of assets they can hold or on the percentage split between bond market sectors. Globally, investment-grade bond markets tend to be more liquid than high-yield bond markets.

2.1.3 *Classification by Maturity*

Fixed-income securities can also be classified by the original maturity of the bonds when they are issued. Securities may be issued with maturities at issuance (original maturity) ranging from overnight to 30, 40, 50, and up to 100 years. For example, large companies including The Walt Disney Company and The Coca-Cola Company have issued 100-year bonds in the past. Some governments have done the same. These bonds are attractive to the issuers because they secure long-term funding. They are also attractive for many investors, such as insurance companies, that have long-term liabilities and need to buy long-duration assets.

2.1.4 *Classification by Currency Denomination*

One of the critical ways to distinguish among fixed-income securities is by currency denomination. The currency denomination of the bond's cash flows influences the yield that must be offered to investors to compensate for the potential impact of currency movements on bond investment returns. For example, if a bond is denominated in Brazilian real, its price will be primarily driven by the credit quality of the issuer, the general level of interest rates in Brazil, and the economic fundamentals of Brazil to the extent that they will affect the exchange rate, and thus *ex post* bond returns in the investor's domestic currency.

Exhibit 3 presents data on debt as a percentage of GDP for several emerging market countries for local currency (LC) and foreign currency (FC) issuances at the end of the third quarter of 2019. Government debt is almost entirely in the domestic currency, whereas corporate and bank debt differs considerably by country.

Exhibit 3 Debt as a Percentage of GDP by Local and Foreign Currency for Several Emerging Markets

	Non-financial Corporates		Government		Financial Sector	
	LC	FC	LC	FC	LC	FC
Brazil	28.1%	14.9%	87.2%	0.6%	31.9%	8.5%
India	37.9	6.3	66.3	1.9	0.9	3.1
Mexico	9.0	17.3	28.3	7.0	13.9	2.7
Turkey	30.8	36.5	17.9	17.1	5.3	22.2

Note: LC stands for "local currency"; FC stands for "foreign currency".
Source: Institute of International Finance Global Debt Monitor, 13 January 2020.

2.1.5 *Classification by Type of Coupon*

Another way of classifying fixed-income markets is by type of coupon. Some bonds pay a fixed rate of interest; others, called floating-rate bonds, floating-rate notes (FRNs), or floaters, pay a rate of interest that adjusts to market interest rates at regular, short-term intervals (e.g., quarterly).

2.1.5.1 Demand and Supply of Fixed-Rate vs. Floating-Rate Debt Balance sheet risk management considerations explain much of the demand and supply of floating-rate debt. For instance, the funding of banks—that is, the money banks raise to make loans to companies and individuals—is often short term and issued at rates that change or reset frequently. When a mismatch arises between the interest paid on the liabilities (money the bank borrowed) and the interest received on the assets (money the bank lent or invested), banks are exposed to interest rate risk—that is, the risk associated with a change in interest rate. In an effort to limit the volatility of their net worth resulting from interest rate risk, banks that issue floating-rate debt often prefer to make floating-rate loans and invest in floating-rate bonds or in other adjustable-rate assets. In addition to institutions with short-term funding needs, demand for floating-rate bonds comes from investors who believe that interest rates will rise. In that case, investors will benefit from holding floating-rate investments compared with fixed-rate ones.

On the supply side, issuance of floating-rate debt comes from institutions needing to finance short-term loans, such as consumer finance companies. Corporate borrowers also view floating-rate bonds as an alternative to using bank liquidity facilities (e.g., lines of credit), particularly when they are a lower-cost option, and as an alternative to borrowing long term at fixed rates when they expect interest rates will fall.

2.1.5.2 Reference Rates The coupon rate of a floating-rate bond is typically expressed as a reference rate plus a constant spread or margin. It is primarily a function of the issuer's credit risk at issuance: The lower the issuer's credit quality (the higher its credit risk), the higher the spread. The reference rate, however, resets periodically. Therefore, the coupon rate adjusts to the level of market interest rates each time the reference rate is reset.

Different reference rates are used depending on where the bonds are issued and their currency denomination. In the past, the **London interbank offered rate (Libor)** has long been the reference rate for many floating-rate bonds. Libor reflected the rate at which a panel of banks believed they could borrow unsecured funds from other banks. It was gathered by way of survey, and subjectivity is believed to have made its way into the process. The sidebar describes how Libor is being phased out and set to be replaced by new money market reference rates.

The Phaseout of Libor

Starting in 1986, daily Libor quotations for various currencies and maturities were set by the British Bankers' Association (BBA). Every business day, a select group of banks would submit to the BBA the rates at which they believed they could borrow from other banks in the London interbank market. Over time, the coverage grew to include 10 currencies and 15 borrowing periods. The submitted rates would be ranked from highest to lowest, and the upper and lower four submissions would be discarded. The arithmetic mean of the remaining rates became the Libor setting for a particular combination of currency and maturity. The 150 Libor quotations would then be communicated to market participants for use as reference rates in many different types of debt, including floating-rate bonds.

Problems with Libor first arose during the 2007–2009 global financial crisis, when the perceived default and liquidity risks of major international banks rose significantly. Some banks were allegedly submitting lowered rates to influence the market perception of their credit quality. In addition, some banks were allegedly altering their submitted rates to improve the valuation of their derivatives positions in contracts tied to Libor. In 2014, UK regulatory authorities transferred the administration of Libor from BBA to the International Currency Exchange (ICE). Under ICE, the size for Libor quotations was reduced to five currencies and seven maturities.

By 2017, it became apparent that the lack of activity in interbank borrowing and lending meant that Libor submissions were based more on judgment of market conditions than actual trades. Therefore, the decision was made that the panel of banks would no longer be required to submit quotations after 2021. In anticipation of the eventual demise of Libor, market participants and regulators have been working to develop new alternative money market reference rates. In the United States, it appears that the new rate will be the secured overnight financing rate (SOFR). SOFR, which has begun to be reported daily by the US Federal Reserve, is based on actual transactions in the sale and repurchase agreement ("repo") market. (Repos are described later.) In July 2018, Fannie Mae, a major institution in the US secondary mortgage market, issued USD6 billion in floating-rate notes at SOFR plus a small spread. (Note: based on "Libor: Its Astonishing Ride and How to Plan for Its End," by Liang Wu, a Numerix white paper, February 2018.)

The use of these money market reference rates extends beyond setting coupon rates for floating-rate debt. These rates are also used as reference rates for other debt instruments, including mortgages, derivatives such as interest rate and cross-currency swaps, and many other financial contracts and products. In 2018, nearly USD350 trillion in financial instruments were estimated to be tied to Libor. A major task for the future will be the transition of the terms on these contracts to the new Market Reference Rate (MRR) as Libor is phased out.

2.1.6 *Classification by Geography*

A distinction is very often made among the domestic bond, foreign bond, and Eurobond markets. Bonds issued in a specific country, denominated in the currency of that country, and sold in that country are classified as domestic bonds if they are issued by an issuer domiciled in that country and foreign bonds if they are issued by an issuer domiciled in another country. Domestic and foreign bonds are subject to the legal, regulatory, and tax requirements that apply in that particular country. In contrast, a Eurobond is issued internationally, outside the jurisdiction of the country in whose currency the bond is denominated. The Eurobond market has traditionally been characterized by fewer reporting, regulatory, and tax constraints than domestic and foreign bond markets. This less constrained environment explains why approximately 80% of entities that issue bonds outside their country of origin choose to do so in the Eurobond market rather than in a foreign bond market. In addition, Eurobonds are attractive for issuers because it enables them to reach out to more investors globally. Access to a wider pool of investors often allows issuers to raise more capital, usually at a lower cost. Along these lines, Eurobond markets are attractive for investors because they offer investors the ability to take on sovereign credit risk without being exposed to risks that are germane to the individual jurisdiction of the issuer countries, such as potential changes to capital control rules, which would restrict the return of capital, and other country risks.

Investors make a distinction between countries with established capital markets (developed markets) and countries where the capital markets are in earlier stages of development (emerging markets). For emerging bond markets, a further distinction is made between bonds issued in local currency and bonds issued in a foreign currency, such as the euro or the US dollar. All else equal, the yield offered by an individual emerging market country for a bond issued denominated in its domestic currency will be higher than the yield offered if the bond were denominated in euros or US dollars. Emerging market currencies are perceived to be riskier, and thus higher yields need to be offered to investors as compensation.

Emerging bond markets are much smaller than developed bond markets. But as demand from local and international investors has increased, issuance and trading of emerging market bonds have risen in volume and value. International investors' interest in emerging market bonds has been triggered by a desire to diversify risk across several jurisdictions in the belief that investment returns across markets are not closely correlated. In addition, emerging market bonds usually offer higher yields (return) than developed market bonds because of the higher perceived risk. Emerging countries typically lag developed countries in the areas of political stability, property rights, and contract enforcement, which often leads to a higher credit risk and higher yields. Many emerging countries, however, are less indebted than their developed counterparts and benefit from higher growth prospects, which appeals to many investors.

Composition of the bond markets varies widely by geography, with many developed and particularly emerging fixed-income markets dominated by sovereign fixed-income securities, with few bonds issued by the private sector and a predominance of bank lending. Potential bond investors may have access to private sector debt securities only indirectly by investing in a bank that holds such loans on its balance sheet. This scenario stands in contrast to markets such as the United States.

2.1.7 *Other Classifications of Fixed-Income Markets*

There are various other ways of classifying fixed-income markets. Specific market sectors that are of interest to some investors are inflation-linked bonds and, in some jurisdictions, tax-exempt bonds. Issuance of either type of bond tends to be limited to certain types of issuers. Inflation-linked bonds or linkers are typically issued by governments, government-related entities, and corporate issuers that have an investment-grade rating. They offer investors protection against inflation by linking the coupon payment and/or the principal repayment to an index of consumer prices.

Tax-exempt bonds can be issued only in those jurisdictions that recognize such tax exemption. In the United States for example, there is an income tax exemption for some of the bonds issued by governments or by some non-profit organizations. In particular, local governments can issue **municipal bonds** (or **munis**) that are tax exempt (they can also issue taxable municipal bonds, although tax-exempt munis are more frequently issued than taxable munis). Tax-exempt bonds also exist in other jurisdictions. For example, the National Highways Authority of India issues tax-exempt bonds. In countries that implement a capital gains tax, there may be tax exemptions for some types of bonds. In the United Kingdom, for example, government gilts are not subject to capital gains tax.

EXAMPLE 1

Classification of Fixed-Income Markets

1 Which of the following is *most likely* an issuer of bonds?

- **A** Hedge fund
- **B** Pension fund
- **C** Local government

2 A bond issued by a city would *most likely* be classified as a:

- **A** supranational bond.
- **B** quasi-government bond.

C non-sovereign government bond.

3 A fixed-income security issued with a maturity at issuance of nine months is *most likely* classified as a:

A capital market security.

B money market security.

C securitized debt instrument.

4 The price of a bond issued in the United States by a British company and denominated in US dollars is *most likely* to:

A change as US interest rates change.

B change as British interest rates change.

C be unaffected by changes in US and British interest rates.

5 Interbank offered rates are *best* described as the rates at which a panel of banks can:

A issue short-term debt.

B borrow unsecured funds from other major banks.

C borrow from other major banks against some form of collateral.

6 A company issues floating-rate bonds. The coupon rate is expressed as the three-month Libor plus a spread. The coupon payments are *most likely* to increase as:

A Libor increases.

B the spread increases.

C the company's credit quality decreases.

Solution to 1:

C is correct. Major issuers of bonds include sovereign (national) governments, non-sovereign (local) governments, quasi-government agencies, supranational organizations, and financial and non-financial companies. A and B are incorrect because hedge funds and pension funds are typically investors in, not issuers of, bonds.

Solution to 2:

C is correct. Non-sovereign (local) government bond issuers include provinces, regions, states, and cities. A is incorrect because supranational bonds are issued by international organizations. B is incorrect because quasi-government bonds are issued by agencies that are either owned or sponsored by governments.

Solution to 3:

B is correct. Money market securities are issued with a maturity at issuance (original maturity) ranging from overnight to one year. A is incorrect because capital market securities are issued with an original maturity longer than one year. C is incorrect because securitization, which leads to the creation of securitized debt instruments, does not relate to a bond's maturity but to the process that transforms private transactions between borrowers and lenders into securities traded in public markets.

Solution to 4:

A is correct. The currency denomination of a bond's cash flows influences which country's interest rates affect a bond's price. The price of a bond issued by a British company and denominated in US dollars will be affected by US interest rates.

Solution to 5:

B is correct. Interbank offered rates represent a set of interest rates at which major banks believe they could borrow unsecured funds from other major banks in the interbank money market for different currencies and different borrowing periods ranging from overnight to one year.

Solution to 6:

A is correct. The coupon payments of a floating-rate bond that is tied to three-month Libor will reset every three months, based on changes in Libor. Thus, as Libor increases, so will the coupon payments. B is incorrect because the spread on a floating-rate bond is typically constant; it is set when the bond is issued and does not change afterward. C is incorrect because the issuer's credit quality affects the spread and thus the coupon rate that serves as the basis for the calculation of the coupon payments, but only when the spread is set—that is, at issuance.

2.2 Fixed-Income Indexes

A fixed-income index is a multi-purpose tool used by investors and investment managers to describe a given bond market or sector, as well as to track and quantify the performance of investments and investment managers. Most fixed-income indexes are constructed as portfolios of securities that reflect a particular bond market or sector. The index construction—that is, the security selection and the index weighting—varies among indexes. Index weighting may be based on price, value (market capitalization), or total return.

There are dozens of fixed-income indexes globally, capturing different aspects of the fixed-income markets discussed earlier. One of the most popular set of indexes is the Bloomberg Barclays Global Aggregate Bond Index, which represents a broad-based measure of the global investment-grade fixed-rate bond market. It has an index history beginning on 1 January 1990 and contains three important components: the US Aggregate Bond Index (formerly Lehman Aggregate Bond Index), the Pan-European Aggregate Bond Index, and the Asian-Pacific Aggregate Bond Index. These indexes reflect the investment-grade sectors of the US, European, and Asian-Pacific bond markets, respectively.

With respect to emerging markets, one of the most widely followed indexes is the J.P. Morgan Emerging Market Bond Index (EMBI) Global, which includes US dollar-denominated Brady bonds (bonds issued primarily by Latin American countries in the late 1980s under a debt restructuring plan aimed at converting bank loans into tradable securities), Eurobonds, and loans issued by sovereign and quasi-sovereign entities in several emerging markets.

Another popular set of indexes is the FTSE Global Bond Index Series, set up to provide coverage of different classes of securities related to the government and corporate bond markets. It includes indexes of global government bonds, euro-denominated government bonds from emerging markets, sterling- and euro-denominated investment-grade corporate bonds, and covered bonds from Germany and other European Union issuers. Covered bonds are debt obligations issued by banks and backed (secured) by a segregated pool of assets.

Many other fixed-income indexes are available to investors and investment managers to measure and report performance.

It is important to note that constructing a fixed-income index is more challenging than constructing an equity index, because fixed-income markets feature far more securities of different types and more frequent changes in the index as new bonds are issued while others reach maturity.

2.3 Investors in Fixed-Income Securities

The overview of fixed-income markets has so far focused on the supply side. Before discussing bond issuers in greater detail, it is important to consider the demand side because demand for a particular type of bond or issuer may affect supply. After all, market prices are the result of the interaction between demand and supply; neither one can be considered in isolation. For example, an increase in demand for inflation-linked bonds as a result of investors' desire to protect the value of their portfolios against inflation risk may lead governments to issue a greater quantity of this type of bond. By issuing relatively more inflation-linked bonds for which there is demand, a government not only manages to sell its bond issue and obtain the funds required, but it may also benefit from a lower cost of financing.

There are different types of investors in fixed-income securities. Major categories of bond investors include central banks, institutional investors, and retail investors. The first two typically invest directly in fixed-income securities. In contrast, retail investors often invest indirectly, traditionally through fixed-income mutual funds and, during the last two decades, through exchange-traded funds (ETFs), described later.

Central banks use open market operations to implement monetary policy. **Open market operations** refer to the purchase or sale of bonds, usually but not always sovereign bonds issued by the national government. By purchasing (selling) domestic bonds, central banks increase (decrease) the monetary base in the economy. Central banks may also purchase and sell bonds denominated in foreign currencies as part of their efforts to manage the relative value of the domestic currency and their country's foreign reserves.

Institutional investors, including pension funds, hedge funds, charitable foundations and endowments, insurance companies, and banks, represent the largest groups of investors in fixed-income securities. Another major group of investors is sovereign wealth funds, which are state-owned investment funds that tend to have very long investment horizons and aim to preserve or create wealth for future generations.

Finally, retail investors often invest heavily in fixed-income securities because of the attractiveness of relatively stable prices and steady income production.

Fixed-income markets are dominated by institutional investors, in part because of the high informational barriers to entry and high minimum transaction sizes. Fixed-income securities are far more diverse than equity securities because of the variety of types of issuers and securities. In addition, unlike common shares that are primarily issued and traded in organized markets, the issuance and trading of bonds very often occurs in **over-the-counter (OTC) markets**. Thus, fixed-income securities are more difficult to access than equity securities. For these reasons, institutional investors tend to invest directly in bonds, whereas most retail investors prefer to use investment vehicles, such as mutual funds and ETFs.

EXAMPLE 2

Investors in Fixed-Income Securities

1 Open market operations describe the process used by central banks to buy and sell bonds to:

- **A** implement fiscal policy.
- **B** control the monetary base.
- **C** issue and repay government debt.

2 Retail investors *most often*:

- **A** do not invest in fixed-income securities.

B invest directly in fixed-income securities.

C invest indirectly in fixed-income securities through mutual funds or exchange-traded funds.

Solution to 1:

B is correct. Open market operations refer to the purchase or sale of bonds, usually sovereign bonds issued by the national government, as a means of implementing monetary policy. By purchasing (selling) bonds, central banks increase (decrease) the monetary base in the economy, thus controlling the money supply. A is incorrect because open market operations help facilitate monetary policy, not fiscal policy (which is the taxing and spending by the national government). C is incorrect because although Treasury departments and some central banks may facilitate the issuance and repayment of government debt, open market operations specifically refer to the implementation of monetary policy.

Solution to 2:

C is correct. Retail investors often invest in fixed-income securities because of the attractiveness of relatively stable prices and steady income production. Because most retail investors lack the expertise to value fixed-income securities and are not large enough investors to buy and sell them directly, however, they usually invest in fixed income indirectly through mutual funds and exchange-traded funds.

PRIMARY BOND MARKETS

3

c describe mechanisms available for issuing bonds in primary markets

Primary bond markets are markets in which issuers initially sell bonds to investors to raise capital. In contrast, **secondary bond markets** are markets in which existing bonds are subsequently traded among investors. As with all financial markets, primary and secondary bond markets are regulated within the framework of the overall financial system. An established independent regulatory authority is usually responsible for overseeing both the structure of the markets and the credentials of market participants.

3.1 Primary Bond Markets

Issuances in primary bond markets are frequent. Different bond issuing mechanisms are used depending on the type of issuer and the type of bond issued. A bond issue can be sold via a **public offering** (or **public offer**), in which any member of the public may buy the bonds, or via a **private placement**, in which only a selected investor, or group of investors, may buy the bonds.

3.1.1 *Public Offerings*

Investment banks play a critical role in bond issuance by assisting the issuer in accessing the primary market and by providing an array of financial services. The most common bond issuing mechanisms are underwritten offerings, best-efforts offerings, and auctions. In an **underwritten offering**, also called a **firm commitment offering**, the investment bank guarantees the sale of the bond issue at an offering price that is negotiated with the issuer. Thus, the investment bank, called the **underwriter**, takes the risk associated with selling the bonds. In contrast, in a **best-efforts offering**, the investment bank serves only as a broker. It tries to sell the bond issue at the negotiated offering price only if it is able to for a commission. Thus, the investment bank has less risk and correspondingly less incentive to sell the bonds in a best-efforts offering than in an underwritten offering. An **auction** is a bond issuing mechanism that involves bidding.

3.1.1.1 Underwritten Offerings Underwritten offerings are typical bond-issuing mechanisms for corporate bonds, some local government bonds (such as municipal bonds in the United States), and some asset-backed securities (such as mortgage-backed securities). The underwriting process typically includes six phases.

The underwriting process starts with the determination of the funding needs. Often with the help of an adviser or advisers, the issuer must determine how much money must be raised, the type of bond offering, and whether the bond issue should be underwritten.

Once the issuer has decided that the bond issue should be underwritten, it must select the underwriter, which is typically an investment bank. The underwriter of a bond issue takes the risk of buying the newly issued bonds from the issuer, and it then resells them to investors or to dealers that then sell them to investors. The difference between the purchase price of the new bond issue and the reselling price to investors is the underwriter's revenue. A relatively small-size bond issue may be underwritten by a single investment bank. It is more common for larger bond issues, however, to be underwritten by a group, or syndicate, of investment banks. In this case, the bond issue is referred to as a **syndicated offering**. A lead underwriter invites other investment banks to join the syndicate and coordinates the effort. The syndicate is collectively responsible for determining the pricing of the bond issue and for placing (selling) the bonds with investors.

The third phase of an underwritten offering is to structure the transaction. Before the bond issue is announced, the issuer and the lead underwriter discuss the terms of the bond issue, such as the bond's notional principal (total amount), the coupon rate, and the expected offering price. The underwriter or the syndicate typically organizes the necessary regulatory filings and prepares the offering circular or prospectus that provides information about the terms of the bond issue. The issuer must also choose a trustee, which is typically a trust company or the trust department of a bank, to oversee the master bond agreement. The bond offering is formally launched the day the transaction is announced, usually in the form of a press release. The announcement specifies the new bond issue's terms and conditions, including the bond's features, such as the maturity date, the currency denomination, and the expected coupon range, as well as the expected offering price. The issuer also releases the offering circular or prospectus. The final terms may differ from these terms because of changes in market conditions between the announcement day and the pricing day.

The success of the bond issue depends on the underwriter or syndicate's discernment in assessing market conditions and in pricing the bond issue accordingly. The pricing of the bond issue is, therefore, an important phase of an underwritten offering. Ideally, the bond issue should be priced so that the amount of bonds available equals investors' demand for the bonds. If the offering price is set too high, the offering will be undersubscribed—that is, there will be insufficient demand for the bond issue. As a consequence, the underwriter or syndicate will fail to sell the entire bond issue. Alternatively, if the offering price is set too low, the offering will be oversubscribed. Underwriters may aim at a small oversubscription because it reduces the risk of being unable to sell the entire bond issue. But a large oversubscription indicates that the offering terms were probably unfavorable to the issuer, in that the issuer might have raised the desired amount of capital at a lower coupon rate.

Between the announcement of a bond issue and the end of the subscription period, the underwriter or syndicate must gauge the demand for the bond issue and the price at which the bond should be offered to ensure that the entire bond issue is placed without running the risk of a large oversubscription. There are different ways for underwriters to do so. The bond issue is usually marketed to potential investors, either through an indirect approach—such as an advertisement in a newspaper, a commonly used approach for bond issued by household names—or through direct marketing and road shows, aimed at institutional investors such as pension funds and insurance companies. The underwriter or syndicate may also approach large institutional investors and discuss with them the kind of bond issues they are willing to buy. These buyers are known as the "anchor." For some, but not all, bond issues, the grey market is another way for underwriters to gauge investors' interest. The **grey market**, also called the "when issued" market, is a forward market for bonds about to be issued. Trading in the grey market helps underwriters determine what the final offering price should be.

The pricing day is the last day when investors can commit to buy the bond issue, and it is also the day when the final terms of the bond issue are agreed on. The following day, called the "offering day," the underwriting agreement that includes the bond issue's final terms is signed. The underwriting process then enters the issuing phase. The underwriter or the syndicate purchases the entire bond issue from the issuer, delivers the proceeds, and starts reselling the bonds through its sales network.

The underwriting process comes to an end about 14 days later, on the closing day, when the bonds are delivered to investors. Investors no longer receive a paper settlement; instead, the bond itself is represented by a global note that is typically held by the paying agent.

3.1.1.2 Shelf Registrations A **shelf registration** allows certain authorized issuers to offer additional bonds to the general public without having to prepare a new and separate offering circular for each bond issue. Rather, the issuer prepares a single, all-encompassing offering circular that describes a range of future bond issuances, all under the same document. This master prospectus may be in place for years before it is replaced or updated, and it may be used to cover multiple bond issuances.

Under a shelf registration, each individual offering is prefaced with a short issue announcement document. This document must confirm that there has been no change to material elements of the issuer's business, or otherwise describe any changes to the issuer's financial condition since the master prospectus was filed. Because shelf issuances are subject to a lower level of scrutiny compared with standard public offerings, they are an option only for well-established issuers that have convinced the regulatory authorities of their financial strength. Additionally, certain jurisdictions may allow shelf registrations to be purchased only by "qualified" institutional investors—that is, institutional investors that meet a set of criteria set forth by the regulators.

3.1.1.3 Auctions An auction is a method that involves bidding. It is helpful in providing price discovery (i.e., it facilitates supply and demand in determining prices) and in allocating securities. In many countries, most sovereign bonds are sold to the public via a public auction. For example, in 2017, the United States conducted 277 public auctions and issued approximately $8.5 trillion of new securities such as Treasury bills, notes, bonds, and Treasury Inflation-Protected Securities (TIPS). The public auction process used in the United States is a single-price auction through which all the winning bidders pay the same price and receive the same coupon rate for the bonds. In contrast, the public auction process used in Canada and Germany is a multiple-price auction process, which generates multiple prices and yields for the same bond issue.

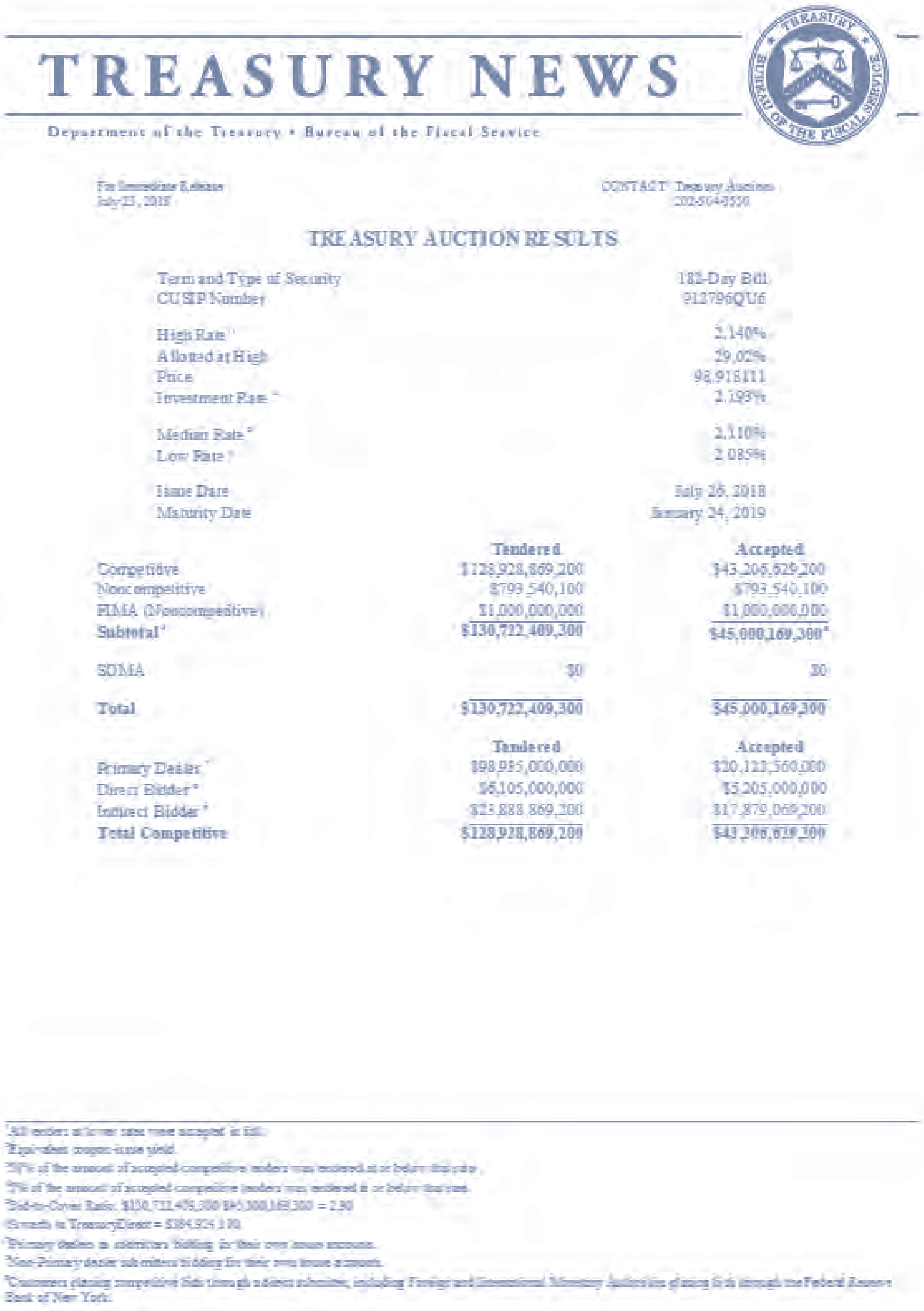
The US sovereign bond market is one of the largest and most liquid bond markets globally, so we will illustrate the US single-price auction process. This process includes three phases: announcement, bidding, and issuance. First, the US Treasury announces the auction and provides information about the bond issue, such as the amount of securities being offered, the auction date, the issue date, the maturity date, bidding close times, and other pertinent information.

After the auction announcement is made, dealers, institutional investors, and individual investors may enter competitive or non-competitive bids. With competitive bids, a bidder specifies the rate (yield) that is considered acceptable; if the rate determined at auction is lower than the rate specified in the competitive bid, the investor will not be offered any securities. In contrast, with non-competitive bids, a bidder agrees to accept the rate determined at auction; non-competitive bidders always receive their securities. At the close of the auction, the US Treasury accepts all non-competitive bids and competitive bids in ascending order of their rates (lowest to highest) until the amount of bids is equal to the amount the issuer requires. All bidders receive the same rate, based on the highest accepted bid. This single-price auction process encourages aggressive bidding and potentially results in a lower cost of funds (i.e., lower coupon rate) for the US Treasury because all the winning bidders pay the same price.

On the issue day, the US Treasury delivers the securities to the winning bidders and collects the proceeds from investors. After the auction process is complete, the securities are traded in secondary markets like other securities.

Exhibit 4 shows the results of a US Treasury public auction.

Exhibit 4 Results of a US Treasury Public Auction on 23 July 2018



TREASURY NEWS

Department of the Treasury • Bureau of the Fiscal Service

For Immediate Release
July 23, 2018

CONTACT: Treasury Auctions
202-504-3550

TREASURY AUCTION RESULTS

Term and Type of Security	182-Day Bill
CUSIP Number	912796QU6
High Rate[1]	2.140%
Allotted at High	29.02%
Price	98.918111
Investment Rate[2]	2.193%
Median Rate[3]	2.110%
Low Rate[4]	2.085%
Issue Date	July 26, 2018
Maturity Date	January 24, 2019

	Tendered	**Accepted**
Competitive	$128,928,869,200	$43,206,629,200
Noncompetitive	$793,540,100	$793,540,100
FIMA (Noncompetitive)	$1,000,000,000	$1,000,000,000
Subtotal[5]	**$130,722,409,300**	**$45,000,169,300**[6]
SOMA	$0	$0
Total	**$130,722,409,300**	**$45,000,169,300**

	Tendered	**Accepted**
Primary Dealer[7]	$98,935,000,000	$20,122,560,000
Direct Bidder[8]	$6,105,000,000	$5,205,000,000
Indirect Bidder[9]	$23,888,869,200	$17,879,069,200
Total Competitive	**$128,928,869,200**	**$43,206,629,200**

Source: Based on information from www.treasurydirect.gov.

The rate determined at auction was 2.140%. T-bills are pure discount bonds; they are issued at a discount to par value and redeemed at par. Investors paid 98.918111% of par—that is, $9,891.81 per $10,000 in par value. The US Treasury received bids for $130.7 billion but raised only $45.0 billion. All the non-competitive bids ($793.5 million) were accepted, but only a third ($43.2 of the $128.9 billion) of competitive bids were accepted. Note that half the competitive bids were submitted with a rate lower than 2.110% (the median rate). All successful bidders, however, received the rate of 2.140%, which is the essential feature of a single-price auction.

Exhibit 4 also identifies the types of bidders. Most US Treasury securities are bought at auction by primary dealers. **Primary dealers** are financial institutions that are authorized to deal in new issues of US Treasury securities. They have established business relationships with the Federal Reserve Bank of New York (New York Fed), which implements US monetary policy. Primary dealers serve primarily as trading counterparties of the New York Fed and are required to participate meaningfully in open market operations and in all auctions of US Treasury securities. They also provide the New York Fed with market information. Institutional investors and central banks are the largest investors in US Treasury securities; only a very small amount of these bonds is purchased directly by individual investors.

3.1.2 *Private Placements*

A private placement is typically a non-underwritten, unregistered offering of bonds that are sold only to an investor or a small group of investors. Typical investors in privately placed bonds are large institutional investors. A private placement can be accomplished directly between the issuer and the investor(s) or through an investment bank. Because privately placed bonds are unregistered and may be restricted securities that can be purchased by only some types of investors, there is usually no active secondary market to trade them. However, trading may be possible under certain conditions. For example, restricted securities issued under Rule 144A in the United States cannot be sold to the public, but they can be traded among qualified institutional investors. Even if trading is possible, privately placed bonds typically exhibit lower liquidity than publicly issued bonds. Insurance companies and pension funds are major buyers of privately placed bonds because they do not need every security in their portfolios to be liquid and they often value the additional yield offered by these bonds.

Private placements sometimes represent a step in the company's financing evolution between **syndicated loans** (loans from a group of lenders to a single borrower further discussed later) and public offerings. Privately placed bonds are often issued in small aggregate amounts, at times by unknown issuers. Many investors may be unwilling to undertake the credit analysis that is required for a new name, in particular if the offering amount is small. Unlike in a public offering in which the bonds are often sold to investors on a take-it-or-leave-it basis, investors in a private placement can influence the structure of the bond issue, including such considerations as asset and collateral backing, credit enhancements, and covenants. It is common for privately placed bonds to have more customized and restrictive covenants than publicly issued ones. In addition to being able to negotiate the terms of the bonds and align those terms with their needs, investors in private placements are rewarded by receiving the bonds, which is not always the case in public offerings in which investors cannot know for sure when the issue will become available and how many securities they will be allocated.

Private placements are also offered by regular bond issuers, in particular for smaller amounts of capital raised in major currencies, such as US dollars, euros, or sterling. Private placements are usually more flexible than public offerings and allow regular issuers to tailor the bond issue to their own needs.

4

SECONDARY BOND MARKETS

d describe secondary markets for bonds

Secondary markets, also called the "aftermarket," are where existing securities are traded among investors. Securities can be traded directly from investor to investor, through a broker or dealer, or directly from issuer to investor by way of bond tender offer. The major participants in secondary bond markets globally are large institutional investors and central banks. The direct presence of retail investors in secondary bonds markets is limited, unlike in secondary equity markets. Retail investors do participate in secondary bond markets, albeit indirectly, through mutual and exchange-traded funds (ETFs) offered by fund management companies.

There are three main ways for secondary markets to be structured: as an organized exchange, as an over-the-counter market, or as a bond tender offer. An **organized exchange** provides a place where buyers and sellers can meet to arrange their trades. Although buy or sell orders may come from anywhere, the transaction must take place at the exchange according to the rules imposed by the exchange. In contrast, with **over-the-counter (OTC) markets**, buy and sell orders initiated from various locations are matched through a communications network. Thus, OTC markets need electronic trading platforms over which users submit buy and sell orders. Bloomberg Fixed Income Electronic Trading platform is an example of such a platform through which dealers stand ready to trade in multiple bond markets globally. Although there is some trading of government bonds and very active corporate bonds on many stock exchanges around the world, the vast majority of bonds are traded in OTC markets. A bond tender offer involves an issuer making an offer to existing bondholders of the company to repurchase a specified number of bonds at a particular price and a specified time. Companies use such offers to restructure or refinance their current capital structure. Investors generally receive a price that is above the market value of the bond, thereby enticing them to participate; yet this price is likely below par, and hence also a win for the issuer. Note, however, that the price at which the bonds are repurchased is not determined in advance at the time of the original issuance of the bond. It merely reflects the market conditions at the time of the repurchase.

The liquidity demands of fixed-income investors have evolved since the early 1990s. The type of investors that would buy and hold a bond to maturity, which once dominated the fixed-income markets, has been supplanted by institutional investors who trade actively. The dynamics of global fixed-income markets reflect this change in the relative demand for liquidity.

We will illustrate how secondary markets work by using the example of Eurobonds. The most important Eurobond trading center by volume is in London, although a large number of market participants are also based in Brussels, Frankfurt, Zurich, and Singapore. Liquidity is supplied by Eurobond market makers, of which approximately 35 are registered with the International Capital Market Association (ICMA). ICMA is an association of banks and other financial institutions that provides a regulatory framework for international bond markets and that is behind much of the established uniform practices observed by all market participants in the Eurobond market.

The key to understanding how secondary bond markets are structured and function is to understand liquidity. Liquidity refers to the ability to trade (buy or sell) securities quickly and easily at prices close to their fair market value. Liquidity involves much more than "how quickly one can turn a bond into cash." This statement implicitly assumes a long position, but some market participants need to buy quickly when covering a short position. The other aspect of liquidity that is often ignored is that speed of trading alone does not constitute a liquid market. One can always buy something quickly by offering a very high price or sell something quickly by accepting a very low price. In a liquid market, trading takes place quickly at prices close to the security's fair market value.

A key indicator and measurement of liquidity is the **bid–offer spread** or **bid–ask spread**, which reflects the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It can be as low as 5 bps for very liquid bond issues, such as issues of the World Bank, to no price quoted for illiquid issues. A reasonable spread is of the order of 10 bps to 12 bps, whereas an illiquid spread may be in excess of 50 bps. When there is no bid or offer price, the issue is completely illiquid for trading purposes.

Settlement is the process that occurs after the trade is made. The bonds are passed to the buyer, and payment is received by the seller. Secondary market settlement for government and quasi-government bonds typically takes place either on a cash basis or on a $T + 1$ basis. With cash settlement, trading and settlement occur on the same day. With $T + 1$ settlement, settlement takes place the day after the trade date. In contrast, corporate bonds usually settle on a $T + 2$ or $T + 3$ basis, although settlement can extend to $T + 7$ in some jurisdictions. Trades clear within either or both of the two main clearing systems, Euroclear and Clearstream. Settlement occurs by means of a simultaneous exchange of bonds for cash on the books of the clearing system. An electronic bridge connecting Euroclear and Clearstream allows transfer of bonds from one system to the other, so it is not necessary to have accounts at both systems. Both systems operate on a paperless, computerized book-entry basis, although a bond issue is still represented by a physical document, the global note mentioned earlier. All participants in either system will have their own internal account set up, and they may also act as agent for buyers or sellers who do not possess an account.

EXAMPLE 3

Bond Markets

1 Which of the following *best* describes a primary market for bonds? A market:

- **A** in which bonds are issued for the first time to raise capital.
- **B** that has a specific location where the trading of bonds takes place.
- **C** in which existing bonds are traded among individuals and institutions.

2 US Treasury bonds are typically sold to the public via a(n):

- **A** auction.
- **B** primary dealer.
- **C** secondary bond market.

3 In a single-price bond auction, an investor who places a competitive bid and specifies a rate that is above the rate determined at auction will *most likely*:

- **A** not receive any bonds.
- **B** receive the bonds at the rate determined at auction.
- **C** receive the bonds at the rate specified in the investor's competitive bid.

4 A bond purchased in a secondary market is *most likely* purchased from:

- **A** the bond's issuer.
- **B** the bond's lead underwriter.
- **C** another investor in the bond.

5 Corporate bonds will *most likely* settle:

- **A** on the trade date.
- **B** on the trade date plus one day.
- **C** by the trade date plus three days.

Solution to 1:

A is correct. Primary bond markets are markets in which bonds are issued for the first time to raise capital. B is incorrect because having a specific location where the trading of bonds takes place is not a requirement for a primary bond market. C is incorrect because a market in which existing bonds are traded among individuals and institutions is the definition of a secondary, not primary, market.

Solution to 2:

A is correct. US Treasury bonds are typically sold to the public via an auction. B is incorrect because primary dealers are often bidders in the auction; they are financial institutions that are active in trading US Treasury bonds. C is incorrect because any bond issue coming directly to the market is considered to be in the primary, not the secondary, market.

Solution to 3:

A is correct. In a single-price bond auction, a bidder that enters a competitive bid specifies the rate (yield) that is considered acceptable. If the rate specified in the competitive bid is above the rate (yield) determined at auction, the investor will not be offered any securities.

Solution to 4:

C is correct. Secondary bond markets are where bonds are traded among investors. A and B are incorrect because a bond purchased from the bond's issuer or from the bond's lead underwriter would happen in the primary, not secondary, market.

Solution to 5:

C is correct. Corporate bonds typically settle on a $T + 2$ or $T + 3$ basis—that is, two or three days after the trade date—although settlement can extend to $T + 7$ in some jurisdictions. A and B are incorrect because it is government and quasi-government bonds, not corporate bonds, that typically settle either on a cash basis or on a $T + 1$ basis.

SOVEREIGN BONDS

5

e describe securities issued by sovereign governments

National governments issue bonds primarily for fiscal reasons—to fund spending when tax revenues are insufficient to cover expenditures. To meet their spending goals, national governments issue bonds in various types and amounts. This section discusses bonds issued by national governments, often referred to as **sovereign bonds** or sovereigns.

5.1 Characteristics of Sovereign Bonds

Sovereign bonds denominated in local currency have different names in different countries. For example, they are named US Treasuries in the United States, Japanese government bonds (JGBs) in Japan, gilts in the United Kingdom, Bunds in Germany, and obligations assimilables du Trésor (OATs) in France. Some investors or market participants may refer to sovereign bonds as Treasury securities or Treasuries for short, on the principle that the national Treasury department is often in charge of managing a national government's funding needs.

Names may also vary depending on the original maturity of the sovereign bond. For example, US government bonds are named Treasury bills (T-bills) when the original maturity is one year or shorter, Treasury notes (T-notes) when the original maturity is longer than one year and up to 10 years, and Treasury bonds (T-bonds) when the original maturity is longer than 10 years. In Spain, the sovereigns issued by Tesoro Público are named letras del Tesoro, bonos del Estado, and obligaciones del Estado, depending on the sovereign's original maturity: one year or shorter, longer than one year and up to five years, or longer than five years, respectively. Although very rare, some bonds, such as the consols in the United Kingdom, have no stated maturity date.

The majority of the trading in secondary markets is of sovereign securities that were most recently issued, called **on-the-run** securities. The latest sovereign bond issue for a given maturity is also referred to as a **benchmark issue** because it serves as a benchmark against which to compare bonds that have the same features (i.e., maturity, coupon type and frequency, and currency denomination) but that are issued by another type of issuer (e.g., non-sovereign, corporate). As a general rule, as sovereign securities age, they trade less frequently.

One salient difference between money market securities, such as T-bills, and capital market securities, such as T-notes and T-bonds, is the interest provision. T-bills are pure discount bonds; they are issued at a discount to par value and redeemed at par. The difference between the par value and the issue price is the interest paid on the borrowing. In contrast, capital market securities are typically coupon (or coupon-bearing) bonds; these securities make regular coupon payments and repay the par value at maturity. Bunds pay coupons annually, whereas US Treasuries, JGBs, gilts, and OATs make semi-annual coupon payments.

5.2 Credit Quality of Sovereign Bonds

Sovereign bonds are usually unsecured obligations of the sovereign issuer—that is, they are backed not by collateral but by the taxing authority of the national government. When a national government runs a budget surplus, excess tax revenues over expenditures are the primary source of funds for making interest payments and repaying the principal. In contrast, when a country runs a budget deficit, the source of the funds used for the payment of interest and repayment of principal comes either from tax revenues and/or by "rolling over" (refinancing) existing debt into new debt.

Highly rated sovereign bonds denominated in local currency are virtually free of credit risk. Credit rating agencies assign ratings to sovereign bonds, and these ratings are called sovereign ratings. The highest rating (i.e., highest credit quality and lowest credit risk) is AAA by S&P and Fitch and Aaa by Moody's. In 2018, only a handful of sovereign issuers were rated at this (theoretically) risk-free level by these three credit rating agencies, including Germany, Singapore, Canada, Sweden, Norway, Denmark, Luxembourg, Australia, Switzerland, and the Netherlands. Notably, S&P downgraded the United States from AAA to AA+ in 2011 and downgraded the United Kingdom from AAA to AA, two notches, in 2016 following the referendum that approved the United Kingdom's exit from the European Union.

Credit rating agencies distinguish between bonds issued in the sovereign's local currency and bonds issued in a foreign currency. In theory, a government can make interest payments and repay the principal by generating cash flows from its unlimited power (in the short run at least) to tax its citizens. A national government also has the ability to print its own currency, whereas it is restricted in being able to pay in a foreign currency only what it earns in exports or can exchange in financial markets. Thus, it is common to observe a higher credit rating for sovereign bonds issued in local currency than for those issued in a foreign currency. But there are limits to a government's ability to reduce the debt burden. As the sovereign debt crisis that followed the global financial crisis has shown, taxing citizens can only go so far in paying down debt before the taxation becomes an economic burden. Additionally, printing money only serves to weaken a country's currency relative to other currencies over time.

The national government of a country that has a strong domestic savings base has the luxury of being able to issue bonds in its local currency and sell them to domestic investors. If the local currency is liquid and freely traded, the sovereign issuer may also attract international investors who may want to hold that sovereign issuer's bonds and have exposure to that country's local currency. A national government may also issue debt in a foreign currency when there is demand for the sovereign issuer's bonds but not necessarily in the sovereign's local currency. For example, demand from overseas investors has caused national governments such as Switzerland and Sweden to issue sovereign bonds in US dollars and euros. Emerging market countries may also have to issue in major currencies because international investors may be willing to accept the credit risk but not the foreign exchange (currency) risk associated with emerging market bonds. The yield offered for bonds issued in major currencies will be lower because of the lower risk profile. When a sovereign issuer raises debt in a foreign currency, it usually swaps the proceeds into its local currency for use and deployment.

5.3 Types of Sovereign Bonds

National governments issue different types of bonds, some of them paying a fixed rate of interest and others paying a floating rate, including inflation-linked bonds.

5.3.1 *Fixed-Rate Bonds*

Fixed-rate bonds (i.e., bonds that pay a fixed rate of interest) are by far the most common type of sovereign bond. National governments routinely issue two types of fixed-rate bonds: zero-coupon bonds (or pure discount bonds) and coupon bonds. A zero-coupon bond does not pay interest. Instead, it is issued at a discount to par value and redeemed at par at maturity. Coupon bonds are issued with a stated rate of interest and make interest payments periodically, such as semi-annually or annually. They have a terminal cash flow equal to the final interest payment plus the par value. As mentioned earlier, most sovereign bonds with an original maturity of one year or less are zero-coupon bonds, whereas bonds with an original maturity longer than one year are typically issued as coupon bonds.

5.3.2 *Floating-Rate Bonds*

The price of a bond changes in the opposite direction from the change in interest rates, a relationship that is fully explained later when discussing the risk and return of fixed-income securities. Thus, investors who hold fixed-rate bonds are exposed to interest rate risk: As interest rates increase, bond prices decrease, which lowers the value of their portfolio. In response to public demand for less interest rate risk, some national governments around the world issue bonds with a floating rate of interest that resets periodically based on changes in the level of a reference rate such as Libor. Although interest rate risk still exists on floating-rate bonds, it is far less pronounced than that on fixed-rate bonds.

Examples of countries where the national government issues floating-rate bonds include Germany, Spain, and Belgium in developed markets and Brazil, Turkey, Mexico, Indonesia, and Poland in emerging markets. The largest sovereign issuer, the United States, began issuing floating-rate bonds in January 2014. Two other large sovereign issuers, Japan and the United Kingdom, have never issued bonds whose coupon rate is tied to a reference rate.

5.3.3 *Inflation-Linked Bonds*

Fixed-income investors are exposed to inflation risk. The cash flows of fixed-rate bonds are fixed by contract. If a particular country experiences an inflationary episode, the purchasing power of the fixed cash flows erodes over time. Thus, to respond to the demand for less inflation risk, many national governments issue inflation-linked bonds, or linkers, whose cash flows are adjusted for inflation. First issuers of inflation-linked bonds were the governments of Argentina, Brazil, and Israel. The United States introduced inflation-linked securities in January 1997, calling them Treasury Inflation-Protected Securities (TIPS). Other countries where the national government has issued inflation-linked bonds include the United Kingdom, Sweden, Australia, and Canada in developed markets and Brazil, South Africa, and Chile in emerging markets.

As explained in the earlier coverage of the defining elements of fixed-income securities, the index to which the coupon payments and/or principal repayments are linked is typically an index of consumer prices. Inflation-linked bonds can be structured a variety of ways: The inflation adjustment can be made via the coupon payments, the principal repayment, or both. In the United States, the index used is the Consumer Price Index for All Urban Consumers (CPI-U). In the United Kingdom, it is the Retail Price Index (RPI) (All Items). In France, there are two inflation-linked bonds with two different indexes: the French consumer price index (CPI) (excluding tobacco) and the Eurozone's Harmonized Index of Consumer Prices (HICP) (excluding tobacco). Although linking the cash flow payments to a consumer price index reduces inflation risk, it does not necessarily eliminate the effect of inflation completely because the consumer price index may be an imperfect proxy for inflation.

EXAMPLE 4

Sovereign Bonds

1 Sovereign debt with a maturity at issuance shorter than one year *most likely* consists of:

- **A** floating-rate instruments.
- **B** zero-coupon instruments.
- **C** coupon-bearing instruments.

2 Floating-rate bonds are issued by national governments as the *best* way to reduce:

- **A** credit risk.
- **B** inflation risk.
- **C** interest rate risk.

3 Sovereign bonds whose coupon payments and/or principal repayments are adjusted by a consumer price index are *most likely* known as:

- **A** linkers.
- **B** floaters.
- **C** consols.

Solution to 1:

B is correct. Most debt issued by national governments with a maturity at issuance (original maturity) shorter than one year takes the form of zero-coupon instruments. A and C are incorrect because floating-rate and coupon-bearing instruments are typically types of sovereign debt with maturities longer than one year.

Solution to 2:

C is correct. The coupon rates of floating-rate bonds are reset periodically based on changes in the level of a reference rate such as Libor, which reduces interest rate risk. A is incorrect because credit risk, although low for sovereign bonds, cannot be reduced by linking the coupon rate to a reference rate. B is incorrect because although inflation risk is lower for floating-rate bonds than for fixed-rate bonds, floating-rate bonds are not as good as inflation-linked bonds to reduce inflation risk.

Solution to 3:

A is correct because sovereign bonds whose coupon payments and/or principal repayment are adjusted by a consumer price index are known as inflation-linked bonds or linkers. B is incorrect because floaters describe floating-rate bonds that have a coupon rate tied to a reference rate such as Libor. C is incorrect because consols are sovereign bonds with no stated maturity date issued by the UK government.

NON-SOVEREIGN, QUASI-GOVERNMENT, AND SUPRANATIONAL BONDS

6

f describe securities issued by non-sovereign governments, quasi-government entities, and supranational agencies

This section covers the bonds issued by local governments and by government-related entities.

6.1 Non-Sovereign Bonds

Levels of government below the national level, such as provinces, regions, states, and cities, issue bonds called non-sovereign government bonds or **non-sovereign bonds**. These bonds are typically issued to finance public projects, such as schools, motorways, hospitals, bridges, and airports. The sources for paying interest and repaying the principal include the taxing authority of the local government, the cash flows of the project the bond issue is financing, or special taxes and fees established specifically for making interest payments and principal repayments. Non-sovereign bonds are typically not guaranteed by the national government.

As mentioned earlier, bonds issued by state and local governments in the United States are known as municipal bonds, and they often offer income tax exemptions. In the United Kingdom, non-sovereign bonds are known as local authority bonds. Other non-sovereign bonds include those issued by state authorities such as the 16 Länder in Germany.

Credit ratings for non-sovereign bonds vary widely because of the differences in credit and collateral quality. Because default rates of non-sovereign bonds are historically low, they very often receive high credit ratings. However, non-sovereign bonds usually trade at a higher yield and lower price than sovereign bonds with similar characteristics. The additional yield depends on the credit quality, the liquidity of the bond issue, and the implicit or explicit level of guarantee or funding commitment from the national government. The additional yield is the lowest for non-sovereign bonds that have high credit quality, are liquid, and are guaranteed by the national government.

6.2 Quasi-Government Bonds

National governments establish organizations that perform various functions for them. These organizations often have both public and private sector characteristics, but they are not actual governmental entities. They are referred to as quasi-government entities, although they take different names in different countries. These quasi-government entities often issue bonds to fund specific financing needs. These bonds are known as **quasi-government bonds** or **agency bonds**.

Examples of quasi-government entities include government-sponsored enterprises (GSEs) in the United States, such as the Federal National Mortgage Association ("Fannie Mae"), the Federal Home Loan Mortgage Corporation ("Freddie Mac"), and the Federal Home Loan Bank (FHLB). Other examples of quasi-government entities that issue bonds include Hydro Quebec in Canada or the Japan Bank for International Cooperation (JBIC). In the case of JBIC's bonds, timely payments of interest and repayment of principal are guaranteed by the Japanese government. Most quasi-government bonds, however, do not offer an explicit guarantee by the national government, although investors often perceive an implicit guarantee.

Because a quasi-government entity typically does not have direct taxing authority, bonds are repaid from the cash flows generated by the entity or from the project the bond issue is financing. Quasi-government bonds may be backed by collateral, but this is not always the case. Quasi-government bonds are usually rated very high by the credit rating agencies because historical default rates are extremely low. Bonds that are guaranteed by the national government receive the highest ratings and trade at a lower yield and higher price than otherwise similar bonds that are not backed by the sovereign government's guarantee.

6.3 Supranational Bonds

A form of often highly rated bonds is issued by supranational agencies, also referred to as multilateral agencies. The most well-known supranational agencies are the International Bank for Reconstruction and Development (the World Bank), the International Monetary Fund (IMF), the European Investment Bank (EIB), the Asian Development Bank (ADB), and the African Development Bank (AFDB). Bonds issued by supranational agencies are called **supranational bonds**.

Supranational bonds are typically plain-vanilla bonds, although floating-rate bonds and callable bonds are sometimes issued. Highly rated supranational agencies, such as the World Bank, frequently issue large-size bond issues that are often used as benchmarks issues when there is no liquid sovereign bond available.

EXAMPLE 5

Non-Sovereign Government, Quasi-Government, and Supranational Bonds

1 Relative to sovereign bonds, non-sovereign bonds with similar characteristics *most likely* trade at a yield that is:

A lower.

B the same.

C higher.

2 Bonds issued by a governmental agency are *most likely*:

A repaid from the cash flows generated by the agency.

B guaranteed by the national government that sponsored the agency.

C backed by the taxing power of the national government that sponsored the agency.

Solution to 1:

C is correct. Non-sovereign bonds usually trade at a higher yield and lower price than sovereign bonds with similar characteristics. The higher yield is because of the higher credit risk associated with non-sovereign issuers relative to sovereign issuers, although default rates of local governments are historically low and their credit quality is usually high. The higher yield may also be a consequence of non-sovereign bonds being less liquid than sovereign bonds with similar characteristics.

Solution to 2:

A is correct. Most bonds issued by a governmental agency are repaid from the cash flows generated by the agency or from the project the bond issue is financing. B and C are incorrect because although some bonds issued by governmental agencies are guaranteed by the national government or are backed by the taxing power of the national government that sponsored the agency, bonds are most likely repaid first from the cash flows generated by the agency.

CORPORATE DEBT: BANK LOANS, SYNDICATED LOANS, AND COMMERCIAL PAPER

7

g describe types of debt issued by corporations

Companies differ from governments and government-related entities in that their primary goal is profit; profitability is an important consideration when companies make decisions, including financing decisions. Companies routinely raise debt as part of their overall capital structure, both to fund short-term spending needs (e.g., working capital) as well as long-term capital investments. We have so far focused on publicly issued debt, but loans from banks and other financial institutions are a significant part of the debt raised by companies. For example, it is estimated that European companies meet 75% of their borrowing needs from banks and only 25% from financial markets. In Japan, these percentages are 80% and 20%, respectively. In the United States, however, debt capital is much more significant: 80% is from financial markets and just 20% from bank lending (SIFMA 2018 Outlook: Through the Looking Glass).

7.1 Bank Loans and Syndicated Loans

A **bilateral loan** is a loan from a single lender to a single borrower. Companies routinely use bilateral loans from their banks, and these bank loans are governed by the bank loan documents. Bank loans are the primary source of debt financing for small and medium-size companies as well as for large companies in countries where bond markets are either underdeveloped or where most bond issuances are from government, government-related entities, and financial institutions. Access to bank loans depends not only on the characteristics and financial health of the company but also on market conditions and bank capital availability.

A syndicated loan is a loan from a group of lenders, called the "syndicate," to a single borrower. A syndicated loan is a hybrid between relational lending and publicly traded debt. Syndicated loans are primarily originated by banks, and the loans are extended to companies but also to governments and government-related entities. The coordinator, or lead bank, originates the loan, forms the syndicate, and processes the payments. In addition to banks, a variety of lenders participate in the syndicate, such as pension funds, insurance companies, and hedge funds. Syndicated loans are a way for these institutional investors to participate in corporate lending while diversifying the credit risk among a group of lenders.

There is also a secondary market in syndicated loans. These loans are often packaged and securitized, and the securities created are then sold in secondary markets to investors.

Most bilateral and syndicated loans are floating-rate loans, and the interest rate is based on a reference rate plus a spread. The reference rate may be MRR, formerly Libor, a sovereign rate (e.g., the T-bill rate), or the prime lending rate, also called the "prime rate." The prime rate reflects the rate at which banks lend to their most creditworthy customers, which tends to vary based on the overnight rate at which banks lend to each other plus a spread. Bank loans can be customized to the borrower's needs. They can have different maturities, as well as different interest payment and principal repayment structures. The frequency of interest payments varies among bank loans. Some loans are bullet loans, in which the entire payment of principal occurs at maturity, and others are amortizing loans, in which the principal is repaid over time.

For highly rated companies, both bilateral and syndicated loans can be more expensive than bonds issued in financial markets. Thus, companies often turn to money and capital markets to raise funds, which allows them to diversify their sources of financing.

7.2 Commercial Paper

Commercial paper is a short-term, unsecured promissory note issued in the public market or via a private placement that represents a debt obligation of the issuer. Commercial paper was first issued in the United States more than a century ago. It later appeared in the United Kingdom, in other European countries, and then in the rest of the world.

7.2.1 *Characteristics of Commercial Paper*

Commercial paper is a valuable source of flexible, readily available, and relatively low-cost short-term financing. It is a source of funding for working capital and seasonal demands for cash. It is also a source of **bridge financing**—that is, interim financing that provides funds until permanent financing can be arranged. Suppose a company wants to build a new distribution center in southeast China and wants to finance this investment with an issuance of long-term bonds. The market conditions for issuing long-term bonds may currently be volatile, which would translate into a higher cost of borrowing. Rather than issuing long-term bonds immediately, the company may opt to raise funds with commercial paper and wait for a more favorable environment in which to sell long-term bonds.

The largest issuers of commercial paper are financial institutions, but some non-financial companies are also regular issuers of commercial paper. Although the focus of this section is on corporate borrowers, sovereign governments and supranational agencies routinely issue commercial paper as well.

The maturity of commercial paper can range from overnight to one year, but a typical issue matures in less than three months.

7.2.2 *Credit Quality of Commercial Paper*

Traditionally, only the largest, most stable companies issued commercial paper. Although only the strongest, highest-rated companies issue low-cost commercial paper, issuers from across the risk spectrum can issue commercial paper with higher yields than higher-rated companies. Thus, investors in commercial paper are exposed to various levels of credit risk depending on the issuer's creditworthiness. Many investors perform their own credit analysis, but most investors also assess a commercial paper's credit quality by using the ratings provided by credit rating agencies. Exhibit 5 presents the range of commercial paper ratings from the main credit rating agencies. Commercial paper rated adequate or above (the shaded area of Exhibit 5) is called "prime paper," and it is typically considered investment grade by investors.

Exhibit 5 Commercial Paper Ratings

Credit Quality	Moody's	S&P	Fitch
Superior	P1	A1+/A1	F1+/F1
Satisfactory	P2	A2	F2
Adequate	P3	A3	F3
Speculative	NP	B/C	F4
Defaulted	NP	D	F5

In most cases, maturing commercial paper is paid with the proceeds of new issuances of commercial paper, a practice referred to as "rolling over the paper." This practice creates a risk that the issuer will be unable to issue new paper at maturity, referred to as rollover risk. As a safeguard against rollover risk, credit rating agencies often require that commercial paper issuers secure a **backup line of credit** from banks. The purpose of the backup lines of credit is to ensure that the issuer will have access to sufficient liquidity to repay maturing commercial paper if rolling over the paper is not a viable option. Therefore, backup lines of credit are sometimes called "liquidity enhancement" or "backup liquidity lines." Issuers of commercial paper may be unable to roll over the paper because of either market-wide or company-specific events. For example, financial markets could be in the midst of a financial crisis that would make it difficult to roll over the paper. A company could also experience some sort of financial distress such that it could only issue new commercial paper at significantly higher rates. In this case, the company could draw on its credit lines instead of rolling over its paper. Most commercial paper issuers maintain 100% backing, although some large, high-credit-quality issues carry less than 100% backing. Backup lines of credit typically contain a "material adverse change" provision that allows the bank to cancel the backup line of credit if the financial condition of the issuer deteriorates substantially.

Historically, defaults on commercial paper have been relatively rare, primarily because commercial paper has a short maturity. Each time existing paper matures, investors have another opportunity to assess the issuer's financial position, and they can refuse to buy the new paper if they estimate that the issuer's credit risk is too high. Thus, the commercial paper markets adapt more quickly to a change in an issuer's credit quality than do the markets for longer-term securities. This flexibility reduces the exposure of the commercial paper market to defaults. In addition, corporate managers realize that defaulting on commercial paper would likely prevent any future issuance of this valuable financing alternative.

The combination of short-dated maturity, relatively low credit risk, and a large number of issuers makes commercial paper attractive to a diverse range of investors, including money market mutual funds, bank liquidity desks, corporate treasury departments, and other institutional investors that have liquidity constraints. Most commercial paper investors hold their position to maturity. The result is little secondary market trading except for the largest issues. Investors who wish to sell commercial paper prior to maturity can sell it either back to the dealer, to another investor, or in some cases, directly back to the issuer.

The yield on commercial paper is typically higher than that on short-term sovereign bonds of the same maturity for two main reasons. First, commercial paper is exposed to credit risk unlike most highly rated sovereign bonds. Second, commercial paper markets are generally less liquid than short-term sovereign bond markets. Thus, investors require higher yields to compensate for the lower liquidity. In the United States, the yield on commercial paper also tends to be higher than that on short-term municipal bonds for tax reasons. Income generated by investments in commercial paper is usually subject to income taxes, whereas income from many municipal bonds is tax exempt. Thus, to attract taxable investors, bonds that are subject to income taxes must offer higher yields than those that are tax exempt.

7.2.3 *US Commercial Paper vs. Eurocommercial Paper*

The US commercial paper (USCP) market is the largest commercial paper market in the world, although there are other active commercial paper markets in other countries. Commercial paper issued in the international market is known as Eurocommercial paper (ECP). Although ECP is a similar instrument to USCP, there are some differences between the two. These differences are shown in Exhibit 6.

Exhibit 6 USCP vs. ECP

Feature	US Commercial Paper	Eurocommercial Paper
Currency	US dollar	Any currency
Maturity	Overnight to 270 days[a]	Overnight to 364 days
Interest	Discount basis (instrument is issued at a discount to par value)	Interest-bearing (instrument issued at par and pays interest) or discount basis
Settlement	$T + 0$ (trade date)	$T + 2$ (trade date plus two days)
Negotiable	Can be sold to another party	Can be sold to another party

[a] In the United States, securities with an original maturity greater than 270 days must be registered with the Securities and Exchange Commission (SEC). To avoid the time and expense associated with a SEC registration, issuers of US commercial paper rarely offer maturities longer than 270 days.

A difference between USCP and ECP is related to the interest provision. USCP is typically issued on a discount basis—that is, USCP is issued at a discount to par value and pays full par value at maturity. The difference between the par value and the issue price is the interest paid on the borrowing. In contrast, ECP may be issued at, and trade on, an interest-bearing or yield basis or a discount basis. The distinction between the discount and the interest-bearing basis is explored later in the discussion on fixed-income valuation.

Typical transaction sizes in ECP are also much smaller than in USCP, and it is difficult to place longer-term ECP with investors. The ECP market also exhibits less liquidity than the USCP market.

CORPORATE DEBT: NOTES AND BONDS

8

Companies are active participants in global capital markets and regularly issue corporate notes and bonds. These securities can be placed directly with specific investors via private placements or sold in public securities markets. This section discusses various characteristics of corporate notes and bonds.

8.1 Maturities

There is no universally accepted taxonomy as to what constitutes short-, medium-, and long-term maturities. For our purposes, short term refers to original maturities of 5 years or less; intermediate term to original maturities longer than 5 years and up to 12 years; and long term to original maturities longer than 12 years. Those securities with maturities between 1 and 12 years are often considered notes, whereas securities with maturities greater than 12 years are considered bonds. It is not uncommon, however, to refer to bonds for all securities, irrespective of their original maturity.

In practice, most corporate bonds range in term to maturity between 1 and 30 years. In Europe, however, some bond issues have maturities of 40 or 50 years. In addition, companies and sovereigns have issued 100-year bonds; these are called "century bonds." For example, in 2017 Austria issued EUR 3.5 billion in bonds that will mature in 2117.

The first century bond was issued by the Walt Disney Company in 1993 as part of its **medium-term note** program. Medium-term note (MTN) is a misnomer because, as the Disney example illustrates, MTNs can have very long maturities. From the issuer's perspective, the initial purpose of MTNs was to fill the funding gap between commercial paper and long-term bonds. It is for this reason that they are referred to as "medium term." The MTN market can be broken into three segments: short-term securities that carry floating or fixed rates, medium- to long-term securities that primarily bear a fixed rate of interest, and structured notes.

MTNs have the unique characteristic of being securities that are offered continuously to investors by an agent of the issuer. This feature gives the borrower maximum flexibility for issuing securities on a continuous basis. Financial institutions are the primary issuers of MTNs, in particular short-term ones. Life insurance companies, pension funds, and banks are among the largest buyers of MTNs because they can customize the bond issue to their needs and stipulate the amount and characteristics of the securities they want to purchase. These investors are often willing to accept less liquidity than they would receive with a comparable publicly issued bond because the yield is slightly higher. The cost savings in registration and underwriting often makes MTNs a lower-cost option for the issuer.

8.2 Coupon Payment Structures

Corporate notes and bonds have a range of coupon payment structures. Financial and non-financial companies issue conventional coupon bonds that pay a fixed periodic coupon during the bond's life. They also issue bonds for which the periodic coupon payments adjust to changes in market conditions and/or changes to the issuer's credit quality. Such bonds typically offer investors the opportunity to reduce their exposure to a particular type of risk. For example, FRNs, whose coupon payments adjust to changes in the level of market interest rates, are a way to limit interest rate risk; some of the inflation-linked bonds whose coupon payments adjust to changes in the level of a consumer price index offer a protection against inflation risk; credit-linked coupon bonds, whose coupon payments adjust to changes in the issuer's credit quality, are a way to reduce credit risk. Whether the periodic coupon is fixed or not, coupon payments can be made quarterly, semi-annually, or annually depending on the type of bond and where the bonds are issued and traded.

Other coupon payment structures exist. Zero-coupon bonds pay no coupon. Deferred coupon bonds pay no coupon initially but then offer a higher coupon. Payment-in-kind (PIK) coupon bonds make periodic coupon payments, but not necessarily in cash; the issuer may pay interest in the form of securities, such as bonds or common shares. These types of coupon payment structures increase issuers' flexibility regarding the servicing of their debt.

8.3 Principal Repayment Structures

Corporate note or bond issues have either a serial or a term maturity structure. With a **serial maturity structure**, the maturity dates are spread out during the bond's life; a stated number of bonds mature and are paid off each year before final maturity. With a **term maturity structure**, the bond's notional principal is paid off in a lump sum at maturity. Because there is no regular repayment of the principal outstanding throughout the bond's life, a term maturity structure carries more credit risk than a serial maturity structure.

A sinking fund arrangement is a way to reduce credit risk by making the issuer set aside funds over time to retire the bond issue. For example, a corporate bond issue may require a specified percentage of the bond's outstanding principal amount to be retired each year. The issuer may satisfy this requirement in one of two ways. The most common approach is for the issuer to make a random call for the specified percentage of bonds that must be retired and to pay the bondholders whose bonds are called the sinking fund price, which is typically par. Alternatively, the issuer can deliver bonds to the trustee with a total amount equal to the amount that must be retired. To do so, the issuer may purchase the bonds in the open market. The sinking fund arrangement on a term maturity structure accomplishes the same goal as the serial maturity structure—that is, both result in a portion of the bond issue being paid off each year. With a serial maturity structure, however, the bondholders know which bonds will mature and will thus be paid off each year. In contrast, the bonds retired annually with a sinking fund arrangement are designated by a random drawing.

8.4 Asset or Collateral Backing

Unlike most highly rated sovereign bonds, all corporate debt is exposed to varying degrees of credit risk. Thus, corporate debt is structured with this risk in mind. An important consideration for investors is seniority ranking—that is, the systematic way in which lenders are repaid if the issuer defaults. In the case of secured debt, there is some form of collateral pledged to ensure payment of the debt. In contrast, in the case of unsecured debt, claims are settled by the general assets of the company in accordance with the priority of payments that applies either legally or contractually and as described in the bond indenture. Within each category of debt (secured and unsecured), there are finer gradations of rankings, which are discussed in the later in the curriculum coverage of credit analysis.

There is a wide range of bonds that are secured by some form of collateral. Companies that need to finance equipment or physical assets may issue equipment trust certificates. Corporate issuers also sell collateral trust bonds that are secured by securities, such as common shares, bonds, or other financial assets. Banks, particularly in Europe, may issue covered bonds, which are a type of debt obligation that is secured by a segregated pool of assets. Asset-backed securities are also secured forms of debt.

Companies can and do default on their debt. Debt secured by collateral may still experience losses, but during a company's bankruptcy proceedings, investors in secured debt usually fare better than investors in unsecured debt. Investors who face a higher level of credit risk typically require a higher yield than investors exposed to very little credit risk.

8.5 Contingency Provisions

Contingency provisions are clauses in the indenture that provide the issuer or the bondholders rights that affect the disposal or redemption of the bond. The three commonly used contingency provisions are call, put, and conversion provisions.

Callable bonds give issuers the ability to retire debt prior to maturity. The most compelling reason for them to do so is to take advantage of lower borrowing rates. By calling the bonds before their maturity date, the issuer can substitute a new, lower-cost bond issue for an older, higher-cost one. In addition, companies may also retire debt to eliminate restrictive covenants or to alter their capital structure to improve flexibility. Because the call provision is a valuable option for the issuer, investors demand compensation *ex ante* (before investing in the bond). Thus, other things equal, investors require a higher yield (and thus pay a lower price) for a callable bond than for an otherwise similar non-callable bond.

Companies also issue putable bonds, which give the bondholders the right to sell the bond back to the issuer at a pre-determined price on specified dates before maturity. Most putable bonds pay a fixed rate of interest, although some bonds may have step-up coupons that increase by specified margins at specified dates. Because the put provision is a valuable option for the bondholders, putable bonds offer a lower yield (and thus have a higher price) than otherwise similar non-putable bonds. The main corporate issuers of putable bonds are investment-grade companies. Putable bonds may offer them a cheaper way of raising capital, especially if the company estimates that the benefit of a lower coupon outweighs the risk associated with the put provision.

A convertible bond is a hybrid security that lies on a continuum between debt and equity. It consists of a long position in an option-free bond and a conversion option that gives the bondholder the right to convert the bond into a specified number of shares of the issuer's common shares. From the issuer's point of view, convertible bonds make it possible to raise funds that may not be possible without the incentive associated with the conversion option. The more common issuers of convertibles bonds are newer companies that have not established a presence in debt capital markets but who are able to present a more attractive package to institutional investors by including an equity upside potential. Established issuers of bonds may also prefer to issue convertible bonds because they are usually sold at a lower coupon rate than otherwise similar non-convertible bonds as a result of investors' attraction to the conversion provision. There is a potential equity dilution effect, however, if the bonds are converted. From the investor's point of view, convertible bonds represent a means of accessing the equity upside potential of the issuer but at a lower risk–reward profile, because there is the floor of the coupon payments in the meantime.

8.6 Issuance, Trading, and Settlement

In the era before electronic settlement, there were some differences in the processes of issuing and settling corporate bonds depending on where the securities were registered. This is no longer the case; the processes of issuing and settling bonds are now essentially the same globally. New corporate bond issues are usually sold to investors by investment banks acting as underwriters in the case of underwritten offerings or brokers in the case of best-efforts offerings. They are then settled via the local settlement system. These local systems typically possess a "bridge" to the two Eurobond systems, Euroclear and Clearstream. As for Eurobonds from the corporate sector, they are all issued, traded, and settled in the same way, irrespective of the issuer and its local jurisdiction.

Most bond prices are quoted in basis points. The vast majority of corporate bonds are traded in OTC markets through dealers that "make a market" in bonds and sell from their inventory. Dealers do not typically charge a commission or a transaction fee. Instead, they earn a profit from the bid–offer spread.

For corporate bonds, settlement differences exist primarily between new bond issues and the secondary trading of bonds. The issuing phase for an underwritten offering usually takes several days. Thus, settlement takes longer for new bond issued than for the secondary trading of bonds, for which settlement is typically on a $T + 2$ or $T + 3$ basis.

EXAMPLE 6

Corporate Debt

1 A loan made by a group of banks to a private company is *most likely*:

  A a bilateral loan.

  B a syndicated loan.

  C a securitized loan.

2 Which of the following statements relating to commercial paper is *most accurate*? Companies issue commercial paper:

  A only for funding working capital.

  B only as an interim source of financing.

  C both for funding working capital and as an interim source of funding.

3 Maturities of Eurocommercial paper range from:

  A overnight to three months.

  B overnight to one year.

  C three months to one year.

4 A bond issue that has a stated number of bonds that mature and are paid off each year before final maturity *most likely* has a:

  A term maturity.

  B serial maturity.

  C sinking fund arrangement.

Solution to 1:

B is correct. A loan from a group of lenders to a single borrower is a syndicated loan. A is incorrect because a bilateral loan is a loan from a single lender to a single borrower. C is incorrect because securitization involves moving assets, such as loans, from the owner of the assets into a special legal entity.

Solution to 2:

C is correct. Companies use commercial paper as a source of funding working capital and seasonal demand for cash, as well as an interim source of financing until permanent financing can be arranged.

Solution to 3:

B is correct. Eurocommercial paper ranges in maturity from overnight to 364 days.

Solution to 4:

B is correct. With a serial maturity structure, a stated number of bonds mature and are paid off each year before final maturity. A is incorrect because a bond issue with a term maturity structure is paid off in one lump sum at maturity. C is incorrect because a sinking fund arrangement, such as a serial maturity structure, results in a portion of the bond issue being paid off every year. With a serial maturity structure, however, the bonds are paid off because the maturity dates are spread out during the life of the bond and the bonds that are retired are maturing; the bondholders know in advance which bonds will be retired. In contrast, the bonds retired annually with a sinking fund arrangement are designated by a random drawing.

STRUCTURED FINANCIAL INSTRUMENTS

9

h describe structured financial instruments

Structured financial instruments represent a broad sector of financial instruments. This sector includes asset-backed securities (ABS) and collateralized debt obligations (CDOs). CDOs are securities backed by a diversified pool of one or more debt obligations, and like ABS, they are discussed later in curriculum coverage on asset-backed securities. A common attribute of all these financial instruments is that they repackage and redistribute risks.

Our focus in this section is on structured financial instruments apart from ABS and CDOs. These instruments typically have customized structures that often combine a bond and at least one derivative. Some of these instruments are called structured products. The use of derivatives gives the holder of the structured financial instrument exposure to one or more underlying assets, such as equities, bonds, and commodities. The redemption value and often the coupons of structured financial instruments are linked via a formula to the performance of the underlying asset(s). Thus, the bond's payment features are replaced with non-traditional payoffs that are derived not from the issuer's cash flows but from the performance of the underlying asset(s). Although no universally accepted taxonomy exists to categorize structured financial instruments, we will present four broad categories of instruments: capital protected, yield enhancement, participation, and leveraged instruments.

9.1 Capital Protected Instruments

Suppose an investor has \$100,000 to invest. The investor buys zero-coupon bonds issued by a sovereign issuer that will pay off \$100,000 one year from now. Also suppose the cost of buying the zero-coupon bonds is \$99,000. The investor can use the \$1,000 left over from the purchase of the zero-coupon bond to buy a call or put option on some underlying asset that expires one year from now. In this context, buying a call (put) option gives the investor the right to buy (sell) the underlying asset in one year at a pre-determined price. The investor will receive \$100,000 when the zero-coupon bond matures and may also gain from the upside potential of the option, if any. This combination of the zero-coupon bond and the call option can be prepackaged as a structured financial instrument called a **guarantee certificate**. The zero-coupon bond provides the investor capital protection; at maturity, the investor will receive 100% of the capital invested even if the option expires worthless. The call (put) option provides upside potential if the price of the underlying asset rises (falls) and a limited downside if the price of the underlying asset falls (rises) or remains unchanged. The downside is limited to the price, often called the premium, paid for the call or put option. In our example, the maximum loss the investor faces is $1,000, which is the price paid for the option.

Capital protected instruments offer different levels of capital protection. A guarantee certificate offers full capital protection. Other structured financial instruments may offer only partial capital protection. Note that the capital protection is only as good as the issuer of the instrument. Should the issuer of guarantee certificates go bankrupt, investors may lose their entire capital. These instruments are naturally more prevalent when the general level of interest rates is high and/or implied volatilities (the key driver of option prices, as we will explain later) are low. The cost of the options in the structure must be absorbed by the forfeited interest.

9.2 Yield Enhancement Instruments

Yield enhancement refers to increasing risk exposure in the hope of realizing a higher expected return. A **credit-linked note (CLN)** is an example of a yield enhancement instrument. Specifically, it is a type of bond that pays regular coupons but whose redemption value depends on the occurrence of a well-defined credit event, such as a rating downgrade or the default of an underlying asset, called the reference asset. If the specified credit event does not occur, the investor receives the par value of the CLN at maturity. But if the specified credit event occurs, the investor receives the par value of the CLN minus the nominal value of the reference asset to which the CLN is linked.

A CLN allows the issuer to transfer the effect of a credit event to investors. Thus, the issuer is the protection buyer and the investor is the protection seller. Investors are willing to buy CLNs because these securities offer higher coupons than otherwise similar bonds. In addition, CLNs are usually issued at a discount. Thus, if the specified credit event does not occur, investors will realize a significant capital gain on the purchase of the CLN.

9.3 Participation Instruments

As the name suggests, a participation instrument is one that allows investors to participate in the return of an underlying asset. Floating-rate bonds can be viewed as a type of participation instrument. As discussed earlier, floaters differ from fixed-rate bonds in that their coupon rate adjusts periodically according to a pre-specified formula. The coupon formula is usually expressed as a reference rate adjusted for a spread. A floater has almost zero interest rate risk because changes in the cash flows limit the effect of changes in interest rates on the floater’s price. Thus, floaters give investors the opportunity to participate in movements of interest rates. For example, the Italian government issued in June 2005 floaters set to mature in June 2020. The coupon payments are delivered annually and determined by the formula of 85% of the 10-year constant maturity swap rate, a widely used type of interest rate. Thus, investors who hold these floaters participate partially in movements of the 10-year constant maturity swap rate.

Most participation instruments are designed to give investors indirect exposure to a specific index or asset price. For example, investors who are precluded from investing in equity directly may gain indirect equity exposure by investing in participation instruments that are linked via a formula to the performance of equity indexes. Many structured products sold to individuals are participation instruments linked to an equity index. In contrast to capital protected instruments that offer equity exposure, these participation instruments usually do not offer capital protection.

9.4 Leveraged Instruments

Leveraged instruments are structured financial instruments created to magnify returns and offer the possibility of high payoffs from small investments. An **inverse floater** is an example of a leveraged instrument. As the name suggests, an inverse floater is the opposite of a traditional floater. The cash flows are adjusted periodically and move in the opposite direction of changes in the reference rate. So, when the reference rate decreases, the coupon payment of an inverse floater increases.

A general formula for an inverse floater's coupon rate is as follows:

Inverse floater coupon rate = $C - (L \times R)$

where C is the maximum coupon rate reached if the reference rate is equal to zero, L is the coupon leverage, and R is the reference rate on the reset date. Note that the coupon leverage indicates the multiple that the coupon rate will change in response to a 100-bp change in the reference rate. For example, if the coupon leverage is three, the inverse floater's coupon rate will decrease by 300 bps when the reference rate increases by 100 bps.

Inverse floaters with a coupon leverage greater than zero but lower than one are called deleveraged inverse floaters. Inverse floaters with a coupon leverage greater than one are called leveraged inverse floaters. For example, the Barclays Bank PLC issued a 15-year bond in January 2010 having quarterly payments. The coupon rate was fixed at 7.50% for the first three years and then in January 2013 transformed into a leveraged inverse floater paying 7.50% *minus* the euro three-month Libor. In this case, the coupon leverage is one. Thus, for a 100 bps increase in the euro three-month Libor, the coupon rate of the leveraged inverse floater will decrease by 100 bps. Inverse floaters often have a floor that specifies a minimum coupon rate; for example, a floor may be set at zero to avoid the possibility of a negative interest rate. This inverse floater does not have a maximum coupon rate. At the July 2018 reset, euro three-month Libor was –0.32%, so the coupon rate for the following quarter was 7.82%. Note that the example involves a bond issued in the past, and thus Libor was the reference rate. Future bonds issued will reference a different rate, including prime or MRR.

EXAMPLE 7

Structured Financial Instruments

1 If an investor holds a credit-linked note and the credit event does not occur, the investor receives:

A all promised cash flows as scheduled.

B all coupon payments as scheduled but not the par value at maturity.

C all coupon payments as scheduled and the par value minus the nominal value of the reference asset to which the credit-linked note is linked at maturity.

2 A structured financial instrument whose coupon rate is determined by the formula 5% – (0.5 × Libor) is *most likely*:

A a leveraged inverse floater.

B a participation instrument.

C a deleveraged inverse floater.

Solution to 1:

A is correct. If the credit event does not occur, the issuer must make all promised cash flows as scheduled—that is, the regular coupon payments and the par value at maturity.

Solution to 2:

C is correct. A structured financial instrument whose coupon rate moves in the opposite direction of the reference rate is called an inverse floater. Because the coupon leverage (0.5) is greater than zero but lower than one, the structured financial instrument is a deleveraged inverse floater. In this example, if the reference rate increases by 100 bps, the coupon rate decreases by 50 bps. A is incorrect because the coupon leverage would have to be higher than one for the structured financial instrument to be a leveraged inverse floater. B is incorrect because a participation instrument is designed to give investors indirect exposure to a particular underlying asset.

10

SHORT-TERM BANK FUNDING ALTERNATIVES

i. describe short-term funding alternatives available to banks

Funding refers to the amount of money or resources necessary to finance some specific project or enterprise. Accordingly, funding markets are markets in which debt issuers borrow to meet their financial needs. Companies have a range of funding alternatives, including bank loans, commercial paper, notes, and bonds. Financial institutions such as banks have larger financing needs than non-financial companies because of the nature of their operations. This section discusses the additional funding alternatives available to them. The majority of these funding alternatives have short maturities.

Banks, such as deposit-taking (or depository) institutions, typically have access to funds obtained from the retail market—that is, deposit accounts from their customers. It is quite common, however, for banks to originate more loans than they have retail deposits. Thus, whenever the amount of retail deposits is insufficient to meet their financial needs, banks also need to raise funds from the wholesale market. Wholesale funds include central bank funds, interbank deposits, and certificates of deposit. In addition to filling the gaps between loans and deposits, banks raise wholesale funds to minimize their funding cost. At the margin, wholesale funds may be less expensive (in terms of interest expense) than deposit funding. Finally, financial institutions may raise wholesale funds as a balance sheet risk management tool to reduce interest rate risk, as discussed earlier.

10.1 Retail Deposits

One of the primary sources of funding for deposit-taking banks is their retail deposit base, which includes funds from both individual and commercial depositors. There are several types of retail deposit accounts. Demand deposits, also known as checking accounts, are available to customers "on demand." Depositors have immediate access to the funds in their deposit accounts and use the funds as a form of payment for transactions. Because the funds are available immediately, deposit accounts typically pay no interest. In contrast, savings accounts pay interest and allow depositors to accumulate wealth in a very liquid form, but they do not offer the same transactional convenience as demand deposits. Money market accounts were originally designed to compete with money market mutual funds. They offer money market rates of return and depositors can access funds at short or no notice. Thus, money market accounts are, for depositors, an intermediate between demand deposit and savings accounts.

10.2 Short-Term Wholesale Funds

Wholesale funds available for banks include reserve funds, interbank funds, and certificates of deposit.

10.2.1 *Reserve Funds*

Many countries require deposit-taking banks to place a reserve balance with the national central bank. The reserve funds help to ensure sufficient liquidity should depositors require withdrawal of funds. When a bank cannot obtain short-term funding, most countries allow that bank to borrow from the central bank. In aggregate, the reserve funds act as a liquidity buffer, providing comfort to depositors and investors that the central bank can act as lender of last resort.

Treatment of interest on reserve funds varies among countries, from a low interest payment, to no interest payment, to charges for keeping reserve funds. Additionally, there is an opportunity cost to the banks for holding reserves with the central bank, in that these funds cannot be invested with higher interest or loaned out to consumers or commercial enterprises. Some banks have an excess over the minimum required funds to be held in reserve. At the same time, other banks run short of required reserves. This imbalance is solved through the **central bank funds market**, which allows banks that have a surplus of funds to lend money to banks that need funds for maturities of up to one year. These funds are known as central bank funds and are called “overnight funds” when the maturity is one day and “term funds” when the maturity ranges from two days to one year. The interest rates at which central bank funds are bought (i.e., borrowed) and sold (i.e., lent) are short-term interest rates determined by the markets but influenced by the central bank’s open market operations. These rates are termed the **central bank funds rates**.

In the United States, the central bank is the Federal Reserve (Fed). The central bank funds and funds rate are called Fed funds and the Fed funds rate, respectively. Other short-term interest rates, such as the yields on Treasury bills, are highly correlated with the Fed funds rate. The most widely followed rate is known as the Fed funds effective rate, which is the volume-weighted average of rates for Fed fund trades arranged throughout the day by the major New York City brokers. Fed funds are traded between banks and other financial institutions globally and may be transacted directly or through money market brokers.

10.2.2 *Interbank Funds*

The **interbank market** is the market of loans and deposits between banks. The term to maturity of an interbank loan or deposit ranges from overnight to one year. The rate on an interbank loan or deposit can be quoted relative to a reference rate, such as an interbank offered rate, or as a fixed interest rate. An interbank deposit is unsecured, so banks placing deposits with another bank need to have an interbank line of credit in place for that institution. Usually, a large bank will make a two-way price, indicating the rate at which it will lend funds and the rate at which it will borrow funds for a specific maturity, on demand. Interest on the deposit is payable at maturity. Much interbank dealing takes place on the Reuters electronic dealing system, so that the transaction is done without either party speaking to the other.

Because the market is unsecured, it is essentially based on confidence in the banking system. At times of stress, such as in the aftermath of the Lehman Brothers’ bankruptcy in 2008, the market is prone to “dry up” as banks withdraw from funding other banks.

10.2.3 *Large-Denomination Negotiable Certificates of Deposit*

A **certificate of deposit** (CD) is an instrument that represents a specified amount of funds on deposit for a specified maturity and interest rate. CDs are an important source of funds for financial institutions. A CD may take one of two forms: non-negotiable or negotiable. If the CD is non-negotiable, the deposit plus the interest are paid to the initial depositor at maturity. A withdrawal penalty is imposed if the depositor withdraws funds prior to the maturity date.

Alternatively, a negotiable CD allows any depositor (initial or subsequent) to sell the CD in the open market prior to the maturity date. Negotiable CDs were introduced in the United States in the early 1960s when various types of deposits were constrained by interest rate ceilings. At the time, bank deposits were not an attractive investment because investors earned a below-market interest rate unless they were prepared to commit their capital for an extended period. The introduction of negotiable CDs enabled bank customers to buy a three-month or longer negotiable instrument yielding a market interest rate and to recover their investment by selling it in the market. This innovation helped banks increase the amount of funds raised in the money markets. It also fostered competition among deposit-taking institutions.

There are two types of negotiable CDs: large-denomination CDs and small-denomination CDs. Thresholds between small- and large-denomination CDs vary among countries. For example, in the United States, large-denomination CDs are usually issued in denominations of $1 million or more. Small-denomination CDs are a retail-oriented product, and they are of secondary importance as a funding alternative. Large-denomination CDs, in contrast, are an important source of wholesale funds and are typically traded among institutional investors.

Like other money market securities, CDs are available in domestic bond markets as well as in the Eurobond market. Most CDs have maturities shorter than one year and pay interest at maturity. CDs with longer maturities are called “term CDs.”

Yields on CDs are driven primarily by the credit risk of the issuing bank and to a lesser extent by the term to maturity. The spread attributable to credit risk will vary with economic conditions and confidence in the banking system in general and in the issuing bank in particular. As with all debt instruments, spreads widen during times of financial turmoil in response to increased risk aversion.

11 REPURCHASE AND REVERSE REPURCHASE AGREEMENTS

j describe repurchase agreements (repos) and the risks associated with them

Repurchase agreements are another important source of funding not only for banks but also for other market participants. A **repurchase agreement** or **repo** is the sale of a security with a simultaneous agreement by the seller to buy the same security back from the purchaser at an agreed-on price and future date. In practical terms, a repurchase agreement can be viewed as a collateralized loan in which the security sold and subsequently repurchased represents the collateral posted. One party is borrowing money and providing collateral for the loan at an interest rate that is typically lower than on an otherwise similar bank loan. The other party is lending money while accepting a security as collateral for the loan.

Repurchase agreements are a common source of money market funding for dealer firms in many countries. An active market in repurchase agreements underpins every liquid bond market. Financial and non-financial companies participate actively in the market as both sellers and buyers of collateral depending on their circumstances.

Central banks are also active users of repurchase agreements in their daily open market operations; they either lend to the market to increase the supply of funds or withdraw surplus funds from the market.

11.1 Structure of Repurchase and Reverse Repurchase Agreements

Suppose a government securities dealer purchases a 2.25% UK gilt that matures in three years. The dealer wants to fund the position overnight through the end of the next business day. The dealer could finance the transaction with its own funds, which is what other market participants, such as insurance companies or pension funds, may do in similar circumstances. But a securities dealer typically uses leverage (debt) to fund the position. Rather than borrowing from a bank, the dealer uses a repurchase agreement to obtain financing by using the gilt as collateral for the loan.

A repurchase agreement may be constructed as follows: The dealer sells the 2.25% UK gilt that matures in three years to a counterparty for cash today. At the same time, the dealer makes a promise to buy the same gilt the next business day for an agreed-on price. The price at which the dealer repurchases the gilt is known as the **repurchase price**. The date when the gilt is repurchased, the next business day in this example, is called the **repurchase date**. When the term of a repurchase agreement is one day, it is called an "overnight repo." When the agreement is for more than one day, it is called a "term repo." An agreement lasting until the final maturity date is known as a "repo to maturity."

As in any borrowing or lending transaction, the interest rate of the loan must be negotiated in the agreement. The interest rate on a repurchase agreement is called the **repo rate**. Several factors affect the repo rate:

- The *risk* associated with the collateral. Repo rates are typically lower for highly rated collaterals, such as highly rated sovereign bonds. They increase with the level of credit risk associated with the collateral underlying the transaction.
- The *term* of the repurchase agreement. Repo rates generally increase with maturity because long-term rates are typically higher than short-term rates in normal circumstances.
- The *delivery requirement* for the collateral. Repo rates are usually lower when delivery to the lender is required.
- The *supply and demand conditions* of the collateral. The more in demand a specific piece of collateral is, the lower the repo rate against it because the borrower has a security that lenders of cash want for specific reasons, perhaps because the underlying issue is in great demand. The demand for such collateral means that it considered to be "on special." Collateral that is not special is known as "general collateral." The party that has a need for collateral that is on special is typically required to lend funds at a below-market repo rate to obtain the collateral.
- The *interest rates of alternative financing* in the money market.

The interest on a repurchase agreement is paid on the repurchase date—that is, at the termination of the agreement. Note that any coupon paid by the security during the repurchase agreement belongs to the seller of the security (i.e., the borrower of cash).

When a repurchase agreement is viewed through the lens of the cash lending counterparty, the transaction is referred to as a **reverse repurchase agreement** or **reverse repo**. In the foregoing example, the counterparty agrees to buy the 2.25% UK gilt that matures in three years and promises to sell it back the next business day at the agreed-on price. The counterparty is making a collateralized loan to the dealer. Reverse repurchase agreements are very often used to borrow securities to cover short positions.

The question of whether a particular transaction is labeled a repurchase agreement or a reverse repurchase agreement depends on one's point of view. Standard practice is to view the transaction from the dealer's perspective. If the dealer is borrowing cash from a counterparty and providing securities as collateral, the transaction is termed a repurchase agreement. If the dealer is borrowing securities and lending cash to the counterparty, the transaction is termed a reverse repurchase agreement.

11.2 Credit Risk Associated with Repurchase Agreements

Each market participant in a repurchase agreement is exposed to the risk that the counterparty defaults, regardless of the collateral exchanged. Credit risk is present even if the collateral is a highly rated sovereign bond. Suppose that a dealer (i.e., the borrower of cash) defaults and is not in a position to repurchase the collateral on the specified repurchase date. The lender of funds takes possession of the collateral and retains any income owed to the borrower. The risk is that the price of the collateral has fallen following the inception of the repurchase agreement, causing the market value of the collateral to be lower than the unpaid repurchase price. Conversely, suppose the investor (i.e., the lender of cash) defaults and is unable to deliver the collateral on the repurchase date. The risk is that the price of the collateral has risen since the inception of the repurchase agreement, resulting in the dealer now holding an amount of cash lower than the market value of the collateral. In this case, the investor is liable for any excess of the price paid by the dealer for replacement of the securities over the repurchase price.

Although both parties to a repurchase agreement are subject to credit risk, the agreement is structured as if the lender of funds is the most vulnerable party. Specifically, the amount lent is lower than the collateral's market value. The difference between the market value of the security used as collateral and the value of the loan is known as the **repo margin**, although the term **haircut** is more commonly used, particularly in the United States. The repo margin allows for some worsening in market value and thus provides the cash lender a margin of safety if the collateral's market value declines. Repo margins vary by transaction and are negotiated bilaterally between the counterparties. The level of margin is a function of the following factors:

- The *length* of the repurchase agreement. The longer the repurchase agreement, the higher the repo margin.
- The *quality* of the collateral. The higher the quality of the collateral, the lower the repo margin.
- The *credit quality* of the counterparty. The higher the creditworthiness of the counterparty, the lower the repo margin.
- The *supply and demand conditions* of the collateral. Repo margins are lower if the collateral is in short supply or if there is a high demand for it.

EXAMPLE 8

Short-Term Funding Alternatives Available to Banks

1 Which of the following are **not** considered wholesale funds?

A Interbank funds

B Central bank funds

   C Repurchase agreements

2 A large-denomination negotiable certificate of deposit *most likely*:

   A is traded in the open market.

   B is purchased by retail investors.

   C has a penalty for early withdrawal of funds.

3 From the dealer's viewpoint, a repurchase agreement is *best* described as a type of:

   A collateralized short-term lending.

   B collateralized short-term borrowing.

   C uncollateralized short-term borrowing.

4 The interest on a repurchase agreement is known as the:

   A repo rate.

   B repo yield.

   C repo margin.

5 The level of repo margin is higher:

   A the higher the quality of the collateral.

   B the higher the credit quality of the counterparty.

   C the longer the length of the repurchase agreement.

Solution to 1:

C is correct. Wholesale funds refer to the funds that financial institutions lend to and borrow from each other. They include central bank funds, interbank funds, and certificates of deposit. Although repurchase agreements are an important source of funding for banks, they are not considered wholesale funds.

Solution to 2:

A is correct. Large-denomination negotiable CDs can be traded in the open market. B is incorrect because it is small-denomination, not large-denomination, negotiable CDs that are primarily purchased by retail investors. C is incorrect because it is non-negotiable, not negotiable, CDs that have a penalty for early withdrawal of funds.

Solution to 3:

B is correct. In a repurchase agreement, a security is sold with a simultaneous agreement by the seller to buy the same security back from the purchaser later at a higher price. Thus, a repurchase agreement is similar to a collateralized short-term borrowing in which the security sold and subsequently repurchased represents the collateral posted. A is incorrect because collateralized short-term lending is a description of a reverse repurchase agreement. C is incorrect because a repurchase agreement involves collateral. Thus, it is a collateralized, not uncollateralized, short-term borrowing.

Solution to 4:

A is correct. The repo rate is the interest rate on a repurchase agreement. B is incorrect because the interest on a repurchase agreement is known as the repo rate, not repo yield. C is incorrect because the repo margin refers to the difference between the market value of the security used as collateral and the value of the loan.

Solution to 5:

C is correct. The longer the length of the repurchase agreement, the higher the repo margin (haircut). A is incorrect because the higher the quality of the collateral, the lower the repo margin. B is incorrect because the higher the credit quality of the counterparty, the lower the repo margin.

SUMMARY

Debt financing is an important source of funds for households, governments, government-related entities, financial institutions, and non-financial companies. Well-functioning fixed-income markets help ensure that capital is allocated efficiently to its highest and best use globally. Important points include the following:

- The most widely used ways of classifying fixed-income markets include the type of issuer; the bonds' credit quality, maturity, currency denomination, and type of coupon; and where the bonds are issued and traded.
- Based on the type of issuer, the four major bond market sectors are the household, non-financial corporate, government, and financial institution sectors.
- Investors distinguish between investment-grade and high-yield bond markets based on the issuer's credit quality.
- Money markets are where securities with original maturities ranging from overnight to one year are issued and traded, whereas capital markets are where securities with original maturities longer than one year are issued and traded.
- The majority of bonds are denominated in either euros or US dollars.
- Investors distinguish between bonds that pay a fixed rate versus a floating rate of interest. The coupon rate of floating-rate bonds is often expressed as a reference rate plus a spread. Interbank offered rates, such as Libor, historically have been the most commonly used reference rates for floating-rate debt and other financial instruments but are being phased out to be replaced by alternative reference rates.
- Based on where the bonds are issued and traded, investors distinguish between domestic and international bond markets. The latter includes the Eurobond market, which falls outside the jurisdiction of any single country and is characterized by less reporting, regulatory, and tax constraints. Investors also distinguish between developed and emerging bond markets.
- Investors and investment managers use fixed-income indexes to describe bond markets or sectors and to evaluate performance of investments and investment managers.
- The largest investors in bonds include central banks; institutional investors, such as pension funds, hedge funds, charitable foundations and endowments, insurance companies, mutual funds and ETFs, and banks; and retail investors, typically by means of indirect investments.
- Primary markets are markets in which issuers first sell bonds to investors to raise capital. Secondary markets are markets in which existing bonds are subsequently traded among investors.

- There are two mechanisms for issuing a bond in primary markets: a public offering, in which any member of the public may buy the bonds, or a private placement, in which only an investor or small group of investors may buy the bonds either directly from the issuer or through an investment bank.
- Public bond issuing mechanisms include underwritten offerings, best-efforts offerings, shelf registrations, and auctions.
- When an investment bank underwrites a bond issue, it buys the entire issue and takes the risk of reselling it to investors or dealers. In contrast, in a best-efforts offering, the investment bank serves only as a broker and sells the bond issue only if it is able to do so. Underwritten and best-efforts offerings are frequently used in the issuance of corporate bonds.
- The underwriting process typically includes six phases: the determination of the funding needs, the selection of the underwriter, the structuring and announcement of the bond offering, pricing, issuance, and closing.
- A shelf registration is a method for issuing securities in which the issuer files a single document with regulators that describes and allows for a range of future issuances.
- An auction is a public offering method that involves bidding and is helpful both in providing price discovery and in allocating securities. Auctions are frequently used in the issuance of sovereign bonds.
- Most bonds are traded in OTC markets, and institutional investors are the major buyers and sellers of bonds in secondary markets.
- Sovereign bonds are issued by national governments primarily for fiscal reasons. These bonds take different names and forms depending on where they are issued, their maturities, and their coupon types. Most sovereign bonds are fixed-rate bonds, although some national governments also issue floating-rate bonds and inflation-linked bonds.
- Local governments, quasi-government entities, and supranational agencies issue bonds, which are named non-sovereign, quasi-government, and supranational bonds, respectively.
- Companies raise debt in the form of bilateral loans, syndicated loans, commercial paper, notes, and bonds.
- Commercial paper is a short-term unsecured security that companies use as a source of short-term and bridge financing. Investors in commercial paper are exposed to credit risk, although defaults are rare. Many issuers roll over their commercial paper on a regular basis.
- Corporate bonds and notes take different forms depending on the maturities, coupon payment, and principal repayment structures. Important considerations also include collateral backing and contingency provisions.
- Medium-term notes are securities that are offered continuously to investors by an agent of the issuer. They can have short-term or long-term maturities.
- The structured finance sector includes asset-backed securities, collateralized debt obligations, and other structured financial instruments. All of these seemingly disparate financial instruments share the common attribute of repackaging risks.
- Many structured financial instruments are customized instruments that often combine a bond and at least one derivative. The redemption and often the coupons of these structured financial instruments are linked via a formula to the performance of the underlying asset(s). Thus, the bond's payment features are replaced with non-traditional payoffs derived not from the issuer's cash flows

but from the performance of the underlying asset(s). Capital protected, yield enhancement, participation and leveraged instruments are typical examples of structured financial instruments.

- Financial institutions have access to additional sources of funds, such as retail deposits, central bank funds, interbank funds, large-denomination negotiable certificates of deposit, and repurchase agreements.
- A repurchase agreement is similar to a collateralized loan. It involves the sale of a security (the collateral) with a simultaneous agreement by the seller (the borrower) to buy back the same security from the purchaser (the lender) at an agreed-on price in the future. Repurchase agreements are a common source of funding for dealer firms and are also used to borrow securities to implement short positions.

PRACTICE PROBLEMS

1 The distinction between investment-grade debt and non-investment-grade debt is *best* described by differences in:

A tax status.

B credit quality.

C maturity dates.

2 A bond issued internationally, outside the jurisdiction of the country in whose currency the bond is denominated, is *best* described as a:

A Eurobond.

B foreign bond.

C municipal bond.

3 When classified by type of issuer, asset-backed securities are part of the:

A corporate sector.

B structured finance sector.

C government and government-related sector.

4 Compared with developed market bonds, emerging market bonds *most likely*:

A offer lower yields.

B exhibit higher risk.

C benefit from lower growth prospects.

5 With respect to floating-rate bonds, a reference rate (such as MRR) is *most likely* used to determine the bond's:

A spread.

B coupon rate.

C frequency of coupon payments.

6 The variability of the coupon rate on a Libor-based floating-rate bond is *most likely* caused by:

A periodic resets of the reference rate.

B market-based reassessments of the issuer's creditworthiness.

C changing estimates by the Libor administrator of borrowing capacity.

7 Which of the following statements is *most accurate*? An interbank offered rate:

A is a single reference rate.

B applies to borrowing periods of up to 10 years.

C is used as a reference rate for interest rate swaps.

8 An investment bank that underwrites a bond issue *most likely*:

A buys and resells the newly issued bonds to investors or dealers.

B acts as a broker and receives a commission for selling the bonds to investors.

C incurs less risk associated with selling the bonds than in a best-efforts offering.

9 In major developed bond markets, newly issued sovereign bonds are *most* often sold to the public via a(n):

A auction.

© 2017 CFA Institute. All rights reserved.

B private placement.

C best-efforts offering.

10 Which of the following describes privately placed bonds?

A They are non-underwritten and unregistered.

B They usually have active secondary markets.

C They are less customized than publicly offered bonds.

11 A mechanism by which an issuer may be able to offer additional bonds to the general public without preparing a new and separate offering circular *best* describes:

A the grey market.

B a shelf registration.

C a private placement.

12 Which of the following statements related to secondary bond markets is *most accurate*?

A Newly issued corporate bonds are issued in secondary bond markets.

B Secondary bond markets are where bonds are traded between investors.

C The major participants in secondary bond markets globally are retail investors.

13 A bond market in which a communications network matches buy and sell orders initiated from various locations is *best* described as an:

A organized exchange.

B open market operation.

C over-the-counter market.

14 A liquid secondary bond market allows an investor to sell a bond at:

A the desired price.

B a price at least equal to the purchase price.

C a price close to the bond's fair market value.

15 Corporate bond secondary market trading *most often* occurs:

A on a book-entry basis.

B on organized exchanges.

C prior to settlement at $T + 1$.

16 Sovereign bonds are *best* described as:

A bonds issued by local governments.

B secured obligations of a national government.

C bonds backed by the taxing authority of a national government.

17 Which factor is associated with a more favorable quality sovereign bond credit rating?

A Issued in local currency, only

B Strong domestic savings base, only

C Issued in local currency of country with strong domestic savings base

18 Which type of sovereign bond has the lowest interest rate risk for an investor?

A Floaters

B Coupon bonds

C Discount bonds

19 Agency bonds are issued by:

A local governments.
B national governments.
C quasi-government entities.

20 The type of bond issued by a multilateral agency such as the International Monetary Fund (IMF) is *best* described as a:
A sovereign bond.
B supranational bond.
C quasi-government bond.

21 A bond issued by a local government authority, typically without an explicit funding commitment from the national government, is *most likely* classified as a:
A sovereign bond.
B quasi-government bond
C non-sovereign government bond.

22 Which of the following statements relating to commercial paper is *most accurate*?
A There is no secondary market for trading commercial paper.
B Only the strongest, highly rated companies issue commercial paper.
C Commercial paper is a source of interim financing for long-term projects.

23 Eurocommercial paper is *most likely*:
A negotiable.
B denominated in euros.
C issued on a discount basis.

24 For the issuer, a sinking fund arrangement is *most similar* to a:
A term maturity structure.
B serial maturity structure.
C bondholder put provision.

25 When issuing debt, a company may use a sinking fund arrangement as a means of reducing:
A credit risk.
B inflation risk.
C interest rate risk.

26 Which of the following is a source of wholesale funds for banks?
A Demand deposits
B Money market accounts
C Negotiable certificates of deposit

27 A characteristic of negotiable certificates of deposit is:
A they are mostly available in small denominations.
B they can be sold in the open market prior to maturity.
C a penalty is imposed if the depositor withdraws funds prior to maturity.

28 A repurchase agreement is *most* comparable to a(n):
A interbank deposit.
B collateralized loan.
C negotiable certificate of deposit.

29 The repo margin is:

- **A** negotiated between counterparties.
- **B** established independently of market-related conditions.
- **C** structured on an agreement assuming equal credit risks to all counterparties.

30 The repo margin on a repurchase agreement is *most likely* to be lower when:

- **A** the underlying collateral is in short supply.
- **B** the maturity of the repurchase agreement is long.
- **C** the credit risk associated with the underlying collateral is high.

SOLUTIONS

1 B is correct. The distinction between investment-grade and non-investment-grade debt relates to differences in credit quality, not tax status or maturity dates. Debt markets are classified based on the issuer's creditworthiness as judged by the credit ratings agencies. Ratings of Baa3 or above by Moody's Investors Service or BBB– or above by Standard & Poor's and Fitch Ratings are considered investment grade, whereas ratings below these levels are referred to as non–investment grade (also called high yield, speculative, or junk).

2 A is correct. Eurobonds are issued internationally, outside the jurisdiction of any single country. B is incorrect because foreign bonds are considered international bonds, but they are issued in a specific country, in the currency of that country, by an issuer domiciled in another country. C is incorrect because municipal bonds are US domestic bonds issued by a state or local government.

3 B is correct. Asset-backed securities are securitized debt instruments created by securitization, a process that involves transferring ownership of assets from the original owners to a special legal entity. The special legal entity then issues securities backed by the transferred assets. The assets' cash flows are used to pay interest and repay the principal owed to the holders of the securities. Assets that are typically used to create securitized debt instruments include loans (such as mortgage loans) and receivables (such as credit card receivables). The structured finance sector includes such securitized debt instruments (also called asset-backed securities).

4 B is correct. Many emerging countries lag developed countries in the areas of political stability, property rights, and contract enforcement. Consequently, emerging market bonds usually exhibit higher risk than developed market bonds. A is incorrect because emerging market bonds typically offer higher (not lower) yields than developed market bonds to compensate investors for the higher risk. C is incorrect because emerging market bonds usually benefit from higher (not lower) growth prospects than developed market bonds.

5 B is correct. The coupon rate of a floating-rate bond is expressed as a reference rate plus a spread. Different reference rates are used depending on where the bond is issued and its currency denomination, but one of the most widely used set of reference rates is Libor. A and C are incorrect because a bond's spread and frequency of coupon payments are typically set when the bond is issued and do not change during the bond's life.

6 A is correct. Changes in the coupon rate of interest on a floating-rate bond that uses a Libor reference rate result from changes in the reference rate (for example, 90-day Libor), which resets periodically. Therefore, the coupon rate adjusts to the level of market interest rates (plus the spread) each time the reference rate is reset.

7 C is correct. Interbank offered rates are used as reference rates not only for floating-rate bonds but also for other debt instruments, including mortgages, derivatives such as interest rate and currency swaps, and many other financial contracts and products. A and B are incorrect because an interbank offered rate such as Libor or Euribor is a set of reference rates (not a single reference rate) for different borrowing periods of up to one year (not 10 years).

8 A is correct. In an underwritten offering (also called firm commitment offering), the investment bank (called the underwriter) guarantees the sale of the bond issue at an offering price that is negotiated with the issuer. Thus, the underwriter takes the risk of buying the newly issued bonds from the issuer and

then reselling them to investors or to dealers, which then sell them to investors. B and C are incorrect because the bond issuing mechanism in which an investment bank acts as a broker and receives a commission for selling the bonds to investors, and incurs less risk associated with selling the bonds, is a best-efforts offering (not an underwritten offering).

9 A is correct. In major developed bond markets, newly issued sovereign bonds are sold to the public via an auction. B and C are incorrect because sovereign bonds are rarely issued via private placements or best-efforts offerings.

10 A is correct. Private placements are typically non-underwritten, unregistered bond offerings that are sold only to a single investor or a small group of investors.

11 B is correct. A shelf registration allows certain authorized issuers to offer additional bonds to the general public without having to prepare a new and separate offering circular. The issuer can offer multiple bond issuances under the same master prospectus and only has to prepare a short document when additional bonds are issued. A is incorrect because the grey market is a forward market for bonds about to be issued. C is incorrect because a private placement is a non-underwritten, unregistered offering of bonds that are not sold to the general public but directly to an investor or a small group of investors.

12 B is correct. In secondary bond markets, bonds are traded between investors. A is incorrect because newly issued bonds (whether from corporate issuers or other types of issuers) are issued in primary (not secondary) bond markets. C is incorrect because the major participants in secondary bond markets globally are large institutional investors and central banks (not retail investors).

13 C is correct. In over-the-counter (OTC) markets, buy and sell orders are initiated from various locations and then matched through a communications network. Most bonds are traded in OTC markets. A is incorrect because on organized exchanges, buy and sell orders may come from anywhere, but the transactions must take place at the exchange according to the rules imposed by the exchange. B is incorrect because open market operations refer to central bank activities in secondary bond markets. Central banks buy and sell bonds, usually sovereign bonds issued by the national government, as a means to implement monetary policy.

14 C is correct. Liquidity in secondary bond markets refers to the ability to buy or sell bonds quickly at prices close to their fair market value. A and B are incorrect because a liquid secondary bond market does not guarantee that a bond will sell at the price sought by the investor, or that the investor will not face a loss on his or her investment.

15 A is correct. The vast majority of corporate bonds are traded in over-the-counter (OTC) markets that use electronic trading platforms through which users submit buy and sell orders. Settlement of trades in the OTC markets occurs by means of a simultaneous exchange of bonds for cash on the books of the clearing system "on a paperless, computerized book-entry basis."

16 C is correct. Sovereign bonds are usually unsecured obligations of the national government issuing the bonds; they are backed not by collateral but by the taxing authority of the national government. A is incorrect because bonds issued by local governments are non-sovereign (not sovereign) bonds. B is incorrect because sovereign bonds are typically unsecured (not secured) obligations of a national government.

17 C is correct. Bonds issued in the sovereign's currency and a strong domestic savings base are both favorable sovereign rating factors. It is common to observe a higher credit rating for sovereign bonds issued in local currency because of the sovereign's ability to tax its citizens and print its own currency. Although there are practical limits to the sovereign's taxing and currency-printing capacities, each tends to support a sovereign's ability to repay debt. A strong domestic savings base is advantageous because it supports the sovereign's ability to issue debt in local currency to domestic investors.

18 A is correct. Floaters are bonds with a floating rate of interest that resets periodically based on changes in the level of a reference rate, such as Libor. Because changes in the reference rate reflect changes in market interest rates, price changes of floaters are far less pronounced than those of fixed-rate bonds, such as coupon bonds and discount bonds. Thus, investors holding floaters are less exposed to interest rate risk than investors holding fixed-rate discount or coupon bonds.

19 C is correct. Agency bonds are issued by quasi-government entities. These entities are agencies and organizations usually established by national governments to perform various functions for them. A and B are incorrect because local and national governments issue non-sovereign and sovereign bonds, respectively.

20 B is correct. The IMF is a multilateral agency that issues supranational bonds. A and C are incorrect because sovereign bonds and quasi-government bonds are issued by national governments and by entities that perform various functions for national governments, respectively.

21 C is correct. Bonds issued by levels of government below the national level—such as provinces, regions, states, cities, and local government authorities—are classified as non-sovereign government bonds. These bonds are typically not guaranteed by the national government.

22 C is correct. Companies use commercial paper not only as a source of funding working capital and seasonal demand for cash but also as a source of interim financing for long-term projects until permanent financing can be arranged. A is incorrect because there is a secondary market for trading commercial paper, although trading is limited except for the largest issues. B is incorrect because commercial paper is issued by companies across the risk spectrum, although only the strongest, highly rated companies issue *low-cost* commercial paper.

23 A is correct. Commercial paper, whether US commercial paper or Eurocommercial paper, is negotiable—that is, investors can buy and sell commercial paper on secondary markets. B is incorrect because Eurocommercial paper can be denominated in any currency. C is incorrect because Eurocommercial paper may be issued on an interest-bearing (or yield) basis or a discount basis.

24 B is correct. With a serial maturity structure, a stated number of bonds mature and are paid off on a pre-determined schedule before final maturity. With a sinking fund arrangement, the issuer is required to set aside funds over time to retire the bond issue. Both result in a pre-determined portion of the issue being paid off according to a pre-determined schedule.

25 A is correct. A sinking fund arrangement is a way to reduce credit risk by making the issuer set aside funds over time to retire the bond issue. B and C are incorrect because a sinking fund arrangement has no effect on inflation risk or interest rate risk.

26 C is correct. Wholesale funds available for banks include central bank funds, interbank funds, and negotiable certificates of deposit. A and B are incorrect because demand deposits (also known as checking accounts) and money market accounts are retail deposits, not wholesale funds.

27 B is correct. A negotiable certificate of deposit (CD) allows any depositor (initial or subsequent) to sell the CD in the open market prior to maturity. A is incorrect because negotiable CDs are mostly available in large (not small) denominations. Large-denomination negotiable CDs are an important source of wholesale funds for banks, whereas small-denomination CDs are not. C is incorrect because a penalty is imposed if the depositor withdraws funds prior to maturity for non-negotiable (instead of negotiable) CDs.

28 B is correct. A repurchase agreement (repo) can be viewed as a collateralized loan in which the security sold and subsequently repurchased represents the collateral posted. A and C are incorrect because interbank deposits and negotiable certificates of deposit are unsecured deposits—that is, there is no collateral backing the deposit.

29 A is correct. Repo margins vary by transaction and are negotiated bilaterally between the counterparties.

30 A is correct. The repo margin (the difference between the market value of the underlying collateral and the value of the loan) is a function of the supply and demand conditions of the collateral. The repo margin is typically lower if the underlying collateral is in short supply or if there is a high demand for it. B and C are incorrect because the repo margin is usually higher (not lower) when the maturity of the repurchase agreement is long and when the credit risk associated with the underlying collateral is high.

READING

41

Introduction to Fixed-Income Valuation

by James F. Adams, PhD, CFA, and Donald J. Smith, PhD

James F. Adams, PhD, CFA, is at New York University (USA). Donald J. Smith, PhD, is at Boston University Questrom School of Business (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** calculate a bond's price given a market discount rate;
☐	**b.** identify the relationships among a bond's price, coupon rate, maturity, and market discount rate (yield-to-maturity);
☐	**c.** define spot rates and calculate the price of a bond using spot rates;
☐	**d.** describe and calculate the flat price, accrued interest, and the full price of a bond;
☐	**e.** describe matrix pricing;
☐	**f.** calculate annual yield on a bond for varying compounding periods in a year;
☐	**g.** calculate and interpret yield measures for fixed-rate bonds and floating-rate notes;
☐	**h.** calculate and interpret yield measures for money market instruments;
☐	**i.** define and compare the spot curve, yield curve on coupon bonds, par curve, and forward curve;
☐	**j.** define forward rates and calculate spot rates from forward rates, forward rates from spot rates, and the price of a bond using forward rates;
☐	**k.** compare, calculate, and interpret yield spread measures.

INTRODUCTION

1

Globally, the fixed-income market is a key source of financing for businesses and governments. In fact, the total market value outstanding of corporate and government bonds is significantly larger than that of equity securities. Similarly, the fixed-income

© 2021 CFA Institute. All rights reserved.

market, which is also called the debt market or bond market, represents a significant investing opportunity for institutions as well as individuals. Pension funds, mutual funds, insurance companies, and sovereign wealth funds, among others, are major fixed-income investors. Retirees who desire a relatively stable income stream often hold fixed-income securities. Clearly, understanding how to value fixed-income securities is important to investors, issuers, and financial analysts. We focus on the valuation of traditional (option-free) fixed-rate bonds, although other debt securities, such as floating-rate notes and money market instruments, are also covered.

We first describe and illustrate basic bond valuation, which includes pricing a bond using a market discount rate for each of the future cash flows and pricing a bond using a series of spot rates. Valuation using spot rates allows for each future cash flow to be discounted at a rate associated with its timing. This valuation methodology for future cash flows has applications well beyond the fixed-income market. Relationships among a bond's price, coupon rate, maturity, and market discount rate (yield-to-maturity) are also described and illustrated.

We then turn our attention to how bond prices and yields are quoted and calculated in practice. When bonds are actively traded, investors can observe the price and calculate various yield measures. However, these yield measures differ by the type of bond. In practice, different measures are used for fixed-rate bonds, floating-rate notes, and money market instruments.

We then discuss the maturity or term structure of interest rates, involving an analysis of yield curves, which illustrates the relationship between yields-to-maturity and times-to-maturity on bonds with otherwise similar characteristics. Lastly, we describe yield spreads, measures of how much additional yield over the benchmark security (usually a government bond) investors expect for bearing additional risk.

2

BOND PRICES AND THE TIME VALUE OF MONEY

a calculate a bond's price given a market discount rate

Bond pricing is an application of discounted cash flow analysis. The complexity of the pricing depends on the particular bond's features and the rate (or rates) used to do the discounting. This section starts with using a single discount factor for all future cash flows and concludes with the most general approach to bond valuation. The general approach to bond valuation is to use a series of spot rates that correspond to the timing of the future cash flows.

2.1 Bond Pricing with a Market Discount Rate

On a traditional (option-free) fixed-rate bond, the promised future cash flows are a series of coupon interest payments and repayment of the full principal at maturity. The coupon payments occur on regularly scheduled dates; for example, an annual payment bond might pay interest on 15 June of each year for five years. The final coupon typically is paid together with the full principal on the maturity date. The price of the bond at issuance is the present value of the promised cash flows. The **market discount rate** is used in the time-value-of-money calculation to obtain the present value. The market discount rate is the rate of return required by investors given the risk of the investment in the bond. It is also called the **required yield**, or the **required rate of return**.

For example, suppose the coupon rate on a bond is 4% and the payment is made once a year. If the time-to-maturity is five years and the market discount rate is 6%, the price of the bond is 91.575 per 100 of **par value**. The par value is the amount of principal on the bond.

$$\frac{4}{(1.06)^1}+\frac{4}{(1.06)^2}+\frac{4}{(1.06)^3}+\frac{4}{(1.06)^4}+\frac{104}{(1.06)^5}=$$

$$3.774+3.560+3.358+3.168+77.715=91.575.$$

The final cash flow of 104 is the redemption of principal (100) plus the coupon payment for that date (4). The price of the bond is the sum of the present values of the five cash flows. The price per 100 of par value may be interpreted as the percentage of par value. If the par value is USD100,000, the coupon payments are USD4,000 each year and the price of the bond is USD91,575. Its price is 91.575% of par value. This bond is described as trading at a **discount** because the price is below par value.

Suppose that another five-year bond has a coupon rate of 8% paid annually. If the market discount rate is again 6%, the price of the bond is 108.425.

$$\frac{8}{(1.06)^1}+\frac{8}{(1.06)^2}+\frac{8}{(1.06)^3}+\frac{8}{(1.06)^4}+\frac{108}{(1.06)^5}=$$

$$7.547+7.120+6.717+6.337+80.704=108.425.$$

This bond is trading at a **premium** because its price is above par value.

If another five-year bond pays a 6% annual coupon and the market discount rate still is 6%, the bond would trade at par value.

$$\frac{6}{(1.06)^1}+\frac{6}{(1.06)^2}+\frac{6}{(1.06)^3}+\frac{6}{(1.06)^4}+\frac{106}{(1.06)^5}=$$

$$5.660+5.340+5.038+4.753+79.209=100.000.$$

The coupon rate indicates the amount the issuer promises to pay the bondholders each year in interest. The market discount rate reflects the amount investors need to receive in interest each year in order to pay full par value for the bond. Therefore, assuming that these three bonds have the same risk, which is consistent with them having the same market discount rate, the 4% bond offers a “deficient” coupon rate. The amount of the discount below par value is the present value of the deficiency, which is 2% of par value each year. The present value of the deficiency, discounted using the market discount rate, is –8.425.

$$\frac{-2}{(1.06)^1}+\frac{-2}{(1.06)^2}+\frac{-2}{(1.06)^3}+\frac{-2}{(1.06)^4}+\frac{-2}{(1.06)^5}=-8.425.$$

The price of the 4% coupon bond is 91.575 (= 100 – 8.425). In the same manner, the 8% bond offers an “excessive” coupon rate given the risk because investors require only 6%. The amount of the premium is the present value of the excess cash flows, which is +8.425. The price of the 8% bond is 108.425 (= 100 + 8.425).

These examples demonstrate that the price of a fixed-rate bond, relative to par value, depends on the relationship of the coupon rate to the market discount rate. Here is a summary of the relationships:

- When the coupon rate is less than the market discount rate, the bond is priced at a discount below par value.
- When the coupon rate is greater than the market discount rate, the bond is priced at a premium above par value.
- When the coupon rate is equal to the market discount rate, the bond is priced at par value.

At this point, it is assumed that the bond is priced on a coupon payment date. If the bond is between coupon payment dates, the price paid will include accrued interest, which is interest that has been earned but not yet paid. Accrued interest is discussed in detail later.

Equation 1 is a general formula for calculating a bond price given the market discount rate:

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \cdots + \frac{PMT + FV}{(1+r)^N}, \tag{1}$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
r = market discount rate, or required rate of return per period
N = number of evenly spaced periods to maturity

The examples so far have been for an annual payment bond, which is the convention for most European bonds. Asian and North American bonds generally make semiannual payments, and the stated rate is the annual coupon rate. Suppose the coupon rate on a bond is stated to be 8% and the payments are made twice a year (semiannually) on 15 June and 15 December. For each 100 in par value ($FV = 100$), the coupon payment per period is 4 ($PMT = 4$). If there are three years to maturity, there are six evenly spaced semiannual periods ($N = 6$). If the market discount rate is 3% per semiannual period ($r = 0.03$), the price of the bond is 105.417 per 100 of par value.

$$\frac{4}{(1.03)^1} + \frac{4}{(1.03)^2} + \frac{4}{(1.03)^3} + \frac{4}{(1.03)^4} + \frac{4}{(1.03)^5} + \frac{104}{(1.03)^6} = 105.417.$$

This bond is trading at a premium above par value because the coupon rate of 4% *per period* is greater than the market discount rate of 3% *per period*. Usually, those interest rates are annualized by multiplying the rate per period by the number of periods in a year. Therefore, an equivalent statement is that the bond is priced at a premium because its stated *annual* coupon rate of 8% is greater than the stated *annual* market discount rate of 6%. Interest rates, unless stated otherwise, are typically quoted as annual rates.

EXAMPLE 1

Bonds Trading at a Discount, at a Premium, and at Par

Identify whether each of the following bonds is trading at a discount, at par value, or at a premium. Calculate the prices of the bonds per 100 in par value using Equation 1. If the coupon rate is deficient or excessive compared with the market discount rate, calculate the amount of the deficiency or excess per 100 of par value.

Bond	Coupon Payment per Period	Number of Periods to Maturity	Market Discount Rate per Period
A	2	6	3%
B	6	4	4%
C	5	5	5%
D	0	10	2%

Solution:

Bond A

$$\frac{2}{(1.03)^1} + \frac{2}{(1.03)^2} + \frac{2}{(1.03)^3} + \frac{2}{(1.03)^4} + \frac{2}{(1.03)^5} + \frac{102}{(1.03)^6} = 94.583.$$

Bond A is trading at a discount. Its price is below par value because the coupon rate per period (2%) is less than the required yield per period (3%). The deficiency per period is the coupon rate minus the market discount rate, times the par value: (0.02 – 0.03) × 100 = –1. The present value of deficiency is –5.417, discounted using the required yield (market discount rate) per period.

$$\frac{-1}{(1.03)^1} + \frac{-1}{(1.03)^2} + \frac{-1}{(1.03)^3} + \frac{-1}{(1.03)^4} + \frac{-1}{(1.03)^5} + \frac{-1}{(1.03)^6} = -5.417.$$

The amount of the deficiency can be used to calculate the price of the bond; the price is 94.583 (= 100 – 5.417).

Bond B

$$\frac{6}{(1.04)^1} + \frac{6}{(1.04)^2} + \frac{6}{(1.04)^3} + \frac{106}{(1.04)^4} = 107.260.$$

Bond B is trading at a premium because the coupon rate per period (6%) is greater than the market discount rate per period (4%). The excess per period is the coupon rate minus the market discount rate, times the par value: (0.06 – 0.04) × 100 = +2. The present value of excess is +7.260, discounted using the required yield per period.

$$\frac{2}{(1.04)^1} + \frac{2}{(1.04)^2} + \frac{2}{(1.04)^3} + \frac{2}{(1.04)^4} = 7.260.$$

The price of the bond is 107.260 (= 100 + 7.260).

Bond C

$$\frac{5}{(1.05)^1} + \frac{5}{(1.05)^2} + \frac{5}{(1.05)^3} + \frac{5}{(1.05)^4} + \frac{105}{(1.05)^5} = 100.000.$$

Bond C is trading at par value because the coupon rate is equal to the market discount rate. The coupon payments are neither excessive nor deficient given the risk of the bond.

Bond D

$$\frac{100}{(1.02)^{10}} = 82.035.$$

Bond D is a zero-coupon bond, which always will trade at a discount below par value (as long as the required yield is greater than zero). The deficiency in the coupon payments is –2 per period: (0 – 0.02) × 100 = –2.

$$\frac{-2}{(1.02)^1} + \frac{-2}{(1.02)^2} + \frac{-2}{(1.02)^3} + \frac{-2}{(1.02)^4} + \frac{-2}{(1.02)^5} +$$
$$\frac{-2}{(1.02)^6} + \frac{-2}{(1.02)^7} + \frac{-2}{(1.02)^8} + \frac{-2}{(1.02)^9} + \frac{-2}{(1.02)^{10}} = -17.965.$$

The price of the bond is 82.035 (= 100 – 17.965).

2.2 Yield-to-Maturity

If the market price of a bond is known, Equation 1 can be used to calculate its **yield-to-maturity** (YTM). The yield-to-maturity is the internal rate of return on the cash flows—the uniform interest rate such that when the future cash flows are discounted at that rate, the sum of the present values equals the price of the bond. It is the *implied* market discount rate.

The yield-to-maturity is the rate of return on the bond to an investor given three critical assumptions:

1 The investor holds the bond to maturity.

2 The issuer makes all the coupon and principal payments in the full amount on the scheduled dates. Therefore, the yield-to-maturity is the *promised* yield—the yield assuming the issuer does not default on any of the payments.

3 The investor is able to reinvest coupon payments at that same yield. This is a characteristic of an internal rate of return.

For example, suppose that a four-year, 5% annual coupon payment bond is priced at 105 per 100 of par value. The yield-to-maturity is the solution for the rate, r, in this equation:

$$105 = \frac{5}{(1+r)^1} + \frac{5}{(1+r)^2} + \frac{5}{(1+r)^3} + \frac{105}{(1+r)^4}.$$

Solving by trial-and-error search or using the time-value-of-money keys on a financial calculator obtains the result that $r = 0.03634$. The bond trades at a premium because its coupon rate (5%) is greater than the yield that is required by investors (3.634%). The yield is implied by the market price.

The yield-to-maturity on a bond may be positive or negative. In fact, following the launch of unconventional monetary policies in many regions after the global financial crisis of 2008–2009, yields on many government bonds fell into the negative territory. The market value of negative-yielding bonds reached $17 trillion in August 2019, amounting to roughly 1 in 3 of all investment-grade bonds, according to the Bloomberg Barclays Global-Aggregate Index. Bonds with a negative yield-to-maturity include those issued at higher yields in the past that have experienced significant price appreciation, as well as newly issued zero-coupon sovereign government bonds priced at a premium to par value. Suppose the four-year bond in the previous example were trading at a price of 122.50 per 100 of par value, instead of 105. Using Equation 1, the yield-to-maturity, r, for this bond would be –0.549%.

Yield-to-maturity does not depend on the actual amount of par value in a fixed-income portfolio. For example, suppose a Japanese institutional investor owns a three-year, 2.5% semiannual payment bond having a par value of JPY100 million. The bond currently is priced at JPY98,175,677. The yield per semiannual period can be obtained by solving this equation for r:

$$98.175677 = \frac{1.25}{(1+r)^1} + \frac{1.25}{(1+r)^2} + \frac{1.25}{(1+r)^3} + \frac{1.25}{(1+r)^4} + \frac{1.25}{(1+r)^5} + \frac{101.25}{(1+r)^6}.$$

The yield per semiannual period turns out to be 1.571% ($r = 0.01571$), which can be annualized to be 3.142% ($0.01571 \times 2 = 0.03142$). In general, a three-year, 2.5% semiannual bond for *any* amount of par value has an annualized yield-to-maturity of 3.142% if it is priced at 98.175677% of par value.

EXAMPLE 2

Yields-to-Maturity for a Premium, Discount, and Zero-Coupon Bond

Calculate the yields-to-maturity for the following bonds. The prices are stated per 100 of par value.

Bond	Coupon Payment per Period	Number of Periods to Maturity	Price
A	3.5	4	103.75
B	2.25	6	96.50
C	0	60	22.375

Solution:

Bond A

$$103.75 = \frac{3.5}{(1+r)^1} + \frac{3.5}{(1+r)^2} + \frac{3.5}{(1+r)^3} + \frac{103.5}{(1+r)^4}; r = 0.02503.$$

Bond A is trading at a premium, so its yield-to-maturity per period (2.503%) must be lower than its coupon rate per period (3.5%).

Bond B

$$96.50 = \frac{2.25}{(1+r)^1} + \frac{2.25}{(1+r)^2} + \frac{2.25}{(1+r)^3} + \frac{2.25}{(1+r)^4} + \frac{2.25}{(1+r)^5} + \frac{102.25}{(1+r)^6}; r = 0.02894.$$

Bond B is trading at a discount, so the yield-to-maturity per period (2.894%) must be higher than the coupon rate per period (2.25%).

Bond C

$$22.375 = \frac{100}{(1+r)^{60}}; r = 0.02527.$$

Bond C is a zero-coupon bond trading at a significant discount below par value. Its yield-to-maturity is 2.527% per period.

2.3 Relationships between the Bond Price and Bond Characteristics

b identify the relationships among a bond's price, coupon rate, maturity, and market discount rate (yield-to-maturity)

The price of a fixed-rate bond will change whenever the market discount rate changes. The following relationships pertain to the change in the bond price given the market discount rate:

1 The bond price is inversely related to the market discount rate. When the market discount rate increases, the bond price decreases (the inverse effect).

2 For the same coupon rate and time-to-maturity, the percentage price change is greater (in absolute value, meaning without regard to the sign of the change) when the market discount rate goes down than when it goes up (the convexity effect).

3 For the same time-to-maturity, a lower-coupon bond has a greater percentage price change than a higher-coupon bond when their market discount rates change by the same amount (the coupon effect).

4 Generally, for the same coupon rate, a longer-term bond has a greater percentage price change than a shorter-term bond when their market discount rates change by the same amount (the maturity effect).

Exhibit 1 illustrates these relationships using nine annual coupon payment bonds. The bonds have different coupon rates and times-to-maturity but otherwise are the same in terms of risk. The coupon rates are 10%, 20%, and 30% for bonds having 10, 20, and 30 years to maturity. At first, the bonds are all priced at a market discount rate of 20%. Equation 1 is used to determine the prices. Going across columns, the market discount rate is decreased by 1 percentage point, from 20% to 19%, and next, it is increased from 20% to 21%.

Exhibit 1 Relationships between Bond Prices and Bond Characteristics

				Discount Rates Go Down		Discount Rates Go Up	
Bond	**Coupon Rate**	**Maturity**	**Price at 20%**	**Price at 19%**	**% Change**	**Price at 21%**	**% Change**
A	10.00%	10	58.075	60.950	4.95%	55.405	–4.60%
B	20.00%	10	100.000	104.339	4.34%	95.946	–4.05%
C	30.00%	10	141.925	147.728	4.09%	136.487	–3.83%
D	10.00%	20	51.304	54.092	5.43%	48.776	–4.93%
E	20.00%	20	100.000	105.101	5.10%	95.343	–4.66%
F	30.00%	20	148.696	156.109	4.99%	141.910	–4.56%
G	10.00%	30	50.211	52.888	5.33%	47.791	–4.82%
H	20.00%	30	100.000	105.235	5.23%	95.254	–4.75%
I	30.00%	30	149.789	157.581	5.20%	142.716	–4.72%

The first relationship is that the bond price and the market discount rate move inversely. All bond prices in Exhibit 1 go up when the rates go down from 20% to 19%, and all prices go down when the rates go up from 20% to 21%. This happens because of the fixed cash flows on a fixed-rate bond. The numerators in Equation 1 do not change when the market discount rate in the denominators rises or falls. Therefore, the price (*PV*) moves inversely with the market discount rate (*r*).

The second relationship reflects the convexity effect. In Exhibit 1, the percentage price changes are calculated using this equation:

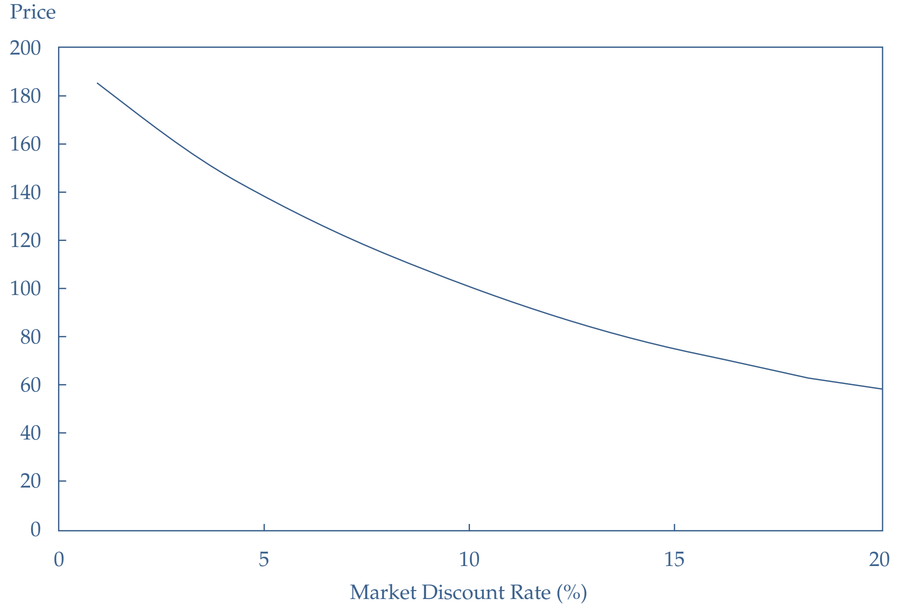
$$\%\ \text{Change} = \frac{\text{New price} - \text{Old price}}{\text{Old price}}.$$

For example, when the market discount rate on Bond A falls, the price rises from 58.075 to 60.950. The percentage price increase is 4.95%.

$$\% \text{ Change} = \frac{60.950 - 58.075}{58.075} = 0.0495.$$

For each bond, the percentage price increases are greater in *absolute value* than the percentage price decreases. This implies that the relationship between bond prices and the market discount rate is not linear; instead, it is curved. It is described as being "convex." The convexity effect is shown in Exhibit 2 for a 10%, 10-year bond.

Exhibit 2 The Convex Relationship between the Market Discount Rate and the Price of a 10-Year, 10% Annual Coupon Payment Bond



The third relationship is the coupon effect. Consider Bonds A, B, and C, which have 10 years to maturity. For both the decrease and increase in the yield-to-maturity, Bond A has a larger percentage price change than Bond B and Bond B has a larger change than C. The same pattern holds for the 20-year and 30-year bonds. Therefore, lower-coupon bonds have more price volatility than higher-coupon bonds, other things being equal.

The fourth relationship is the maturity effect. Compare the results for Bonds A and D, for Bonds B and E, and for Bonds C and F. The 20-year bonds have greater percentage price changes than the 10-year bonds for either an increase or a decrease in the market discount rate. In general, longer-term bonds have more price volatility than shorter-term bonds, other things being equal.

There are exceptions to the maturity effect. That is why the word "generally" appears in the statement of the relationship at the beginning of this section. Compare the results in Exhibit 1 for Bonds D and G, for Bonds E and H, and for Bonds F and I. For the higher-coupon bonds trading at a premium, Bonds F and I, the usual property holds: The 30-year bonds have greater percentage price changes than the 20-year bonds. The same pattern holds for Bonds E and H, which are priced initially at par value. The exception is illustrated in the results for Bonds D and G, which are priced at a discount because the coupon rate is lower than the market discount rate. The 20-year, 10% bond has a greater percentage price change than the 30-year, 10%

bond. Exceptions to the maturity effect are rare in practice. They occur only for low-coupon (but not zero-coupon), long-term bonds trading at a discount. The maturity effect always holds on zero-coupon bonds, as it does for bonds priced at par value or at a premium above par value.

One final point to note in Exhibit 1 is that Bonds B, E, and H, which have coupon rates of 20%, all trade at par value when the market discount rate is 20%. A bond having a coupon rate equal to the market discount rate is priced at par value on a coupon payment date, regardless of the number of years to maturity.

EXAMPLE 3

Bond Percentage Price Changes Based on Coupon and Time-to-Maturity

An investor is considering the following six annual coupon payment government bonds:

Bond	Coupon Rate	Time-to-Maturity	Yield-to-Maturity
A	0%	2 years	5.00%
B	5%	2 years	5.00%
C	8%	2 years	5.00%
D	0%	4 years	5.00%
E	5%	4 years	5.00%
F	8%	4 years	5.00%

1. Based on the relationships between bond prices and bond characteristics, which bond will go up in price the *most* on a percentage basis if all yields go down from 5.00% to 4.90%?
2. Based on the relationships between the bond prices and bond characteristics, which bond will go down in price the *least* on a percentage basis if all yields go up from 5.00% to 5.10%?

Solution to 1:

Bond D will go up in price the most on a percentage basis because it has the lowest coupon rate (the coupon effect) and the longer time-to-maturity (the maturity effect). There is no exception to the maturity effect in these bonds because there are no low-coupon bonds trading at a discount.

Solution to 2:

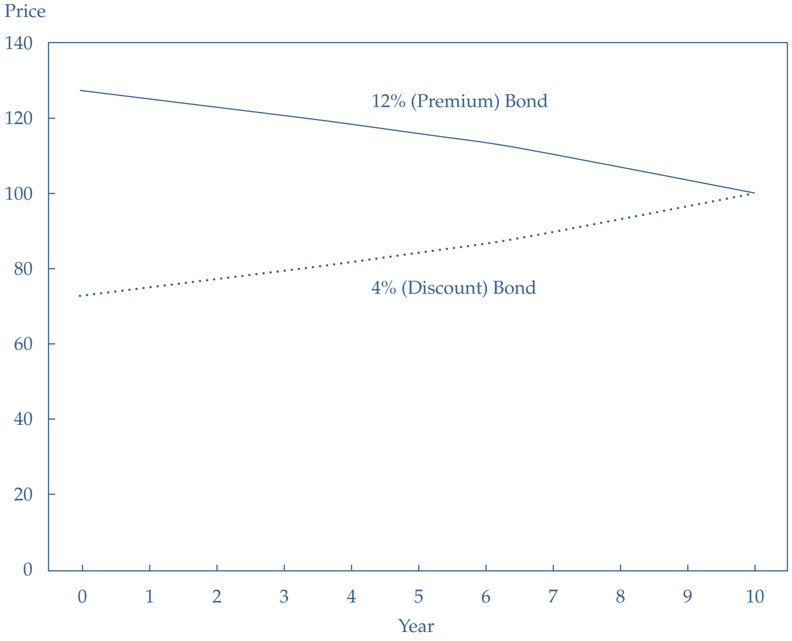
Bond C will go down in price the least on a percentage basis because it has the highest coupon rate (the coupon effect) and the shorter time-to-maturity (the maturity effect). There is no exception to the maturity effect because Bonds C and F are priced at a premium above par value.

Exhibit 2 demonstrates the impact on a bond price assuming the time-to-maturity does not change. It shows an *instantaneous* change in the market discount rate from one moment to the next.

In practice, bond prices change even if the market discount rate remains the same. As time passes, the bondholder comes closer to receiving the par value at maturity. The **constant-yield price trajectory** illustrates the change in the price of a fixed-income bond over time. This trajectory shows the "pull to par" effect on the price of a bond trading at a premium or a discount to par value. If the issuer does not default, the price of a bond approaches par value as its time-to-maturity approaches zero.

Exhibit 3 shows the constant-yield price trajectories for 4% and 12% annual coupon payment 10-year bonds. Both bonds have a market discount rate of 8%. The 4% bond's initial price is 73.160 per 100 of par value. The price increases each year and approaches par value as the maturity date nears. The 12% bond's initial price is 126.840, and it decreases each year, approaching par value as the maturity date nears. Both prices are "pulled to par."

Exhibit 3 Constant-Yield Price Trajectories for 4% and 12% Annual Coupon Payment 10-Year Bonds at a Market Discount Rate of 8%



Discount Bond	73.160	75.012	77.013	79.175	81.508	84.029	86.751	89.692	92.867	96.296	100.00
Premium Bond	126.84	124.98	122.98	120.82	118.49	115.97	113.24	110.30	107.13	103.70	100.00

2.4 Pricing Bonds Using Spot Rates

c define spot rates and calculate the price of a bond using spot rates

When a fixed-rate bond is priced using the market discount rate, the same discount rate is used for each cash flow. A more fundamental approach to calculate the price of a bond is to use a sequence of market discount rates that correspond to the cash flow dates. These market discount rates are called **spot rates**. Spot rates are yields-to-maturity on zero-coupon bonds maturing at the date of each cash flow. Sometimes these are called "zero rates." Bond price (or value) determined using the spot rates is sometimes referred to as the bond's "no-arbitrage value." If a bond's price differs from its no-arbitrage value, an arbitrage opportunity exists in the absence of transaction costs.

Suppose that the one-year spot rate is 2%, the two-year spot rate is 3%, and the three-year spot rate is 4%. Then, the price of a three-year bond that makes a 5% annual coupon payment is 102.960.

$$\frac{5}{(1.02)^1} + \frac{5}{(1.03)^2} + \frac{105}{(1.04)^3} =$$

$$4.902 + 4.713 + 93.345 = 102.960.$$

This three-year bond is priced at a premium above par value, so its yield-to-maturity must be less than 5%. Using Equation 1, the yield-to-maturity is 3.935%.

$$102.960 = \frac{5}{(1+r)^1} + \frac{5}{(1+r)^2} + \frac{105}{(1+r)^3}; r = 0.03935.$$

When the coupon and principal cash flows are discounted using the yield-to-maturity, the same price is obtained.

$$\frac{5}{(1.03935)^1} + \frac{5}{(1.03935)^2} + \frac{105}{(1.03935)^3} =$$

$$4.811 + 4.629 + 93.520 = 102.960.$$

Notice that the present values of the individual cash flows discounted using spot rates differ from those using the yield-to-maturity. The present value of the first coupon payment is 4.902 when discounted at 2%, but it is 4.811 when discounted at 3.935%. The present value of the final cash flow, which includes the redemption of principal, is 93.345 at 4% and 93.520 at 3.935%. Nevertheless, the sum of the present values using either approach is 102.960.

Equation 2 is a general formula for calculating a bond price given the sequence of spot rates:

$$PV = \frac{PMT}{(1+Z_1)^1} + \frac{PMT}{(1+Z_2)^2} + \cdots + \frac{PMT+FV}{(1+Z_N)^N}, \quad (2)$$

where

Z_1 = spot rate, or zero-coupon yield, or zero rate, for Period 1
Z_2 = spot rate, or zero-coupon yield, or zero rate, for Period 2
Z_N = spot rate, or zero-coupon yield, or zero rate, for Period N

EXAMPLE 4

Bond Prices and Yields-to-Maturity Based on Spot Rates

Calculate the price (per 100 of par value) and the yield-to-maturity for a four-year, 3% annual coupon payment bond given the following two sequences of spot rates.

Time-to-Maturity	Spot Rates A	Spot Rates B
1 year	0.39%	4.08%
2 years	1.40%	4.01%
3 years	2.50%	3.70%
4 years	3.60%	3.50%

Solution:

Spot Rates A

$$\frac{3}{(1.0039)^1} + \frac{3}{(1.0140)^2} + \frac{3}{(1.0250)^3} + \frac{103}{(1.0360)^4} =$$

$$2.988 + 2.918 + 2.786 + 89.412 = 98.104.$$

Given Spot Rates A, the four-year, 3% bond is priced at 98.104.

$$98.104 = \frac{3}{(1+r)^1} + \frac{3}{(1+r)^2} + \frac{3}{(1+r)^3} + \frac{103}{(1+r)^4}; r = 0.03516.$$

The yield-to-maturity is 3.516%.

Spot Rates B

$$\frac{3}{(1.0408)^1} + \frac{3}{(1.0401)^2} + \frac{3}{(1.0370)^3} + \frac{103}{(1.0350)^4} =$$

$$2.882 + 2.773 + 2.690 + 89.759 = 98.104.$$

$$98.104 = \frac{3}{(1+r)^1} + \frac{3}{(1+r)^2} + \frac{3}{(1+r)^3} + \frac{103}{(1+r)^4}; r = 0.03516.$$

Given Spot Rates B, the four-year, 3% bond is again priced at 98.104 to yield 3.516%.

This example demonstrates that two very different sequences of spot rates can result in the same bond price and yield-to-maturity. Spot Rates A are increasing for longer maturities, whereas Spot Rates B are decreasing.

PRICES AND YIELDS: CONVENTIONS FOR QUOTES AND CALCULATIONS

3

d describe and calculate the flat price, accrued interest, and the full price of a bond

When investors purchase shares, they pay the quoted price. For bonds, however, there can be a difference between the quoted price and the price paid. This section explains why this difference occurs and how to calculate the quoted price and the price that will be paid. It also describes how prices are estimated for bonds that are not actively traded and demonstrates how yield measures are calculated for fixed-rate bonds, floating-rate notes, and money market instruments.

3.1 Flat Price, Accrued Interest, and the Full Price

When a bond is between coupon payment dates, its price has two parts: the **flat price** (PV^{Flat}) and the **accrued interest** (AI). The sum of the parts is the **full price** (PV^{Full}), which also is called the invoice or "dirty" price. The flat price, which is the full price minus the accrued interest, is also called the quoted or "clean" price.

$$PV^{Full} = PV^{Flat} + AI. \quad (3)$$

The flat price usually is quoted by bond dealers. If a trade takes place, the accrued interest is added to the flat price to obtain the full price paid by the buyer and received by the seller on the **settlement date**. The settlement date is when the bond buyer makes cash payment and the seller delivers the security.

The reason for using the flat price for quotation is to avoid misleading investors about the market price trend for the bond. If the full price were to be quoted by dealers, investors would see the price rise day after day even if the yield-to-maturity did not change. That is because the amount of accrued interest increases each day. Then, after the coupon payment is made, the quoted price would drop dramatically. Using the flat price for quotation avoids that misrepresentation. It is the flat price that is "pulled to par" along the constant-yield price trajectory shown in Exhibit 3.

Accrued interest is the proportional share of the next coupon payment. Assume that the coupon period has "T" days between payment dates and that "t" days have gone by since the last payment. The accrued interest is calculated using Equation 4:

$$AI = \frac{t}{T} \times PMT, \tag{4}$$

where

t = number of days from the last coupon payment to the settlement date

T = number of days in the coupon period

t/T = fraction of the coupon period that has gone by since the last payment

PMT = coupon payment per period

Notice that the accrued interest part of the full price does not depend on the yield-to-maturity. Therefore, it is the flat price that is affected by a market discount rate change.

There are different conventions used in bond markets to count days. The two most common day-count conventions are actual/actual and 30/360. For the actual/actual method, the actual number of days is used, including weekends, holidays, and leap days. For example, a semiannual payment bond pays interest on 15 May and 15 November of each year. The accrued interest for settlement on 27 June would be the actual number of days between 15 May and 27 June (t = 43 days) divided by the actual number of days between 15 May and 15 November (T = 184 days), times the coupon payment. If the stated coupon rate is 4.375%, the accrued interest is 0.511209 per 100 of par value.

$$AI = \frac{43}{184} \times \frac{4.375}{2} = 0.511209.$$

Day-count conventions vary from market to market. However, actual/actual is most common for government bonds.

The 30/360 day-count convention often is used for corporate bonds. It *assumes* that each month has 30 days and that a full year has 360 days. Therefore, for this method, there are assumed to be 42 days between 15 May and 27 June: 15 days between 15 May and 30 May and 27 days between 1 June and 27 June. There are assumed to be 180 days in the six-month period between 15 May and 15 November. The accrued interest on a 4.375% semiannual payment corporate bond is 0.510417 per 100 of par value.

$$AI = \frac{42}{180} \times \frac{4.375}{2} = 0.510417.$$

The full price of a fixed-rate bond between coupon payments given the market discount rate per period (r) can be calculated with Equation 5:

$$PV^{Full} = \frac{PMT}{(1+r)^{1-t/T}} + \frac{PMT}{(1+r)^{2-t/T}} + \cdots + \frac{PMT + FV}{(1+r)^{N-t/T}}. \tag{5}$$

This is very similar to Equation 1. The difference is that the next coupon payment (PMT) is discounted for the remainder of the coupon period, which is $1 - t/T$. The second coupon payment is discounted for that fraction plus another full period, $2 - t/T$.

Equation 5 is simplified by multiplying the numerator and denominator by the expression $(1 + r)^{t/T}$. The result is Equation 6:

$$\begin{aligned} PV^{Full} &= \left[\frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \cdots + \frac{PMT + FV}{(1+r)^N}\right] \times (1+r)^{t/T} \\ &= PV \times (1+r)^{t/T}. \end{aligned} \qquad (6)$$

An advantage to Equation 6 is that PV, the expression in the brackets, is easily obtained using the time-value-of-money keys on a financial calculator because there are N evenly spaced periods. PV here is identical to Equation 1 and is not the same as PV^{Flat}.

For example, consider a 5% semiannual coupon payment government bond that matures on 15 February 2028. Accrued interest on this bond uses the actual/actual day-count convention. The coupon payments are made on 15 February and 15 August of each year. The bond is to be priced for settlement on 14 May 2019. That date is 88 days into the 181-day period. There are actually 88 days from the last coupon on 15 February to 14 May and 181 days between 15 February and the next coupon on 15 August. The annual yield-to-maturity is stated to be 4.80%. That corresponds to a market discount rate of 2.40% per semiannual period. As of the beginning of the coupon period on 15 February 2019, there would be 18 evenly spaced semiannual periods until maturity. The first step is to solve for PV using Equation 1, whereby $PMT = 2.5$, $N = 18$, $FV = 100$, and $r = 0.0240$.

$$PV = \frac{2.5}{(1.0240)^1} + \frac{2.5}{(1.0240)^2} + \cdots + \frac{102.5}{(1.0240)^{18}} = 101.447790.$$

The price of the bond would be 101.447790 per 100 of par value if its yield-to-maturity is 2.40% per period on the last coupon payment date. This is not the actual price for the bond on that date. It is a "what-if" price using the required yield that corresponds to the settlement date of 14 May 2019.

Equation 6 can be used to get the full price for the bond.

$$PV^{Full} = 101.447790 \times (1.0240)^{88/181} = 102.624323.$$

The full price is 102.624323 per 100 of par value. The accrued interest is 1.215470 per 100 of par value.

$$AI = \frac{88}{181} \times 2.5 = 1.215470.$$

The flat price is 101.408853 per 100 of par value.

$$PV^{Flat} = PV^{Full} - AI = 102.624323 - 1.215470 = 101.408853.$$

Microsoft Excel users can obtain the flat price using the PRICE financial function: PRICE(DATE(2019,5,14),DATE(2028,2,15),0.05,0.048,100,2,1). The inputs are the settlement date, maturity date, annual coupon rate as a decimal, annual yield-to-maturity as a decimal, par value, number of periods in the year, and code for the day count (0 for 30/360, 1 for actual/actual).

EXAMPLE 5

Calculating the Full Price, Accrued Interest, and Flat Price for a Bond

A 6% German corporate bond is priced for settlement on 18 June 2019. The bond makes semiannual coupon payments on 19 March and 19 September of each year and matures on 19 September 2030. The corporate bond uses the 30/360 day-count convention for accrued interest. Calculate the full price, the accrued interest, and the flat price per EUR100 of par value for three stated annual yields-to-maturity: (A) 5.80%, (B) 6.00%, and (C) 6.20%.

Solution:

Given the 30/360 day-count convention assumption, there are 89 days between the last coupon on 19 March 2015 and the settlement date on 18 June 2019 (11 days between 19 March and 30 March, plus 60 days for the full months of April and May, plus 18 days in June). Therefore, the fraction of the coupon period that has gone by is assumed to be 89/180. At the beginning of the period, there are 11.5 years (and 23 semiannual periods) to maturity.

(A) ***Stated annual yield-to-maturity of 5.80%, or 2.90% per semiannual period:***

The price at the beginning of the period is 101.661589 per 100 of par value.

$$PV = \frac{3}{(1.0290)^1} + \frac{3}{(1.0290)^2} + \cdots + \frac{103}{(1.0290)^{23}} = 101.661589.$$

The full price on 18 June is EUR103.108770.

$$PV^{Full} = 101.661589 \times (1.0290)^{89/180} = 103.108770.$$

The accrued interest is EUR1.483333, and the flat price is EUR101.625437.

$$AI = \frac{89}{180} \times 3 = 1.4833333.$$

$$PV^{Flat} = 103.108770 - 1.483333 = 101.625437.$$

(B) ***Stated annual yield-to-maturity of 6.00%, or 3.00% per semiannual period:***

The price at the beginning of the period is par value, as expected, because the coupon rate and the market discount rate are equal.

$$PV = \frac{3}{(1.0300)^1} + \frac{3}{(1.0300)^2} + \cdots + \frac{103}{(1.0300)^{23}} = 100.000000.$$

The full price on 18 June is EUR101.472251.

$$PV^{Full} = 100.000000 \times (1.0300)^{89/180} = 101.472251.$$

The accrued interest is EUR1.483333, and the flat price is EUR99.988918.

$$AI = \frac{89}{180} \times 3 = 1.4833333.$$

$$PV^{Flat} = 101.472251 - 1.483333 = 99.988918.$$

The flat price of the bond is a little below par value, even though the coupon rate and the yield-to-maturity are equal, because the accrued interest does not take into account the time value of money. The accrued interest is the interest earned by the owner of the bond for the time between the last coupon payment and the settlement date, 1.483333 per 100 of par value. However, that interest income is not received until the next coupon date. In theory, the accrued interest should be the *present value* of 1.483333. In practice, however, accounting and financial reporting need to consider issues of practicality and materiality. For those reasons, the calculation of accrued interest in practice neglects the time value of money. Therefore, compared with theory, the reported accrued interest is a little “too high” and the flat price is a little “too low.” The full price, however, is correct because it is the sum of the present values of the future cash flows, discounted using the market discount rate.

(C) ***Stated annual yield-to-maturity of 6.20%, or 3.10% per semiannual period:***

The price at the beginning of the period is 98.372607 per 100 of par value.

$$PV = \frac{3}{(1.0310)^1} + \frac{3}{(1.0310)^2} + \cdots + \frac{103}{(1.0310)^{23}} = 98.372607.$$

The full price on 18 June is EUR99.868805.

$$PV^{Full} = 98.372607 \times (1.0310)^{89/180} = 99.868805.$$

The accrued interest is EUR1.483333, and the flat price is EUR98.385472.

$$AI = \frac{89}{180} \times 3 = 1.4833333.$$

$$PV^{Flat} = 99.868805 - 1.483333 = 98.385472.$$

The accrued interest is the same in each case because it does not depend on the yield-to-maturity. The differences in the flat prices indicate the differences in the rate of return that is required by investors.

3.2 Matrix Pricing

e describe matrix pricing

Some fixed-rate bonds are not actively traded. Therefore, there is no market price available to calculate the rate of return required by investors. The same problem occurs for bonds that are not yet issued. In these situations, it is common to estimate the market discount rate and price based on the quoted or flat prices of more frequently traded comparable bonds. These comparable bonds have similar times-to-maturity, coupon rates, and credit quality. This estimation process is called **matrix pricing**.

For example, suppose that an analyst needs to value a three-year, 4% semiannual coupon payment corporate bond, Bond X. Assume that Bond X is not actively traded and that there are no recent transactions reported for this particular security. However, there are quoted prices for four corporate bonds that have very similar credit quality:

- Bond A: Two-year, 3% semiannual coupon payment bond trading at a price of 98.500
- Bond B: Two-year, 5% semiannual coupon payment bond trading at a price of 102.250
- Bond C: Five-year, 2% semiannual coupon payment bond trading at a price of 90.250
- Bond D: Five-year, 4% semiannual coupon payment bond trading at a price of 99.125

The bonds are displayed in a matrix according to the coupon rate and the time-to-maturity. This matrix is shown in Exhibit 4.

Exhibit 4 Matrix Pricing Example

	2% Coupon	3% Coupon	4% Coupon	5% Coupon
Two Years		98.500 3.786%		102.250 3.821%
Three Years			Bond X	
Four Years				
Five Years	90.250 4.181%		99.125 4.196%	

In Exhibit 4, below each bond price is the yield-to-maturity. It is stated as the yield per semiannual period times two. For example, the yield-to-maturity on the two-year, 3% semiannual coupon payment corporate bond is 3.786%.

$$98.500 = \frac{1.5}{(1+r)^1} + \frac{1.5}{(1+r)^2} + \frac{1.5}{(1+r)^3} + \frac{101.5}{(1+r)^4}; r = 0.01893; \times 2 = 0.03786.$$

Next, the analyst calculates the average yield for each year: 3.8035% for the two-year bonds and 4.1885% for the five-year bonds.

$$\frac{0.03786 + 0.03821}{2} = 0.038035.$$

$$\frac{0.04181 + 0.04196}{2} = 0.041885.$$

The estimated three-year market discount rate can be obtained with linear interpolation. The interpolated yield is 3.9318%.

$$0.038035 + \left(\frac{3-2}{5-2}\right) \times (0.041885 - 0.038035) = 0.039318.$$

Using 3.9318% as the estimated three-year annual market discount rate, the three-year, 4% semiannual coupon payment corporate bond has an estimated price of 100.191 per 100 of par value.

$$\frac{2}{(1.019659)^1} + \frac{2}{(1.019659)^2} + \frac{2}{(1.019659)^3} + \frac{2}{(1.019659)^4} + \frac{2}{(1.019659)^5} + \frac{102}{(1.019659)^6} = 100.191.$$

Notice that 3.9318% is the stated annual rate. It is divided by two to get the yield per semiannual period: (0.039318/2 = 0.019659).

Matrix pricing also is used in underwriting new bonds to get an estimate of the **required yield spread** over the **benchmark rate**. The benchmark rate typically is the yield-to-maturity on a government bond having the same or close to the same time-to-maturity. The spread is the difference between the yield-to-maturity on the new bond and the benchmark rate. The yield spread is the additional compensation required by investors for the difference in the credit risk, liquidity risk, tax status, and other risk premiums of the bond relative to the government bond. This spread is sometimes called the **spread over the benchmark**. Yield spreads are often stated in terms of basis points (bps), where one **basis point** equals one-hundredth of a percentage point. For example, if a yield-to-maturity is 2.25% and the benchmark rate is 1.50%, the yield spread is 0.75%, or 75 bps. Yield spreads are covered in more detail later.

Suppose that a corporation is about to issue a five-year bond. The corporate issuer currently has a four-year, 3% annual coupon payment debt liability on its books. The price of that bond is 102.400 per 100 of par value. This is the full price, which is the same as the flat price because the accrued interest is zero. This implies that the coupon payment has just been made and there are four full years to maturity. The four-year rate of return required by investors for this bond is 2.36%.

$$102.400 = \frac{3}{(1+r)^1} + \frac{3}{(1+r)^2} + \frac{3}{(1+r)^3} + \frac{103}{(1+r)^4};\ r = 0.0236.$$

Suppose that there are no four-year government bonds to calculate the yield spread on this security. However, there are three-year and five-year government bonds that have yields-to-maturity of 0.75% and 1.45%, respectively. The average of the two yields-to-maturity is 1.10%, which is the estimated yield for the four-year government bond. Therefore, the estimated yield spread is 126 bps over the implied benchmark rate (0.0236 – 0.0110 = 0.0126).

There usually is a different yield spread for each maturity and for each credit rating. The term structure of "risk-free" rates, which is further discussed later, is the relationship between yields-to-maturity on "risk-free" bonds and times-to-maturity. The quotation marks around "risk-free" indicate that no bond is truly without risk. The primary component of the yield spread for many bonds is compensation for credit risk, not for time-to-maturity, and as a result, the yield spreads reflect the **term structure of credit spreads**. The term structure of credit spreads is the relationship between the spreads over the "risk-free" (or benchmark) rates and times-to-maturity. These term structures are covered in more detail later.

The issuer now has an estimate of the four-year yield spread: 126 bps. This spread is a reference point for estimating the five-year spread for the newly issued bond. Suppose that the term structure of credit spreads for bonds of the corporate issuer's quality indicates that five-year spreads are about 25 bps higher than four-year spreads. Therefore, the estimated five-year required yield spread is 151 bps (0.0126 + 0.0025 = 0.0151). Given the yield-to-maturity of 1.45% on the five-year government bond, the expected market discount rate for the newly issued bond is 2.96% (0.0145 + 0.0151 = 0.0296). The corporation might set the coupon rate to be 3% and expect that the bond can be sold for a small premium above par value.

EXAMPLE 6

Using Matrix Pricing to Estimate Bond Price

An analyst needs to assign a value to an illiquid four-year, 4.5% annual coupon payment corporate bond. The analyst identifies two corporate bonds that have similar credit quality: One is a three-year, 5.50% annual coupon payment bond priced at 107.500 per 100 of par value, and the other is a five-year, 4.50% annual coupon payment bond priced at 104.750 per 100 of par value. Using matrix pricing, the estimated price of the illiquid bond per 100 of par value is *closest* to:

A 103.895.

B 104.991.

C 106.125.

Solution:

B is correct. The first step is to determine the yields-to-maturity on the observed bonds. The required yield on the three-year, 5.50% bond priced at 107.500 is 2.856%.

$$107.500 = \frac{5.50}{(1+r)^1} + \frac{5.50}{(1+r)^2} + \frac{105.50}{(1+r)^3};\ r = 0.02856.$$

The required yield on the five-year, 4.50% bond priced at 104.750 is 3.449%.

$$104.750 = \frac{4.50}{(1+r)^1} + \frac{4.50}{(1+r)^2} + \frac{4.50}{(1+r)^3} + \frac{4.50}{(1+r)^4} + \frac{104.50}{(1+r)^5};\ r = 0.03449.$$

The estimated market discount rate for a four-year bond having the same credit quality is the average of two required yields:

$$\frac{0.02856 + 0.03449}{2} = 0.031525.$$

Given an estimated yield-to-maturity of 3.1525%, the estimated price of the illiquid four-year, 4.50% annual coupon payment corporate bond is 104.991 per 100 of par value.

$$\frac{4.50}{(1.031525)^1} + \frac{4.50}{(1.031525)^2} + \frac{4.50}{(1.031525)^3} + \frac{104.50}{(1.031525)^4} = 104.991.$$

3.3 Annual Yields for Varying Compounding Periods in the Year

f calculate annual yield on a bond for varying compounding periods in a year

There are many ways to measure the rate of return on a fixed-rate bond investment. Consider a five-year, zero-coupon government bond. The purchase price today is 80. The investor receives 100 at redemption in five years. One possible yield measure is 25%—the gain of 20 divided by the amount invested, 80. However, investors want a yield measure that is *standardized* to allow for comparison between bonds that have different times-to-maturity. Therefore, yield measures typically are *annualized*. A possible annual rate for this zero-coupon bond is 5% per year—25% divided by five years. But for bonds maturing in more than one year, investors want an *annualized and compounded* yield-to-maturity. Money market rates on instruments maturing in one year or less typically are *annualized but not compounded*. They are stated on a simple interest basis. This concept is covered later.

In general, an annualized and compounded yield on a fixed-rate bond depends on the assumed number of periods in the year, which is called the **periodicity** of the annual rate. Typically, the periodicity matches the frequency of coupon payments. A bond that pays semiannual coupons has a stated annual yield-to-maturity for a periodicity of two—the rate per semiannual period times two. A bond that pays quarterly coupons has a stated annual yield for a periodicity of four—the rate per quarter times four. It is always important to know the periodicity of a stated annual rate.

The periodicity of the annual market discount rate for a zero-coupon bond is arbitrary because there are no coupon payments. For semiannual compounding, the annual yield-to-maturity on the five-year, zero-coupon bond priced at 80 per 100 of par value is stated to be 4.5130%. This annual rate has a periodicity of two.

$$80 = \frac{100}{(1+r)^{10}};\ r = 0.022565;\ \times 2 = 0.045130.$$

For quarterly compounding, the annual yield-to-maturity is stated to be 4.4880%. This annual rate has a periodicity of four.

$$80 = \frac{100}{(1+r)^{20}};\ r = 0.011220; \times 4 = 0.044880.$$

For monthly compounding, the annual yield-to-maturity is stated to be 4.4712%. This annual rate has a periodicity of 12.

$$80 = \frac{100}{(1+r)^{60}};\ r = 0.003726; \times 12 = 0.044712.$$

For annual compounding, the yield-to-maturity is stated to be 4.5640%. This annual rate has a periodicity of one.

$$80 = \frac{100}{(1+r)^{5}};\ r = 0.045640; \times 1 = 0.045640.$$

This is known as an **effective annual rate**. An effective annual rate has a periodicity of one because there is just one compounding period in the year.

In this zero-coupon bond example, 2.2565% compounded two times a year, 1.1220% compounded four times a year, and 0.3726% compounded 12 times a year are all equivalent to an effective annual rate of 4.5640%. The compounded total return is the same for each expression for the annual rate. They differ in terms of the number of compounding periods per year—that is, in terms of the *periodicity* of the annual rate. For a given pair of cash flows, the stated annual rate and the periodicity are inversely related.

The most common periodicity for US dollar-denominated bond yields is two because most bonds in the US dollar market make semiannual coupon payments. An annual rate having a periodicity of two is known as a **semiannual bond basis yield**, or **semiannual bond equivalent yield**. Therefore, a semiannual bond basis yield is the yield per semiannual period times two. It is important to remember that "semiannual bond basis yield" and "yield per semiannual period" have different meanings. For example, if a bond yield is 2% per semiannual period, its annual yield is 4% when stated on a semiannual bond basis.

An important tool used in fixed-income analysis is to convert an annual yield from one periodicity to another. These are called periodicity, or compounding, conversions. A general formula to convert an annual percentage rate for m periods per year, denoted APR_m, to an annual percentage rate for n periods per year, APR_n, is Equation 7:

$$\left(1 + \frac{APR_m}{m}\right)^m = \left(1 + \frac{APR_n}{n}\right)^n. \tag{7}$$

For example, suppose that a three-year, 5% semiannual coupon payment corporate bond is priced at 104 per 100 of par value. Its yield-to-maturity is 3.582%, quoted on a semiannual bond basis for a periodicity of two: 0.01791 × 2 = 0.03582.

$$104 = \frac{2.5}{(1+r)^1} + \frac{2.5}{(1+r)^2} + \frac{2.5}{(1+r)^3} + \frac{2.5}{(1+r)^4} + \frac{2.5}{(1+r)^5} + \frac{102.5}{(1+r)^6};\ r = 0.01791.$$

To compare this bond with others, an analyst converts this annualized yield-to-maturity to quarterly and monthly compounding. Doing so entails using Equation 7 to convert from a periodicity of $m = 2$ to periodicities of $n = 4$ and $n = 12$.

$$\left(1 + \frac{0.03582}{2}\right)^2 = \left(1 + \frac{APR_4}{4}\right)^4;\ APR_4 = 0.03566.$$

$$\left(1+\frac{0.03582}{2}\right)^2=\left(1+\frac{APR_{12}}{12}\right)^{12};\ APR_{12}=0.03556.$$

An annual yield-to-maturity of 3.582% for semiannual compounding provides the same rate of return as annual yields of 3.566% and 3.556% for quarterly and monthly compounding, respectively. A general rule for these periodicity conversions is *compounding more frequently at a lower annual rate corresponds to compounding less frequently at a higher annual rate*. This rule can be used to check periodicity conversion calculations.

Equation 7 also applies to *negative* bond yields. Government bond yields have been negative in several countries, including Switzerland, Germany, Sweden, and Japan. For a simple example, consider a five-year, zero-coupon bond priced at 105 (% of par value). Its yield-to-maturity is –0.971% stated as an effective annual rate for a periodicity of 1.

$$105=\frac{100}{(1+r)^5};\ r=-0.00971.$$

Converting that to semiannual and monthly compounding for periodicities of 2 and 12 results in annual yields of –0.973% and –0.975%, respectively.

$$\left(1+\frac{-0.00971}{1}\right)^1=\left(1+\frac{APR_2}{2}\right)^2;\ APR_2=-0.00973.$$

$$\left(1+\frac{-0.00971}{1}\right)^1=\left(1+\frac{APR_{12}}{12}\right)^{12};\ APR_{12}=-0.00975.$$

Compounding more frequently within the year results in a lower (more negative) yield-to-maturity.

EXAMPLE 7

Yield Conversion Based on Periodicity

A five-year, 4.50% semiannual coupon payment government bond is priced at 98 per 100 of par value. Calculate the annual yield-to-maturity stated on a semiannual bond basis, rounded to the nearest basis point. Convert that annual yield to

A an annual rate that can be used for direct comparison with otherwise comparable bonds that make *quarterly* coupon payments and

B an annual rate that can be used for direct comparison with otherwise comparable bonds that make *annual* coupon payments.

Solution:

The stated annual yield-to-maturity on a semiannual bond basis is 4.96% (0.0248 × 2 = 0.0496).

$$98 = \frac{2.25}{(1+r)^1} + \frac{2.25}{(1+r)^2} + \frac{2.25}{(1+r)^3} + \frac{2.25}{(1+r)^4} + \frac{2.25}{(1+r)^5} + \frac{2.25}{(1+r)^6} + \frac{2.25}{(1+r)^7} + \frac{2.25}{(1+r)^8} + \frac{2.25}{(1+r)^9} + \frac{102.25}{(1+r)^{10}}; r = 0.0248.$$

A Convert 4.96% from a periodicity of two to a periodicity of four:

$$\left(1 + \frac{0.0496}{2}\right)^2 = \left(1 + \frac{APR_4}{4}\right)^4; APR_4 = 0.0493.$$

The annual percentage rate of 4.96% for compounding semiannually compares with 4.93% for compounding quarterly. That makes sense because increasing the frequency of compounding lowers the annual rate.

B Convert 4.96% from a periodicity of two to a periodicity of one:

$$\left(1 + \frac{0.0496}{2}\right)^2 = \left(1 + \frac{APR_1}{1}\right)^1; APR_1 = 0.0502.$$

The annual rate of 4.96% for compounding semiannually compares with an effective annual rate of 5.02%. Converting from more frequent to less frequent compounding entails raising the annual percentage rate.

3.4 Yield Measures for Fixed-Rate Bonds

g calculate and interpret yield measures for fixed-rate bonds and floating-rate notes

An important concern for quoting and calculating bond yields-to-maturity is the actual timing of the cash flows. Consider a 6% semiannual payment corporate bond that matures on 15 March 2028. Suppose that for settlement on 23 January 2020, the bond is priced at 98.5 per 100 of par value to yield 6.236% quoted on a semiannual bond basis. Its coupon payments are scheduled for 15 March and 15 September of each year. The yield calculation implicitly assumes that the payments are made on those dates. It neglects the reality that 15 March 2020 is a Sunday and 15 March 2025 is a Saturday. In fact, the coupon payments will be made to investors on the following Monday.

Yield measures that neglect weekends and holidays are quoted on what is called **street convention**. The street convention yield-to-maturity is the internal rate of return on the cash flows assuming the payments are made on the scheduled dates. This assumption simplifies bond price and yield calculations and commonly is used in practice. Sometimes the **true yield** is also quoted. The true yield-to-maturity is the internal rate of return on the cash flows using the actual calendar of weekends and bank holidays. The true yield is never higher than the street convention yield because weekends and holidays delay the time to payment. The difference is typically small, no more than a basis point or two. Therefore, the true yield is not commonly used in practice. Sometimes, a **government equivalent yield** is quoted for a corporate bond. A government equivalent yield restates a yield-to-maturity based on a 30/360 day count to one based on actual/actual. The government equivalent yield on a corporate bond can be used to obtain the spread over the government yield. Doing so keeps the yields stated on the same day-count convention basis.

Another yield measure that is commonly quoted for fixed-income bonds is the **current yield**, also called the income or running yield. The current yield is the sum of the coupon payments received over the year divided by the flat price. For example, a 10-year, 2% semiannual coupon payment bond is priced at 95 per 100 of par value. Its current yield is 2.105%.

$$\frac{2}{95} = 0.02105.$$

The current yield is a crude measure of the rate of return to an investor because it neglects the frequency of coupon payments in the numerator and any accrued interest in the denominator. It focuses only on interest income. In addition to collecting and reinvesting coupon payments, the investor has a gain if the bond is purchased at a discount and is redeemed at par value. The investor has a loss if the bond is purchased at a premium and is redeemed at par value. Sometimes the **simple yield** on a bond is quoted. It is the sum of the coupon payments plus the straight-line amortized share of the gain or loss, divided by the flat price. Simple yields are used mostly to quote Japanese government bonds, known as "JGBs."

EXAMPLE 8

Comparing Yields for Different Periodicities

An analyst observes these reported statistics for two bonds.

	Bond A	Bond B
Annual Coupon Rate	8.00%	12.00%
Coupon Payment Frequency	Semiannually	Quarterly
Years to Maturity	5 Years	5 Years
Price (per 100 of par value)	90	105
Current Yield	8.889%	11.429%
Yield-to-Maturity	10.630%	10.696%

1 Confirm the calculation of the two yield measures for the two bonds.

2 The analyst believes that Bond B has a little more risk than Bond A. How much additional compensation, in terms of a higher yield-to-maturity, does a buyer of Bond B receive for bearing this risk compared with Bond A?

Solution to 1:

Current yield for Bond A:

$$\frac{8}{90} = 0.08889.$$

Yield-to-maturity for Bond A:

$$90 = \frac{4}{(1+r)^1} + \frac{4}{(1+r)^2} + \cdots + \frac{104}{(1+r)^{10}};\ r = 0.05315;\ \times 2 = 0.10630.$$

Current yield for Bond B:

$$\frac{12}{105} = 0.11429.$$

Yield-to-maturity for Bond B:

$$105 = \frac{3}{(1+r)^1} + \frac{3}{(1+r)^2} + \cdots + \frac{103}{(1+r)^{20}};\ r = 0.02674; \times 4 = 0.10696.$$

Solution to 2:

The yield-to-maturity on Bond A of 10.630% is an annual rate for compounding semiannually. The yield-to-maturity on Bond B of 10.696% is an annual rate for compounding quarterly. The difference in the yields is *not* 6.6 bps (0.10696 – 0.10630 = 0.00066). It is essential to compare the yields for the same periodicity to make a statement about relative value.

A yield-to-maturity of 10.630% for a periodicity of two converts to 10.492% for a periodicity of four:

$$\left(1 + \frac{0.10630}{2}\right)^2 = \left(1 + \frac{APR_4}{4}\right)^4;\ APR_4 = 0.10492.$$

A yield-to-maturity of 10.696% for a periodicity of four converts to 10.839% for a periodicity of two:

$$\left(1 + \frac{0.10696}{4}\right)^4 = \left(1 + \frac{APR_2}{2}\right)^2;\ APR_2 = 0.10839.$$

The additional compensation for the greater risk in Bond B is 20.9 bps (0.10839 – 0.10630 = 0.00209) when the yields are stated on a semiannual bond basis. The additional compensation is 20.4 bps (0.10696 – 0.10492 = 0.00204) when both are annualized for quarterly compounding.

If a fixed-rate bond contains an **embedded option**, other yield measures are used. An embedded option is part of the security and cannot be removed and sold separately. For example, a **callable bond** contains an embedded call option that gives the issuer the right to buy the bond back from the investor at specified prices on predetermined dates. The preset dates usually coincide with coupon payment dates after a **call protection** period. A call protection period is the time during which the issuer of the bond is not allowed to exercise the call option.

Suppose that a seven-year, 8% annual coupon payment bond is first callable in four years. That gives the investor four years of protection against the bond being called. After the call protection period, the issuer might exercise the call option if interest rates decrease or the issuer's credit quality improves. Those circumstances allow the issuer to refinance the debt at a lower cost of funds. The preset prices that the issuer pays if the bond is called often are at a premium above par. For example, the "call schedule" for this bond might be that it is callable first at 102 (per 100 of par value) on the coupon payment date in four years, at 101 in five years, and at par value on coupon payment dates thereafter.

The yield-to-maturity on this seven-year, 8% callable bond is just one of several traditional yield measures for the investment. Others are yield-to-first-call, yield-to-second-call, and so on. If the current price for the bond is 105 per 100 of par value, the yield-to-first-call in four years is 6.975%.

$$105 = \frac{8}{(1+r)^1} + \frac{8}{(1+r)^2} + \frac{8}{(1+r)^3} + \frac{8+102}{(1+r)^4},\quad r = 0.06975$$

The yield-to-second-call in five years is 6.956%.

$$105 = \frac{8}{(1+r)^1} + \frac{8}{(1+r)^2} + \frac{8}{(1+r)^3} + \frac{8}{(1+r)^4} + \frac{8+101}{(1+r)^5}; r = 0.06956.$$

The yield-to-third-call is 6.953%.

$$105 = \frac{8}{(1+r)^1} + \frac{8}{(1+r)^2} + \frac{8}{(1+r)^3} + \frac{8}{(1+r)^4} + \frac{8}{(1+r)^5} + \frac{8+100}{(1+r)^6}; r = 0.06953.$$

Finally, the yield-to-maturity is 7.070%.

$$105 = \frac{8}{(1+r)^1} + \frac{8}{(1+r)^2} + \frac{8}{(1+r)^3} + \frac{8}{(1+r)^4} + \frac{8}{(1+r)^5} + \frac{8}{(1+r)^6} +$$
$$\frac{8+100}{(1+r)^7}; r = 0.07070.$$

Each calculation is based on Equation 1, whereby the call price (or par value) is used for FV. The lowest of the sequence of yields-to-call and the yield-to-maturity is known as the **yield-to-worst**. In this case, it is the yield-to-third-call of 6.953%. The intent of this yield measure is to provide to the investor the most conservative assumption for the rate of return.

The yield-to-worst is a commonly cited yield measure for fixed-rate callable bonds used by bond dealers and investors. However, a more precise approach is to use an option pricing model and an assumption about future interest rate volatility to value the embedded call option. The value of the embedded call option is added to the flat price of the bond to get the **option-adjusted price**. The investor bears the call risk (the bond issuer has the option to call), so the embedded call option reduces the value of the bond from the investor's perspective. The investor pays a lower price for the callable bond than if it were option-free. If the bond were non-callable, its price would be higher. The option-adjusted price is used to calculate the **option-adjusted yield**. The option-adjusted yield is the required market discount rate whereby the price is adjusted for the value of the embedded option. The value of the call option is the price of the option-free bond minus the price of the callable bond.

3.5 Yield Measures for Floating-Rate Notes

Floating-rate notes are very different from fixed-rate bonds. The interest payments on a floating-rate note, which often is called a floater or an FRN, are not fixed. Instead, they vary from period to period depending on the current level of a reference interest rate. The interest payments could go up or down; that is why they are called "floaters." The intent of an FRN is to offer the investor a security that has less market price risk than a fixed-rate bond when market interest rates fluctuate. In principle, a floater has a stable price even in a period of volatile interest rates. With a traditional fixed-income security, interest rate volatility affects the price because the future cash flows are constant. With a floating-rate note, interest rate volatility affects future interest payments.

The reference rate on a floating-rate note usually is a short-term money market rate. The three-month Libor (London Interbank Offered Rate) had long been the reference rate for many floating-rate notes. As discussed earlier, Libor is being phased out. We refer to the new rate as the Market Reference Rate (MRR). The principal on the floater typically is non-amortizing and is redeemed in full at maturity. As of June 2020, the reference rate is determined at the beginning of the period, and the interest payment is made at the end of the period. This payment structure is called "in arrears." The most common day-count conventions for calculating accrued interest on floaters are actual/360 and actual/365.

Although there are many varieties of FRNs, only the most common and traditional floaters are covered here. On these floaters, a specified yield spread is added to or subtracted from the reference rate. For example, the floater might reset its interest rate quarterly at MRR plus 0.50%. This specified yield spread over the reference rate is called the **quoted margin** on the FRN. The role of the quoted margin is to compensate the investor for the difference in the credit risk of the issuer and that implied by the reference rate. For example, a company with a stronger credit rating than that of the banks or financial institutions from which MRR is computed may be able to obtain a "sub-MRR" cost of borrowed funds, which results in a negative quoted margin. An AAA rated company might be able to issue an FRN that pays MRR minus 0.25%.

The **required margin** is the yield spread over or under the reference rate such that the FRN is priced at par value on a rate reset date. Suppose that a traditional floater is issued at par value and pays MRR plus 0.50%. The quoted margin is 50 bps. If there is no change in the credit risk of the issuer, the required margin remains at 50 bps. On each quarterly reset date, the floater will be priced at par value. Between coupon dates, its flat price will be at a premium or discount to par value if MRR goes, respectively, down or up. However, if the required margin continues to be the same as the quoted margin, the flat price is "pulled to par" as the next reset date nears. At the reset date, any change in MRR is included in the interest payment for the next period.

Changes in the required margin usually come from changes in the issuer's credit risk. Changes in liquidity or tax status also could affect the required margin. Suppose that on a reset date, the required margin goes up to 75 bps because of a downgrade in the issuer's credit rating. A floater having a quoted margin of 50 bps now pays its investors a "deficient" interest payment. This FRN will be priced at a discount below par value. The amount of the discount is the present value of the deficient future cash flows. That annuity is 25 bps per period for the remaining life of the bond. It is the difference between the required and quoted margins. If the required margin goes down from 50 bps to 40 bps, the FRN will be priced at a premium. The amount of the premium is the present value of the 10 bp annuity for the "excess" interest payment each period.

Fixed-rate and floating-rate bonds are essentially the same with respect to changes in credit risk. With fixed-rate bonds, the premium or discount arises from a difference in the fixed coupon rate and the required yield-to-maturity. With floating-rate bonds, the premium or discount arises from a difference in the fixed quoted margin and the required margin. However, fixed-rate and floating-rate bonds are very different with respect to changes in benchmark interest rates.

The valuation of a floating-rate note needs a pricing model. Equation 8 is a simplified FRN pricing model. Following market practice, the required margin is called the **discount margin**.

$$PV = \frac{\frac{(\text{Index} + QM) \times FV}{m}}{\left(1 + \frac{\text{Index} + DM}{m}\right)^1} + \frac{\frac{(\text{Index} + QM) \times FV}{m}}{\left(1 + \frac{\text{Index} + DM}{m}\right)^2} + \cdots + \frac{\frac{(\text{Index} + QM) \times FV}{m} + FV}{\left(1 + \frac{\text{Index} + DM}{m}\right)^N}, \tag{8}$$

where

PV = present value, or the price of the floating-rate note
Index = reference rate, stated as an annual percentage rate
QM = quoted margin, stated as an annual percentage rate
FV = future value paid at maturity, or the par value of the bond
m = periodicity of the floating-rate note, the number of payment periods per year
DM = discount margin, the required margin stated as an annual percentage rate
N = number of evenly spaced periods to maturity

This equation is similar to Equation 1, which is the basic pricing formula for a fixed-rate bond given the market discount rate. In Equation 1, PMT is the coupon payment *per period*. Here, *annual* rates are used. The first interest payment is the annual rate for the period (Index + QM) times the par value (FV) and divided by the number of periods in the year (m). In Equation 1, the market discount rate per period (r) is used to discount the cash flows. Here, the discount rate per period is the reference rate plus the discount margin (Index + DM) divided by the periodicity (m).

This is a simplified FRN pricing model, for several reasons. First, PV is for a rate reset date when there are N evenly spaced periods to maturity. There is no accrued interest so that the flat price is the full price. Second, the model assumes a 30/360 day-count convention so that the periodicity is an integer. Third, and most important, the same reference rate (Index) is used for all payment periods in both the numerators and denominators. More complex FRN pricing models use projected future rates for Index in the numerators and spot rates in the denominators. Therefore, the calculation for DM depends on the simplifying assumptions in the pricing model.

Suppose that a two-year FRN pays MRR plus 0.50%. Currently, MRR is 1.25%. In Equation 8, Index = 0.0125, QM = 0.0050, and m = 2. The numerators in Equation 8, ignoring the repayment of principal, are 0.875.

$$\frac{(\text{Index} + QM) \times FV}{m} = \frac{(0.0125 + 0.0050) \times 100}{2} = 0.875.$$

Suppose that the yield spread required by investors is 40 bps over the reference rate; DM = 0.0040. The assumed discount rate per period is 0.825%.

$$\frac{\text{Index} + DM}{m} = \frac{0.0125 + 0.0040}{2} = 0.00825.$$

Using Equation 8 for N = 4, the FRN is priced at 100.196 per 100 of par value.

$$\frac{0.875}{(1 + 0.00825)^1} + \frac{0.875}{(1 + 0.00825)^2} + \frac{0.875}{(1 + 0.00825)^3} + \frac{0.875 + 100}{(1 + 0.00825)^4} = 100.196.$$

This floater is priced at a premium above par value because the quoted margin is greater than the discount margin.

A similar calculation is to estimate the discount margin given the market price of the floating-rate note. Suppose that a five-year FRN pays MRR plus 0.75% on a quarterly basis. Currently, MRR is 1.10%. The price of the floater is 95.50 per 100 of par value, a discount below par value because of a downgrade in the issuer's credit rating.

$$\frac{(\text{Index} + QM) \times FV}{m} = \frac{(0.0110 + 0.0075) \times 100}{4} = 0.4625.$$

In Equation 8, use $PV = 95.50$ and $N = 20$.

$$95.50 = \frac{0.4625}{\left(1 + \frac{0.0110 + DM}{4}\right)^1} + \frac{0.4625}{\left(1 + \frac{0.0110 + DM}{4}\right)^2} + \cdots + \frac{0.4625 + 100}{\left(1 + \frac{0.0110 + DM}{4}\right)^{20}}.$$

This has the same format as Equation 1, which can be used to solve for the market discount rate per period, $r = 0.7045\%$.

$$95.50 = \frac{0.4625}{(1 + r)^1} + \frac{0.4625}{(1 + r)^2} + \cdots + \frac{0.4625 + 100}{(1 + r)^{20}}; r = 0.007045.$$

This can be used to solve for $DM = 1.718\%$.

$$0.007045 = \frac{0.0110 + DM}{4}; DM = 0.01718.$$

If this FRN was issued at par value, investors required at that time a spread of only 75 bps over MRR. Now, after the credit downgrade, investors require an *estimated* discount margin of 171.8 bps. The floater trades at a discount because the quoted margin remains fixed at 75 bps. The calculated discount margin is an estimate because it is based on a simplified FRN pricing model.The FRN pricing model in Equation 8 similarly applies to adjustable-rate loans made by banks and other non-securitized fixed-income instruments. Because a large portion of a bank's funding comes from short-duration investor deposits, banks prefer to make floating-rate unsecured loans to businesses and individuals, as opposed to fixed-rate loans, in order to help maintain a match between assets and liabilities on the balance sheet. This model is used to determine the appropriate interest rate to pay on demand or certificates of deposit (CDs), as well as to develop scorecards for risk-based floating-rate loan pricing.

EXAMPLE 9

Calculating the Discount Margin for a Floating-Rate Note

A four-year French floating-rate note pays three-month Euribor (Euro Interbank Offered Rate, an index produced by the European Banking Federation) plus 1.25%. The floater is priced at 98 per 100 of par value. Calculate the discount margin for the floater assuming that three-month Euribor is constant at 2%. Assume the 30/360 day-count convention and evenly spaced periods.

Solution:

By assumption, the interest payment each period is 0.8125 per 100 of par value.

$$\frac{(\text{Index} + QM) \times FV}{m} = \frac{(0.0200 + 0.0125) \times 100}{4} = 0.8125.$$

The discount margin can be estimated by solving for DM in this equation.

$$98 = \frac{0.8125}{\left(1 + \frac{0.0200 + DM}{4}\right)^1} + \frac{0.8125}{\left(1 + \frac{0.0200 + DM}{4}\right)^2} + \cdots + \frac{0.8125 + 100}{\left(1 + \frac{0.0200 + DM}{4}\right)^{16}}.$$

The solution for the discount rate per period is 0.9478%.

$$98 = \frac{0.8125}{(1+r)^1} + \frac{0.8125}{(1+r)^2} + \cdots + \frac{0.8125 + 100}{(1+r)^{16}};\ r = 0.009478.$$

Therefore, $DM = 1.791\%$.

$$0.009478 = \frac{0.0200 + DM}{4};\ DM = 0.01791.$$

The quoted margin is 125 bps over the Euribor reference rate. Using the simplified FRN pricing model, it is estimated that investors require a 179.1 bp spread for the floater to be priced at par value.

3.6 Yield Measures for Money Market Instruments

h calculate and interpret yield measures for money market instruments

Money market instruments are short-term debt securities. They range in time-to-maturity from overnight sale and repurchase agreements (repos) to one-year bank certificates of deposit. Money market instruments also include commercial paper, government issues of less than one year, bankers' acceptances, and time deposits based on such indexes as Libor and Euribor. Money market mutual funds are a major investor in such securities. These mutual funds can invest only in certain eligible money market securities.

There are several important differences in yield measures between the money market and the bond market:

1. Bond yields-to-maturity are annualized and compounded. Yield measures in the money market are annualized but not compounded. Instead, the rate of return on a money market instrument is stated on a simple interest basis.
2. Bond yields-to-maturity can be calculated using standard time-value-of-money analysis and with formulas programmed into a financial calculator. Money market instruments often are quoted using non-standard interest rates and require different pricing equations than those used for bonds.
3. Bond yields-to-maturity usually are stated for a common periodicity for all times-to-maturity. Money market instruments having different times-to-maturity have different periodicities for the annual rate.

In general, quoted money market rates are either **discount rates** or **add-on rates**. Although market conventions vary around the world, commercial paper, Treasury bills (US government securities issued with a maturity of one year or less), and bankers' acceptances often are quoted on a discount rate basis. Bank certificates of deposit, repos, and such indexes as Libor and Euribor are quoted on an add-on rate basis. It is important to understand that "discount rate" has a unique meaning in the money market. In general, discount rate means "interest rate used to calculate a present value"—for instance, "market discount rate" as used in our discussion. In the money market, however, discount rate is a specific type of quoted rate. Some examples will clarify this point.

Equation 9 is the pricing formula for money market instruments quoted on a discount rate basis.

$$PV = FV \times \left(1 - \frac{\text{Days}}{\text{Year}} \times DR\right), \tag{9}$$

where

PV = present value, or the price of the money market instrument
FV = future value paid at maturity, or the face value of the money market instrument
Days = number of days between settlement and maturity
Year = number of days in the year
DR = discount rate, stated as an annual percentage rate

Suppose that a 91-day US Treasury bill (T-bill) with a face value of USD10 million is quoted at a discount rate of 2.25% for an assumed 360-day year. Enter FV = 10,000,000, Days = 91, Year = 360, and DR = 0.0225. The price of the T-bill is USD9,943,125.

$$PV = 10{,}000{,}000 \times \left(1 - \frac{91}{360} \times 0.0225\right) = 9{,}943{,}125.$$

The unique characteristics of a money market discount rate can be examined with Equation 10, which transforms Equation 9 algebraically to isolate the DR term.

$$DR = \left(\frac{\text{Year}}{\text{Days}}\right) \times \left(\frac{FV - PV}{FV}\right). \tag{10}$$

The first term, Year/Days, is the periodicity of the annual rate. The second term reveals the odd character of a money market discount rate. The numerator, $FV - PV$, is the interest earned on the T-bill, USD56,875 (= 10,000,000 – 9,943,125), over the 91 days to maturity. However, the denominator is FV, not PV. In theory, an interest rate is the amount earned divided by the investment amount (PV)—not divided by the total return at maturity, which includes the earnings (FV). Therefore, by design, a money market discount rate *understates* the rate of return to the investor, and it *understates* the cost of borrowed funds to the issuer. That is because PV is less than FV (as long as DR is greater than zero).

Equation 11 is the pricing formula for money market instruments quoted on an add-on rate basis.

$$PV = \frac{FV}{\left(1 + \frac{\text{Days}}{\text{Year}} \times AOR\right)}, \tag{11}$$

where

PV = present value, principal amount, or the price of the money market instrument
FV = future value, or the redemption amount paid at maturity including interest
Days = number of days between settlement and maturity
Year = number of days in the year
AOR = add-on rate, stated as an annual percentage rate

Suppose that a Canadian pension fund buys a 180-day banker's acceptance (BA) with a quoted add-on rate of 4.38% for a 365-day year. If the initial principal amount is CAD10 million, the redemption amount due at maturity is found by re-arranging Equation 11 and entering PV = 10,000,000, Days = 180, Year = 365, and AOR = 0.0438.

$$FV = 10{,}000{,}000 + \left(10{,}000{,}000 \times \frac{180}{365} \times 0.0438\right) = 10{,}216{,}000.$$

At maturity, the pension fund receives CAD10,216,000, the principal of CAD10 million plus interest of CAD216,000. The interest is calculated as the principal times the fraction of the year times the annual add-on rate. It is added to the principal to determine the redemption amount.

Suppose that after 45 days, the pension fund sells the BA to a dealer. At that time, the quoted add-on rate for a 135-day BA is 4.17%. The sale price for the BA can be calculated using Equation 11 for FV = 10,216,000, Days = 135, Year = 365, and AOR = 0.0417. The sale price is CAD10,060,829.

$$PV = \frac{10{,}216{,}000}{\left(1 + \frac{135}{365} \times 0.0417\right)} = 10{,}060{,}829.$$

The characteristics of an add-on rate can be examined with Equation 12, which transforms Equation 11 algebraically to isolate the AOR term.

$$AOR = \left(\frac{\text{Year}}{\text{Days}}\right) \times \left(\frac{FV - PV}{PV}\right). \tag{12}$$

This equation indicates that an add-on rate is a reasonable yield measure for a money market investment. The first term, Year/Days, is the periodicity of the annual rate. The second term is the interest earned, $FV - PV$, divided by PV, the amount invested.

The pension fund's rate of return on its 45-day investment in the banker's acceptance can be calculated with Equation 12. Enter Year = 365, Days = 45, FV = 10,060,829, and PV = 10,000,000. Notice that FV here is the sale price, not the redemption amount.

$$AOR = \left(\frac{365}{45}\right) \times \left(\frac{10{,}060{,}829 - 10{,}000{,}000}{10{,}000{,}000}\right) = 0.04934.$$

The rate of return, stated on a 365-day add-on rate basis, is 4.934%. This result is an annual rate for a periodicity of 8.11 (= 365/45). Implicitly, this assumes that the investment can be replicated 8.11 times over the year.

Investment analysis is made difficult for money market securities because (1) some instruments are quoted on a discount rate basis and others on an add-on rate basis and (2) some are quoted for a 360-day year and others for a 365-day year. Another difference is that the "amount" of a money market instrument quoted on a discount rate basis typically is the face value paid at maturity. However, the "amount" when quoted on an add-on rate basis usually is the principal, the price at issuance. To make money market investment decisions, it is essential to compare instruments on a common basis. An example illustrates this point.

Suppose that an investor is comparing two money market instruments: (A) 90-day commercial paper quoted at a discount rate of 5.76% for a 360-day year and (B) a 90-day bank time deposit quoted at an add-on rate of 5.90% for a 365-day year. Which offers the higher expected rate of return assuming that the credit risks are the same? The price of the commercial paper is 98.560 per 100 of face value, calculated using Equation 9 and entering FV = 100, Days = 90, Year = 360, and DR = 0.0576.

$$PV = 100 \times \left(1 - \frac{90}{360} \times 0.0576\right) = 98.560.$$

Next, use Equation 12 to solve for the AOR for a 365-day year, whereby Year = 365, Days = 90, FV = 100, and PV = 98.560.

$$AOR = \left(\frac{365}{90}\right) \times \left(\frac{100 - 98.560}{98.560}\right) = 0.05925.$$

The 90-day commercial paper discount rate of 5.76% converts to an add-on rate for a 365-day year of 5.925%. This converted rate is called a **bond equivalent yield**, or sometimes just an "investment yield." A bond equivalent yield is a money market rate stated on a 365-day add-on rate basis. If the risks are the same, the commercial paper offers 2.5 bps more in annual return than the bank time deposit.

EXAMPLE 10

Comparing Money Market Instruments Based on Bond Equivalent Yields

Suppose that a money market investor observes quoted rates on the following four 180-day money market instruments:

Money Market Instrument	Quotation Basis	Assumed Number of Days in the Year	Quoted Rate
A	Discount Rate	360	4.33%
B	Discount Rate	365	4.36%
C	Add-On Rate	360	4.35%
D	Add-On Rate	365	4.45%

Calculate the bond equivalent yield for each instrument. Which instrument offers the investor the highest rate of return if the credit risks are the same?

Solution:

A Use Equation 9 to get the price per 100 of par value, where $FV = 100$, Days = 180, Year = 360, and $DR = 0.0433$.

$$PV = 100 \times \left(1 - \frac{180}{360} \times 0.0433\right) = 97.835.$$

Use Equation 12 to get the bond equivalent yield, where Year = 365, Days = 180, $FV = 100$, and $PV = 97.835$.

$$AOR = \left(\frac{365}{180}\right) \times \left(\frac{100 - 97.835}{97.835}\right) = 0.04487.$$

The bond equivalent yield for Bond A is 4.487%.

B Use Equation 9 to get the price per 100 of face value, where $FV = 100$, Days = 180, Year = 365, and $DR = 0.0436$.

$$PV = 100 \times \left(1 - \frac{180}{365} \times 0.0436\right) = 97.850.$$

Use Equation 12 to get the bond equivalent yield, where Year = 365, Days = 180, $FV = 100$, and $PV = 97.850$.

$$AOR = \left(\frac{365}{180}\right) \times \left(\frac{100 - 97.850}{97.850}\right) = 0.04456.$$

The bond equivalent yield for Bond B is 4.456%.

C First, determine the redemption amount per 100 of principal ($PV = 100$), where Days = 180, Year = 360, and $AOR = 0.0435$.

$$FV = 100 + \left(100 \times \frac{180}{360} \times 0.0435\right) = 102.175.$$

Use Equation 12 to get the bond equivalent yield, where Year = 365, Days = 180, FV = 102.175, and PV = 100.

$$AOR = \left(\frac{365}{180}\right) \times \left(\frac{102.175 - 100}{100}\right) = 0.04410.$$

The bond equivalent yield for Bond C is 4.410%.

Another way to get the bond equivalent yield for Bond C is to observe that the AOR of 4.35% for a 360-day year can be obtained using Equation 12 for Year = 360, Days = 180, FV = 102.175, and PV = 100.

$$AOR = \left(\frac{360}{180}\right) \times \left(\frac{102.175 - 100}{100}\right) = 0.0435.$$

Therefore, an add-on rate for a 360-day year only needs to be multiplied by the factor of 365/360 to get the 365-day year bond equivalent yield.

$$\frac{365}{360} \times 0.0435 = 0.04410.$$

D The quoted rate for Bond D of 4.45% is a bond equivalent yield, which is defined as an add-on rate for a 365-day year.

If the risks of these money market instruments are the same, Bond A offers the highest rate of return on a bond equivalent yield basis, 4.487%.

The third difference between yield measures in the money market and the bond market is the periodicity of the annual rate. Because bond yields-to-maturity are computed using interest rate compounding, there is a well-defined periodicity. For instance, bond yields-to-maturity for semiannual compounding are annualized for a periodicity of two. Money market rates are computed using simple interest without compounding. In the money market, the periodicity is the number of days in the year divided by the number of days to maturity. Therefore, money market rates for different times-to-maturity have different periodicities.

Suppose that an analyst prefers to convert money market rates to a semiannual bond basis so that the rates are directly comparable to yields on bonds that make semiannual coupon payments. The quoted rate for a 90-day money market instrument is 10%, quoted as a bond equivalent yield, which means its periodicity is 365/90. Using Equation 7, the conversion is from m = 365/90 to n = 2 for $APR_{365/90}$ = 0.10.

$$\left(1 + \frac{0.10}{365/90}\right)^{365/90} = \left(1 + \frac{APR_2}{2}\right)^2; APR_2 = 0.10127.$$

Therefore, 10% for a periodicity of 365/90 corresponds to 10.127% for a periodicity of two. The difference is significant—12.7 bps. In general, the difference depends on the level of the annual percentage rate. When interest rates are lower, the difference between the annual rates for any two periodicities is reduced.

4 THE MATURITY STRUCTURE OF INTEREST RATES

i. define and compare the spot curve, yield curve on coupon bonds, par curve, and forward curve

j. define forward rates and calculate spot rates from forward rates, forward rates from spot rates, and the price of a bond using forward rates

There are many reasons why the yields-to-maturity on any two bonds are different. Suppose that the yield-to-maturity is higher on Bond X than on Bond Y. The following are some possible reasons for the difference between the yields:

- Currency—Bond X could be denominated in a currency with a higher expected rate of inflation than the currency in which Bond Y is denominated.
- Credit risk—Bond X could have a non-investment-grade rating of BB, and Bond Y could have an investment-grade rating of AA.
- Liquidity—Bond X could be illiquid, and Bond Y could be actively traded.
- Tax status—Interest income on Bond X could be taxable, whereas interest income on Bond Y could be exempt from taxation.
- Periodicity—Bond X could make a single annual coupon payment, and its yield-to-maturity could be quoted for a periodicity of one. Bond Y could make monthly coupon payments, and its yield-to-maturity could be annualized for a periodicity of 12.

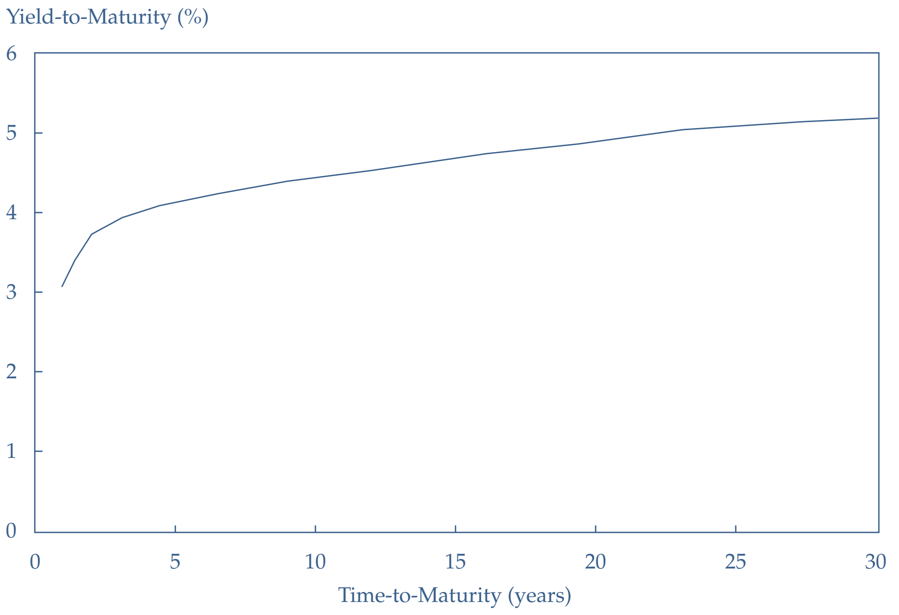
Obviously, another reason is that Bond X and Bond Y could have different times-to-maturity. This factor explaining the differences in yields is called the **maturity structure**, or **term structure**, of interest rates. It involves the analysis of yield curves, which are relationships between yields-to-maturity and times-to-maturity. There are different types of yield curves, depending on the characteristics of the underlying bonds.

In theory, maturity structure should be analyzed for bonds that have the same properties other than time-to-maturity. The bonds should be denominated in the same currency and have the same credit risk, liquidity, and tax status. Their annual rates should be quoted for the same periodicity. Also, they should have the same coupon rate so that they each have the same degree of coupon reinvestment risk. In practice, maturity structure is analyzed for bonds for which these strong assumptions rarely hold.

The ideal dataset would be yields-to-maturity on a series of *zero-coupon* government bonds for a full range of maturities. This dataset is the government bond **spot curve**, sometimes called the zero or “strip” curve (because the coupon payments are “stripped” off the bonds). The spot, zero, or strip curve is a sequence of yields-to-maturity on zero-coupon bonds. Often, these government spot rates are interpreted as the “risk-free” yields; in this context, “risk-free” refers only to default risk. There still could be a significant amount of inflation risk to the investor, as well as liquidity risk.

A government bond spot curve is illustrated in Exhibit 5 for maturities ranging from 1 to 30 years. The annual yields are stated on a semiannual bond basis, which facilitates comparison to coupon-bearing bonds that make semiannual payments.

Exhibit 5 A Government Bond Spot Curve



This spot curve is upward sloping and flattens for longer times-to-maturity. Longer-term government bonds usually have higher yields than shorter-term bonds. This pattern is typical under normal market conditions. Sometimes, a spot curve is downward sloping in that shorter-term yields are higher than longer-term yields. Such downward-sloping spot curves are called inverted yield curves. The theories that attempt to explain the shape of the yield curve and its implications for future financial market conditions are covered later.

This hypothetical spot curve is ideal for analyzing maturity structure because it best meets the "other things being equal" assumption. These government bonds presumably have the same currency, credit risk, liquidity, and tax status. Most importantly, they have no coupon reinvestment risk because there are no coupons to reinvest. However, most actively traded government and corporate bonds make coupon payments. Therefore, analysis of maturity structure usually is based on price data on government bonds that make coupon payments. These coupon bonds might not have the same liquidity and tax status. Older ("seasoned") bonds tend to be less liquid than newly issued debt because they are owned by "buy-and-hold" institutional and retail investors. Governments issue new debt for regular times-to-maturity—for instance, 5-year and 10-year bonds. The current 6-year bond could be a 10-year bond that was issued four years ago. Also, as interest rates fluctuate, older bonds are priced at a discount or premium to par value, which can lead to tax differences. In some countries, capital gains have different tax treatment than capital losses and interest income.

Analysts usually use only the most recently issued and actively traded government bonds to build a yield curve. These bonds have similar liquidity, and because they are priced closer to par value, they have fewer tax effects. A problem is that there are limited data for the full range of maturities. Therefore, it is necessary to *interpolate* between observed yields. Exhibit 6 illustrates a yield curve for a government that issues 2-year, 3-year, 5-year, 7-year, 10-year, and 30-year bonds that make semiannual coupon payments. Straight-line interpolation is used between those points on the yield curve for coupon bonds.

Exhibit 6 A Government Bond Yield Curve

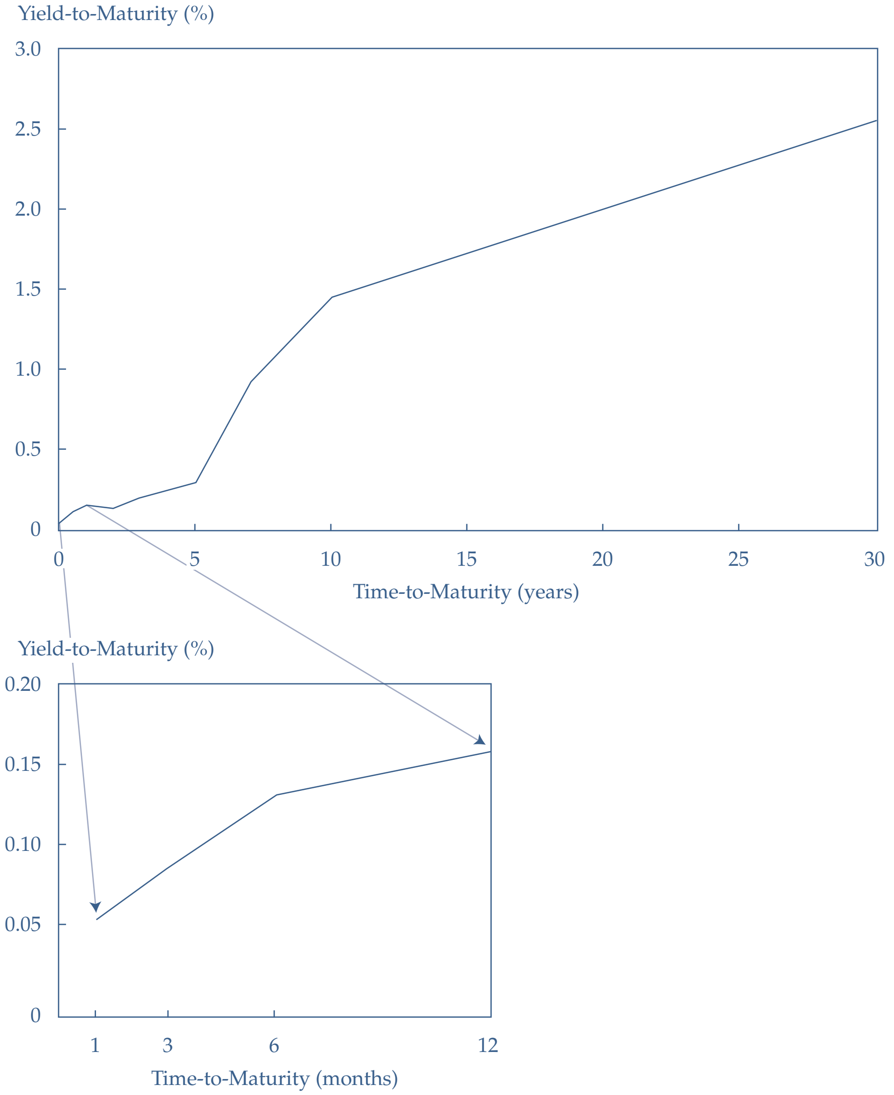


Exhibit 6 also includes yields for short-term government securities having 1 month, 3 months, 6 months, and 12 months to maturity. Although these money market instruments might have been issued and traded on a discount rate basis, they typically are reported as bond equivalent yields. It is important for the analyst to know whether they have been converted to the same periodicity as the longer-term government bonds. If not, the observed yield curve can be misleading because the number of periods in the year is not the same.

In addition to the yield curve on coupon bonds and the spot curve on zero-coupon bonds, maturity structure can be assessed using a **par curve**. A par curve is a sequence of yields-to-maturity such that each bond is priced at par value. The bonds, of course, are assumed to have the same currency, credit risk, liquidity, tax status, and annual yields stated for the same periodicity. Between coupon payment dates, the flat price (not the full price) is assumed to be equal to par value.

The par curve is obtained from a spot curve. On a coupon payment date, the following equation can be used to calculate a par rate given the sequence of spot rates.

$$100 = \frac{PMT}{(1+z_1)^1} + \frac{PMT}{(1+z_2)^2} + \cdots + \frac{PMT + 100}{(1+z_N)^N}. \tag{13}$$

This equation is very similar to Equation 2, whereby $PV = FV = 100$. The problem is to solve for PMT algebraically. Then, $PMT/100$ is equal to the par rate *per period*.

An example illustrates the calculation of the par curve given a spot curve. Suppose the spot rates on government bonds are 5.263% for one year, 5.616% for two years, 6.359% for three years, and 7.008% for four years. These are effective annual rates. The one-year par rate is 5.263%.

$$100 = \frac{PMT + 100}{(1.05263)^1};\ PMT = 5.263.$$

The two-year par rate is 5.606%.

$$100 = \frac{PMT}{(1.05263)^1} + \frac{PMT + 100}{(1.05616)^2};\ PMT = 5.606.$$

The three-year and four-year par rates are 6.306% and 6.899%, respectively.

$$100 = \frac{PMT}{(1.05263)^1} + \frac{PMT}{(1.05616)^2} + \frac{PMT + 100}{(1.06359)^3};\ PMT = 6.306.$$

$$100 = \frac{PMT}{(1.05263)^1} + \frac{PMT}{(1.05616)^2} + \frac{PMT}{(1.06359)^3} + \frac{PMT + 100}{(1.07008)^4};\ PMT = 6.899.$$

The fixed-income securities covered so far have been **cash market securities**. Money market securities often are settled on a "same day," or "cash settlement," basis. Other securities have a difference between the trade date and the settlement date. For instance, if a government bond trades on a $T + 1$ basis, there is a one-day difference between the trade date and the settlement date. If a corporate bond trades on a $T + 3$ basis, the seller delivers the bond and the buyer makes payment in three business days. Cash markets are also called spot markets, which can be confusing because "spot rate" can have two meanings. It can mean the rate on a bond traded in the spot, or cash, market. It can also mean yield on a zero-coupon bond, which is the meaning used in our discussion.

A **forward market** is for future delivery, beyond the usual settlement time period in the cash market. Agreement to the terms for the transaction is on the trade date, but delivery of the security and payment for it are deferred to a future date. A **forward rate** is the interest rate on a bond or money market instrument traded in a forward market. For example, suppose that in the cash market, a five-year zero-coupon bond is priced at 81 per 100 of par value. Its yield-to-maturity is 4.2592%, stated on a semiannual bond basis.

$$81 = \frac{100}{(1+r)^{10}};\ r = 0.021296;\ \times 2 = 0.042592.$$

Suppose that a dealer agrees to deliver a five-year bond two years into the future for a price of 75 per 100 of par value. The credit risk, liquidity, and tax status of this bond traded in the forward market are the same as those for this bond in the cash market. The forward rate is 5.8372%.

$$75 = \frac{100}{(1+r)^{10}};\ r = 0.029186;\ \times 2 = 0.058372.$$

The notation for forward rates is important to understand. Although finance textbook authors use varying notation, the most common market practice is to name this forward rate "the 2y5y." This term is pronounced "the two-year into five-year rate," or simply "the 2's, 5's." The idea is that the first number (two years) refers to the length of the forward period in years from today and the second number (five years) refers to the **tenor** of the underlying bond. The tenor is the remaining time-to-maturity for a bond (or a derivative contract). Therefore, 5.8372% is the "2y5y" forward rate for the zero-coupon bond—the five-year yield two years into the future. Note that the bond that will be a five-year zero in two years currently has seven years to maturity. In the money market, the forward rate usually refers to months. For instance, an analyst might inquire about the "1m6m" forward rate on Euribor, which is the rate on six-month Euribor one month into the future.

Implied forward rates (also known as forward yields) are calculated from spot rates. An implied forward rate is a breakeven reinvestment rate. It links the return on an investment in a shorter-term zero-coupon bond to the return on an investment in a longer-term zero-coupon bond. Suppose that the shorter-term bond matures in A periods and the longer-term bond matures in B periods. The yields-to-maturity per period on these bonds are denoted z_A and z_B. The first is an A-period zero-coupon bond trading in the cash market. The second is a B-period zero-coupon cash market bond. The implied forward rate between period A and period B is denoted $IFR_{A,B-A}$. It is a forward rate on a security that starts in period A and ends in period B. Its tenor is $B - A$ periods.

Equation 14 is a general formula for the relationship between the two spot rates and the implied forward rate.

$$(1 + z_A)^A \times (1 + \text{IFR}_{A,B-A})^{B-A} = (1 + z_B)^B. \quad \text{(14)}$$

Suppose that the yields-to-maturity on three-year and four-year zero-coupon bonds are 3.65% and 4.18%, respectively, stated on a semiannual bond basis. An analyst would like to know the "3y1y" implied forward rate, which is the implied one-year forward yield three years into the future. Therefore, $A = 6$ (periods), $B = 8$ (periods), $B - A = 2$ (periods), $z_6 = 0.0365/2$ (per period), and $z_8 = 0.0418/2$ (per period).

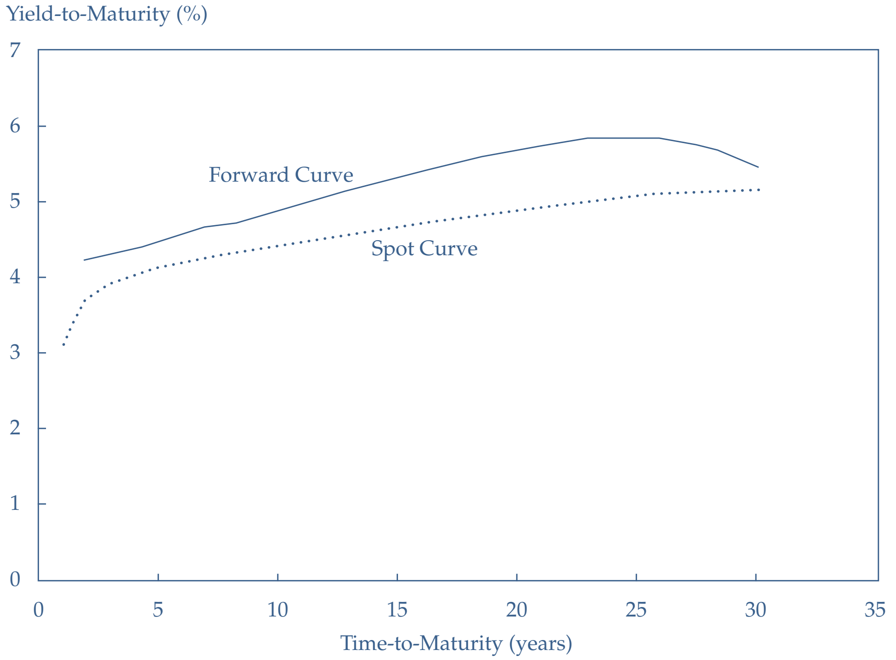
$$\left(1 + \frac{0.0365}{2}\right)^6 \times \left(1 + IFR_{6,2}\right)^2 = \left(1 + \frac{0.0418}{2}\right)^8; IFR_{6,2} = 0.02889;$$

$$\times 2 = 0.05778.$$

The "3y1y" implied forward yield is 5.778%, annualized for a periodicity of two.

Equation 14 can be used to construct a **forward curve**. A forward curve is a series of forward rates, each having the same time frame. These forward rates might be observed on transactions in the derivatives market. Often, the forward rates are implied from transactions in the cash market. Exhibit 7 displays the forward curve that is calculated from the government bond spot curve shown in Exhibit 5. These are one-year forward rates stated on a semiannual bond basis.

Exhibit 7 A Government Bond Spot Curve and Forward Curve



A forward rate can be interpreted as an incremental, or marginal, return for extending the time-to-maturity for an additional time period. Suppose an investor has a four-year investment horizon and is choosing between buying a three-year zero-coupon bond that is priced to yield 3.65% and a four-year zero that is priced to yield 4.18%. The incremental, or marginal, return for the fourth year is 5.778%, the "3y1y" implied forward rate. If the investor's view on future bond yields is that the one-year yield in three years is likely to be less than 5.778%, the investor might prefer to buy the four-year bond. However, if the investor's view is that the one-year yield will be more than the implied forward rate, the investor might prefer the three-year bond and the opportunity to reinvest at the expected higher rate. That explains why an implied forward rate is the *breakeven reinvestment rate*. Implied forward rates are very useful to investors as well as bond issuers in making maturity decisions.

EXAMPLE 11

Computing Forward Rates

Suppose that an investor observes the following prices and yields-to-maturity on zero-coupon government bonds:

Maturity	Price	Yield-to-Maturity
1 year	97.50	2.548%
2 years	94.25	2.983%
3 years	91.75	2.891%

The prices are per 100 of par value. The yields-to-maturity are stated on a semiannual bond basis.

1 Compute the "1y1y" and "2y1y" implied forward rates, stated on a semiannual bond basis.

2 The investor has a three-year investment horizon and is choosing between (1) buying the two-year zero and reinvesting in another one-year zero in two years and (2) buying and holding to maturity the three-year zero. The investor decides to buy the two-year bond. Based on this decision, which of the following is the minimum yield-to-maturity the investor expects on one-year zeros two years from now?

A 2.548%

B 2.707%

C 2.983%

Solution to 1:

The "1y1y" implied forward rate is 3.419%. In Equation 14, $A = 2$ (periods), $B = 4$ (periods), $B - A = 2$ (periods), $z_2 = 0.02548/2$ (per period), and $z_4 = 0.02983/2$ (per period).

$$\left(1+\frac{0.02548}{2}\right)^2 \times \left(1+IFR_{2,2}\right)^2 = \left(1+\frac{0.02983}{2}\right)^4;\ IFR_{2,2} = 0.017095;$$
$$\times\ 2 = 0.03419.$$

The "2y1y" implied forward rate is 2.707%. In Equation 14, $A = 4$ (periods), $B = 6$ (periods), $B - A = 2$ (periods), $z_4 = 0.02983/2$ (per period), and $z_6 = 0.02891/2$ (per period).

$$\left(1+\frac{0.02983}{2}\right)^4 \times \left(1+IFR_{4,2}\right)^2 = \left(1+\frac{0.02891}{2}\right)^6;\ IFR_{4,2} = 0.013536;$$
$$\times\ 2 = 0.02707.$$

Solution to 2:

B is correct. The investor's view is that the one-year yield in two years will be greater than or equal to 2.707%.

The "2y1y" implied forward rate of 2.707% is the breakeven reinvestment rate. If the investor expects the one-year rate in two years to be less than that, the investor would prefer to buy the three-year zero. If the investor expects the one-year rate in two years to be greater than 2.707%, the investor might prefer to buy the two-year zero and reinvest the cash flow.

The forward curve has many applications in fixed-income analysis. Forward rates are used to make maturity decisions. They are used to identify arbitrage opportunities between transactions in the cash market for bonds and in derivatives markets. Forward rates are important in the valuation of derivatives, especially interest rate swaps and options. Those applications for the forward curve are covered elsewhere.

Forward rates can be used to value a fixed-income security in the same manner as spot rates because they are interconnected. The spot curve can be calculated from the forward curve, and the forward curve can be calculated from the spot curve. Either curve can be used to value a fixed-rate bond. An example will illustrate this process.

Suppose the current forward curve for one-year rates is the following:

Time Period	Forward Rate
0y1y	1.88%
1y1y	2.77%
2y1y	3.54%
3y1y	4.12%

These are annual rates stated for a periodicity of one. They are effective annual rates. The first rate, the "0y1y," is the one-year spot rate. The others are one-year forward rates. Given these rates, the spot curve can be calculated as the *geometric average* of the forward rates.

The two-year implied spot rate is 2.3240%.

$(1.0188 \times 1.0277) = (1 + z_2)^2$; $z_2 = 0.023240$.

The following are the equations for the three-year and four-year implied spot rates.

$(1.0188 \times 1.0277 \times 1.0354) = (1 + z_3)^3$; $z_3 = 0.027278$.

$(1.0188 \times 1.0277 \times 1.0354 \times 1.0412) = (1 + z_4)^4$; $z_4 = 0.030741$.

The three-year implied spot rate is 2.7278%, and the four-year spot rate is 3.0741%.

Suppose that an analyst needs to value a four-year, 3.75% annual coupon payment bond that has the same risks as the bonds used to obtain the forward curve. Using the implied spot rates, the value of the bond is 102.637 per 100 of par value.

$$\frac{3.75}{(1.0188)^1} + \frac{3.75}{(1.023240)^2} + \frac{3.75}{(1.027278)^3} + \frac{103.75}{(1.030741)^4} = 102.637.$$

The bond also can be valued using the forward curve.

$$\frac{3.75}{(1.0188)} + \frac{3.75}{(1.0188 \times 1.0277)} + \frac{3.75}{(1.0188 \times 1.0277 \times 1.0354)} + \frac{103.75}{(1.0188 \times 1.0277 \times 1.0354 \times 1.0412)} = 102.637.$$

5 YIELD SPREADS

k compare, calculate, and interpret yield spread measures

A yield spread, in general, is the difference in yield between different fixed-income securities. This section describes a number of yield spread measures.

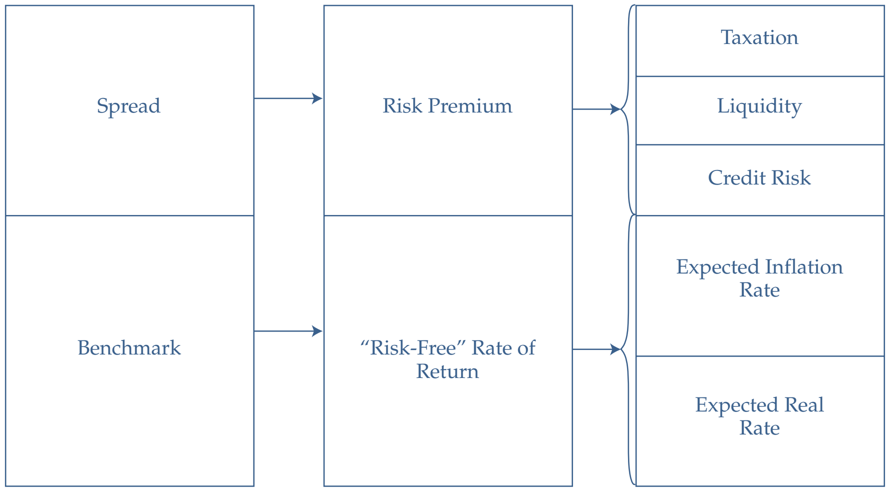
5.1 Yield Spreads over Benchmark Rates

In fixed-income security analysis, it is important to understand *why* bond prices and yields-to-maturity change. To do this, it is useful to separate a yield-to-maturity into two components: the **benchmark** and the **spread**. The benchmark yield for a fixed-income security with a given time-to-maturity is the base rate, often a government bond yield. The spread is the difference between the yield-to-maturity and the benchmark.

The reason for this separation is to distinguish between macroeconomic and microeconomic factors that affect the bond price and, therefore, its yield-to-maturity. The benchmark captures the macroeconomic factors: the expected rate of inflation in the currency in which the bond is denominated, general economic growth and the business cycle, foreign exchange rates, and the impact of monetary and fiscal policy. Changes in those factors impact all bonds in the market, and the effect is seen mostly in changes in the benchmark yield. The spread captures the microeconomic factors specific to the bond issuer and the bond itself: credit risk of the issuer and changes in the quality rating on the bond, liquidity and trading in comparable securities, and the tax status of the bond. It should be noted, however, that general yield spreads across issuers can widen and narrow with changes in macroeconomic factors.

Exhibit 8 illustrates the building blocks of the yield-to-maturity, starting with the benchmark and the spread. The benchmark is often called the risk-free rate of return. Also, the benchmark can be broken down into the expected real rate and the expected inflation rate in the economy. The yield spread is called the risk premium over the "risk-free" rate of return. The risk premium provides the investor with compensation for the credit and liquidity risks and possibly the tax impact of holding a specific bond.

Exhibit 8 Yield-to-Maturity Building Blocks



The benchmark varies across financial markets. Fixed-rate bonds often use a government benchmark security with the same time-to-maturity as, or the closest time-to-maturity to, the specified bond. This benchmark is usually the most recently issued government bond and is called the **on-the-run** security. The on-the-run government bond is the most actively traded security and has a coupon rate closest to the current market discount rate for that maturity. That implies that it is priced close to par value. Seasoned government bonds are called **off-the-run**. On-the-run bonds typically trade at slightly lower yields-to-maturity than off-the-run bonds having the same or similar times-to-maturity because of differences in demand for the securities and, sometimes, differences in the cost of financing the government security in the repo market.

A frequently used benchmark for floating-rate notes has long been Libor, which is due to be phased out, with the Market Reference Rate to take its place. As a composite interbank rate, it is not a risk-free rate. The yield spread over a specific benchmark is referred to as the **benchmark spread** and is usually measured in basis points. If no benchmark exists for a specific bond's tenor or a bond has an unusual maturity, interpolation is used to derive an implied benchmark. Also, bonds with very long tenors are priced over the longest available benchmark bond. For example, 100-year bonds (often called "century bonds") in the United States are priced over the 30-year US Treasury benchmark rate.

In the United Kingdom, the United States, and Japan, the benchmark rate for fixed-rate bonds is a government bond yield. The yield spread in basis points over an actual or interpolated government bond is known as the **G-spread**. The spread over a government bond is the return for bearing greater credit, liquidity, and other risks relative to the sovereign bond. Euro-denominated corporate bonds are priced over a EUR interest rate swap benchmark. For example, a newly issued five-year EUR bond might be priced at a rate of "mid-swaps" plus 150 bps, where "mid-swaps" is the average of the bid and offered swap rates. The yield spread is over a five-year EUR swap rate rather than a government benchmark. Note that the government bond yield or swap rate used as the benchmark for a specific corporate bond will change over time as the remaining time-to-maturity changes.

The yield spread of a specific bond over the standard swap rate in that currency of the same tenor is known as the **I-spread** or **interpolated spread** to the swap curve. This yield spread over MRR allows comparison of bonds with differing credit and liquidity risks against an interbank lending benchmark. Issuers will use the spread above MRR to determine the relative cost of fixed-rate bonds versus floating-rate alternatives, such as an FRN or commercial paper. Investors use the spread over MRR as a measure of a bond's credit risk. Whereas a standard interest rate swap involves an exchange of fixed for floating cash flows based on a floating index, an **asset swap** converts the periodic fixed coupon of a specific bond to an MRR plus or minus a spread. If the bond is priced close to par, this conversion approximates the price of a bond's credit risk over the MRR index. Exhibit 9 illustrates these yield spreads using the Bloomberg Fixed Income Relative Value (FIRV) page.

This example is for the 3.75% Apple bond that matures on 13 November 2047. The spreads are in the top-left corner of the page. The bond's flat asked price was 96.461 per 100 of par value on 12 July 2018, and its yield-to-maturity was 3.955%. On that date, the yield spread over a particular Treasury benchmark was 104 bps. Its G-spread over an interpolated government bond yield was also 104 bps. These two spreads sometimes differ by a few basis points, especially if the benchmark maturity differs from that of the underlying bond. The bond's I-spread was 109 bps. That Libor spread (to be replaced by MRR in the future) was a little larger than the G-spread because 30-year Treasury yields were slightly higher than 30-year Libor swap rates at that time. The use of these spreads in investor strategies will be covered in more detail later. In general, an analyst will track these spreads relative to their averages and historical highs and lows in an attempt to identify relative value.

5.2 Yield Spreads over the Benchmark Yield Curve

A yield curve shows the relationship between yields-to-maturity and times-to-maturity for securities with the same risk profile. For example, the government bond yield curve is the relationship between the yields of on-the-run government bonds and their times-to-maturity. The swap yield curve shows the relationship between fixed MRR swap rates and their times-to-maturity.

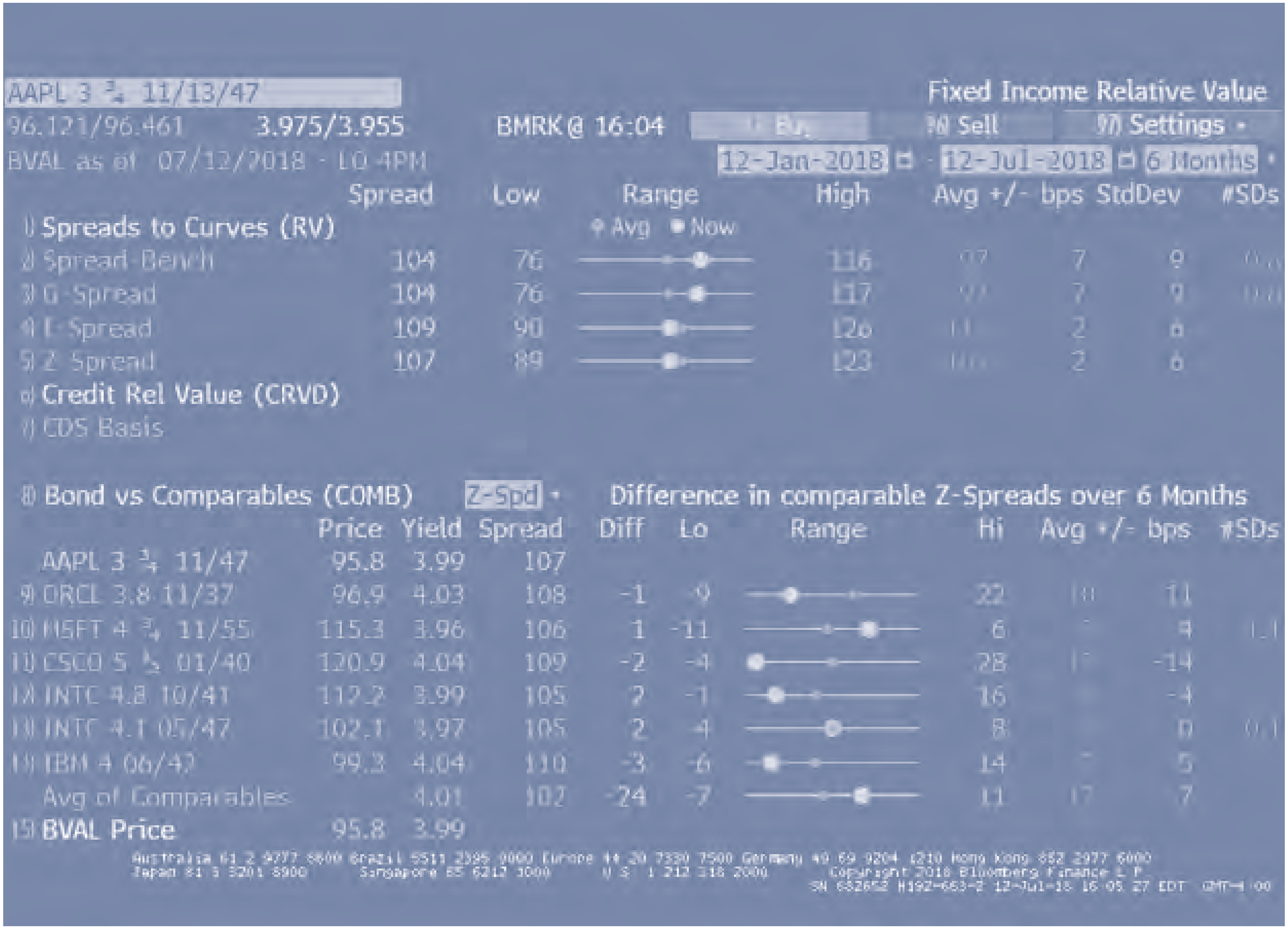
Each of these yield curves represents the term structure of benchmark interest rates, whether for "risk-free" government yields or "risky" fixed swap rates. Benchmark yield curves tend to be upward-sloping because investors typically demand a premium for holding longer-term securities. In general, investors face greater price risk for a given change in yield for longer-term bonds. This concept is covered further in the discussion of the topic of "Fixed-Income Risk and Return." The term structure of interest rates is dynamic, with short-term rates driven by central bank policy and longer-term rates affected by long-term growth and inflation expectations.

Isolating credit risk over varying times-to-maturity gives rise to a term structure of credit spreads that is distinct for each borrower. The G-spread and I-spread each use the same discount rate for each cash flow. Another approach is to calculate a constant yield spread over a government (or interest rate swap) spot curve instead. This spread is known as the **zero-volatility spread (Z-spread)** of a bond over the benchmark rate. In Exhibit 9, the Z-spread for the Apple bond was reported to be 107 bps.

The Z-spread over the benchmark spot curve can be calculated with Equation 15:

$$PV = \frac{PMT}{\left(1 + z_1 + Z\right)^1} + \frac{PMT}{\left(1 + z_2 + Z\right)^2} + \cdots + \frac{PMT + FV}{\left(1 + z_N + Z\right)^N}. \tag{15}$$

Exhibit 9 Bloomberg FIRV Page for the 3.75% Apple Bond



AAPL 3 ¾ 11/13/47 Fixed Income Relative Value

96.121/96.461 3.975/3.955 BMRK @ 16:04

BVAL as of 07/12/2018 - LO 4PM 12-Jan-2018 - 12-Jul-2018 6 Months

	Spread	Low	Range	High	Avg	+/- bps	StdDev	#SDs
1) Spreads to Curves (RV)			◆ Avg ● Now					
2) Spread-Bench	104	76		116	97	7	9	[illegible]
3) G-Spread	104	76		117	97	7	9	[illegible]
4) I-Spread	109	90		126	111	2	6	
5) Z-Spread	107	89		123	[illegible]	2	6	

6) Credit Rel Value (CRVD)

7) CDS Basis

8) Bond vs Comparables (COMB) Z-Spd Difference in comparable Z-Spreads over 6 Months

	Price	Yield	Spread	Diff	Lo	Range	Hi	Avg	+/- bps	#SDs
AAPL 3 ¾ 11/47	95.8	3.99	107							
9) ORCL 3.8 11/37	96.9	4.03	108	-1	-9		22	[illegible]	11	
10) MSFT 4 ¾ 11/55	115.3	3.96	106	1	-11		6	[illegible]	9	[illegible]
11) CSCO 5 ½ 01/40	120.9	4.04	109	-2	-4		28	[illegible]	-14	
12) INTC 4.8 10/41	112.2	3.99	105	2	-1		16	[illegible]	-4	
13) INTC 4.1 05/47	102.1	3.97	105	2	-4		8	[illegible]	0	[illegible]
14) IBM 4 06/42	99.3	4.04	110	-3	-6		14	[illegible]	5	
Avg of Comparables		4.01	102	-24	-7		11	17	7	
15) BVAL Price	95.8	3.99								

Australia 61 2 9777 8600 Brazil 5511 2395 9000 Europe 44 20 7330 7500 Germany 49 69 9204 1210 Hong Kong 852 2977 6000
Japan 81 3 3201 8900 Singapore 65 6212 1000 U.S. 1 212 318 2000 Copyright 2018 Bloomberg Finance L.P.
SN 632652 H192-663-2 12-Jul-18 16:05:27 EDT GMT-4:00

© 2019 Bloomberg L.P. All rights reserved. Reproduced with permission.

The benchmark spot rates—z_1, z_2, . . ., z_N—are derived from the government yield curve (or from fixed rates on interest rate swaps). Z is the Z-spread per period and is the same for all time periods. In Equation 15, N is an integer, so the calculation is on a coupon date when the accrued interest is zero. Sometimes, the Z-spread is called the "static spread" because it is constant (and has zero volatility). In practice, the Z-spread is usually calculated in a spreadsheet using a goal seek function or similar solver function.

The Z-spread is also used to calculate the **option-adjusted spread** (OAS) on a callable bond. The OAS, like the option-adjusted yield, is based on an option-pricing model and an assumption about future interest rate volatility. Then, the value of the embedded call option, which is stated in basis points per year, is subtracted from the yield spread. In particular, it is subtracted from the Z-spread:

$$\text{OAS} = \text{Z-spread} - \text{Option value (in basis points per year)}.$$

This important topic is covered later.

EXAMPLE 12

The G-Spread and the Z-Spread

A 6% annual coupon corporate bond with two years remaining to maturity is trading at a price of 100.125. The two-year, 4% annual payment government benchmark bond is trading at a price of 100.750. The one-year and two-year government spot rates are 2.10% and 3.635%, respectively, stated as effective annual rates.

1 Calculate the G-spread, the spread between the yields-to-maturity on the corporate bond and the government bond having the same maturity.

2 Demonstrate that the Z-spread is 234.22 bps.

Solution to 1:

The yield-to-maturity for the corporate bond is 5.932%.

$$100.125 = \frac{6}{(1+r)^1} + \frac{106}{(1+r)^2};\ r = 0.05932.$$

The yield-to-maturity for the government benchmark bond is 3.605%.

$$100.750 = \frac{4}{(1+r)^1} + \frac{104}{(1+r)^2};\ r = 0.03605.$$

The G-spread is 232.7 bps: 0.05932 – 0.03605 = 0.02327.

Solution to 2:

Solve for the value of the corporate bond using $z_1 = 0.0210$, $z_2 = 0.03635$, and $Z = 0.023422$:

$$\frac{6}{(1+0.0210+0.023422)^1} + \frac{106}{(1+0.03635+0.023422)^2}$$

$$= \frac{6}{(1.044422)^1} + \frac{106}{(1.059772)^2} = 100.125.$$

SUMMARY

We have covered the principles and techniques that are used in the valuation of fixed-rate bonds, as well as floating-rate notes and money market instruments. These building blocks are used extensively in fixed-income analysis. The following are the main points made:

- The market discount rate is the rate of return required by investors given the risk of the investment in the bond.
- A bond is priced at a premium above par value when the coupon rate is greater than the market discount rate.
- A bond is priced at a discount below par value when the coupon rate is less than the market discount rate.
- The amount of any premium or discount is the present value of the "excess" or "deficiency" in the coupon payments relative to the yield-to-maturity.
- The yield-to-maturity, the internal rate of return on the cash flows, is the implied market discount rate given the price of the bond.
- A bond price moves inversely with its market discount rate.
- The relationship between a bond price and its market discount rate is convex.
- The price of a lower-coupon bond is more volatile than the price of a higher-coupon bond, other things being equal.
- Generally, the price of a longer-term bond is more volatile than the price of a shorter-term bond, other things being equal. An exception to this phenomenon can occur on low-coupon (but not zero-coupon) bonds that are priced at a discount to par value.
- Assuming no default, premium and discount bond prices are "pulled to par" as maturity nears.
- A spot rate is the yield-to-maturity on a zero-coupon bond.
- A yield-to-maturity can be approximated as a weighted average of the underlying spot rates.
- Between coupon dates, the full (or invoice, or "dirty") price of a bond is split between the flat (or quoted, or "clean") price and the accrued interest.
- Flat prices are quoted to not misrepresent the daily increase in the full price as a result of interest accruals.
- Accrued interest is calculated as a proportional share of the next coupon payment using either the actual/actual or 30/360 methods to count days.
- Matrix pricing is used to value illiquid bonds by using prices and yields on comparable securities having the same or similar credit risk, coupon rate, and maturity.
- The periodicity of an annual interest rate is the number of periods in the year.
- A yield quoted on a semiannual bond basis is an annual rate for a periodicity of two. It is the yield per semiannual period times two.
- The general rule for periodicity conversions is that compounding more frequently at a lower annual rate corresponds to compounding less frequently at a higher annual rate.
- Street convention yields assume payments are made on scheduled dates, neglecting weekends and holidays.

- The current yield is the annual coupon payment divided by the flat price, thereby neglecting as a measure of the investor's rate of return the time value of money, any accrued interest, and the gain from buying at a discount or the loss from buying at a premium.
- The simple yield is like the current yield but includes the straight-line amortization of the discount or premium.
- The yield-to-worst on a callable bond is the lowest of the yield-to-first-call, yield-to-second-call, and so on, calculated using the call price for the future value and the call date for the number of periods.
- The option-adjusted yield on a callable bond is the yield-to-maturity after adding the theoretical value of the call option to the price.
- A floating-rate note (floater, or FRN) maintains a more stable price than a fixed-rate note because interest payments adjust for changes in market interest rates.
- The quoted margin on a floater is typically the specified yield spread over or under the reference rate, which we refer to as the Market Reference Rate.
- The discount margin on a floater is the spread required by investors, and to which the quoted margin must be set, for the FRN to trade at par value on a rate reset date.
- Money market instruments, having one year or less time-to-maturity, are quoted on a discount rate or add-on rate basis.
- Money market discount rates understate the investor's rate of return (and the borrower's cost of funds) because the interest income is divided by the face value or the total amount redeemed at maturity, and not by the amount of the investment.
- Money market instruments need to be converted to a common basis for analysis.
- A money market bond equivalent yield is an add-on rate for a 365-day year.
- The periodicity of a money market instrument is the number of days in the year divided by the number of days to maturity. Therefore, money market instruments with different times-to-maturity have annual rates for different periodicities.
- In theory, the maturity structure, or term structure, of interest rates is the relationship between yields-to-maturity and times-to-maturity on bonds having the same currency, credit risk, liquidity, tax status, and periodicity.
- A spot curve is a series of yields-to-maturity on zero-coupon bonds.
- A frequently used yield curve is a series of yields-to-maturity on coupon bonds.
- A par curve is a series of yields-to-maturity assuming the bonds are priced at par value.
- In a cash market, the delivery of the security and cash payment is made on a settlement date within a customary time period after the trade date—for example, "T + 3."
- In a forward market, the delivery of the security and cash payment are made on a predetermined future date.
- A forward rate is the interest rate on a bond or money market instrument traded in a forward market.
- An implied forward rate (or forward yield) is the breakeven reinvestment rate linking the return on an investment in a shorter-term zero-coupon bond to the return on an investment in a longer-term zero-coupon bond.
- An implied forward curve can be calculated from the spot curve.

- Implied spot rates can be calculated as geometric averages of forward rates.
- A fixed-income bond can be valued using a market discount rate, a series of spot rates, or a series of forward rates.
- A bond yield-to-maturity can be separated into a benchmark and a spread.
- Changes in benchmark rates capture macroeconomic factors that affect all bonds in the market—inflation, economic growth, foreign exchange rates, and monetary and fiscal policy.
- Changes in spreads typically capture microeconomic factors that affect the particular bond—credit risk, liquidity, and tax effects.
- Benchmark rates are usually yields-to-maturity on government bonds or fixed rates on interest rate swaps.
- A G-spread is the spread over or under a government bond rate, and an I-spread is the spread over or under an interest rate swap rate.
- A G-spread or an I-spread can be based on a specific benchmark rate or on a rate interpolated from the benchmark yield curve.
- A Z-spread (zero-volatility spread) is based on the entire benchmark spot curve. It is the constant spread that is added to each spot rate such that the present value of the cash flows matches the price of the bond.
- An option-adjusted spread (OAS) on a callable bond is the Z-spread minus the theoretical value of the embedded call option.

PRACTICE PROBLEMS

1 A portfolio manager is considering the purchase of a bond with a 5.5% coupon rate that pays interest annually and matures in three years. If the required rate of return on the bond is 5%, the price of the bond per 100 of par value is *closest* to:

A 98.65.

B 101.36.

C 106.43.

2 A bond with two years remaining until maturity offers a 3% coupon rate with interest paid annually. At a market discount rate of 4%, the price of this bond per 100 of par value is *closest* to:

A 95.34.

B 98.00.

C 98.11.

3 An investor who owns a bond with a 9% coupon rate that pays interest semiannually and matures in three years is considering its sale. If the required rate of return on the bond is 11%, the price of the bond per 100 of par value is *closest* to:

A 95.00.

B 95.11.

C 105.15.

4 A bond offers an annual coupon rate of 4%, with interest paid semiannually. The bond matures in two years. At a market discount rate of 6%, the price of this bond per 100 of par value is *closest* to:

A 93.07.

B 96.28.

C 96.33.

5 A bond offers an annual coupon rate of 5%, with interest paid semiannually. The bond matures in seven years. At a market discount rate of 3%, the price of this bond per 100 of par value is *closest* to:

A 106.60.

B 112.54.

C 143.90.

6 A zero-coupon bond matures in 15 years. At a market discount rate of 4.5% per year and assuming annual compounding, the price of the bond per 100 of par value is *closest* to:

A 51.30.

B 51.67.

C 71.62.

7 Consider the following two bonds that pay interest annually:

© 2019 CFA Institute. All rights reserved.

Bond	Coupon Rate	Time-to-Maturity
A	5%	2 years
B	3%	2 years

At a market discount rate of 4%, the price difference between Bond A and Bond B per 100 of par value is *closest* to:

A 3.70.

B 3.77.

C 4.00.

The following information relates to Questions 8 and 9

Bond	Price	Coupon Rate	Time-to-Maturity
A	101.886	5%	2 years
B	100.000	6%	2 years
C	97.327	5%	3 years

8 Which bond offers the lowest yield-to-maturity?

A Bond A

B Bond B

C Bond C

9 Which bond will *most likely* experience the smallest percent change in price if the market discount rates for all three bonds increase by 100 bps?

A Bond A

B Bond B

C Bond C

---

10 Suppose a bond's price is expected to increase by 5% if its market discount rate decreases by 100 bps. If the bond's market discount rate increases by 100 bps, the bond price is *most likely* to change by:

A 5%.

B less than 5%.

C more than 5%.

The following information relates to Questions 11 and 12

Bond	Coupon Rate	Maturity (years)
A	6%	10
B	6%	5
C	8%	5

All three bonds are currently trading at par value.

11 Relative to Bond C, for a 200 bp decrease in the required rate of return, Bond B will *most likely* exhibit a(n):

- **A** equal percentage price change.
- **B** greater percentage price change.
- **C** smaller percentage price change.

12 Which bond will *most likely* experience the greatest percentage change in price if the market discount rates for all three bonds increase by 100 bps?

- **A** Bond A
- **B** Bond B
- **C** Bond C

---

13 An investor considers the purchase of a two-year bond with a 5% coupon rate, with interest paid annually. Assuming the sequence of spot rates shown below, the price of the bond is *closest* to:

Time-to-Maturity	Spot Rates
1 year	3%
2 years	4%

- **A** 101.93.
- **B** 102.85.
- **C** 105.81.

14 A three-year bond offers a 10% coupon rate with interest paid annually. Assuming the following sequence of spot rates, the price of the bond is *closest* to:

Time-to-Maturity	Spot Rates
1 year	8.0%
2 years	9.0%
3 years	9.5%

- **A** 96.98.
- **B** 101.46.
- **C** 102.95.

The following information relates to Questions 15–17

Bond	Coupon Rate	Time-to-Maturity	Time-to-Maturity	Spot Rates
X	8%	3 years	1 year	8%
Y	7%	3 years	2 years	9%
Z	6%	3 years	3 years	10%

All three bonds pay interest annually.

15 Based on the given sequence of spot rates, the price of Bond X is *closest* to:

A 95.02.

B 95.28.

C 97.63.

16 Based on the given sequence of spot rates, the price of Bond Y is *closest* to:

A 87.50.

B 92.54.

C 92.76.

17 Based on the given sequence of spot rates, the yield-to-maturity of Bond Z is *closest* to:

A 9.00%.

B 9.92%.

C 11.93%.

---

18 Bond dealers *most* often quote the:

A flat price.

B full price.

C full price plus accrued interest.

The following information relates to Questions 19–21

Bond G, described in the exhibit below, is sold for settlement on 16 June 2020.

Annual Coupon	5%
Coupon Payment Frequency	Semiannual
Interest Payment Dates	10 April and 10 October
Maturity Date	10 October 2022
Day-Count Convention	30/360
Annual Yield-to-Maturity	4%

19 The full price that Bond G settles at on 16 June 2020 is *closest* to:

A 102.36.

B 103.10.

C 103.65.

20 The accrued interest per 100 of par value for Bond G on the settlement date of 16 June 2020 is *closest* to:

A 0.46.

B 0.73.

C 0.92.

21 The flat price for Bond G on the settlement date of 16 June 2020 is *closest* to:

A 102.18.

B 103.10.

C 104.02.

---

22 Matrix pricing allows investors to estimate market discount rates and prices for bonds:

A with different coupon rates.

B that are not actively traded.

C with different credit quality.

23 When underwriting new corporate bonds, matrix pricing is used to get an estimate of the:

A required yield spread over the benchmark rate.

B market discount rate of other comparable corporate bonds.

C yield-to-maturity on a government bond having a similar time-to-maturity.

24 A bond with 20 years remaining until maturity is currently trading for 111 per 100 of par value. The bond offers a 5% coupon rate with interest paid semiannually. The bond's annual yield-to-maturity is *closest* to:

A 2.09%.

B 4.18%.

C 4.50%.

25 The annual yield-to-maturity, stated for with a periodicity of 12, for a four-year, zero-coupon bond priced at 75 per 100 of par value is *closest* to:

A 6.25%.

B 7.21%.

C 7.46%.

26 A five-year, 5% semiannual coupon payment corporate bond is priced at 104.967 per 100 of par value. The bond's yield-to-maturity, quoted on a semiannual bond basis, is 3.897%. An analyst has been asked to convert to a monthly periodicity. Under this conversion, the yield-to-maturity is *closest to*:

A 3.87%.

B 4.95%.

C 7.67%.

The following information relates to Questions 27–30

A bond with five years remaining until maturity is currently trading for 101 per 100 of par value. The bond offers a 6% coupon rate with interest paid semiannually. The bond is first callable in three years and is callable after that date on coupon dates according to the following schedule:

End of Year	Call Price
3	102
4	101
5	100

27 The bond's annual yield-to-maturity is *closest* to:

- **A** 2.88%.
- **B** 5.77%.
- **C** 5.94%.

28 The bond's annual yield-to-first-call is *closest* to:

- **A** 3.12%.
- **B** 6.11%.
- **C** 6.25%.

29 The bond's annual yield-to-second-call is *closest* to:

- **A** 2.97%.
- **B** 5.72%.
- **C** 5.94%.

30 The bond's yield-to-worst is *closest* to:

- **A** 2.88%.
- **B** 5.77%.
- **C** 6.25%.

---

31 A two-year floating-rate note pays six-month Libor plus 80 bps. The floater is priced at 97 per 100 of par value. The current six-month MRR is 1.00%. Assume a 30/360 day-count convention and evenly spaced periods. The discount margin for the floater in basis points is *closest* to:

- **A** 180 bps.
- **B** 236 bps.
- **C** 420 bps.

32 An analyst evaluates the following information relating to floating-rate notes (FRNs) issued at par value that have three-month MRR as a reference rate:

Floating-Rate Note	Quoted Margin	Discount Margin
X	0.40%	0.32%
Y	0.45%	0.45%
Z	0.55%	0.72%

Based only on the information provided, the FRN that will be priced at a premium on the next reset date is:

A FRN X.

B FRN Y.

C FRN Z.

33 A 365-day year bank certificate of deposit has an initial principal amount of USD96.5 million and a redemption amount due at maturity of USD100 million. The number of days between settlement and maturity is 350. The bond equivalent yield is *closest* to:

A 3.48%.

B 3.65%.

C 3.78%.

34 The bond equivalent yield of a 180-day banker's acceptance quoted at a discount rate of 4.25% for a 360-day year is *closest* to:

A 4.31%.

B 4.34%.

C 4.40%.

35 Which of the following statements describing a par curve is *incorrect*?

A A par curve is obtained from a spot curve.

B All bonds on a par curve are assumed to have different credit risk.

C A par curve is a sequence of yields-to-maturity such that each bond is priced at par value.

36 A yield curve constructed from a sequence of yields-to-maturity on zero-coupon bonds is the:

A par curve.

B spot curve.

C forward curve.

37 The rate interpreted to be the incremental return for extending the time-to-maturity of an investment for an additional time period is the:

A add-on rate.

B forward rate.

C yield-to-maturity.

The following information relates to Questions 38 and 39

Time Period	Forward Rate
"0y1y"	0.80%
"1y1y"	1.12%
"2y1y"	3.94%
"3y1y"	3.28%
"4y1y"	3.14%

All rates are annual rates stated for a periodicity of one (effective annual rates).

38 The three-year implied spot rate is *closest* to:

A 1.18%.

B 1.94%.

C 2.28%.

39 The value per 100 of par value of a two-year, 3.5% coupon bond with interest payments paid annually is *closest* to:

A 101.58.

B 105.01.

C 105.82.

---

40 The spread component of a specific bond's yield-to-maturity is *least likely* impacted by changes in:

A its tax status.

B its quality rating.

C inflation in its currency of denomination.

41 The yield spread of a specific bond over the standard swap rate in that currency of the same tenor is *best* described as the:

A I-spread.

B Z-spread.

C G-spread.

The following information relates to Question 42

Bond	Coupon Rate	Time-to-Maturity	Price
UK Government Benchmark Bond	2%	3 years	100.25
UK Corporate Bond	5%	3 years	100.65

Both bonds pay interest annually. The current three-year EUR interest rate swap benchmark is 2.12%.

42 The G-spread in basis points on the UK corporate bond is *closest* to:

A 264 bps.

B 285 bps.

C 300 bps.

---

43 A corporate bond offers a 5% coupon rate and has exactly three years remaining to maturity. Interest is paid annually. The following rates are from the benchmark spot curve:

Time-to-Maturity	Spot Rate
1 year	4.86%
2 years	4.95%
3 years	5.65%

The bond is currently trading at a Z-spread of 234 bps. The value of the bond is *closest to*:

A 92.38.

B 98.35.

C 106.56.

44 An option-adjusted spread (OAS) on a callable bond is the Z-spread:

A over the benchmark spot curve.

B minus the standard swap rate in that currency of the same tenor.

C minus the value of the embedded call option expressed in basis points per year.

SOLUTIONS

1 B is correct. The bond price is closest to 101.36. The price is determined in the following manner:

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT + FV}{(1+r)^3},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
r = market discount rate, or required rate of return per period

$$PV = \frac{5.5}{(1+0.05)^1} + \frac{5.5}{(1+0.05)^2} + \frac{5.5+100}{(1+0.05)^3}.$$

$PV = 5.24 + 4.99 + 91.13 = 101.36.$

2 C is correct. The bond price is closest to 98.11. The formula for calculating the price of this bond is

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT + FV}{(1+r)^2},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
r = market discount rate, or required rate of return per period

$$PV = \frac{3}{(1+0.04)^1} + \frac{3+100}{(1+0.04)^2} = 2.88 + 95.23 = 98.11.$$

3 A is correct. The bond price is closest to 95.00. The bond has six semiannual periods. Half of the annual coupon is paid in each period with the required rate of return also being halved. The price is determined in the following manner:

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT}{(1+r)^3} + \frac{PMT}{(1+r)^4} + \frac{PMT}{(1+r)^5} + \frac{PMT + FV}{(1+r)^6},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
r = market discount rate, or required rate of return per period

$$PV = \frac{4.5}{(1+0.055)^1} + \frac{4.5}{(1+0.055)^2} + \frac{4.5}{(1+0.055)^3} + \frac{4.5}{(1+0.055)^4} + \frac{4.5}{(1+0.055)^5} + \frac{4.5+100}{(1+0.055)^6}.$$

$PV = 4.27 + 4.04 + 3.83 + 3.63 + 3.44 + 75.79 = 95.00.$

4 B is correct. The bond price is closest to 96.28. The formula for calculating this bond price is

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT}{(1+r)^3} + \frac{PMT+FV}{(1+r)^4},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
r = market discount rate, or required rate of return per period

$$PV = \frac{2}{(1+0.03)^1} + \frac{2}{(1+0.03)^2} + \frac{2}{(1+0.03)^3} + \frac{2+100}{(1+0.03)^4}.$$

$PV = 1.94 + 1.89 + 1.83 + 90.62 = 96.28.$

5 B is correct. The bond price is closest to 112.54. The formula for calculating this bond price is

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT}{(1+r)^3} + \cdots + \frac{PMT+FV}{(1+r)^{14}},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
r = market discount rate, or required rate of return per period

$$PV = \frac{2.5}{(1+0.015)^1} + \frac{2.5}{(1+0.015)^2} + \frac{2.5}{(1+0.015)^3} + \cdots + \frac{2.5}{(1+0.015)^{13}} + \frac{2.5+100}{(1+0.015)^{14}}.$$

$PV = 2.46 + 2.43 + 2.39 + \ldots + 2.06 + 83.21 = 112.54.$

6 B is correct. The price of the zero-coupon bond is closest to 51.67. The price is determined in the following manner:

$$PV = \frac{100}{(1+r)^N},$$

where

PV = present value, or the price of the bond
r = market discount rate, or required rate of return per period
N = number of evenly spaced periods to maturity

$$PV = \frac{100}{(1 + 0.045)^{15}}.$$

$$PV = 51.67.$$

7 B is correct. The price difference between Bonds A and B is closest to 3.77. One method for calculating the price difference between two bonds with identical terms to maturity is to use the following formula:

$$PV = \frac{PMT}{(1 + r)^{1}} + \frac{PMT}{(1 + r)^{2}},$$

where

PV = price difference
PMT = coupon difference per period
r = market discount rate, or required rate of return per period

In this case, the coupon difference is (5% – 3%), or 2%.

$$PV = \frac{2}{(1 + 0.04)^{1}} + \frac{2}{(1 + 0.04)^{2}} = 1.92 + 1.85 = 3.77.$$

8 A is correct. Bond A offers the lowest yield-to-maturity. When a bond is priced at a premium above par value, the yield-to-maturity (YTM), or market discount rate, is less than the coupon rate. Bond A is priced at a premium, so its YTM is below its 5% coupon rate. Bond B is priced at par value, so its YTM is equal to its 6% coupon rate. Bond C is priced at a discount below par value, so its YTM is above its 5% coupon rate.

9 B is correct. Bond B will most likely experience the smallest percentage change in price if market discount rates increase by 100 bps. A higher-coupon bond has a smaller percentage price change than a lower-coupon bond when their market discount rates change by the same amount (the coupon effect). Also, a shorter-term bond generally has a smaller percentage price change than a longer-term bond when their market discount rates change by the same amount (the maturity effect). Bond B will experience a smaller percentage change in price than Bond A because of the coupon effect. Bond B will also experience a smaller percentage change in price than Bond C because of the coupon effect and the maturity effect.

10 B is correct. The bond price is most likely to change by less than 5%. The relationship between bond prices and market discount rate is not linear. The percentage price change is greater in absolute value when the market discount rate goes down than when it goes up by the same amount (the convexity effect). If a 100 bp decrease in the market discount rate will cause the price of the bond to increase by 5%, then a 100 bp increase in the market discount rate will cause the price of the bond to decline by an amount less than 5%.

11 B is correct. Generally, for two bonds with the same time-to-maturity, a lower-coupon bond will experience a greater percentage price change than a higher-coupon bond when their market discount rates change by the same amount. Bond B and Bond C have the same time-to-maturity (five years); however, Bond B offers a lower coupon rate. Therefore, Bond B will likely experience a greater percentage change in price in comparison to Bond C.

12 A is correct. Bond A will likely experience the greatest percentage change in price due to the coupon effect and the maturity effect. For two bonds with the same time-to-maturity, a lower-coupon bond has a greater percentage price change than a higher-coupon bond when their market discount rates change by the same amount. Generally, for the same coupon rate, a longer-term bond has a greater percentage price change than a shorter-term bond when their market discount rates change by the same amount. Relative to Bond C, Bond A and Bond B offer a lower coupon rate of 6%; however, Bond A has a longer time-to-maturity than Bond B. Therefore, Bond A will likely experience the greater percentage change in price if the market discount rates for all three bonds increase by 100 bps.

13 A is correct. The bond price is closest to 101.93. The price is determined in the following manner:

$$PV = \frac{PMT}{(1+Z_1)^1} + \frac{PMT + FV}{(1+Z_2)^2},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
Z_1 = spot rate, or the zero-coupon yield, for Period 1
Z_2 = spot rate, or the zero-coupon yield, for Period 2

$$PV = \frac{5}{(1+0.03)^1} + \frac{5+100}{(1+0.04)^2}.$$

$PV = 4.85 + 97.08 = 101.93.$

14 B is correct. The bond price is closest to 101.46. The price is determined in the following manner:

$$PV = \frac{PMT}{(1+Z_1)^1} + \frac{PMT}{(1+Z_2)^2} + \frac{PMT + FV}{(1+Z_3)^3},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
Z_1 = spot rate, or the zero-coupon yield or zero rate, for Period 1
Z_2 = spot rate, or the zero-coupon yield or zero rate, for Period 2
Z_3 = spot rate, or the zero-coupon yield or zero rate, for Period 3

$$PV = \frac{10}{(1+0.08)^1} + \frac{10}{(1+0.09)^2} + \frac{10+100}{(1+0.095)^3}.$$

$PV = 9.26 + 8.42 + 83.78 = 101.46.$

15 B is correct. The bond price is closest to 95.28. The formula for calculating this bond price is

$$PV = \frac{PMT}{(1 + Z_1)^1} + \frac{PMT}{(1 + Z_2)^2} + \frac{PMT + FV}{(1 + Z_3)^3},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
Z_1 = spot rate, or the zero-coupon yield or zero rate, for Period 1
Z_2 = spot rate, or the zero-coupon yield or zero rate, for Period 2
Z_3 = spot rate, or the zero-coupon yield or zero rate, for Period 3

$$PV = \frac{8}{(1 + 0.08)^1} + \frac{8}{(1 + 0.09)^2} + \frac{8 + 100}{(1 + 0.10)^3}.$$

$PV = 7.41 + 6.73 + 81.14 = 95.28.$

16 C is correct. The bond price is closest to 92.76. The formula for calculating this bond price is

$$PV = \frac{PMT}{(1 + Z_1)^1} + \frac{PMT}{(1 + Z_2)^2} + \frac{PMT + FV}{(1 + Z_3)^3},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
Z_1 = spot rate, or the zero-coupon yield or zero rate, for Period 1
Z_2 = spot rate, or the zero-coupon yield or zero rate, for Period 2
Z_3 = spot rate, or the zero-coupon yield or zero rate, for Period 3

$$PV = \frac{7}{(1 + 0.08)^1} + \frac{7}{(1 + 0.09)^2} + \frac{7 + 100}{(1 + 0.10)^3}.$$

$PV = 6.48 + 5.89 + 80.39 = 92.76.$

17 B is correct. The yield-to-maturity is closest to 9.92%. The formula for calculating the price of Bond Z is

$$PV = \frac{PMT}{(1 + Z_1)^1} + \frac{PMT}{(1 + Z_2)^2} + \frac{PMT + FV}{(1 + Z_3)^3},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
Z_1 = spot rate, or the zero-coupon yield or zero rate, for Period 1
Z_2 = spot rate, or the zero-coupon yield or zero rate, for Period 2
Z_3 = spot rate, or the zero-coupon yield or zero rate, for Period 3

$$PV = \frac{6}{(1+0.08)^1} + \frac{6}{(1+0.09)^2} + \frac{6+100}{(1+0.10)^3}.$$

$$PV = 5.56 + 5.05 + 79.64 = 90.25.$$

Using this price, the bond's yield-to-maturity can be calculated as

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT+FV}{(1+r)^3}.$$

$$90.25 = \frac{6}{(1+r)^1} + \frac{6}{(1+r)^2} + \frac{6+100}{(1+r)^3}.$$

$$r = 9.92\%.$$

18 A is correct. Bond dealers usually quote the flat price. When a trade takes place, the accrued interest is added to the flat price to obtain the full price paid by the buyer and received by the seller on the settlement date. The reason for using the flat price for quotation is to avoid misleading investors about the market price trend for the bond. If the full price were to be quoted by dealers, investors would see the price rise day after day even if the yield-to-maturity did not change. That is because the amount of accrued interest increases each day. After the coupon payment is made, the quoted price would drop dramatically. Using the flat price for quotation avoids that misrepresentation. The full price, flat price plus accrued interest, is not usually quoted by bond dealers. Accrued interest is included in the full price, and bond dealers do not generally quote the full price.

19 B is correct. The bond's full price is 103.10. The price is determined in the following manner:

As of the beginning of the coupon period on 10 April 2020, there are 2.5 years (five semiannual periods) to maturity. These five semiannual periods occur on 10 October 2020, 10 April 2021, 10 October 2021, 10 April 2022, and 10 October 2022.

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT}{(1+r)^3} + \frac{PMT}{(1+r)^4} + \frac{PMT+FV}{(1+r)^5},$$

where

PV = present value

PMT = coupon payment per period

FV = future value paid at maturity, or the par value of the bond

r = market discount rate, or required rate of return per period

$$PV = \frac{2.5}{(1+0.02)^1} + \frac{2.5}{(1+0.02)^2} + \frac{2.5}{(1+0.02)^3} + \frac{2.5}{(1+0.02)^4} + \frac{2.5+100}{(1+0.02)^5}.$$

$$PV = 2.45 + 2.40 + 2.36 + 2.31 + 92.84 = 102.36.$$

The accrued interest period is identified as 66/180. The number of days between 10 April 2020 and 16 June 2020 is 66 days, based on the 30/360 day-count convention (20 days remaining in April + 30 days in May + 16 days in June = 66 days total). The number of days between coupon periods is assumed to be 180 days using the 30/360 day convention.

$$PV^{Full} = PV \times (1 + r)^{66/180}.$$

$$PV^{Full} = 102.36 \times (1.02)^{66/180} = 103.10.$$

20 C is correct. The accrued interest per 100 of par value is closest to 0.92. The accrued interest is determined in the following manner: The accrued interest period is identified as 66/180. The number of days between 10 April 2020 and 16 June 2020 is 66 days, based on the 30/360 day-count convention (20 days remaining in April + 30 days in May + 16 days in June = 66 days total). The number of days between coupon periods is assumed to be 180 days using the 30/360 day convention.

$$\text{Accrued interest} = \frac{t}{T} \times PMT,$$

where

t = number of days from the last coupon payment to the settlement date
T = number of days in the coupon period
t/T = fraction of the coupon period that has gone by since the last payment
PMT = coupon payment per period

$$\text{Accrued interest} = \frac{66}{180} \times \frac{5.00}{2} = 0.92.$$

21 A is correct. The flat price of 102.18 is determined by subtracting the accrued interest (from Question 20) from the full price (from Question 19).

$$PV^{Flat} = PV^{Full} - \text{Accrued interest}.$$

$$PV^{Flat} = 103.10 - 0.92 = 102.18.$$

22 B is correct. For bonds not actively traded or not yet issued, matrix pricing is a price estimation process that uses market discount rates based on the quoted prices of similar bonds (similar times-to-maturity, coupon rates, and credit quality).

23 A is correct. Matrix pricing is used in underwriting new bonds to get an estimate of the required yield spread over the benchmark rate. The benchmark rate is typically the yield-to-maturity on a government bond having the same or close to the same time-to-maturity. The spread is the difference between the yield-to-maturity on the new bond and the benchmark rate. The yield spread is the additional compensation required by investors for the difference in the credit risk, liquidity risk, and tax status of the bond relative to the government bond.

In matrix pricing, the market discount rates of comparable bonds and the yield-to-maturity on a government bond having a similar time-to-maturity are not estimated. Rather, they are known and are used to estimate the required yield spread of a new bond.

24 B is correct. The formula for calculating this bond's yield-to-maturity is

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT}{(1+r)^3} + \cdots + \frac{PMT}{(1+r)^{39}} + \frac{PMT + FV}{(1+r)^{40}},$$

where

PV = present value, or the price of the bond

PMT = coupon payment per period

FV = future value paid at maturity, or the par value of the bond

r = market discount rate, or required rate of return per period

$$111 = \frac{2.5}{(1+r)^1} + \frac{2.5}{(1+r)^2} + \frac{2.5}{(1+r)^3} + \cdots + \frac{2.5}{(1+r)^{39}} + \frac{2.5 + 100}{(1+r)^{40}}.$$

$r = 0.0209.$

To arrive at the annualized yield-to-maturity, the semiannual rate of 2.09% must be multiplied by two. Therefore, the yield-to-maturity is equal to 2.09% × 2 = 4.18%.

25 B is correct. The annual yield-to-maturity, stated for a periodicity of 12, is 7.21%. It is calculated as follows:

$$PV = \frac{FV}{(1+r)^N}.$$

$$75 = \left(\frac{100}{(1+r)^{4\times 12}}\right).$$

$$\frac{100}{75} = (1+r)^{48}.$$

$1.33333 = (1 + r)^{48}.$

$(1.33333)^{1/48} = [(1 + r)^{48}]^{1/48}.$

$1.33333^{02083} = (1 + r).$

$1.00601 = (1 + r).$

$1.00601 - 1 = r.$

$0.00601 = r.$

$r \times 12 = 0.07212$, or approximately 7.21%.

26 A is correct. The yield-to-maturity, stated for a periodicity of 12 (monthly periodicity), is 3.87%. The formula to convert an annual percentage rate (annual yield-to-maturity) from one periodicity to another is as follows:

$$\left(1 + \frac{APR_m}{m}\right)^m = \left(1 + \frac{APR_n}{n}\right)^n.$$

$$\left(1 + \frac{0.03897}{2}\right)^2 = \left(1 + \frac{APR_{12}}{12}\right)^{12}.$$

$$(1.01949)^2 = \left(1 + \frac{APR_{12}}{12}\right)^{12}.$$

$$1.03935 = \left(1 + \frac{APR_{12}}{12}\right)^{12}.$$

$$(1.03935)^{1/12} = \left[\left(1 + \frac{APR_{12}}{12}\right)^{12}\right]^{1/12}.$$

$$1.00322 = \left(1 + \frac{APR_{12}}{12}\right).$$

$$1.00322 - 1 = \left(\frac{APR_{12}}{12}\right).$$

$APR_{12} = 0.00322 \times 12 = 0.03865$, or approximately 3.87%.

27 B is correct. The yield-to-maturity is 5.77%. The formula for calculating this bond's yield-to-maturity is

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT}{(1+r)^3} + \cdots + \frac{PMT}{(1+r)^9} + \frac{PMT + FV}{(1+r)^{10}},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = future value paid at maturity, or the par value of the bond
r = market discount rate, or required rate of return per period

$$101 = \frac{3}{(1+r)^1} + \frac{3}{(1+r)^2} + \frac{3}{(1+r)^3} + \cdots + \frac{3}{(1+r)^9} + \frac{3 + 100}{(1+r)^{10}}.$$

$r = 0.02883$.

To arrive at the annualized yield-to-maturity, the semiannual rate of 2.883% must be multiplied by two. Therefore, the yield-to-maturity is equal to 2.883% × 2 = 5.77% (rounded).

28 C is correct. The yield-to-first-call is 6.25%. Given the first call date is exactly three years away, the formula for calculating this bond's yield-to-first-call is

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT}{(1+r)^3} + \cdots + \frac{PMT}{(1+r)^5} + \frac{PMT + FV}{(1+r)^6},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = call price paid at call date
r = market discount rate, or required rate of return per period

$$101 = \frac{3}{(1+r)^1} + \frac{3}{(1+r)^2} + \frac{3}{(1+r)^3} + \cdots + \frac{3}{(1+r)^5} + \frac{3+102}{(1+r)^6}.$$

$r = 0.03123.$

To arrive at the annualized yield-to-first-call, the semiannual rate of 3.123% must be multiplied by two. Therefore, the yield-to-first-call is equal to 3.123% × 2 = 6.25% (rounded).

29 C is correct. The yield-to-second-call is 5.94%. Given the second call date is exactly four years away, the formula for calculating this bond's yield-to-second-call is

$$PV = \frac{PMT}{(1+r)^1} + \frac{PMT}{(1+r)^2} + \frac{PMT}{(1+r)^3} + \cdots + \frac{PMT}{(1+r)^7} + \frac{PMT + FV}{(1+r)^8},$$

where

PV = present value, or the price of the bond
PMT = coupon payment per period
FV = call price paid at call date
r = market discount rate, or required rate of return per period

$$101 = \frac{3}{(1+r)^1} + \frac{3}{(1+r)^2} + \frac{3}{(1+r)^3} + \cdots \frac{3}{(1+r)^7} + \frac{3+101}{(1+r)^8}.$$

$r = 0.0297.$

To arrive at the annualized yield-to-second-call, the semiannual rate of 2.97% must be multiplied by two. Therefore, the yield-to-second-call is equal to 2.97% × 2 = 5.94%.

30 B is correct. The yield-to-worst is 5.77%. The bond's yield-to-worst is the lowest of the sequence of yields-to-call and the yield-to-maturity. From Questions 27–29, we have the following yield measures for this bond:

Yield-to-first-call: 6.25%

Yield-to-second-call: 5.94%

Yield-to-maturity: 5.77%

Thus, the yield-to-worst is 5.77%.

31 B is correct. The discount or required margin is 236 bps. Given the floater has a maturity of two years and is linked to six-month MRR, the formula for calculating the discount margin is

$$PV = \frac{\frac{(\text{Index} + QM) \times FV}{m}}{\left(1 + \frac{\text{Index} + DM}{m}\right)^1} + \frac{\frac{(\text{Index} + QM) \times FV}{m}}{\left(1 + \frac{\text{Index} + DM}{m}\right)^2} + \cdots + \frac{\frac{(\text{Index} + QM) \times FV}{m} + FV}{\left(1 + \frac{\text{Index} + DM}{m}\right)^4},$$

where

PV = present value, or the price of the floating-rate note = 97
Index = reference rate, stated as an annual percentage rate = 0.01
QM = quoted margin, stated as an annual percentage rate = 0.0080

FV = future value paid at maturity, or the par value of the bond = 100
m = periodicity of the floating-rate note, the number of payment periods per year = 2
DM = discount margin, the required margin stated as an annual percentage rate

Substituting the given values:

$$97 = \frac{\frac{(0.01 + 0.0080) \times 100}{2}}{\left(1 + \frac{0.01 + DM}{2}\right)^1} + \frac{\frac{(0.01 + 0.0080) \times 100}{2}}{\left(1 + \frac{0.01 + DM}{2}\right)^2} + \cdots + \frac{\frac{(0.01 + 0.0080) \times 100}{2} + 100}{\left(1 + \frac{0.01 + DM}{2}\right)^4}.$$

$$97 = \frac{0.90}{\left(1 + \frac{0.01 + DM}{2}\right)^1} + \frac{0.90}{\left(1 + \frac{0.01 + DM}{2}\right)^2} + \frac{0.90}{\left(1 + \frac{0.01 + DM}{2}\right)^3} + \frac{0.90 + 100}{\left(1 + \frac{0.01 + DM}{2}\right)^4}.$$

To calculate DM, begin by solving for the discount rate per period:

$$97 = \frac{0.90}{(1+r)^1} + \frac{0.90}{(1+r)^2} + \frac{0.90}{(1+r)^3} + \frac{0.90 + 100}{(1+r)^4}.$$

$$r = 0.0168.$$

Now, solve for DM:

$$\frac{0.01 + DM}{2} = 0.0168.$$

$$DM = 0.0236.$$

The discount margin for the floater is equal to 236 bps.

32 A is correct. FRN X will be priced at a premium on the next reset date because the quoted margin of 0.40% is greater than the discount, or required, margin of 0.32%. The premium amount is the present value of the extra, or "excess," interest payments of 0.08% each quarter (0.40% – 0.32%). FRN Y will be priced at par value on the next reset date since there is no difference between the quoted and discount margins. FRN Z will be priced at a discount since the quoted margin is less than the required margin.

33 C is correct. The bond equivalent yield is closest to 3.78%. It is calculated as

$$AOR = \left(\frac{\text{Year}}{\text{Days}}\right) \times \left(\frac{FV - PV}{PV}\right),$$

where

PV = present value, principal amount, or the price of the money market instrument
FV = future value, or the redemption amount paid at maturity including interest
Days = number of days between settlement and maturity
Year = number of days in the year
AOR = add-on rate, stated as an annual percentage rate (also called bond equivalent yield)

$$AOR = \left(\frac{365}{350}\right) \times \left(\frac{100 - 96.5}{96.5}\right).$$

$$AOR = 1.04286 \times 0.03627.$$

$$AOR = 0.03783\text{, or approximately } 3.78\%.$$

34 C is correct. The bond equivalent yield is closest to 4.40%. The present value of the banker's acceptance is calculated as

$$PV = FV \times \left(1 - \frac{\text{Days}}{\text{Year}} \times DR\right),$$

where

PV = present value, or price of the money market instrument
FV = future value paid at maturity, or face value of the money market instrument
Days = number of days between settlement and maturity
Year = number of days in the year
DR = discount rate, stated as an annual percentage rate

$$PV = 100 \times \left(1 - \frac{\text{Days}}{\text{Year}} \times DR\right).$$

$$PV = 100 \times \left(1 - \frac{180}{360} \times 0.0425\right).$$

$$PV = 100 \times (1 - 0.02125).$$

$$PV = 100 \times 0.97875.$$

$$PV = 97.875.$$

The bond equivalent yield (AOR) is calculated as

$$AOR = \left(\frac{\text{Year}}{\text{Days}}\right) \times \left(\frac{FV - PV}{PV}\right),$$

where

PV = present value, principal amount, or the price of the money market instrument
FV = future value, or the redemption amount paid at maturity including interest
Days = number of days between settlement and maturity
Year = number of days in the year
AOR = add-on rate (bond equivalent yield), stated as an annual percentage rate

$$AOR = \left(\frac{365}{180}\right) \times \left(\frac{100 - PV}{PV}\right).$$

$$AOR = \left(\frac{365}{180}\right) \times \left(\frac{100 - 97.875}{97.875}\right).$$

$AOR = 2.02778 \times 0.02171.$

$AOR = 0.04402$, or approximately 4.40%.

Note that *PV* is calculated using an assumed 360-day year and *AOR* (bond equivalent yield) is calculated using a 365-day year.

35 B is correct. All bonds on a par curve are assumed to have similar, not different, credit risk. Par curves are obtained from spot curves, and all bonds used to derive the par curve are assumed to have the same credit risk, as well as the same periodicity, currency, liquidity, tax status, and annual yields. A par curve is a sequence of yields-to-maturity such that each bond is priced at par value.

36 B is correct. The spot curve, also known as the strip, or zero, curve, is the yield curve constructed from a sequence of yields-to-maturity on zero-coupon bonds. The par curve is a sequence of yields-to-maturity such that each bond is priced at par value. The forward curve is constructed using a series of forward rates, each having the same time frame.

37 B is correct. The forward rate can be interpreted to be the incremental or marginal return for extending the time-to-maturity of an investment for an additional time period. The add-on rate (bond equivalent yield) is a rate quoted for money market instruments, such as bank certificates of deposit, and indexes, such as MRR, Libor and Euribor. Yield-to-maturity is the internal rate of return on the bond's cash flows—the uniform interest rate such that when the bond's future cash flows are discounted at that rate, the sum of the present values equals the price of the bond. It is the implied market discount rate.

38 B is correct. The three-year implied spot rate is closest to 1.94%. It is calculated as the geometric average of the one-year forward rates:

$(1.0080 \times 1.0112 \times 1.0394) = (1 + z_3)^3.$

$1.05945 = (1 + z_3)^3.$

$(1.05945)^{1/3} = [(1 + z_3)^3]^{1/3}.$

$1.01944 = 1 + z_3.$

$1.01944 - 1 = z_3.$

$0.01944 = z_3.$

$z_3 = 1.944\%$, or approximately 1.94%.

39 B is correct. The value per 100 of par value is closest to 105.01. Using the forward curve, the bond price is calculated as follows:

$$\frac{3.5}{1.0080} + \frac{103.5}{(1.0080 \times 1.0112)} = 3.47 + 101.54 = 105.01.$$

40 C is correct. The spread component of a specific bond's yield-to-maturity is least likely impacted by changes in inflation in its currency of denomination. The effect of changes in macroeconomic factors, such as the expected rate of inflation in the currency of denomination, is seen mostly in changes in the benchmark yield. The spread or risk premium component is impacted by microeconomic factors specific to the bond and bond issuer, including tax status and quality rating.

41 A is correct. The I-spread, or interpolated spread, is the yield spread of a specific bond over the standard swap rate in that currency of the same tenor. The yield spread in basis points over an actual or interpolated government bond is known as the G-spread. The Z-spread (zero-volatility spread) is the constant spread that is added to each spot rate such that the present value of the cash flows matches the price of the bond.

42 B is correct. The G-spread is closest to 285 bps. The benchmark rate for UK fixed-rate bonds is the UK government benchmark bond. The euro interest rate spread benchmark is used to calculate the G-spread for euro-denominated corporate bonds, not UK bonds. The G-spread is calculated as follows:

Yield-to-maturity on the UK corporate bond:

$$100.65 = \frac{5}{(1+r)^1} + \frac{5}{(1+r)^2} + \frac{105}{(1+r)^3}; r = 0.04762, \text{ or } 476 \text{ bps.}$$

Yield-to-maturity on the UK government benchmark bond:

$$100.25 = \frac{2}{(1+r)^1} + \frac{2}{(1+r)^2} + \frac{102}{(1+r)^3}; r = 0.01913, \text{ or } 191 \text{ bps.}$$

The G-spread is 476 – 191 = 285 bps.

43 A is correct. The value of the bond is closest to 92.38. The calculation is

$$\begin{aligned} PV &= \frac{PMT}{(1+z_1+Z)^1} + \frac{PMT}{(1+z_2+Z)^2} + \frac{PMT+FV}{(1+z_3+Z)^3} \\ &= \frac{5}{(1+0.0486+0.0234)^1} + \frac{5}{(1+0.0495+0.0234)^2} + \frac{105}{(1+0.0565+0.0234)^3} \\ &= \frac{5}{1.0720} + \frac{5}{1.15111} + \frac{105}{1.25936} = 4.66 + 4.34 + 83.38 = 92.38. \end{aligned}$$

44 C is correct. The option value in basis points per year is subtracted from the Z-spread to calculate the OAS. The Z-spread is the constant yield spread over the benchmark spot curve. The I-spread is the yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

READING

42

Introduction to Asset-Backed Securities

by Frank J. Fabozzi, PhD, CPA, CFA

Frank J. Fabozzi, PhD, CPA, CFA, is at EDHEC Business School (France).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** explain benefits of securitization for economies and financial markets;
☐	**b.** describe securitization, including the parties involved in the process and the roles they play;
☐	**c.** describe typical structures of securitizations, including credit tranching and time tranching;
☐	**d.** describe types and characteristics of residential mortgage loans that are typically securitized;
☐	**e.** describe types and characteristics of residential mortgage-backed securities, including mortgage pass-through securities and collateralized mortgage obligations, and explain the cash flows and risks for each type;
☐	**f.** define prepayment risk and describe the prepayment risk of mortgage-backed securities;
☐	**g.** describe characteristics and risks of commercial mortgage-backed securities;
☐	**h.** describe types and characteristics of non-mortgage asset-backed securities, including the cash flows and risks of each type;
☐	**i.** describe collateralized debt obligations, including their cash flows and risks;
☐	**j.** describe characteristics and risks of covered bonds and how they differ from other asset-backed securities.

INTRODUCTION: BENEFITS OF SECURITIZATION

1

a explain benefits of securitization for economies and financial markets

© 2021 CFA Institute. All rights reserved.

Previous readings examined the risk characteristics of various fixed-income instruments and the relationships between maturity, coupon, and interest rate changes. This reading introduces fixed-income instruments created through a process known as **securitization**. The securitization process transfers ownership of assets such as loans or receivables from the original owners into a special legal entity. The special legal entity then issues securities, using the asset cash flows to pay interest and repay the principal to investors. These securities are referred to generically as **asset-backed securities** (ABS), and the pool of assets from which their cash flows are generated is called collateral or **securitized assets**. These loans and receivables typically include residential mortgage loans (mortgages), commercial mortgages, automobile (auto) loans, student loans, bank loans, accounts receivable, or credit card receivables. While the ABS market in the United States remains the largest in the world, securitization has recently expanded in both Asia and Europe as well as to other income-yielding assets, such as airport landing slots, toll roads, and cell tower leases. The growing depth and breadth of the global ABS market underscores the importance of a solid understanding of securitization among issuers, investors, and financial analysts.

This reading describes the securitization process, discusses its benefits for both issuers and investors, and explains the investment characteristics of different types of ABS. Although investors can invest in some loan types via private credit funds or through secondary markets, securitization creates a more direct link between investors and borrowers for many types of loans and receivables. The terminology regarding ABS varies by jurisdiction. **Mortgage-backed securities** (MBS) are ABS backed by a pool of mortgages, and a distinction is sometimes made between MBS and ABS backed by non-mortgage assets. This distinction is common in the United States, for example, where typically the term "mortgage-backed securities" refers to securities backed by high-quality real estate mortgages and the term "asset-backed securities" refers to securities backed by other types of assets. **Covered bonds** date back to 18th century Europe and are similar to ABS, but they offer investors recourse to both the issuing financial institution and an underlying asset pool.

To underline the importance of securitization from a macroeconomic perspective, we discuss the benefits of securitization for economies and financial markets. Then, we describe the structure of a securitization, identifying the parties involved and their roles as well as typical structures. We further discuss securities backed by mortgages for real estate property, including residential MBS and commercial MBS, and review two common types of non-mortgage ABS—namely, those for auto loans and credit card receivables. The reading concludes with a description of collateralized debt obligations and covered bonds.

1.1 Benefits of Securitization for Economies and Financial Markets

The securitization of pools of loans and receivables into multiple securities provides economies and financial markets with several benefits.

Home or auto purchases have been traditionally financed by loans originated by financial institutions, such as commercial banks. For investors to gain exposure to these relatively illiquid loans, they must hold some combination of deposits, debt, or common equity issued by banks. This situation creates an additional intermediary (that is, the bank) between the borrowers and the investors. In addition, by being constrained to hold bank deposits and securities, investors cannot gain direct exposure to loans but are also affected by economic risks undertaken in other bank activities.

Securitization solves a number of these problems. It allows investors to achieve more direct legal claims on loan and receivables portfolios, enabling them to tailor interest rate and credit risk exposures to suit their specific needs. Disintermediation (that is, lessening the role of intermediaries) can effectively reduce borrower costs and enhance risk-adjusted investor returns. At the same time, banks can separate loan origination from financing, improving their profitability via origination fees and reducing capital requirements for loans that are sold. Securitization enables banks to expand lending origination beyond their balance sheets, ultimately benefiting individuals, governments, and companies that need to borrow.

Securitization also benefits investors by creating access to securities with profiles that match their risk, return, and maturity needs that are otherwise not directly available. For example, a pension fund with a long-term horizon can gain access to long-term real estate loans by investing in residential MBS without having to invest in bank bonds or stocks. Although few institutional or individual investors are willing to make or purchase real estate loans, auto loans, or credit card receivables directly, they may invest in a security backed by such loans or receivables. The ABS that are created by pooling these loans and receivables have characteristics similar to those of a standard bond and do not require the specialized resources and expertise needed to originate, monitor, and collect the payments from the underlying loans and receivables. As a result, investors can increase exposure to the risk–return characteristics of a wider range of underlying assets. Note that in many countries, the sale of ABS and similar instruments is restricted to investors who meet certain qualifications, such as those pertaining to net worth.

Securitization allows for the creation of tradable securities with better liquidity than that of the original loans on the bank's balance sheet. In making loans and receivables tradable, securitization makes financial markets more efficient. It also improves liquidity, which reduces liquidity risk in the financial system, as described later.

An important benefit of securitization for companies is that ABS provide an alternative means of funding operations that can be considered alongside bond, preferred equity, and common equity issuance. Companies that originate loans and receivables that can be securitized often compare and optimize the funding costs associated with each source of financing. As we will see, securitization is often less costly than a corporate bond issue secured by the same collateral.

For these reasons, securitization is beneficial to economies and financial markets and has been embraced by many sovereign governments. For example, the Italian government has used securitization since the late 1990s for privatizing public assets. In emerging markets, securitization is widely used. For example, in South America, companies and banks with high credit ratings have used securitization to sell receivables on exports, such as oil, to lower their funding costs.

Although securitization brings many benefits to economies, it is not without risks, and some of these risks are widely attributed to have precipitated the turmoil in financial markets during 2007–2009. Broadly, those risks fall into two categories: risks that relate primarily to the timing of the ABS's cash flows, such as contraction risk and extension risk, and risks related to the inherent credit risk of the loans and receivables backing the ABS. This reading describes these risks and discusses some of the structures used to mitigate them as well as redistribute them.

HOW SECURITIZATION WORKS

2

b describe securitization, including the parties involved in the process and the roles they play

When assets are securitized, several legal and regulatory conditions must be satisfied. A number of parties participate in the process to facilitate the transaction and ensure these conditions are met. In this section, a typical securitization is described by way of a hypothetical example. The example describes the parties involved in a securitization and their roles. It also introduces the typical structures of securitizations, such as credit tranching and time tranching.

2.1 An Example of a Securitization

Mediquip, a hypothetical company, is a manufacturer of medical equipment that ranges in cost from US$50,000 to US$300,000. The majority of Mediquip's sales are made through loans granted by the company to its customers, and the medical equipment serves as collateral for the loans. These loans, which represent an asset to Mediquip, have maturities of five years and carry a fixed interest rate. They are fully amortizing with monthly payments; that is, the borrowers make equal payments each month consisting of interest payment and principal repayment. The total principal repaid from the 60 loan payments (12 months × 5 years) is such that the amount borrowed is fully repaid at the end of the term.

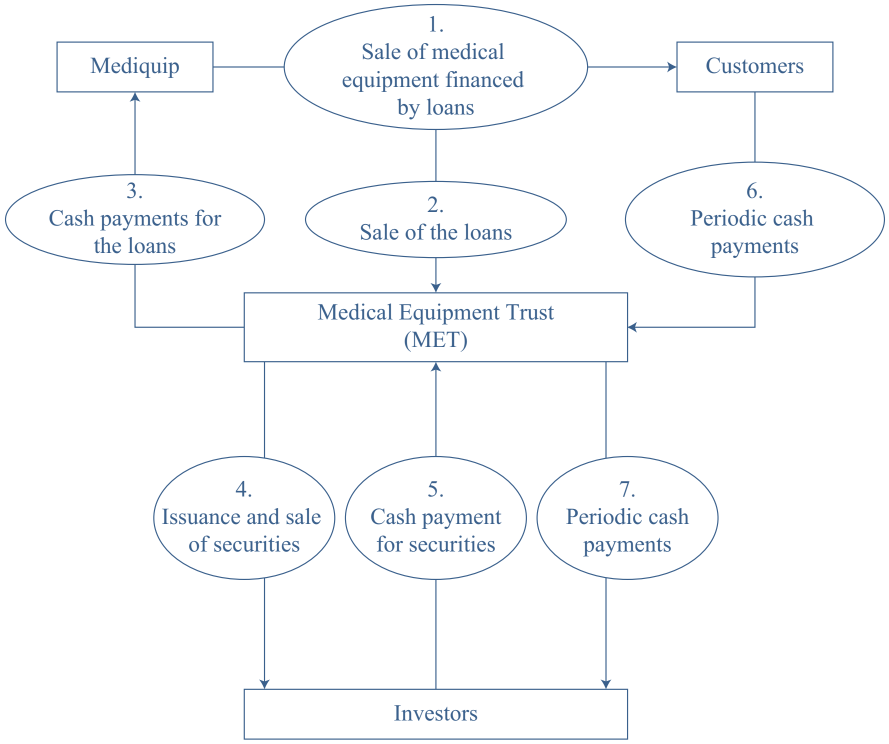
Mediquip's credit department makes the decision about whether to extend credit to customers and services the loans that are made. Loan servicing refers to administering any aspect of a loan, including collecting payments from borrowers, notifying borrowers who may be delinquent, and recovering and disposing of the medical equipment if the borrower does not make the scheduled payments by a specified time. If one of its customers defaults, Mediquip can seize the medical equipment and sell it to try to recoup the remaining principal on the loan. Although the servicer of such loans need not be the originator of the loans, the assumption in this example is that Mediquip is the servicer.

The following is an illustration of how these loans can be securitized. Assume that Mediquip has US$200 million of loans. This amount is shown on Mediquip's balance sheet as an asset. Assume also that Mediquip wants to raise US$200 million, which happens to be the amount of the loans. Because Mediquip's treasurer is aware of the potentially lower funding costs of securitization, he decides to raise the US$200 million by securitizing the loans on the medical equipment rather than by issuing corporate bonds.

To do so, Mediquip sets up a separate legal entity called Medical Equipment Trust (MET), to which it sells the loans on the medical equipment. Such a legal entity is referred to as a **special purpose entity** (SPE) and sometimes also called a special purpose vehicle (SPV) or a special purpose company. The legal form of the SPE varies by jurisdiction, but in almost all cases, the ultimate owner of the loans—MET in our example—is legally independent and is considered bankruptcy remote from the seller of the loans. Setting up a separate legal entity ensures that if Mediquip, the originator of the loans, files for bankruptcy, the loans backing the ABS that are issued by MET are secure within the SPE and creditors of Mediquip have no claim on them. Note that in some jurisdictions, the SPE may, in turn, transfer the loans to a trust or a limited company.

A securitization is diagramed in Exhibit 1. The SPE set up by Mediquip is called MET. The top of Exhibit 1 reflects Mediquip's business model as described above—that is, the sale of medical equipment financed by loans (first oval). Mediquip sells to MET US$200 million of loans (second oval) and receives from MET US$200 million in cash (third oval); in this simplified example, the costs associated with the securitization are ignored. MET issues and sells securities that are backed by the pool of securitized loans (fourth oval) and receives cash (fifth oval). These securities are the ABS mentioned earlier, and the US$200 million of loans represent the collateral. The periodic cash payments that are received from the collateral—that is, the monthly payments made by Mediquip's customers that include both interest payment and principal repayment (sixth oval)—are used to make the periodic cash payments to the security holders—the investors who bought the ABS (seventh oval).

Exhibit 1 Mediquip's Securitization


Mediquip
1.
Sale of medical equipment financed by loans
Customers
3.
Cash payments for the loans
2.
Sale of the loans
6.
Periodic cash payments
Medical Equipment Trust (MET)
4.
Issuance and sale of securities
5.
Cash payment for securities
7.
Periodic cash payments
Investors

2.2 Parties to a Securitization and Their Roles

Securitization requires the publication of a prospectus, a document that contains information about the securitization. The three main parties to a securitization are

- the seller of the collateral, sometimes called the depositor (Mediquip in our example);
- the SPE that purchases the loans or receivables and uses them as collateral to issue the ABS—MET in our example (the SPE is often referred to as the issuer in the prospectus because it is the entity that issues the securities; it may also be called the trust if the SPE is set up as a trust); and
- the servicer of the loans (Mediquip in our example).

Other parties are also involved in a securitization: independent accountants, lawyers/attorneys, trustees, underwriters, rating agencies, and financial guarantors. All these parties, including the servicer when it is different from the seller of the collateral, are referred to as third parties to the securitization.

A significant amount of legal documentation is involved in a securitization. Lawyers/attorneys are responsible for preparing the legal documents. An important legal document is the purchase agreement between the seller of the collateral and the SPE, which sets forth the representations and warranties that the seller makes about the assets sold. These representations and warranties assure investors about the quality of the assets, an important consideration when assessing the risks associated with the ABS.

Another important legal document is the prospectus, which describes the structure of the securitization, including the priority and amount of payments to be made to the servicer, administrators, and the ABS holders. Securitizations often use several forms of credit enhancements, which are documented in the prospectus. Credit enhancements are provisions that are used to reduce the credit risk of a bond issue. They include (1) internal credit enhancements, such as subordination, overcollateralization, and reserve accounts, and (2) external credit enhancements, such as financial guarantees by banks or insurance companies, letters of credit, and cash collateral accounts. Securitizations often use subordination, which will be discussed later. Prior to the 2007–09 credit crisis, many securitizations included financial guarantees by a third party. The most common third-party financial guarantors are monoline insurance companies, or monoline insurers. A monoline insurer is a private insurance company whose business is restricted to providing guarantees for financial instruments such as ABS. Following the financial difficulties and downgrading of the major monoline insurers as a result of the financial crisis that began in the mortgage market in mid-2007, few structures have used financial guarantees from a monoline insurer.

A trustee or trustee agent is typically a financial institution with trust powers that safeguards the assets after they have been sold to the SPE, holds the funds due to the ABS holders until they are paid, and provides periodic information to the ABS holders. The information is provided in the form of remittance reports, which may be issued monthly, quarterly, or as agreed to in the terms of the prospectus.

Underwriters and rating agencies perform the same functions in a securitization as they do in a standard bond offering

EXAMPLE 1

A Used Luxury Auto Securitization

Used Luxury Auto (ULA) is a hypothetical company that has established a nationwide business in buying used luxury autos and then refurbishing them with the latest in electronic equipment (for instance, UBS ports and rear-view cameras). ULA Corp then sells these autos in the retail market, often financing the sales with promissory notes from the buyers via its ULA Credit Corp.

The following information is taken from a theoretical filing by ULA with the Securities and Exchange Commission for a securitization:

Issuer: ULA Trust 2020

Seller and Servicer: ULA Credit Corp

Notes:

$500,000,000 4.00% ULA Trust contract-backed Class A notes, rated AAA

$250,000,000 4.80% ULA Trust contract-backed Class B notes, rated A

Contracts: The assets underlying the notes are fixed-rate promissory notes relating to the purchase of used automobiles refurbished by ULA Corp.

1 The collateral for this securitization is:

- **A** ULA Trust contract-backed Class A and Class B notes.
- **B** Used automobiles refurbished by ULA Corp.
- **C** Fixed-rate promissory notes relating to the purchase of used automobiles refurbished by ULA Corp.

2 The special purpose entity in this securitization is:

- **A** ULA Corp.
- **B** ULA Credit Corp.
- **C** ULA Trust 2020.

3 ULA Credit Corp is responsible for:

- **A** selling the collateral to the SPE and collecting payments from borrowers on the underlying promissory notes.
- **B** refurbishing the used motorcycles and collecting payments from borrowers on the underlying promissory notes.
- **C** selling the contract-backed Class A and Class B notes to investors and making the cash interest and principal payments to them.

Solution to 1:

C is correct. The collateral is the pool of securitized assets from which the cash flows will be generated. It is the debt obligations that have been securitized. These contracts are the loans, called promissory notes, provided to purchasers of the used automobiles that were refurbished by ULA Corp.

Solution to 2:

C is correct. ULA Trust 2020 is the issuer of the ABS and, thus, the SPE. The SPE purchases the contracts that become the collateral from ULA Corp, the automobile refurbisher, but it is ULA Credit Corp that originates the loans and is, therefore, the seller of the collateral. ULA Credit Corp is also the servicer of the debt obligations.

Solution to 3:

A is correct. ULA Credit Corp is the seller of the collateral, the promissory notes. As the servicer, it is responsible for collecting payments from borrowers, notifying borrowers who may be delinquent, and if necessary, recovering and disposing of the automobile if the borrower defaults.

STRUCTURE OF A SECURITIZATION

3

c describe typical structures of securitizations, including credit tranching and time tranching

A simple securitization may involve the sale of only one class of bond or ABS. Let us call this class Bond Class A. Returning to the Mediquip and MET example, MET may raise US$200 million by issuing 200,000 certificates for Bond Class A with a par value of US$1,000 per certificate. Thus, each certificate holder is entitled to 1/200,000 of the payments from the collateral after payment of servicing and other administrative fees.

The structure of the securitization is often more complicated than a single class of ABS. As mentioned earlier, it is common for securitizations to include a form of internal credit enhancement called **subordination**, also referred to as **credit tranching**. In such a structure, there is more than one bond class or tranche, and the bond classes differ as to how they will share any losses resulting from defaults of the borrowers whose loans are in the collateral. The bond classes are classified as senior bond classes or subordinated bond classes—hence, the reason this structure is also referred to as a senior/subordinated structure. The subordinated bond classes are sometimes called "non-senior bond classes" or "junior bond classes." They function as credit protection for the more senior bond classes; that is, losses are realized by the subordinated bond classes before any losses are realized by the senior bond classes. This type of protection is also commonly referred to as a "waterfall" structure because of the cascading flow of payments between bond classes in the event of default.

For example, suppose MET issues two bond classes with a total par value of US$200 million: Bond Class A, the senior bond class, with a par value of US$120 million, and Bond Class B, the subordinated bond class, with a par value of US$80 million. In this senior/subordinated structure, also referred to as credit tranching, Bond Class B will absorb losses up to US$80 million. Thus, if defaults by Mediquip's customers do not exceed US$80 million, Bond Class A will be fully repaid its US$120 million. The purpose of this structure is to redistribute the credit risk associated with the collateral. The creation of a set of bond classes allows investors to choose the level of credit risk that they prefer to bear.

More than one subordinated bond class may be created. Suppose MET issues the following structure:

Bond Class	Par Value (US$ millions)
A (senior)	180
B (subordinated)	14
C (subordinated)	6
Total	200

In this structure, Bond Class A is the senior bond class whereas both Bond Class B and Bond Class C are subordinated bond classes from the perspective of Bond Class A. The rules for the distribution of losses are as follows. All losses on the collateral are absorbed by Bond Class C before any losses are realized by Bond Class B and then Bond Class A. Consequently, if the losses on the collateral do not exceed US$6 million, no losses will be realized by Bond Class A or Bond Class B. If the losses exceed US$6 million, Bond Class B must absorb the losses up to an additional US$14 million. For example, if the total loss on the collateral is US$16 million, Bond Class C loses its entire par value of US$6 million and Bond Class B realizes a loss of US$10 million of its par value of US$14 million. Bond Class A does not realize any loss in this scenario. Clearly, Bond Class A realizes a loss only if the total loss on the collateral exceeds US$20 million.

The structure of a securitization may also allow the redistribution of another type of risk, called "prepayment risk," among bond classes. **Prepayment risk** is the uncertainty that the cash flows will be different from the scheduled cash flows as set forth in the loan agreement because of the borrowers' ability to alter payments, usually to take advantage of interest rate movements. For example, when interest rates decline, borrowers tend to pay off part or all of their loans and refinance at lower interest rates. The creation of bond classes that possess different expected maturities is referred to as **time tranching** and is discussed later in this reading.

It is possible, and quite common, for a securitization to have structures with both credit tranching and time tranching.

EXAMPLE 2

Bond Classes and Tranching

Return to the ULA securitization described in Example 1. Based on the information provided, the form of credit enhancement that the transaction *most likely* includes is:

A time tranching.

B credit tranching.

C a financial guarantee.

Solution:

B is correct. The ULA securitization includes two bond classes: Class A and Class B. Each bond class has a fixed but different interest rate; the interest rate increases from 4.00% for Class A notes to 4.80% for Class B notes. Thus, it is likely that the transaction has credit tranching and that the two bond classes display a senior/subordinated structure, with Class A notes being senior to Class B notes, the subordinated bond class. As the credit risk increases from the Class A notes to the Class B notes, so does the interest rate to reflect the additional compensation investors require for bearing the additional credit risk. The information provided does not give any indication of either time tranching or a financial guarantee.

3.1 Key Role of the Special Purpose Entity

The SPE plays a pivotal role in securitization. In fact, the setup of a legal entity that plays the same role as an SPE in terms of protecting the rights of ABS holders is a prerequisite in any country that wants to allow securitization. Indeed, without a provision in a country's legal system for the equivalent of an SPE, the benefits of using securitization by an entity seeking to raise funds would not exist. Let us explain why by returning to our example involving Mediquip and MET.

Assume that Mediquip has a credit rating from a credit-rating agency, such as Standard & Poor's, Moody's Investors Service, or Fitch Ratings. A credit rating reflects the agency's opinion about the creditworthiness of an entity and/or the debt securities the entity issues. Suppose that the credit rating assigned to Mediquip is BB or Ba2. Such a credit rating means that Mediquip is below what is referred to as an investment-grade credit rating.

Assume again that Mediquip's treasurer wants to raise US$200 million and is contemplating doing so by issuing a five-year corporate bond rather than by securitizing the loans. The treasurer is, of course, concerned about the funding cost and would like the lowest possible interest rate available relative to some benchmark interest rate. The difference between the interest rate paid on the five-year corporate bond and the benchmark interest rate is the spread. The spread reflects the compensation investors require for buying the corporate bond, which is riskier than the bonds issued at the benchmark interest rate. The major factor affecting the spread is the issuer's credit rating—hence, the reason the spread is called a "credit spread."

Another factor that will influence the credit spread is whether the bond is backed by collateral. A corporate bond that has collateral is often referred to as a secured bond. The collateral usually reduces the credit spread, making the credit spread of a secured bond lower than that of an otherwise identical unsecured bond. In our example, Mediquip's treasurer can use the loans on the medical equipment as collateral for the secured corporate bond issue. Thus, if Mediquip issues a five-year corporate bond to raise US$200 million, the credit spread will primarily reflect its credit rating, with a slight benefit for the use of collateral as explained below.

Now suppose that instead of using the loans as collateral for a secured corporate bond issue, Mediquip sells the loan contracts in an arm's length transaction to MET, the SPE. After the sale is completed, it is MET, not Mediquip, that legally owns them. As a result, if Mediquip is forced into bankruptcy while the loans are still outstanding, Mediquip's creditors cannot recover them because they are legally owned by another entity. Note that it is possible, however, that transfers made to bankruptcy-remote vehicles can be challenged as fraudulent conveyances and potentially unwound. The legal implication of setting up MET is that investors contemplating the purchase of any bond class backed by the cash flows from the pool of loans on the medical equipment will evaluate the credit risk associated with collecting the payments due on the receivables independently of Mediquip's credit rating.

Credit ratings are assigned to each bond class created in the securitization. They depend on the quality of the collateral—that is, how the rating agencies evaluate the credit risk of the pool of securitized loans or receivables. Depending on the structure of the securitization, each bond class receives a credit rating that reflects its credit risk, and some of the bond classes may have a better credit rating than the company that is seeking to raise funds. As a result, in aggregate, the funding cost of a securitization may be lower than that of a corporate bond issue. Access to lower funding cost is a key role of the SPE in a securitization.

You may ask why a securitization can be cheaper than a corporate bond issue secured by the same collateral. The reason is that the SPE would not be affected by the bankruptcy of the seller of the collateral. As mentioned above, the assets belong to the SPE, not to the entity that sold the assets to the SPE. In the United States and other countries, when a company is liquidated, creditors receive distributions based on the absolute priority rule to the extent that assets are available. The absolute priority rule is the principle that senior creditors are paid in full before subordinated creditors are paid anything. The absolute priority rule also guarantees the seniority of creditors relative to equityholders.

Whereas the absolute priority rule generally holds in liquidations, it has not always been upheld by the courts in reorganizations. Thus, although investors in the debt of a company may believe they have priority over the equityholders and priority over other classes of creditors, the actual outcome of a reorganization may be far different from the terms stated in the debt agreement; that is, there is no assurance that if the corporate bond has collateral, the rights of the bondholders will be respected. For this reason, the credit spread for a corporate bond backed by collateral does not decrease dramatically.

In the case of a securitization, the courts have in most jurisdictions no discretion to change seniority because the bankruptcy of a company does not affect the SPE. The rules set forth in the legal document, which describes how losses are to be absorbed by each bond class, are unaffected by the company's bankruptcy. This important decoupling of the credit risk of the entity needing funds from the bond classes issued by the SPE explains why the SPE's legal role is critical.

The SPE is bankruptcy remote from the seller of the collateral, which means that the bankruptcy of the seller of the collateral will not affect the holders of securities issued by the SPE and backed by the collateral. The security holders face credit risk only to the extent that the borrowers whose claims the SPE has purchased default on their loans. The SPE's ability to make cash payments to the security holders remains intact if the borrowers make the interest payments and/or the principal repayments on their loans.

In many EU countries, the creditors are protected in the recognition of the securitization as a true sale. The SPE has full legal ownership of the securitized assets, which are de-recognized from the seller's balance sheet. However, it is important to note that not all countries have the same legal framework. Impediments with respect to ABS issuance have arisen in some countries because the concept of trust law is less well developed globally than it is in the United States and other developed countries. Thus, investors should be aware of the legal considerations that apply in the jurisdictions where they purchase ABS.

EXAMPLE 3

Special Purpose Entity and Bankruptcy

Agnelli Industries (Agnelli), a manufacturer of industrial machine tools based in Bergamo, Italy, has €500 million of corporate bonds outstanding. These bonds have a credit rating below investment grade. Agnelli has €400 million of receivables on its balance sheet that it would like to securitize. The receivables represent payments Agnelli expects to receive for machine tools it has sold to various customers in Europe. Agnelli sells the receivables to Agnelli Trust, a special purpose entity. Agnelli Trust then issues ABS, backed by the pool of receivables, with the following structure:

Bond Class	Par Value (€ millions)
A (senior)	280
B (subordinated)	60
C (subordinated)	60
Total	400

Bond Class A is given an investment-grade credit rating by the credit-rating agencies.

1 Why does Bond Class A have a higher credit rating than the corporate bonds?

2 If Agnelli Industries files for bankruptcy after the issuance of the asset-backed security:

- A Bond Classes A, B, and C will be unaffected.
- B Bond Classes A, B, and C will lose their entire par value.
- C Losses will be realized by Bond Class C first, then by Bond Class B, and then by Bond Class A.

3 If one of Agnelli's customers defaults on its €60 million loan:

- A Bond Classes A, B, and C will realize losses of €20 million each.
- B Bond Class C will realize losses of €60 million, but Bond Classes A and B will be unaffected.
- C Bond Classes B and C will realize losses of €30 million each, but Bond Class A will be unaffected.

Solution to 1:

Bond Class A is issued by Agnelli Trust, an SPE that is bankruptcy remote from Agnelli. Thus, the investors who hold Agnelli's bonds and/or common shares have no legal claim on the cash flows from the securitized receivables that are the collateral for the ABS. As long as Agnelli's customers make the interest payments and/or principal repayments on their loans, Agnelli Trust will be able to make cash payments to the ABS investors. Because of the credit tranching, even if some of Agnelli's customers were to default on their loans, the losses would be realized by the subordinated Bond Classes B and C before any losses are realized by the senior Bond Class A. The credit risk associated with Bond Class A is, therefore, lower than that of Bond Classes B and C and the corporate bonds, justifying the investment-grade credit rating.

Solution to 2:

A is correct. The ABS have been issued by Agnelli Trust, an SPE that is bankruptcy remote from Agnelli. If the securitization is viewed as resulting in a true sale, the fact that Agnelli files for bankruptcy does not affect the ABS holders. These ABS holders face credit risk only to the extent that Agnelli's customers who bought the machine tools do not make the obligatory payments on their loans. As long as the customers continue to make payments, all three bond classes will receive their expected cash flows. These cash flows are completely and legally independent of anything that happens to Agnelli itself.

Solution to 3:

B is correct. The rules for the distribution of losses are as follows. All losses on the collateral are absorbed by Bond Class C before any losses are realized by Bond Class B and then Bond Class A. Consequently, if the losses on the collateral are €60 million, which is the par value of Bond Class C, Bond Class C loses its entire par value, but Bond Classes A and B are unaffected.

4

RESIDENTIAL MORTGAGE LOANS

d describe types and characteristics of residential mortgage loans that are typically securitized

Before describing the various types of residential mortgage-backed securities, this section briefly discusses the fundamental features of the underlying assets: residential mortgage loans. The mortgage designs described in this section are those that are typically securitized.

A **mortgage loan**, or simply mortgage, is a loan secured by the collateral of some specified real estate property that obliges the borrower (often someone wishing to buy a home) to make a predetermined series of payments to the lender (often initially a bank or mortgage company). The mortgage gives the lender the right to foreclose on the loan if the borrower defaults; that is, a **foreclosure** allows the lender to take possession of the mortgaged property and then sell it in order to recover funds toward satisfying the debt obligation.

Typically, the amount of the loan advanced to buy the property is less than the property's purchase price. The borrower makes a down payment, and the amount borrowed is the difference between the property's purchase price and the down payment. When the loan is first taken out, the borrower's equity in the property is equal to the down payment. Over time, as the market value of the property changes, the borrower's equity also changes. It also changes as the borrower makes mortgage payments that include principal repayment.

The ratio of the amount of the mortgage to the property's value is called the **loan-to-value ratio** (LTV). The lower the LTV, the higher the borrower's equity. From the lender's perspective, the higher the borrower's equity, the less likely the borrower is to default. Moreover, the lower the LTV, the more protection the lender has for recovering the amount loaned if the borrower does default and the lender repossesses and sells the property.

In the United States, market participants typically identify two types of mortgages based on the credit quality of the borrower: prime loans and subprime loans. Prime loans generally have borrowers of high credit quality with strong employment and credit histories, income sufficient to pay the loan obligation, and substantial equity in the underlying property. Subprime loans have borrowers with lower credit quality and/or are loans without a first lien on the property (that is, another party has a prior claim on the underlying property).

Throughout the world, there are a considerable number of mortgage designs. Mortgage design means the specification of (1) the maturity of the loan, (2) how the interest rate is determined, (3) how the principal is to be repaid (that is, the amortization schedule), (4) whether the borrower has the option to prepay and, in such cases, whether any prepayment penalties might be imposed, and (5) the rights of the lender in a foreclosure.

4.1 Maturity

In the United States, the typical term or number of years to maturity of a mortgage ranges from 15 to 30 years. For most European countries, a residential mortgage typically has a maturity between 20 and 40 years, whereas in France and Spain, it can be as long as 50 years. Japan is an extreme case, with mortgage maturities of 100 years.

4.2 Interest Rate Determination

The interest rate on a mortgage is called the **mortgage rate**, **contract rate**, or **note rate**. Determination of mortgage rates varies considerably among countries but may be categorized in four basic ways:

- *Fixed rate:* The mortgage rate remains the same during the life of the mortgage. The United States and France have a high proportion of this type of interest rate determination. Although fixed-rate mortgages are not the dominant form in Germany, they do exist there.
- *Adjustable or variable rate:* The mortgage rate is reset periodically—daily, weekly, monthly, or annually. The determination of the new mortgage rate for an adjustable-rate mortgage (ARM) at the reset date may be based on some reference rate or index (in which case, it is called an indexed-referenced ARM) or a rate determined at the lender's discretion (in which case, it is called a reviewable ARM). Residential mortgages in Australia, Ireland, South Korea, Spain, and the United Kingdom are dominated by adjustable-rate mortgages. In Australia, Ireland, and the United Kingdom, the reviewable ARM is standard. In South Korea and Spain, the indexed-referenced ARM is the norm. Canada and the United States have ARMs that are typically tied to an index or reference rate, although this type of mortgage rate is not the dominant form of interest rate determination. An important feature of an ARM is that it will usually have a maximum interest rate by which the mortgage rate can change at a reset date and a maximum interest rate that the mortgage rate can reach during the mortgage's life.
- *Initial period fixed rate:* The mortgage rate is fixed for some initial period and is then adjusted. The adjustment may call for a new fixed rate or for a variable rate. When the adjustment calls for a fixed rate, the mortgage is referred to as a rollover or renegotiable mortgage. This mortgage design is dominant in Canada, Denmark, Germany, the Netherlands, and Switzerland. When the mortgage

starts out with a fixed rate and then switches to an adjustable rate after a specified initial term, the mortgage is referred to as a hybrid mortgage. Hybrid mortgages are popular in the United Kingdom.

- *Convertible*: The mortgage rate is initially either a fixed rate or an adjustable rate. At some point, the borrower has the option to convert the mortgage to a fixed rate or an adjustable rate for the remainder of the mortgage's life. Almost half of the mortgages in Japan are convertible.

4.3 Amortization Schedule

In most countries, residential mortgages are **amortizing loans**. The amortization of a loan means the gradual scheduled reduction of the principal over time. Assuming no prepayments are made by the borrower, the periodic mortgage payments made by the borrower consist of interest payments and scheduled principal repayments. As discussed in a previous reading, there are two types of amortizing loans: fully amortizing loans and partially amortizing loans. In a fully amortizing loan, the sum of all the scheduled principal repayments during the mortgage's life is such that when the last mortgage payment is made, the loan is fully repaid. Most residential mortgages in the United States are fully amortizing loans. In a partially amortizing loan, the sum of all the scheduled principal repayments is less than the amount borrowed. The final payment then includes the unpaid mortgage balance, sometimes referred to as a "balloon" payment.

If no scheduled principal repayment is specified for a certain number of years, the loan is said to be an **interest-only mortgage**. Interest-only mortgages have been available in Australia, Denmark, Finland, France, Germany, Greece, Ireland, the Netherlands, Portugal, South Korea, Spain, Switzerland, and the United Kingdom. Interest-only mortgages also have been available to a limited extent in the United States. A special type of interest-only mortgage is one in which there are no scheduled principal repayments over the entire life of the loan. In this case, the balloon payment is equal to the original loan amount. These mortgages, referred to as "interest-only lifetime mortgages" or "bullet mortgages," have been available in Denmark, the Netherlands, and the United Kingdom.

4.4 Prepayment Options and Prepayment Penalties

A prepayment is any payment of principal that exceeds the scheduled principal repayment. A mortgage may entitle the borrower to prepay all or part of the outstanding mortgage principal prior to the scheduled due date. This contractual provision is referred to as a **prepayment option** or an **early repayment option**. From the lender's or investor's viewpoint, the effect of a prepayment option is that the cash flow amounts and timing from a mortgage cannot be known with certainty. This risk was referred to earlier as prepayment risk. Prepayment risk affects all mortgages that allow prepayment, not just the level-payment, fixed-rate, fully amortizing mortgages.

The mortgage may stipulate a monetary penalty when a borrower prepays within a certain time period after the mortgage is originated, which may extend for the full life of the loan. Such mortgage designs are referred to as **prepayment penalty mortgages**. The purpose of the prepayment penalty is to compensate the lender for the difference between the contract rate and the prevailing mortgage rate if the borrower prepays when interest rates decline. Hence, the prepayment penalty is effectively a mechanism that provides for yield maintenance for the lender. The method for calculating the penalty varies. Prepayment penalty mortgages are common in Europe. Although the proportion of prepayment penalty mortgages in the United States is small, they do exist.

4.5 Rights of the Lender in a Foreclosure

A mortgage can be a recourse loan or a non-recourse loan. When the borrower fails to make the contractual loan payments, the lender can repossess the property and sell it, but the proceeds received from the sale of the property may be insufficient to recoup the losses. In a **recourse loan**, the lender has a claim against the borrower for the shortfall between the amount of the outstanding mortgage balance and the proceeds received from the sale of the property. In a **non-recourse loan**, the lender does not have such a claim and thus can look only to the property to recover the outstanding mortgage balance. In the United States, recourse is typically determined by the state, and residential mortgages are non-recourse loans in many states. In contrast, residential mortgages in most European countries are recourse loans.

The recourse/non-recourse feature of a mortgage has implications for projecting the likelihood of defaults by borrowers, particularly in the case of what is sometimes called "underwater mortgages"—that is, mortgages for which the value of the property has declined below the amount owed by the borrower. For example, in the United States, where mortgages are typically non-recourse, the borrower may have an incentive to default on an underwater mortgage and allow the lender to foreclose on the property, even if resources are available to continue to make mortgage payments. This type of default by a borrower is referred to as a "strategic default." A strategic default, however, has negative consequences for the borrower, who will then have a lower credit score and a reduced ability to borrow in the future. Thus, not all borrowers faced with underwater mortgages will default. In countries where residential mortgages are recourse loans, a strategic default is less likely because the lender can seek to recover the shortfall from the borrower's other assets and/or income.

Now that the basics of residential mortgages have been set out, we can turn our attention to how these mortgages are securitized—that is, transformed into MBS. In the following sections, we focus on the US residential mortgage sector because it is the largest in the world and many non-US investors hold US MBS in their portfolios.

EXAMPLE 4

Residential Mortgage Designs

1 In an interest-only mortgage, the borrower:

A does not have to repay the principal as long as she pays the interest.

B does not have to make principal repayments for a certain number of years, after which she starts paying down the original loan amount.

C does not have to make principal repayments over the entire life of the mortgage and pays down the original loan amount as a balloon payment.

2 A bank advertises a mortgage with the following interest rate: 2.99% (12-month Euribor + 2.50%), resetting once a year. The mortgage is *most likely*:

A a hybrid mortgage.

B an adjustable-rate mortgage.

C an initial period fixed-rate mortgage.

3 If the borrower fails to make the contractual mortgage payments on a non-recourse mortgage, the lender:

A cannot foreclose the property.

B can recover the outstanding mortgage balance only through the sale of the property.

C can recover the outstanding mortgage balance through the sale of the property and the borrower's other assets and/or income.

Solution to 1:

B is correct. In an interest-only mortgage, there is no scheduled principal repayment for a certain number of years, so the borrower starts paying down the original loan amount only after an initial period of interest-only payments. Some, but not all, interest-only mortgages do not have scheduled principal repayments over the entire life of the loan. These mortgages are called interest-only lifetime mortgages or bullet mortgages, and they require the borrower to pay back the original loan amount at maturity.

Solution to 2:

B is correct. An adjustable-rate mortgage is one for which the mortgage rate is typically based on some reference rate or index (indexed-referenced ARM) or a rate determined at the lender's discretion (reviewable ARM) and is reset periodically. A mortgage rate of 12-month Euribor + 2.50%, resetting once per year, suggests that the mortgage is an indexed-referenced ARM. The 2.99% rate is the current mortgage rate (that is, 12-month Euribor of 0.49% + 2.50%) and should not be taken as an indication that it is a fixed-rate, initial period fixed-rate, or hybrid mortgage.

Solution to 3:

B is correct. In the case of a non-recourse mortgage, the lender can foreclose the property if the borrower fails to make the contractual mortgage payments. However, the lender can use only the proceeds from the property to recover the outstanding mortgage balance.

5

MORTGAGE PASS-THROUGH SECURITIES

e **describe types and characteristics of residential mortgage-backed securities, including mortgage pass-through securities and collateralized mortgage obligations, and explain the cash flows and risks for each type**

f **define prepayment risk and describe the prepayment risk of mortgage-backed securities**

The bonds created from the securitization of mortgages related to the purchase of residential properties are residential mortgage-backed securities (RMBS). In such countries as the United States, Canada, Japan, and South Korea, a distinction is often made between securities that are guaranteed by the government or a quasi-government entity and securities that are not. Quasi-government entities are usually created by governments to perform various functions for them. Examples of quasi-government entities include government-sponsored enterprises (GSEs) such as Fannie Mae (previously the Federal National Mortgage Association) and Freddie Mac (previously the Federal Home Loan Mortgage Corporation) in the United States and the Japan Housing Finance Agency (JHF).

In the United States, securities backed by residential mortgages are divided into three sectors: (1) those guaranteed by a federal agency, (2) those guaranteed by a GSE, and (3) those issued by private entities and that are not guaranteed by a federal agency or a GSE. The first two sectors are referred to as **agency RMBS**, and the third sector is referred to as **non-agency RMBS**. We devote significant attention to US agency and non-agency RMBS because these securities represent a large sector of the investment-grade bond market and are included in the portfolios of many global investors.

Agency RMBS include securities issued by federal agencies, such as the Government National Mortgage Association, popularly referred to as Ginnie Mae. This entity is a federally related institution because it is part of the US Department of Housing and Urban Development. As a result, the RMBS that it guarantees carry the full faith and credit of the US government with respect to timely payment of interest and repayment of principal.

Agency RMBS also include RMBS issued by GSEs, such as Fannie Mae and Freddie Mac. RMBS issued by GSEs do not carry the full faith and credit of the US government. Agency RMBS issued by GSEs differ from non-agency RMBS in two ways. First, the credit risk of the RMBS issued by Fannie Mae and Freddie Mac is reduced by the guarantee of the GSE itself, which charges a fee for insuring the issue. In contrast, non-agency RMBS use credit enhancements to reduce credit risk. The pool of securitized loans is another way in which RMBS issued by GSEs differ from non-agency RMBS. Loans included in agency RMBS must meet specific underwriting standards established by various government agencies. These standards set forth the maximum size of the loan, the loan documentation required, the maximum loan-to-value ratio, and whether insurance is required. Loans satisfying the underwriting standards for inclusion as collateral for an agency RMBS are called "conforming" mortgages; otherwise, loans are categorized as "non-conforming" mortgages.

This section starts with a discussion of agency RMBS, which include mortgage pass-through securities and collateralized mortgage obligations. Note that the popular Bloomberg Barclays US Aggregate Bond Index includes only agency RMBS that are mortgage pass-through securities in its mortgage sector definition. We then discuss non-agency RMBS.

5.1 Mortgage Pass-Through Securities

A **mortgage pass-through security** is a security created when one or more holders of mortgages form a pool of mortgages and sell shares or participation certificates in the pool. A pool can consist of several thousand or only a few mortgages. When a mortgage is used as collateral for a mortgage pass-through security, the mortgage is said to be securitized.

5.1.1 *Characteristics*

The cash flows of a mortgage pass-through security depend on the cash flows of the underlying pool of mortgages. The cash flows consist of monthly mortgage payments representing interest, the scheduled repayment of principal, and any prepayments. Cash payments are made to security holders each month. Neither the amount nor the timing of the cash flows from the pool of mortgages, however, is necessarily identical to that of the cash flow passed through to the security holders. In fact, the monthly cash flows of a mortgage pass-through security are less than the monthly cash flows of the underlying pool of mortgages by an amount equal to the servicing and other administrative fees.

The servicing fee is the charge related to administrative tasks, such as collecting monthly payments from borrowers, forwarding proceeds to owners of the loan, sending payment notices to borrowers, reminding borrowers when payments are overdue, maintaining records of the outstanding mortgage balance, initiating foreclosure proceedings if necessary, and providing tax information to borrowers when applicable. The servicing fee is typically a portion of the mortgage rate. The other administrative fees are those charged by the issuer or financial guarantor of the mortgage pass-through security for guaranteeing the issue.

A mortgage pass-through security's coupon rate is called the **pass-through rate**. The pass-through rate is lower than the mortgage rate on the underlying pool of mortgages by an amount equal to the servicing and other administrative fees. The pass-through rate that the investor receives is said to be "net interest" or "net coupon."

Not all mortgages included in a pool of securitized mortgages have the same mortgage rate and the same maturity. Consequently, for each mortgage pass-through security, a **weighted average coupon rate** (WAC) and a **weighted average maturity** (WAM) are determined. The WAC is calculated by weighting the mortgage rate of each mortgage in the pool by the percentage of the outstanding mortgage balance relative to the outstanding amount of all the mortgages in the pool. Similarly, the WAM is calculated by weighting the remaining number of months to maturity of each mortgage in the pool by the outstanding mortgage balance relative to the outstanding amount of all the mortgages in the pool. Example 5 illustrates the calculation of the WAC and WAM.

EXAMPLE 5

Weighted Average Coupon Rate and Weighted Average Maturity

Assume that a pool includes three mortgages with the following characteristics:

Mortgage	Outstanding Mortgage Balance (US$)	Coupon Rate (%)	Number of Months to Maturity
1	1,000	5.1	34
2	3,000	5.7	76
3	6,000	5.3	88

The outstanding amount of three mortgages is US$10,000. Thus, the weights of Mortgages 1, 2, and 3 are 10%, 30%, and 60 %, respectively.

The WAC is

$$10\% \times 5.1\% + 30\% \times 5.7\% + 60\% \times 5.3\% = 5.4\%.$$

The WAM is

$$10\% \times 34 + 30\% \times 76 + 60\% \times 88 = 79 \text{ months}.$$

5.1.2 *Prepayment Risk*

Mortgage pass-through security cash flows are uncertain because they depend on actual prepayments. As we noted earlier, this risk is called prepayment risk. Prepayment risk has two components: contraction risk and extension risk, both of which largely reflect changes in the general level of interest rates.

Contraction risk is the risk that when interest rates decline, actual prepayments will be higher than forecasted because homeowners will refinance at now-available lower interest rates. Thus, a security backed by mortgages will have a shorter maturity than was anticipated at the time of purchase. Holding a security whose maturity becomes shorter when interest rates decline has two adverse consequences for investors. First, investors must reinvest the proceeds at lower interest rates. Second, if the security is prepayable or callable, its price appreciation is not as great as that of an otherwise identical bond without a prepayment or call option.

In contrast, **extension risk** is the risk that when interest rates rise, prepayments will be lower than forecasted because homeowners are reluctant to give up the benefits of a contractual interest rate that now looks low. As a result, a security backed by mortgages will typically have a longer maturity than was anticipated at the time of purchase. From the investor's perspective, the value of the security has fallen because the higher interest rates translate into a lower price for the security, and the income investors receive (and can potentially reinvest) is typically limited to the interest payment and scheduled principal repayments.

5.1.3 *Prepayment Rate Measures*

In describing prepayments, market participants refer to the prepayment rate or prepayment speed. The two key prepayment rate measures are the single monthly mortality rate (SMM), a monthly measure, and its corresponding annualized rate, the conditional prepayment rate (CPR).

The SMM reflects the dollar amount of prepayment for the month as a fraction of the balance on the mortgage after accounting for the scheduled principal repayment for the month. It is calculated as follows:

$$\text{SMM} = \frac{\text{Prepayment for the month}}{\begin{pmatrix}\text{Beginning outstanding mortgage balance for the month} - \\ \text{Scheduled principal repayment for the month}\end{pmatrix}} \tag{1}$$

Note that the SMM is typically expressed as a percentage.

When market participants describe the assumed prepayment for a pool of residential mortgages, they refer to the annualized SMM, which is the CPR. A CPR of 6%, for example, means that approximately 6% of the outstanding mortgage balance at the beginning of the year is expected to be prepaid by the end of the year.

A key factor in the valuation of a mortgage pass-through security and other products derived from a pool of mortgages is forecasting the future prepayment rate. This task involves prepayment modeling. Prepayment modeling uses characteristics of the mortgage pool and other factors to develop a statistical model for forecasting future prepayments.

In the United States, market participants describe prepayment rates in terms of a prepayment pattern or benchmark over the life of a mortgage pool. This pattern is the Public Securities Association (PSA) prepayment benchmark, which is produced by the Securities Industry and Financial Markets Association (SIFMA). The PSA prepayment benchmark is expressed as a series of monthly prepayment rates. Based on historical patterns, it assumes that prepayment rates are low for newly originated mortgages and then speed up as the mortgages become seasoned. Slower or faster prepayment rates are then referred to as some percentage of the PSA prepayment benchmark. Rather than going into the details of the PSA prepayment benchmark, this discussion will rely on some PSA assumptions. What is important to remember is that the standard for the PSA model is 100 PSA; that is, at 100 PSA, investors can expect prepayments to follow the PSA prepayment benchmark—for example, an increase of prepayment rates of 0.20% for the first 30 months until they peak at 6% in Month 30. A PSA assumption greater than 100 PSA means that prepayments are assumed to be faster than the standard model. In contrast, a PSA assumption lower than 100 PSA means that prepayments are assumed to be slower than the standard model.

5.1.4 *Cash Flow Construction*

Let us see how to construct the monthly cash flow for a hypothetical mortgage pass-through security. We assume the following:

- The underlying pool of mortgages has a par value of US$800 million.
- The mortgages are fixed-rate, level-payment, and fully amortizing loans.
- The WAC for the mortgages in the pool is 6%.
- The WAM for the mortgages in the pool is 357 months.
- The pass-through rate is 5.5%.

Exhibit 2 shows the cash flows to the mortgage pass-through security holders for selected months assuming a prepayment rate of 165 PSA. The SMM is given in in Column 3, and mortgage payments are given in Column 4. The net interest payment in Column 5 is the amount available to pay security holders after servicing and other administrative fees. This amount is equal to the beginning outstanding mortgage balance in Column 2 multiplied by the pass-through rate of 5.5% and then divided by 12. The scheduled principal repayment in Column 6 is the difference between the mortgage payment in Column 4 and the gross interest payment. The gross interest payment is equal to the beginning outstanding mortgage balance in Column 2 multiplied by the WAC of 6% and then divided by 12. The prepayment in Column 7 is calculated by applying Equation 1, using the SMM provided in Column 3, the beginning outstanding mortgage balance in Column 2, and the scheduled principal repayment in Column 6. The total principal repayment in Column 8 is the sum of the scheduled principal repayment in Column 6 and the prepayments in Column 7. Subtracting this amount from the beginning outstanding mortgage balance for the month gives the beginning outstanding mortgage balance for the following month. Finally, the projected cash flow for this mortgage pass-through security in Column 9 is the sum of the net interest payment in Column 5 and the total principal repayment in Column 8.

Exhibit 2 Monthly Cash Flow to Bondholders for a US$800 Million Mortgage Pass-Through Security with a WAC of 6.0%, a WAM of 357 Months, and a Pass-Through Rate of 5.5%, Assuming a Prepayment Rate of 165 PSA

(1) Month	(2) Beginning Outstanding Mortgage Balance (US$)	(3) SMM (%)	(4) Mortgage Payment (US$)	(5) Net Interest Payment (US$)	(6) Scheduled Principal Repayment (US$)	(7) Prepayment (US$)	(8) Total Principal Repayment (US$)	(9) Projected Cash Flow (US$)
1	800,000,000	0.111	4,810,844	3,666,667	810,844	884,472	1,695,316	5,361,982
2	798,304,684	0.139	4,805,520	3,658,896	813,996	1,104,931	1,918,927	5,577,823
3	796,385,757	0.167	4,798,862	3,650,101	816,933	1,324,754	2,141,687	5,791,788
⋮								
29	674,744,235	0.865	4,184,747	3,092,578	811,026	5,829,438	6,640,464	9,733,042
30	668,103,771	0.865	4,148,550	3,062,142	808,031	5,772,024	6,580,055	9,642,198
⋮								
100	326,937,929	0.865	2,258,348	1,498,466	623,659	2,822,577	3,446,236	4,944,702
101	323,491,693	0.865	2,238,814	1,482,670	621,355	2,792,788	3,414,143	4,896,814
⋮								
200	103,307,518	0.865	947,322	473,493	430,784	889,871	1,320,655	1,794,148

Exhibit 2 (Continued)

(1) Month	(2) Beginning Outstanding Mortgage Balance (US$)	(3) SMM (%)	(4) Mortgage Payment (US$)	(5) Net Interest Payment (US$)	(6) Scheduled Principal Repayment (US$)	(7) Prepayment (US$)	(8) Total Principal Repayment (US$)	(9) Projected Cash Flow (US$)
201	101,986,863	0.865	939,128	467,440	429,193	878,461	1,307,654	1,775,094
⋮								
300	19,963,930	0.865	397,378	91,501	297,559	170,112	467,670	559,172
301	19,496,260	0.865	393,941	89,358	296,460	166,076	462,536	551,893
⋮								
356	484,954	0.865	244,298	2,223	241,873	2,103	243,976	246,199
357	240,978	0.865	242,185	1,104	240,980	0	240,980	242,084

Notes: The SMM in Column 3 is rounded, which results in some rounding error in the calculation of the prepayments in Column 7 and, thus, of the total principal repayments and the projected cash flows in Columns 8 and 9, respectively. Since the WAM is 357 months, the underlying mortgage pool is seasoned an average of three months, and therefore, based on a 165 PSA, the CPR is 0.132% in Month 1 (seasoned Month 4), and the pool seasons at 6% in Month 27.

5.1.5 *Weighted Average Life*

A standard practice in the bond market is to refer to the maturity of a bond. This practice is not followed for MBS because principal repayments (scheduled principal repayments and prepayments) are made over the life of the security. Although an MBS has a "legal maturity," which is the date when the last scheduled principal repayment is due, the legal maturity does not reveal much about the actual principal repayments and the interest rate risk associated with the MBS. For example, a 30-year, option-free, corporate bond and an MBS with a 30-year legal maturity with the same coupon rate are not equivalent in terms of interest rate risk. Effective duration can be calculated for both the corporate bond and the MBS to assess the sensitivity of the securities to interest rate movements. But a measure widely used by market participants for MBS is the **weighted average life**, or simply the **average life** of the MBS. This measure gives investors an indication of how long they can expect to hold the MBS before it is paid off assuming interest rates stay at current levels and, thus, expected prepayments are realized. In other words, the average life of the MBS is the convention-based average time to receipt of all the projected principal repayments (scheduled principal repayments and projected prepayments).

A mortgage pass-through security's average life depends on the prepayment assumption, as illustrated in the following table. This table provides the average life of the mortgage pass-through security used in Exhibit 2 for various prepayment rates. Note that at the assumed prepayment rate of 165 PSA, the mortgage pass-through security has an average life of 8.6 years. The average life extends when the prepayment rate goes down and contracts rapidly as the prepayment rate goes up. So, at a prepayment rate of 600 PSA, the average life of the mortgage pass-through security is only 3.2 years.

PSA assumption	100	125	165	250	400	600
Average life (years)	11.2	10.1	8.6	6.4	4.5	3.2

ENVIRONMENTAL, SOCIAL, AND GOVERNANCE (ESG) RISKS FOR RMBS

Beyond traditional factors that affect RMBS prepayment risk, such as loan-to-value ratios and FICO scores of residential borrowers, rating agencies have more recently incorporated ESG factors into their assessment of these securities. In the first example, the potential loss of mortgage principal in the event of a natural disaster is incorporated into the ratings process for a new securitization, whereas in the second example, several long outstanding RMBS backed by subprime mortgages were downgraded because of a potential decline in interest income as a result of the adverse social impact of the COVID-19 pandemic.

Fitch Ratings Example

As the physical risk of climate change increases, Fitch Ratings has incorporated natural disaster risk into its loan loss expectations by factoring in the impact of potential weather-related damage on the value of residential properties that collateralize the mortgages underlying an RMBS transaction. Fitch Ratings has created ESG Relevance Scores that are integrated into the existing research process (they range from 1 to 5, with 5 representing ESG risk that is "highly relevant" to the credit decision and 1 representing ESG risk that is considered "irrelevant"). In the case of RMBS, Fitch assesses the geographic asset location and concentration within a given pool and compares it to regions that face greater climate risk. Fitch then assesses concentration penalties and adjusts stress assumptions for projected property losses from storm surges in cases of heightened environmental risk. For example, in 2019, Fitch published an ESG Relevance Score of 5 for an RMBS transaction with a large pool concentration in the US Gulf Coast region. Because this region is particularly susceptible to hurricane-related damage and over 75% of pool properties were concentrated in Louisiana and Texas, it faced a 1.16× probability of default adjustment with an increase in expected loan loss of over 100 bps due to the greater natural disaster and catastrophe risk in this pool versus other transactions.

Moody's Investors Service Example

In a 2020 report, Moody's Investor Service announced that ESG factors were a material credit consideration in one-third of its rating actions for private-sector issuers in 2019. The vast majority (88%) cited governance issues, followed by social (20%) and environmental (16%) issues, and many actions referenced more than one ESG consideration. The COVID-19 pandemic has wide-ranging political, economic, and social consequences that can adversely affect the creditworthiness of residential borrowers. For example, in June 2020, Moody's downgraded the ratings of over 40 bonds from 25 different US RMBS transactions. These securities, originally issued between 2004 and 2006 and backed by subprime mortgages, were largely downgraded to B1. In its rating action, Moody's noted the heightened risk of interest loss due to the 2020 slowdown in US economic activity and rising unemployment due to the coronavirus outbreak. Moody's cited unusually high downside risks to its forecast regarding the COVID-19 pandemic as a social risk under its ESG framework, with significant risks to public health and safety. Because many residential borrowers have been offered relief in the form of mortgage payment forbearance owing to the loss of income during the lockdown, these loans also face the possibility of modification in the future depending on the timing and pace of economic recovery.

EXAMPLE 6

Mortgage Pass-Through Securities

1 A non-conforming mortgage:

- **A** cannot be used as collateral in a mortgage-backed security.

  - **B** does not satisfy the underwriting standards for inclusion as collateral for an agency residential mortgage-backed security.
  - **C** does not give the lender a claim against the borrower for the shortfall between the amount of the outstanding mortgage balance and the proceeds from the sale of the property if the borrower defaults on the mortgage.

2 The monthly cash flows of a mortgage pass-through security *most likely*:

  - **A** are constant.
  - **B** change when interest rates decline.
  - **C** are equal to the cash flows of the underlying pool of mortgages.

3 A prepayment rate of 80 PSA means that investors can expect:

  - **A** 80% of the par value of the mortgage pass-through security to be repaid prior to the security's maturity.
  - **B** 80% of the borrowers whose mortgages are included in the collateral backing the mortgage pass-through security to prepay their mortgages.
  - **C** the prepayment rate of the mortgages included in the collateral backing the mortgage pass-through security to be 80% of the monthly prepayment rates forecasted by the PSA model.

4 All else being equal, when interest rates decline:

  - **A** investors in mortgage pass-through securities face extension risk.
  - **B** the weighted average maturity of a mortgage pass-through security lengthens.
  - **C** the increase in the price of a mortgage pass-through security is less than the increase in the price of an otherwise identical bond with no prepayment option.

Solution to 1:

B is correct. A non-conforming mortgage is one that does not satisfy the underwriting standards for inclusion as collateral for an agency RMBS. The standards specify the maximum size of the loan, the loan documentation required, the maximum loan-to-value ratio, and whether or not insurance is required for the loans in the pool.

Solution to 2:

B is correct. The monthly cash flows of a mortgage pass-through security are not equal to but rather depend on the cash flows of the underlying pool of mortgages. That said, their amount and timing cannot be known with certainty because of prepayments. When interest rates decline, borrowers are likely to prepay all or part of their outstanding mortgage balance, which will affect the monthly cash flows of the mortgage pass-through security. Remember that the fees related to servicing and guaranteeing the mortgages reduce the monthly cash flows of a mortgage pass-through security relative to those of the underlying pool of mortgages.

Solution to 3:

C is correct. A prepayment rate of 80 PSA means that investors can expect the prepayment rate of the mortgages included in the collateral backing the mortgage pass-through security to be 80% of the monthly prepayment rates forecasted by the PSA model. For example, if the PSA model forecasts an increase in prepayment rates of 0.20% for the first 30 months until they peak at 6% in Month 30, 80 PSA would assume an increase in prepayment rates of 0.16% (80% × 0.20%) for the first 30 months until they peak at 4.80% (80% × 6%) in Month 30. Thus, investors can expect slower prepayments than the PSA prepayment benchmark.

Solution to 4:

C is correct. When interest rates decline, the prepayment rate on a mortgage pass-through security goes up because homeowners refinance at now-available lower interest rates. As a result, investors face contraction risk; that is, they receive payments faster than anticipated. Investors who decide to retain the security face the prospect of having to reinvest those payments at relatively low interest rates. Investors who decide to sell the security would have to do so at a price lower than that of an otherwise identical bond with no prepayment option and thus no prepayment risk.

6 COLLATERALIZED MORTGAGE OBLIGATIONS AND NON-AGENCY RMBS

e describe types and characteristics of residential mortgage-backed securities, including mortgage pass-through securities and collateralized mortgage obligations, and explain the cash flows and risks for each type

As noted in the previous section, prepayment risk is an important consideration when investing in mortgage pass-through securities. Some institutional investors are concerned with extension risk, and others, with contraction risk. The structuring of a securitization can help redistribute the cash flows of mortgage-related products (mortgage pass-through securities or pools of loans) to different bond classes or tranches, which leads to the creation of securities that have different exposures to prepayment risk and thus different risk–return patterns relative to the mortgage-related product from which they were created.

When the cash flows of mortgage-related products are redistributed to various tranches, the resulting securities are called **collateralized mortgage obligations** (CMOs). The mortgage-related products from which the cash flows are obtained are considered the collateral. Note that in contrast to a mortgage pass-through security, the collateral is not a pool of mortgages but a mortgage pass-through security. In fact, in practice, the collateral is usually a pool of mortgage pass-through securities—hence the reason market participants sometimes use the terms "collateral" and "mortgage pass-through securities" interchangeably.

The creation of a CMO cannot eliminate or change prepayment risk; it can only distribute the various forms of this risk among different bond classes. The CMO's major financial innovation is that securities can be created to closely satisfy the asset/liability needs of institutional investors, thereby broadening the appeal of mortgage-backed products.

A wide range of CMO structures exists. The major ones are reviewed in the following subsections.

6.1 Sequential-Pay CMO Structures

The first CMOs were structured so that each tranche would be retired sequentially. Such structures are called "sequential-pay CMOs." The rule for the monthly distribution of the principal repayments (scheduled principal repayment plus prepayments) to the tranches in this structure is as follows. First, distribute all principal payments to Tranche 1 until the principal balance for Tranche 1 is zero. After Tranche 1 is paid off, distribute all principal payments to Tranche 2 until the principal balance for Tranche 2 is zero, and so on.

To illustrate a sequential-pay CMO, let us use a hypothetical transaction called CMO-01. Assume that the collateral for CMO-01 is the mortgage pass-through security described in Exhibit 2. Recall that the total par value of the collateral is US$800 million, the pass-through coupon rate is 5.5%, the WAC is 6%, and the WAM is 357 months. From this US$800 million of collateral, four tranches are created, as shown in Exhibit 3. In this simple structure, the coupon rate is the same for each tranche and the same as the mortgage pass-through security's coupon rate. This feature is for simplicity; typically, the coupon rate varies by tranche based on the term structure of interest rates, among other factors.

Exhibit 3 CMO-01: Sequential-Pay CMO Structure with Four Tranches

Tranche	Par Amount (US$ millions)	Coupon Rate (%)
A	389	5.5
B	72	5.5
C	193	5.5
D	146	5.5
Total	800	

Payment rules: *For payment of monthly coupon interest:* Disburse monthly coupon interest to each tranche on the basis of the amount of principal outstanding for each tranche at the beginning of the month. *For disbursement of principal payments:* Disburse principal payments to Tranche A until it is completely paid off. After Tranche A is completely paid off, disburse principal payments to Tranche B until it is completely paid off. After Tranche B is completely paid off, disburse principal payments to Tranche C until it is completely paid off. After Tranche C is completely paid off, disburse principal payments to Tranche D until it is completely paid off.

Remember that a CMO is created by redistributing the cash flows—interest payments and principal repayments—to the various tranches based on a set of payment rules. The payment rules at the bottom of Exhibit 3 describe how the cash flows from the mortgage pass-through security are to be distributed to the four tranches. CMO-01 has separate rules for the interest payment and the principal repayment, the latter being the sum of the scheduled principal repayment and the prepayments.

Although the payment rules for the distribution of the principal repayments are known, the precise principal repayment amount in each month is not. This amount will depend on the cash flow of the collateral, which depends on the actual prepayment rate of the collateral. The assumed prepayment rate (165 PSA in Exhibit 2) allows determining only the projected, not the actual, cash flow.

Consider what has been accomplished by creating the sequential-pay CMO-01 structure. Earlier, we saw that with a prepayment rate of 165 PSA, the mortgage pass-through security's average life was 8.6 years. Exhibit 4 reports the average life of the collateral and the four tranches assuming various actual prepayment rates. Note that the four tranches have average lives that are shorter or longer than the collateral, thereby attracting investors who have preferences for different average lives. For example, a pension fund that needs cash only after a few years because it expects a significant increase in the number of retirements after that time may opt for a tranche with a longer average life.

Exhibit 4 Average Life of the Collateral and the Four Tranches of CMO-01 for Various Actual Prepayment Rates

Prepayment Rate	Average Life (years)				
	Collateral	Tranche A	Tranche B	Tranche C	Tranche D
100 PSA	11.2	4.7	10.4	15.1	24.0
125 PSA	10.1	4.1	8.9	13.2	22.4
165 PSA	8.6	3.4	7.3	10.9	19.8
250 PSA	6.4	2.7	5.3	7.9	15.2
400 PSA	4.5	2.0	3.8	5.3	10.3
600 PSA	3.2	1.6	2.8	3.8	7.0

A major problem that remains is the considerable variability of the average lives of the tranches. How this problem can be handled is shown in the next section, but at this point, note that some protection against prepayment risk is provided for each tranche. The protection arises because prioritizing the distribution of principal (that is, establishing the payment rule for the principal repayment) effectively protects the shorter-term tranche (A in this structure) against extension risk. This protection actually comes from the longer-term tranches. Similarly, Tranches C and D provide protection against extension risk for Tranches A and B. At the same time, Tranches C and D benefit because they are provided protection against contraction risk; this protection comes from Tranches A and B. Thus, the sequential-pay CMO-01 structure allows investors concerned about extension risk to invest in Tranches A or B and those concerned about contraction risk to invest in Tranches C or D.

6.2 CMO Structures Including Planned Amortization Class and Support Tranches

A common structure in CMOs is to include planned amortization class (PAC) tranches, which offer greater predictability of the cash flows if the prepayment rate is within a specified band over the collateral's life. Remember that the creation of an MBS, whether it is a mortgage pass-through or a CMO, cannot make prepayment risk disappear. So where does the reduction of prepayment risk (both extension risk and contraction risk) that PAC tranches offer investors come from? The answer is that it comes from the existence of non-PAC tranches, called **support tranches** or companion tranches. The structure of the CMO makes the support tranches absorb prepayment risk first. Because PAC tranches have limited (but not complete) protection against both extension risk and contraction risk, they are said to provide two-sided prepayment protection.

The greater predictability of the cash flows for the PAC tranches occurs because a principal repayment schedule must be satisfied. When the prepayment rate is within the specified band, called the PAC band, all prepayment risk is absorbed by the support tranche. If the collateral prepayments are slower than forecasted, the support tranches do not receive any principal repayment until the PAC tranches receive their scheduled principal repayment. This rule reduces the extension risk of the PAC tranches. Similarly, if the collateral prepayments are faster than forecasted, the support tranches absorb any principal repayments in excess of the scheduled principal repayments. This rule reduces the contraction risk of the PAC tranches. Even if the prepayment rate is outside the PAC band, prepayment risk is first absorbed by the support tranche. Thus, the key to the prepayment protection that PAC tranches offer investors is the amount of support tranches outstanding. If the support tranches are paid off quickly because of faster-than-expected prepayments, they no longer provide any protection for the PAC tranches.

Support tranches expose investors to the highest level of prepayment risk. Therefore, investors must be particularly careful in assessing the cash flow characteristics of support tranches in order to reduce the likelihood of adverse portfolio consequences resulting from prepayments.

To illustrate how to create CMO structures including PAC and support tranches, we use again the US$800 million mortgage pass-through security described in Exhibit 2, with a pass-through coupon rate of 5.5%, a WAC of 6%, and a WAM of 357 months as collateral. The creation of PAC tranches requires the specification of two PSA prepayment rates: a *lower* PSA prepayment assumption and an *upper* PSA prepayment assumption. The lower and upper PSA prepayment assumptions are called the "initial PAC collar" or the "initial PAC band." The PAC collar for a CMO is typically dictated by market conditions. In our example, we assume that the lower and upper PSA prepayment assumptions are 100 PSA and 250 PSA, respectively, so the initial PAC collar is 100–250 PSA.

Exhibit 5 shows a CMO structure called CMO-02 that contains only two tranches: a 5.5% coupon PAC tranche created assuming an initial PAC collar of 100–250 PSA and a support tranche.

Exhibit 5 CMO-02: CMO Structure with One PAC Tranche and One Support Tranche

Tranche	Par Amount (US$ millions)	Coupon Rate (%)
P (PAC)	487.6	5.5
S (support)	312.4	5.5
Total	800.0	

Payment rules: *For payment of monthly coupon interest:* Disburse monthly coupon interest to each tranche based on the amount of principal outstanding for each tranche at the beginning of the month. *For disbursement of principal payments:* Disburse principal payments to Tranche P based on its schedule of principal repayments. Tranche P has priority with respect to current and future principal payments to satisfy the schedule. Any excess principal payments in a month over the amount necessary to satisfy the schedule for Tranche P are paid to Tranche S. When Tranche S is completely paid off, all principal payments are to be made to Tranche P regardless of the schedule.

Exhibit 6 reports the average life of the PAC and support tranches in CMO-02 assuming various actual prepayment rates. Note that between 100 PSA and 250 PSA, the average life of the PAC tranche is constant at 7.7 years. At slower or faster PSA rates, however, the schedule is broken and the average life changes—extending when the prepayment rate is less than 100 PSA and contracting when it is greater than 250 PSA. Even so, there is much less variability for the average life of the PAC tranche compared with that of the support tranche.

Exhibit 6 Average Life of the PAC Tranche and the Support Tranche of CMO-02 for Various Actual Prepayment Rates and an Initial PAC Collar of 100–250 PSA

	Average Life (years)	
Prepayment Rate	**PAC Tranche (P)**	**Support Tranche (S)**
50 PSA	10.2	24.9
75 PSA	8.6	22.7
100 PSA	7.7	20.0
165 PSA	7.7	10.7
250 PSA	7.7	3.3
400 PSA	5.5	1.9
600 PSA	4.0	1.4

Most CMO structures including PAC and support tranches have more than one PAC tranche. A sequence of six PAC tranches (that is, PAC tranches paid off in sequence as specified by a principal repayment schedule) is not uncommon. For example, consider CMO-03 in Exhibit 7, which contains four sequential PAC tranches (P-A, P-B, P-C, and P-D) and one support tranche. The total par amount of the PAC and support tranches is the same as for CMO-02 in Exhibit 5. The difference is that instead of one PAC tranche with a schedule, there are four PAC tranches with schedules. As described in the payment rules, the PAC tranches are paid off in sequence.

Exhibit 7 CMO-03: CMO Structure with Sequential PAC Tranches and One Support Tranche

Tranche	Par Amount (US$ million)	Coupon Rate (%)
P-A (PAC)	287.6	5.5
P-B (PAC)	90.0	5.5
P-C (PAC)	60.0	5.5
P-D (PAC)	50.0	5.5
S (support)	312.4	5.5
Total	800.0	

Payment rules: *For payment of monthly coupon interest:* Disburse monthly coupon interest to each tranche based on the amount of principal outstanding for each tranche at the beginning of the month. *For disbursement of principal payments:* Disburse principal payments to Tranche P-A based on its schedule of principal repayments. Tranche P-A has priority with respect to current and future principal payments to satisfy the schedule. Any excess principal payments in a month over the amount necessary to satisfy the schedule while P-A is outstanding is paid to Tranche S. Once P-A is paid off, disburse principal payments to Tranche P-B based on its schedule of principal repayments. Tranche P-B has priority with respect to current and future principal payments to satisfy the schedule. Any excess principal payments in a month over the amount necessary to satisfy the schedule while P-B is outstanding are paid to Tranche S. The same rule applies for P-C and P-D. When Tranche S is completely paid off, all principal payments are to be made to the outstanding PAC tranches regardless of the schedule.

6.3 Other CMO Structures

Often, there is a demand for tranches that have a floating rate. Although the collateral pays a fixed rate, it is possible to create a tranche with a floating rate. This is done by constructing a floater and an inverse floater combination from any of the fixed-rate tranches in the CMO structure. Because the floating-rate tranche pays a higher rate when interest rates go up and the inverse floater pays a lower rate when interest rates go up, they offset each other. Thus, a fixed-rate tranche can be used to satisfy the demand for a floating-rate tranche.

In a similar vein, other types of tranches to satisfy the various needs of investors are possible.

EXAMPLE 7

Collateralized Mortgage Obligations

1 A collateralized mortgage obligation:

  A eliminates prepayment risk.
  B is created from a pool of conforming loans.
  C redistributes various forms of prepayment risk among different bond classes.

2 The variability in the average life of the PAC tranche of a CMO relative to the average life of the mortgage pass-through securities from which the CMO is created is:

  A lower.
  B the same.
  C higher.

3 Referring to Exhibit 7, the tranche of CMO-03 that is *most suitable* for an investor concerned about contraction risk is:

  A P-A (PAC).
  B P-D (PAC).
  C S (support).

4 The tranche of a collateralized mortgage obligation that is *most suitable* for an investor who expects a fall in interest rates is:

  A a fixed-rate tranche.
  B an inverse floating-rate tranche.
  C a PAC tranche.

5 The investment that is *most suitable* for an investor who is willing and able to accept significant prepayment risk is:

  A a mortgage pass-through security.
  B the support tranche of a collateralized mortgage obligation.
  C the inverse floating-rate tranche of a collateralized mortgage obligation.

Solution to 1:

C is correct. CMOs are created by redistributing the cash flows of mortgage-related products, including mortgage pass-through securities, to different bond classes or tranches on the basis of a set of payment rules.

Solution to 2:

A is correct. The purpose of creating different bond classes in a CMO is to provide a risk–return profile that is more suitable to investors than the risk–return profile of the mortgage pass-through securities from which the CMO is created. The PAC tranche has considerably less variability in average life than the mortgage pass-through securities. In contrast, the support tranche has more variability in average life than the mortgage pass-through securities.

Solution to 3:

B is correct. Contraction risk is the risk that when interest rates decline, prepayments will be higher than expected and the security's maturity will become shorter than was anticipated at the time of purchase. PAC tranches offer investors protection against contraction risk (and extension risk). The PAC tranche that is most suitable for an investor concerned about contraction risk is P-D because it is the latest-payment PAC tranche; that is, any principal repayments in excess of the scheduled principal repayments are absorbed sequentially by the support tranche, then P-A, P-B, and, finally, P-D.

Solution to 4:

B is correct. The tranche of a CMO that is most suitable for an investor who expects a fall in interest rates is an inverse floating-rate tranche. The inverse floater pays a coupon rate that is inversely related to prevailing interest rates. Thus, if interest rates fall, the CMO's coupon rate will rise.

Solution to 5:

B is correct. The investment that is most suitable to an investor who is willing and able to accept significant prepayment risk is the support tranche of a collateralized mortgage obligation. Because the PAC tranche has a stable average life at prepayment rates within the PAC band, all prepayment risk is absorbed by the support tranche for prepayment rates within the band. Even at rates outside the PAC band, prepayment risk is first absorbed by the support tranche. Investors will be compensated for bearing prepayment risk in the sense that, if properly priced, the support tranche will have a higher expected rate of return than the PAC tranche.

6.4 Non-Agency Residential Mortgage-Backed Securities

Agency RMBS are those issued by Ginnie Mae, Fannie Mae, and Freddie Mac. RMBS issued by any other entity are non-agency RMBS. Entities that issue non-agency RMBS are typically thrift institutions, commercial banks, and private conduits. Private conduits may purchase non-conforming mortgages, pool them, and then sell mortgage pass-through securities whose collateral is the underlying pool of non-conforming mortgages. Because they are not guaranteed by the government or by a GSE, credit risk is an important consideration when investing in non-agency RMBS.

Non-agency RMBS share many features and structuring techniques with agency CMOs. However, because non-agency RMBS are not guaranteed by the US government or by a GSE that can provide protection against losses in the pool, some form of internal or external credit enhancement is necessary to make these securities attractive to investors. These credit enhancements allow investors to reduce credit risk or transfer credit risk between bond classes, thus enabling investors to choose the risk–return profile that best suits their needs. Credit enhancements also play an important role in obtaining favorable credit ratings, which make non-agency RMBS more marketable to investors. The level of credit enhancement is usually determined relative to a specific credit rating desired by the issuer for a security. Note that one of the consequences of the 2007–09 credit crisis has been an overall increase in the level of credit enhancement.

As mentioned earlier, subordination, or credit tranching, is a common form of credit enhancement. The subordination levels (that is, the amount of credit protection for a bond class) are set at the time of issuance. However, the subordination levels change over time, as voluntary prepayments and defaults occur. To protect investors in non-agency RMBS, a securitization is designed to keep the amount of credit enhancement from deteriorating over time. If the credit enhancement for senior tranches deteriorates because of poor performance of the collateral, a mechanism called the "shifting interest mechanism" locks out subordinated bond classes from receiving payments for a period. Many non-agency RMBS also include other credit enhancements, such as overcollateralization and reserve accounts.

When forecasting the future cash flows of non-agency RMBS, investors must consider two important components. The first is the assumed default rate for the collateral. The second is the recovery rate, because even though the collateral may default, not all of the outstanding mortgage balance may be lost. The repossession and subsequent sale of the recovered property may provide cash flows that will be available to pay bondholders. That amount is based on the assumed amount that will be recovered.

Now that we have discussed securities backed by a pool of residential mortgages, we turn to securities backed by a pool of commercial mortgages.

COMMERCIAL MORTGAGE-BACKED SECURITIES

7

g describe characteristics and risks of commercial mortgage-backed securities

Commercial mortgage-backed securities (CMBS) are backed by a pool of commercial mortgages on income-producing property, such as multifamily properties (e.g., apartment buildings), office buildings, industrial properties (including warehouses), shopping centers, hotels, and health care facilities (e.g., senior housing care facilities). The collateral is a pool of commercial loans that were originated either to finance a commercial purchase or to refinance a prior mortgage obligation.

7.1 Credit Risk

In the United States and other countries where commercial mortgages are non-recourse loans, the lender can look only to the income-producing property backing the loan for interest payments and principal repayments. If a default occurs, the lender can foreclose the commercial property but it can use only the proceeds from the sale of that property to recover the principal outstanding, and it has no recourse to the borrower's other assets and/or income for any unpaid balance. The lender must view each property individually, and lenders evaluate each property using measures that have been found useful in assessing credit risk.

Two key indicators of potential credit performance are the loan-to-value ratio (LTV), which was discussed earlier, and the debt-service-coverage (DSC) ratio, sometimes referred to as DSCR. The DSC ratio is equal to the property's annual net operating income (NOI) divided by the debt service (that is, the annual amount of interest payments and principal repayments). The NOI is defined as the rental income reduced by cash operating expenses and a non-cash replacement reserve reflecting the depreciation of the property over time. A DSC ratio that exceeds 1.0 indicates that the cash flows from the property are sufficient to cover the debt service while maintaining the property in its initial state of repair. The higher the DSC ratio, the more likely it is that the borrower will be able to meet debt-servicing requirements from the property's cash flows.

7.2 CMBS Structure

A credit-rating agency determines the level of credit enhancement necessary to achieve a desired credit rating. For example, if specific loan-to-value and DSC ratios are needed and those ratios cannot be met at the loan level, subordination is used to achieve the desired credit rating.

Interest on the principal outstanding is paid to all tranches. Losses arising from loan defaults are charged against the outstanding principal balance of the CMBS tranche with the lowest priority. This tranche may not be rated by credit-rating agencies; in this case, this unrated tranche is called the "first-loss piece," "residual tranche," or "equity tranche." The total loss charged includes the amount previously advanced and the actual loss incurred in the sale of the loan's underlying property.

Two characteristics that are usually specific to CMBS structures are the presence of call protection and a balloon maturity provision.

7.2.1 *Call Protection*

A critical investment feature that distinguishes CMBS from RMBS is the protection against early prepayments available to investors known as call protection. An investor in an RMBS is exposed to considerable prepayment risk because the borrower has the right to prepay a loan, in whole or in part, before the scheduled principal repayment date. As explained previously, a borrower in the United States usually does not pay any penalty for prepayment. The discussion of CMOs highlighted how investors can purchase certain types of tranches (e.g., sequential-pay and PAC tranches) to modify or reduce prepayment risk.

CMBS investors have considerable call protection, which results in CMBS trading more like corporate bonds than RMBS. The call protection comes either at the structure level or at the loan level. Structural call protection is achieved when CMBS are structured to have sequential-pay tranches, by credit rating. A lower-rated tranche cannot be paid down until the higher-rated tranche is completely retired, so the AAA rated bonds must be paid off before the AA rated bonds are, and so on. Principal losses resulting from defaults, however, are affected from the bottom of the structure upward.

At the loan level, four mechanisms offer investors call protection:

- A prepayment lockout, which is a contractual agreement that prohibits any prepayments during a specified period.
- Prepayment penalty points, which are predetermined penalties that a borrower who wants to refinance must pay to do so; a point is equal to 1% of the outstanding loan balance.
- A yield maintenance charge, also called a "make-whole charge," which is a penalty paid by the borrower that makes refinancing solely to get a lower mortgage rate uneconomical for the borrower. In its simplest terms, a yield maintenance charge is designed to make the lender indifferent as to the timing of prepayments.
- Defeasance, for which the borrower provides sufficient funds for the servicer to invest in a portfolio of government securities that replicates the cash flows that would exist in the absence of prepayments. The cash payments that must be met by the borrower are projected based on the terms of the loan. Then, a portfolio of government securities is constructed in such a way that the interest

payments and the principal repayments from the portfolio will be sufficient to pay off each obligation when it comes due. When the last obligation is paid off, the value of the portfolio is zero (that is, there are no funds remaining). The cost of assembling such a portfolio is the cost of defeasing the loan that must be repaid by the issuer.

7.2.2 *Balloon Maturity Provision*

Many commercial loans backing CMBS are balloon loans that require a substantial principal repayment at maturity of the loan. If the borrower fails to make the balloon payment, the borrower is in default. The lender may extend the loan over a period known as the "workout period." In doing so, the lender may modify the original terms of the loan and charge a higher interest rate, called the "default interest rate," during the workout period.

The risk that a borrower will not be able to make the balloon payment because either the borrower cannot arrange for refinancing or cannot sell the property to generate sufficient funds to pay off the outstanding principal balance is called "balloon risk." Because the life of the loan is extended by the lender during the workout period, balloon risk is a type of extension risk.

EXAMPLE 8

An Example of a Commercial Mortgage-Backed Security

The following information is taken from a filing with the US Securities and Exchange Commission about a CMBS issued by a special purpose entity established by a major US commercial bank. The collateral for this CMBS was a pool of 72 fixed-rate mortgages secured by first liens (first claims) on various types of commercial, multifamily, and manufactured housing community properties.

Classes of Offered Certificates	Initial Principal Amount (US$)	Initial Pass-Through Rate (%)
A-1	75,176,000	0.754
A-2	290,426,000	1.987
A-3	150,000,000	2.815
A-4	236,220,000	3.093
A-AB	92,911,000	2.690
X-A	948,816,000	1.937
A-S	104,083,000	3.422
B	75,423,000	3.732
C	42,236,000	

The filing included the following statements:

> If you acquire Class B certificates, then your rights to receive distributions of amounts collected or advanced on or in respect of the mortgage loans will be subordinated to those of the holders of the Class A-1, Class A-2, Class A-3, Class A-4, Class A-AB, Class X-A, and Class A-S certificates. If you acquire Class C certificates, then your rights to receive distributions of amounts collected or advanced on or in respect of the mortgage loans will be subordinated to those of the holders of the Class B certificates and all other classes of offered certificates.

> "Prepayment Penalty Description" or "Prepayment Provision" means the number of payments from the first due date through and including the maturity date for which a mortgage loan, as applicable, (i) is locked out from prepayment, (ii) provides for payment of a prepayment premium or yield maintenance charge in connection with a prepayment, (iii) or permits defeasance.

1 Based on the information provided, this CMBS:

A did not include any credit enhancement.

B included an internal credit enhancement.

C included an external credit enhancement.

2 Based on the information provided, investors in this CMBS had prepayment protection at:

A the loan level.

B the structure level.

C both the loan and structure levels.

3 Defeasance can be *best* described as:

A a predetermined penalty that a borrower who wants to refinance must pay to do so.

B a contractual agreement that prohibits any prepayments during a specified period of time.

C funds that the borrower must provide to replicate the cash flows that would exist in the absence of prepayments.

4 A risk that investors typically face when holding CMBS is:

A call risk.

B balloon risk.

C contraction risk.

5 The credit risk of a commercial mortgage-backed security is lower:

A the lower the DSC ratio and the lower the LTV.

B the lower the DSC ratio and the higher the LTV.

C the higher the DSC ratio and the lower the LTV.

Solution to 1:

B is correct. The CMBS included a senior/subordinated structure, which is a form of internal credit enhancement. Class B provided protection for all of the bond classes listed above it. Similarly, Class C provided protection for all other bond classes, including Class B; it was the first-loss piece, also called the residual tranche or equity tranche. Note that because it was the residual tranche, Class C had no specific pass-through rate. Investors in Class C will have priced it on the basis of some expected residual rate of return, but they could have done better or worse than expected depending on how interest rate movements and default rates affected the performance of the other tranches.

Solution to 2:

C is correct. This CMBS offered investors prepayment protection at both the structure and loan levels. The structural call protection was achieved thanks to the sequential-pay tranches. At the loan level, the CMBS included three of the four types of call protection—namely, a prepayment lockout, a yield maintenance charge, and defeasance.

Solution to 3:

C is correct. Defeasance is a call protection at the loan level that requires the borrower to provide sufficient funds for the servicer to invest in a portfolio of government securities that replicates the cash flows that would exist in the absence of prepayments.

Solution to 4:

B is correct. Because many commercial loans backing CMBS require a balloon payment, investors in CMBS typically face balloon risk—that is, the risk that if the borrower cannot arrange for refinancing or cannot sell the property to make the balloon payment, the CMBS may extend in maturity because the lender has to wait to obtain the outstanding principal until the borrower can make the balloon amount. Balloon risk is a type of extension risk.

Solution to 5:

C is correct. The DSC ratio and the LTV are key indicators of potential credit performance and thus allow investors to assess the credit risk of a CMBS. The DSC ratio is equal to the property's annual NOI divided by the annual amount of interest payments and principal repayments. So the higher the DSC ratio, the lower the CMBS's credit risk. The LTV is equal to the amount of the mortgage divided by the property's value. So the lower the LTV, the lower the CMBS's credit risk.

To this point, this reading has addressed the securitization of real estate property, both residential and commercial. We now discuss the securitization of debt obligations in which the underlying asset is not real estate.

NON-MORTGAGE ASSET-BACKED SECURITIES

8

h describe types and characteristics of non-mortgage asset-backed securities, including the cash flows and risks of each type

Numerous types of non-mortgage assets have been used as collateral in securitization. The largest non-mortgage assets in most countries are auto loan and lease receivables, credit card receivables, personal loans, and commercial loans. What is important to keep in mind is that, regardless of the type of asset, ABS that are not guaranteed by a government or a quasi-government entity are subject to credit risk.

ABS can be categorized based on the way the collateral repays—that is, whether the collateral is amortizing or non-amortizing. Traditional residential mortgages and auto loans are examples of amortizing loans. The cash flows for an amortizing loan include interest payments, scheduled principal repayments, and any prepayments, if permissible. If the loan has no schedule for paying down the principal, it is a non-amortizing loan. Because a non-amortizing loan does not involve scheduled principal repayments, an ABS backed by non-amortizing loans is not affected by prepayment risk. Credit card receivable ABS are an example of ABS backed by non-amortizing loans.

Consider an ABS backed by a pool of 1,000 amortizing loans with a total par value of US$100 million. Over time, some of the loans will be paid off; the amounts received from the scheduled principal repayment and any prepayments will be distributed to the bond classes based on the payment rule. Consequently, over time, the number of loans in the collateral will drop from 1,000 and the total par value will fall to less than US$100 million.

Now, what happens if the collateral of the ABS is 1,000 non-amortizing loans? Some of these loans will be paid off in whole or in part before the maturity of the ABS. When those loans are paid off, what happens depends on whether the loans were paid off during or after the lockout period. The lockout or revolving period is the period during which the principal repaid is reinvested to acquire additional loans with a principal equal to the principal repaid. The reinvestment in new loans can result in the collateral including more or less than 1,000 loans, but the loans will still have a total par value of US$100 million. When the lockout period is over, any principal that is repaid will not be used to reinvest in new loans but will instead be distributed to the bond classes.

This reading cannot cover all types of non-mortgage ABS. It focuses on the two popular non-mortgage ABS in most countries: auto loan ABS and credit card receivable ABS.

8.1 Auto Loan ABS

Auto loan ABS are backed by auto loans and lease receivables. The focus in this section is on the largest type of auto securitizations—that is, auto loan-backed securities. In some countries, auto loan-backed securities represent the largest or second largest sector of the securitization market.

The cash flows for auto loan-backed securities consist of scheduled monthly payments (that is, interest payments and scheduled principal repayments) and any prepayments. For securities backed by auto loans, prepayments result from sales and trade-ins requiring full payoff of the loan, repossession and subsequent resale of autos, insurance proceeds received upon loss or destruction of autos, and early payoffs of the loans.

All auto loan-backed securities have some form of credit enhancement, often a senior/subordinated structure. In addition, many auto loan-backed securities come with overcollateralization and a reserve account, often an excess spread account. Recall from a previous reading that the excess spread, sometimes called excess interest cash flow, is an amount that can be retained and deposited into a reserve account and that can serve as a first line of protection against losses.

To illustrate the typical structure of auto loan-backed securities, let us use the example of securities issued by Fideicomiso Financiero Autos VI. The collateral was a pool of 827 auto loans denominated in Argentine pesos (ARS). The loans were originated by BancoFinansur. The structure of the securitization included three bond classes:

Bond Class	Outstanding Principal Balance (ARS)
Class A Floating-Rate Debt Securities	22,706,000
Class B Floating-Rate Debt Securities	1,974,000
Certificates	6,008,581
Total	30,688,581

The certificates provided credit protection for Class B, and Class B provides credit protection for Class A. Further credit enhancement came from overcollateralization and the presence of an excess spread account. The reference rate for the floating-rate debt securities was BADLAR (Buenos Aires Deposits of Large Amount Rate), the benchmark rate for loans in Argentina. This reference rate is the average rate on 30-day deposits of at least ARS1 million. For Class A, the interest rate was BADLAR plus 450 bps, with a minimum rate of 18% and a maximum rate of 26%; for Class B, it was BADLAR plus 650 bps, with 20% and 28% as the minimum and maximum rates, respectively.

EXAMPLE 9

An Example of an Auto Loan ABS

The following information is from the prospectus supplement for US$877,670,000 of auto loan ABS issued by XYZ Credit Automobile Receivables Trust 2019:

> The collateral for this securitization is a pool of subprime automobile loan contracts secured for new and used automobiles and light-duty trucks and vans.
>
> The issuing entity will issue seven sequential-pay classes of asset-backed notes pursuant to the indenture. The notes are designated as the "Class A-1 Notes," the "Class A-2 Notes," the "Class A-3 Notes," the "Class B Notes," the "Class C Notes," the "Class D Notes," and the "Class E Notes." The Class A-1 Notes, the Class A-2 Notes, and the Class A-3 Notes are the "Class A Notes." The Class A Notes, the Class B Notes, the Class C Notes, and the Class D Notes are being offered by this prospectus supplement and are sometimes referred to as the publicly offered notes. The Class E Notes are not being offered by this prospectus supplement and will initially be retained by the depositor or an affiliate of the depositor. The Class E Notes are sometimes referred to as the privately placed notes.
>
> Each class of notes will have the initial note principal balance, interest rate, and final scheduled distribution date listed in the following tables:

Publicly Offered Notes

Class	Initial Note Principal Balance (US$)	Interest Rate (%)	Final Scheduled Distribution Date
A-1 (senior)	168,000,000	0.25	8 August 2020
A-2 (senior)	279,000,000	0.74	8 November 2022
A-3 (senior)	192,260,000	0.96	9 April 2024
B (subordinated)	68,870,000	1.66	10 September 2024
C (subordinated)	85,480,000	2.72	9 September 2025
D (subordinated)	84,060,000	3.31	8 October 2025

Privately Placed Notes

Class	Initial Note Principal Balance (US$)	Interest Rate (%)	Final Scheduled Distribution Date
E (subordinated)	22,330,000	4.01	8 January 2027

> Interest on each class of notes will accrue during each interest period at the applicable interest rate.
>
> The overcollateralization amount represents the amount by which the aggregate principal balance of the automobile loan contracts exceeds the principal balance of the notes. On the closing date, the initial amount of overcollateralization is approximately US$49,868,074, or 5.25% of the aggregate principal balance of the automobile loan contracts as of the cutoff date.

On the closing date, 2.0% of the expected initial aggregate principal balance of the automobile loan contracts will be deposited into the reserve account, which is approximately US$18,997,361.

1 The reference to subprime meant that:
   - **A** the asset-backed notes were rated below investment grade.
   - **B** the automobile (auto) loan contracts were made to borrowers who did not have or could not document strong credit.
   - **C** some of the auto loan contracts were secured by autos of low quality that may have been difficult to sell in case the borrower defaults.

2 Based on the information provided, if on the first distribution date there were losses on the loans of US$10 million:
   - **A** none of the classes of notes will have incurred losses.
   - **B** Class E notes will have incurred losses of US$10 million.
   - **C** Classes B, C, D, and E will have incurred losses pro rata of their initial note principal balances.

3 Based on the information provided, if the first loss on the loans was US$40 million over and above the protection provided by the internal credit enhancements and occurred in January 2020, which class(es) of notes realized losses?
   - **A** Class E and then Class D
   - **B** Each class of subordinated notes in proportion to its principal balance
   - **C** Class E and then each class of subordinated notes in proportion to its principal balance

Solution to 1:

B is correct. A subprime loan is one granted to borrowers with lower credit quality, who have typically experienced prior credit difficulties or who cannot otherwise document strong credit.

Solution to 2:

A is correct. The amount of the loss (US$10 million) was lower than the combined amount of overcollateralization and the reserve account (US$49,868,074 + US$18,997,361 = US$68,865,435). Therefore, none of the classes of notes will have incurred losses.

Solution to 3:

A is correct. Once the amount of losses exceeds the amount of protection provided by the overcollateralization and the reserve account, losses are absorbed by the bond classes. Because it was a sequential-pay structure, Class E notes were the first ones to absorb losses, up to the principal amount of US$22,330,000. It meant that there was still US$17,670,000 to be absorbed by another bond class, which would have been the Class D notes.

8.2 Credit Card Receivable ABS

When a purchase is made on a credit card, the issuer of the credit card (the lender) extends credit to the cardholder (the borrower). Credit cards are issued by banks, credit card companies, retailers, and travel and entertainment companies. At the time of purchase, the cardholder agrees to repay the amount borrowed (that is, the cost of the item purchased) plus any applicable finance charges. The amount that the cardholder agrees to pay the issuer of the credit card is a receivable from the perspective of the issuer of the credit card. Credit card receivables are used as collateral for the issuance of credit card receivable ABS.

For a pool of credit card receivables, the cash flows consist of finance charges collected, fees, and principal repayments. Finance charges collected represent the periodic interest the credit card borrower is charged on the unpaid balance after the grace period, which may be fixed or floating. The floating rate may be capped; that is, it may have an upper limit because some countries have usury laws that impose a maximum interest rate. Fees include late payment fees and any annual membership fees.

Interest is paid to holders of credit card receivable ABS periodically (e.g., monthly, quarterly, or semiannually). As noted earlier, the collateral of credit card receivable ABS is a pool of non-amortizing loans. These loans have lockout periods during which the cash flows that are paid out to security holders are based only on finance charges collected and fees. When the lockout period is over, the principal that is repaid by the cardholders is no longer reinvested but instead is distributed to investors.

Some provisions in credit card receivable ABS require early principal amortization if specific events occur. Such provisions are referred to as "early amortization" or "rapid amortization" provisions and are included to safeguard the credit quality of the issue. The only way the principal cash flows can be altered is by the triggering of the early amortization provision.

To illustrate the typical structure of credit card receivable ABS, consider the Master Note Trust Series 20XX issued in March 20XX by a large, well-known financial institution. The originator of the credit card receivables was the issuer's retail bank, and the servicer was its finance company. The collateral was a pool of credit card receivables from several private-label and co-branded credit card issuers, including JCPenney, Lowe's Home Improvement, Sam's Club, Walmart, Gap, and Chevron. The structure of the US$969,085,000 securitization was as follows: Class A notes for US$800,000,000, Class B notes for US$100,946,373, and Class C notes for US$68,138,802. Thus, the issue had a senior/subordinate structure. The Class A notes were the senior notes and were rated Aaa by Moody's and AAA by Fitch. The Class B notes were rated A2 by Moody's and A+ by Fitch. The Class C notes were rated Baa2 by Moody's and BBB+ by Fitch.

EXAMPLE 10

Credit Card Receivable ABS vs. Auto Loan ABS

Credit card receivable asset-backed securities differ from auto loan ABS in the following way:

A credit card loans are recourse loans, whereas auto loans are non-recourse loans.

B the collateral for credit card receivable-backed securities is a pool of non-amortizing loans, whereas the collateral for auto loan ABS is a pool of amortizing loans.

C credit card receivable-backed securities have regular principal repayments, whereas auto loan ABS include a lockout period during which the cash proceeds from principal repayments are reinvested in additional loan receivables.

Solution:

B is correct. A main difference between credit card receivable ABS and auto loan ABS is the type of loans that back the securities. For credit card receivable ABS, the collateral is a pool of non-amortizing loans. During the lockout period, the cash proceeds from principal repayments are reinvested in additional credit card receivables. When the lockout period is over, principal repayments are used to pay off the outstanding principal. For auto loan-backed securities, the collateral is a pool of amortizing loans. Security holders receive regular principal repayments. As a result, the outstanding principal balance declines over time.

9

COLLATERALIZED DEBT OBLIGATIONS

i. describe collateralized debt obligations, including their cash flows and risks

Collateralized debt obligation (CDO) is a generic term used to describe a security backed by a diversified pool of one or more debt obligations: CDOs backed by corporate and emerging market bonds are collateralized bond obligations (CBOs); CDOs backed by leveraged bank loans are collateralized loan obligations (CLOs); CDOs backed by ABS, RMBS, CMBS, and other CDOs are structured finance CDOs; CDOs backed by a portfolio of credit default swaps for other structured securities are synthetic CDOs.

9.1 CDO Structure

A CDO involves the creation of an SPE. In a CDO, there is a need for a CDO manager, also called "**collateral manager**," to buy and sell debt obligations for and from the CDO's collateral (that is, the portfolio of assets) to generate sufficient cash flows to meet the obligations to the CDO bondholders.

The funds to purchase the collateral assets for a CDO are obtained from the issuance of debt obligations. These debt obligations are bond classes or tranches and include senior bond classes, mezzanine bond classes (that is, bond classes with credit ratings between senior and subordinated bond classes), and subordinated bond classes, often referred to as the residual or equity tranches. The motivation for investors to invest in senior or mezzanine bond classes is to earn a potentially higher yield than that on a comparably rated corporate bond by gaining exposure to debt products that they may not otherwise be able to purchase. Investors in equity tranches have the potential to earn an equity-type return, thereby offsetting the increased risk from investing in the subordinated class. The key to whether a CDO is viable depends upon whether a structure can be created that offers a competitive return for the subordinated tranche.

The basic economics of the CDO is that the funds are raised by the sale of the bond classes and the CDO manager invests those funds in assets. The CDO manager seeks to earn a rate of return higher than the aggregate cost of the bond classes. The return in excess of what is paid out to the bond classes accrues to the holders of the equity tranche and to the CDO manager. In other words, a CDO is a leveraged transaction in which those who invest in the equity tranche use borrowed funds (the bond classes issued) to generate a return above the funding cost.

As with ABS, each CDO bond class is structured to provide a specific level of risk for investors. CDOs impose restrictions on the CDO manager via various tests and limits that must be satisfied for the CDO to meet investors' varying risk appetites while still providing adequate protection for the senior bond class. If the CDO manager fails pre-specified tests, a provision is triggered that requires the payoff of the principal to the senior bond class until the tests are satisfied. This process effectively deleverages the CDO because the cheapest funding source for the CDO, the senior bond class, is reduced.

The ability of the CDO manager to make the interest payments and principal repayments depends on collateral performance. The proceeds to meet the obligations to the CDO bond classes can come from one or more of the following sources: interest payments from collateral assets, maturing of collateral assets, and sale of collateral assets. The cash flows and credit risks of CDOs are best illustrated by an example.

9.2 An Example of a CDO Transaction

Although various motivations may prompt a sponsor to create a CDO, the following example uses a CDO for which the purpose is to capture what market participants mistakenly label a CDO arbitrage transaction. The term "arbitrage" is not used here in the traditional sense—that is, a risk-free transaction that earns an expected positive net profit but requires no net investment of money. In this context, arbitrage is used in a loose sense to describe a transaction in which the motivation is to capture a spread between the potential collateral return and the funding cost.

To understand the structure of a CDO transaction and its risks, consider the following US$100 million issue:

Tranche	**Par Value (US$ million)**	**Coupon Rate**
Senior	80	MRR[a] + 70 bps
Mezzanine	10	10-year US Treasury rate + 200 bps
Equity	10	—

[a] MRR is the Market Reference Rate.

Suppose that the collateral consists of bonds that all mature in 10 years and that the coupon rate for every bond is the 10-year US Treasury rate plus 400 bps. Because the collateral pays a fixed rate (the 10-year US Treasury rate plus 400 bps) but the senior tranche requires a floating-rate payment (MRR plus 70 bps), the CDO manager enters into an interest rate swap agreement with another party. This interest rate swap is simply an agreement to periodically exchange fixed for floating interest payments based on a notional amount used to determine the interest payment for each party. The notional amount of the interest rate swap is the par value of the senior tranche—that is, US$80 million in this example. Let us suppose that through the interest rate swap, the CDO manager agrees to do the following: (1) Pay a fixed rate each year equal to the 10-year US Treasury rate plus 100 bps, and (2) receive MRR.

Assume that the 10-year US Treasury rate at the time this CDO is issued is 7%. Now, consider the annual cash flow for the first year. First, let us look at the collateral. Assuming no default, the collateral will pay an interest rate equal to the 10-year US Treasury rate of 7% plus 400 bps—that is, 11%. So, the interest payment is 11% × US$100,000,000 = US$11,000,000.

Now, let us determine the interest that must be paid to the senior and mezzanine tranches. For the senior tranche, the interest payment is US$80,000,000 × (MRR + 70 bps). For the mezzanine tranche, the coupon rate is the 10-year US Treasury rate plus 200 bps—that is, 9%. So, the interest payment for the mezzanine tranche is 9% × US$10,000,000 = US$900,000.

Finally, consider the interest rate swap. In this agreement, the CDO manager agreed to pay the swap counterparty the 10-year US Treasury rate plus 100 bps—that is, 8%—based on a notional amount of US$80 million. So, the amount paid to the swap counterparty the first year is 8% × US$80,000,000 = US$6,400,000. The amount received from the swap counterparty is MRR based on a notional amount of US$80 million—that is, MRR × US$80,000,000.

Combining this information, the CDO cash inflows are as follows:

Interest from collateral	\$11,000,000
Interest from swap counterparty	\$80,000,000 × MRR
Total interest received	\$11,000,000 + \$80,000,000 × MRR

The CDO cash outflows are as follows:

Interest to senior tranche	\$80,000,000	× (MRR + 70 bps)
Interest to mezzanine tranche	\$900,000	
Interest to swap counterparty	\$6,400,000	
Total interest paid	\$7,300,000	+ \$80,000,000 × (MRR + 70 bps)

Netting the total interest received (\$11,000,000 + \$80,000,000 × MRR) and the total interest paid—\$7,300,000 + \$80,000,000 × (MRR + 70 bps)—leaves a net interest of \$3,700,000 – \$80,000,000 × 70 bps = US\$3,140,000. From this amount, any fees—including the CDO manager's fees—must be paid. The balance is then the amount available to pay the equity tranche. Suppose the CDO manager's fees are US\$640,000. The cash flow available to the equity tranche for the first year is US\$2.5 million (\$3,140,000 – \$640,000). Because the equity tranche has a par value of US\$10 million and is assumed to be sold at par, the annual return is 25%.

This example includes some simplifying assumptions. For instance, it is assumed that no defaults would occur. Furthermore, it is assumed that all securities purchased by the CDO manager are non-callable and, thus, that the coupon rate would not decline because of securities being called. Despite these simplifying assumptions, the example demonstrates the economics of an arbitrage CDO transaction, the need for an interest rate swap, and how the equity tranche realizes a return.

In practice, CDOs are subject to additional risks. For example, in the case of collateral defaults, the manager may fail to earn a return sufficient to pay off the investors in the senior and mezzanine tranches, resulting in a loss. Equity tranche investors may lose their entire investment. Even if payments are made to these investors, the realized return may be below the return expected at the time of purchase.

Moreover, after some period, the CDO manager must begin repaying principal to the senior and mezzanine tranches. The interest rate swap must be structured to take this requirement into account because the entire amount of the senior tranche is not outstanding for the life of the collateral.

EXAMPLE 11

Collateralized Debt Obligations

An additional risk of an investment in an arbitrage collateralized debt obligation relative to an investment in an asset-backed security is:

A the default risk on the collateral assets.

B the risk that the CDO manager will fail to earn a return sufficient to pay off the investors in the senior and the mezzanine tranches.

C the risk due to the mismatch between the collateral making fixed-rate payments and the bond classes making floating-rate payments.

Solution:

B is correct. In addition to the risks associated with investments in ABS, such as the default risk on the collateral assets and the risk due to the potential mismatch between the collateral making fixed-rate payments and the bond classes making floating-rate payments, investors in CDOs face the risk that the CDO manager will fail to earn a return sufficient to pay off the investors in the senior and the mezzanine tranches. With an ABS, the cash flows from the collateral are used to pay off the holders of the bond classes without the active management of the collateral—that is, without a manager altering the composition of the debt obligations in the pool that is backing the securitization. In contrast, in an arbitrage CDO, a CDO manager buys and sells debt obligations with the dual purpose of not only paying off the holders of the bond classes but also generating an attractive/competitive return for the equity tranche and for the manager.

COVERED BONDS

10

j describe characteristics and risks of covered bonds and how they differ from other asset-backed securities

As outlined in an earlier reading, covered bonds are senior debt obligations issued by a financial institution and backed by a segregated pool of assets that typically consist of commercial or residential mortgages or public sector assets. The covered bond (or *Pfandbrief* in German) originated over 250 years ago in Prussia and has been adopted by issuers in Europe, Asia, and Australia. Each country or jurisdiction specifies eligible collateral and structures permissible in the covered bond market, and the European Union is taking steps to harmonize these elements among its member states.

Covered bonds are similar to ABS but offer bondholders dual recourse—that is, to both the issuing financial institution and the underlying asset pool. In the case of ABS, the financial institution that originates loans transfers securitized assets to a bankruptcy-remote special legal entity, but the pool of underlying assets in a covered bond (or "cover pool") remains on the financial institution's balance sheet, against which covered bondholders retain a top-priority claim.

Whereas ABS often use credit tranching to create bond classes with different borrower default exposures, covered bonds usually consist of one bond class per cover pool. Another difference lies in the dynamic nature of the cover pool. In contrast to a static pool of mortgage loans that expose investors to prepayment risk as in the case of US mortgage-backed securities, cover pool sponsors must replace any prepaid or non-performing assets (i.e., assets that do not generate the promised cash flows) in the cover pool to ensure sufficient cash flows until the maturity of the covered bond.

Redemption regimes exist to align the covered bond's cash flows as closely as possible with the original maturity schedule in the event of default of a covered bond's financial sponsor. For example, in the case of **hard-bullet covered bonds**, if payments do not occur according to the original schedule, a bond default is triggered and bond payments are accelerated. **Soft-bullet covered bonds** delay the bond default and payment acceleration of bond cash flows until a new final maturity date, which is usually up to a year after the original maturity date. **Conditional pass-through covered bonds**, in contrast, convert to pass-through securities after the original maturity date if all bond payments have not yet been made.

Covered bonds have remained a relatively stable and reliable source of funding over time because of their dual recourse nature, strict eligibility criteria, dynamic cover pool, and redemption regimes in the event of sponsor default. As a result, covered bonds usually carry lower credit risks and offer lower yields than otherwise similar ABS.

EXAMPLE 12

Covered Bonds

Which of the following statements about covered bonds and asset-backed securities is *most* accurate?

A Both covered bonds and ABS pass prepayment and extension risk of the underlying asset pool through to investors.

B Both covered bonds and ABS offer investors recourse to both the bond's issuer and the underlying asset pool.

C Covered bonds have a dynamic cover pool in which sponsors must replace prepaid or non-performing assets, whereas ABS, such as mortgage-backed securities, pass through default and prepayment risk to investors.

Solution:

C is correct. In contrast to a static pool of mortgage loans that expose investors to prepayment risk as in the case of US mortgage-backed securities, cover pool sponsors must replace any prepaid or non-performing assets in the cover pool to ensure sufficient cash flows until the covered bond matures. A is incorrect because the dynamic cover pool insulates investors from prepayment risk, whereas the static mortgage loan pool does not protect ABS investors from prepayment risk. B is incorrect since only covered bonds offer investors recourse to both the bond's issuer and the underlying asset pool.

SUMMARY

- Securitization involves pooling debt obligations, such as loans or receivables, and creating securities backed by the pool of debt obligations called asset-backed securities (ABS). The cash flows of the debt obligations are used to make interest payments and principal repayments to the holders of the ABS.
- Securitization has several benefits. It allows investors direct access to liquid investments and payment streams that would be unattainable if all the financing were performed through banks. It enables banks to increase loan originations at economic scales greater than if they used only their own in-house loan portfolios. Thus, securitization contributes to lower costs of borrowing for entities raising funds, higher risk-adjusted returns to investors, and greater efficiency and profitability for the banking sector.
- The parties to a securitization include the seller of the collateral (pool of loans), the servicer of the loans, and the special purpose entity (SPE). The SPE is bankruptcy remote, which plays a pivotal role in the securitization.
- A common structure in a securitization is subordination, which leads to the creation of more than one bond class or tranche. Bond classes differ as to how they will share any losses resulting from defaults of the borrowers whose loans are in the collateral. The credit ratings assigned to the various bond classes depend on how the credit-rating agencies evaluate the credit risks of the collateral and any credit enhancements.

- The motivation for the creation of different types of structures is to redistribute prepayment risk and credit risk efficiently among different bond classes in the securitization. Prepayment risk is the uncertainty that the actual cash flows will be different from the scheduled cash flows as set forth in the loan agreements because borrowers may choose to repay the principal early to take advantage of interest rate movements.
- Because of the SPE, the securitization of a company's assets may include some bond classes that have better credit ratings than the company itself or its corporate bonds. Thus, the company's funding cost is often lower when raising funds through securitization than when issuing corporate bonds.
- A mortgage is a loan secured by the collateral of some specified real estate property that obliges the borrower to make a predetermined series of payments to the lender. The cash flow of a mortgage includes (1) interest, (2) scheduled principal payments, and (3) prepayments (any principal repaid in excess of the scheduled principal payment).
- The various mortgage designs throughout the world specify (1) the maturity of the loan; (2) how the interest rate is determined (i.e., fixed rate versus adjustable or variable rate); (3) how the principal is repaid (i.e., whether the loan is amortizing and if it is, whether it is fully amortizing or partially amortizing with a balloon payment); (4) whether the borrower has the option to prepay and if so, whether any prepayment penalties might be imposed; and (5) the rights of the lender in a foreclosure (i.e., whether the loan is a recourse or non-recourse loan).
- In the United States, there are three sectors for securities backed by residential mortgages: (1) those guaranteed by a federal agency (Ginnie Mae) whose securities are backed by the full faith and credit of the US government, (2) those guaranteed by a GSE (e.g., Fannie Mae and Freddie Mac) but not by the US government, and (3) those issued by private entities that are not guaranteed by a federal agency or a GSE. The first two sectors are referred to as agency residential mortgage-backed securities (RMBS), and the third sector, as non-agency RMBS.
- A mortgage pass-through security is created when one or more holders of mortgages form a pool of mortgages and sell shares or participation certificates in the pool. The cash flow of a mortgage pass-through security depends on the cash flow of the underlying pool of mortgages and consists of monthly mortgage payments representing interest, the scheduled repayment of principal, and any prepayments, net of servicing and other administrative fees.
- Market participants measure the prepayment rate using two measures: the single monthly mortality rate (SMM) and its corresponding annualized rate—namely, the conditional prepayment rate (CPR). For MBS, a measure widely used by market participants to assess effective duration is the weighted average life or simply the average life of the MBS.
- Market participants use the Public Securities Association (PSA) prepayment benchmark to describe prepayment rates. A PSA assumption greater than 100 PSA means that prepayments are assumed to occur faster than the benchmark, whereas a PSA assumption lower than 100 PSA means that prepayments are assumed to occur slower than the benchmark.
- Prepayment risk includes two components: contraction risk and extension risk. The former is the risk that when interest rates decline, the security will have a shorter maturity than was anticipated at the time of purchase because homeowners will refinance at the new, lower interest rates. The latter is the risk that

when interest rates rise, fewer prepayments will occur than what was anticipated at the time of purchase because homeowners are reluctant to give up the benefits of a contractual interest rate that now looks low.

- The creation of a collateralized mortgage obligation (CMO) can help manage prepayment risk by distributing the various forms of prepayment risk among different classes of bondholders. The CMO's major financial innovation is that the securities created more closely satisfy the asset/liability needs of institutional investors, thereby broadening the appeal of mortgage-backed products.
- The most common types of CMO tranche are sequential-pay tranches, planned amortization class (PAC) tranches, support tranches, and floating-rate tranches.
- Non-agency RMBS share many features and structuring techniques with agency CMOs. However, they typically include two complementary mechanisms. First, the cash flows are distributed by rules that dictate the allocation of interest payments and principal repayments to tranches with various degrees of priority/seniority. Second, there are rules for the allocation of realized losses that specify that subordinated bond classes have lower payment priority than senior classes.
- In order to obtain favorable credit ratings, non-agency RMBS and non-mortgage ABS often require one or more credit enhancements. The most common forms of internal credit enhancement are senior/subordinated structures, reserve funds, and overcollateralization. In external credit enhancement, credit support in the case of defaults resulting in losses in the pool of loans is provided in the form of a financial guarantee by a third party to the transaction.
- Commercial mortgage-backed securities (CMBS) are securities backed by a pool of commercial mortgages on income-producing property.
- Two key indicators of the potential credit performance of CMBS are the debt-service-coverage (DSC) ratio and the loan-to-value ratio (LTV). The DSC ratio is the property's annual net operating income divided by the debt service.
- CMBS have considerable call protection, which allows CMBS to trade in the market more like corporate bonds than like RMBS. This call protection comes in two forms: at the structure level and at the loan level. The creation of sequential-pay tranches is an example of call protection at the structure level. At the loan level, four mechanisms offer investors call protection: prepayment lockouts, prepayment penalty points, yield maintenance charges, and defeasance.
- ABS are backed by a wide range of asset types. The most popular non-mortgage ABS are auto loan ABS and credit card receivable ABS. The collateral is amortizing for auto loan ABS and non-amortizing for credit card receivable ABS. As with non-agency RMBS, these ABS must offer credit enhancement to be appealing to investors.
- A collateralized debt obligation (CDO) is a generic term used to describe a security backed by a diversified pool of one or more debt obligations (e.g., corporate and emerging market bonds, leveraged bank loans, ABS, RMBS, and CMBS).
- A CDO involves the creation of an SPE. The funds necessary to pay the bond classes come from a pool of loans that must be serviced. A CDO requires a collateral manager to buy and sell debt obligations for and from the CDO's portfolio of assets to generate sufficient cash flows to meet the obligations of the CDO bondholders and to generate a fair return for the equityholders.

- The structure of a CDO includes senior, mezzanine, and subordinated/equity bond classes.
- Covered bonds are similar to ABS, but they differ because of their dual recourse nature, strict eligibility criteria, dynamic cover pool, and redemption regimes in the event of sponsor default.

PRACTICE PROBLEMS

1 Securitization is beneficial for banks because it:
- **A** repackages bank loans into simpler structures.
- **B** increases the funds available for banks to lend.
- **C** allows banks to maintain ownership of their securitized assets.

2 Securitization benefits financial markets by:
- **A** increasing the role of intermediaries.
- **B** establishing a barrier between investors and originating borrowers.
- **C** allowing investors to tailor credit risk and interest rate risk exposures to meet their individual needs.

3 A benefit of securitization is the:
- **A** reduction in disintermediation.
- **B** simplification of debt obligations.
- **C** creation of tradable securities with greater liquidity than the original loans.

4 Securitization benefits investors by:
- **A** providing more direct access to a wider range of assets.
- **B** reducing the inherent credit risk of pools of loans and receivables.
- **C** eliminating cash flow timing risks of an ABS, such as contraction and extension risks.

5 In a securitization, the special purpose entity (SPE) is responsible for the:
- **A** issuance of the asset-backed securities.
- **B** collection of payments from the borrowers.
- **C** recovery of underlying assets from delinquent borrowers.

6 In a securitization, the collateral is initially sold by the:
- **A** issuer.
- **B** depositor.
- **C** underwriter.

7 A special purpose entity issues asset-backed securities in the following structure.

Bond Class	Par Value (€ millions)
A (senior)	200
B (subordinated)	20
C (subordinated)	5

At which of the following amounts of default in par value would Bond Class A experience a loss?
- **A** €20 million
- **B** €25 million
- **C** €26 million

8 In a securitization, time tranching provides investors with the ability to choose between:

© 2019 CFA Institute. All rights reserved.

A extension and contraction risks.

B senior and subordinated bond classes.

C fully amortizing and partially amortizing loans.

9 The creation of bond classes with a waterfall structure for sharing losses is referred to as:

A time tranching.

B credit tranching.

C overcollateralization.

10 Which of the following statements related to securitization is correct?

A Time tranching addresses the uncertainty of a decline in interest rates.

B Securitizations are rarely structured to include both credit tranching and time tranching.

C Junior and senior bond classes differ in that junior classes can be paid off only at the bond's set maturity.

11 A goal of securitization is to:

A separate the seller's collateral from its credit ratings.

B uphold the absolute priority rule in bankruptcy reorganizations.

C account for collateral's primary influence on corporate bond credit spreads.

12 The last payment in a partially amortizing residential mortgage loan is *best* referred to as a:

A waterfall.

B principal repayment.

C balloon payment.

13 If a mortgage borrower makes prepayments without penalty to take advantage of falling interest rates, the lender will *most likely* experience:

A extension risk.

B contraction risk.

C yield maintenance.

14 Which of the following characteristics of a residential mortgage loan would *best* protect the lender from a strategic default by the borrower?

A Recourse

B A prepayment option

C Interest-only payments

15 William Marolf obtains a EUR5 million mortgage loan from Bank Nederlandse. A year later, the principal on the loan is EUR4 million and Marolf defaults on the loan. Bank Nederlandse forecloses, sells the property for EUR2.5 million, and is entitled to collect the EUR1.5 million shortfall from Marolf. Marolf *most likely* had a:

A bullet loan.

B recourse loan.

C non-recourse loan.

16 Fran Martin obtains a non-recourse mortgage loan for $500,000. One year later, when the outstanding balance of the mortgage is $490,000, Martin cannot make his mortgage payments and defaults on the loan. The lender forecloses on the loan and sells the house for $315,000. What amount is the lender entitled to claim from Martin?

A $0.

B $175,000.

C $185,000.

17 A balloon payment equal to a mortgage's original loan amount is a characteristic of a:

A bullet mortgage.

B fully amortizing mortgage.

C partially amortizing mortgage.

18 Which of the following statements is correct concerning mortgage loan defaults?

A A non-recourse jurisdiction poses higher default risks for lenders.

B In a non-recourse jurisdiction, strategic default will not affect the defaulting borrower's future access to credit.

C When a recourse loan defaults, the mortgaged property is the lender's sole source for recovery of the outstanding mortgage balance.

19 Which of the following describes a typical feature of a non-agency residential mortgage-backed security (RMBS)?

A Senior/subordinated structure

B A pool of conforming mortgages as collateral

C A guarantee by a government-sponsored enterprise

20 If interest rates increase, an investor who owns a mortgage pass-through security is *most likely* affected by:

A credit risk.

B extension risk.

C contraction risk.

21 Which of the following is *most likely* an advantage of collateralized mortgage obligations (CMOs)? CMOs can

A eliminate prepayment risk.

B be created directly from a pool of mortgage loans.

C meet the asset/liability requirements of institutional investors.

22 The longest-term tranche of a sequential-pay CMO is *most likely* to have the lowest:

A average life.

B extension risk.

C contraction risk.

23 The tranches in a collateralized mortgage obligation that are *most likely* to provide protection for investors against both extension and contraction risk are:

A planned amortization class (PAC) tranches.

B support tranches.

C sequential-pay tranches.

24 Support tranches are *most* appropriate for investors who are:

A concerned about their exposure to extension risk.

B concerned about their exposure to concentration risk.

C willing to accept prepayment risk in exchange for higher returns.

25 In the context of mortgage-backed securities, a conditional prepayment rate (CPR) of 8% means that approximately 8% of the outstanding mortgage pool balance at the beginning of the year is expected to be prepaid:

A in the current month.

B by the end of the year.

C over the life of the mortgages.

26 For a mortgage pass-through security, which of the following risks *most likely* increases as interest rates decline?

A Balloon

B Extension

C Contraction

27 Compared with the weighted average coupon rate of its underlying pool of mortgages, the pass-through rate on a mortgage pass-through security is:

A lower.

B the same.

C higher.

28 The single monthly mortality rate (SMM) *most likely:*

A increases as extension risk rises.

B decreases as contraction risk falls.

C stays fixed over time when the standard prepayment model remains at 100 PSA.

29 Credit risk is an important consideration for commercial mortgage-backed securities (CMBS) if the CMBS are backed by mortgage loans that:

A are non-recourse.

B have call protection.

C have prepayment penalty points.

30 Which commercial mortgage-backed security characteristic causes a CMBS to trade more like a corporate bond than a residential mortgage-backed security?

A Call protection

B Internal credit enhancement

C Debt-service-coverage ratio level

31 A commercial mortgage-backed security does not meet the debt-to-service coverage at the loan level necessary to achieve a desired credit rating. Which of the following features would *most likely* improve the credit rating of the CMBS?

A Subordination

B Call protection

C Balloon payments

32 If a default occurs in a non-recourse commercial mortgage-backed security, the lender will *most likely*:

A recover prepayment penalty points paid by the borrower to offset losses.

B use only the proceeds received from the sale of the property to recover losses.

C initiate a claim against the borrower for any shortfall resulting from the sale of the property.

33 Which of the following investments is least subject to prepayment risk?

A Auto loan receivable–backed securities

B Commercial mortgage-backed securities

C Non-agency residential mortgage-backed securities

34 An excess spread account incorporated into a securitization is designed to limit:

A credit risk.

B extension risk.

C contraction risk.

35 Which of the following *best* describes the cash flow that owners of credit card receivable asset-backed securities receive during the lockout period?

A No cash flow

B Only principal payments collected

C Only finance charges collected and fees

36 Which type of asset-backed security is not affected by prepayment risk?

A Auto loan ABS

B Residential MBS

C Credit card receivable ABS

37 In auto loan ABS, the form of credit enhancement that *most likely* serves as the first line of loss protection is the:

A excess spread account.

B sequential-pay structure.

C proceeds from repossession sales.

38 In credit card receivable ABS, principal cash flows can be altered only when the:

A lockout period expires.

B excess spread account is depleted.

C early amortization provision is triggered.

39 The CDO tranche with a credit-rating status between senior and subordinated bond classes is called the:

A equity tranche.

B residual tranche.

C mezzanine tranche.

40 The key to a CDO's viability is the creation of a structure with a competitive return for the:

A senior tranche.

B mezzanine tranche.

C subordinated tranche.

41 When the collateral manager fails pre-specified risk tests, a CDO is:

A deleveraged by reducing the senior bond class.

B restructured to reduce its most expensive funding source.

C liquidated by paying off the bond classes in order of seniority.

42 Collateralized mortgage obligations are designed to:

A eliminate contraction risk in support tranches.

B distribute prepayment risk to various tranches.

C eliminate extension risk in planned amortization tranches.

43 Which statement about covered bonds is *least* accurate?

A Covered bonds provide investors with dual recourse, to the cover pool and also to the issuer.

B Covered bonds usually carry higher credit risks and offer higher yields than otherwise similar ABS.

C Covered bonds have a dynamic cover pool, meaning sponsors must replace any prepaid or non-performing assets.

SOLUTIONS

1 B is correct. Securitization increases the funds available for banks to lend because it allows banks to remove loans from their balance sheets and issue bonds that are backed by those loans. Securitization repackages relatively simple debt obligations, such as bank loans, into more complex, not simpler, structures. Securitization involves transferring ownership of assets from the original owner—in this case, the banks—into a special legal entity. As a result, banks do not maintain ownership of the securitized assets.

2 C is correct. By removing the wall between ultimate investors and originating borrowers, investors can achieve better legal claims on the underlying mortgages and portfolios of receivables. This transparency allows investors to tailor interest rate risk and credit risk to their specific needs.

3 C is correct. Securitization allows for the creation of tradable securities with greater liquidity than the original loans on a bank's balance sheet. Securitization results in lessening the roles of intermediaries, which increases disintermediation. Securitization is a process in which relatively simple debt obligations, such as loans, are repackaged into more complex structures.

4 A is correct. Securitization allows investors to achieve more direct legal claims on loans and portfolios of receivables. As a result, investors can add to their portfolios exposure to the risk–return characteristics provided by a wider range of assets.

B is incorrect because securitization does not reduce credit risk but, rather, provides a structure to mitigate and redistribute the inherent credit risks of pools of loans and receivables.

C is incorrect because securitization does not eliminate the timing risks associated with ABS cash flows but, rather, provides a structure to mitigate and redistribute those risks, such as contraction risk and extension risk.

5 A is correct. In a securitization, the special purpose entity is the special legal entity responsible for the issuance of the asset-backed securities. The servicer, not the SPE, is responsible for both the collection of payments from the borrowers and the recovery of underlying assets if the borrowers default on their loans.

6 B is correct. In a securitization, the loans or receivables are initially sold by the depositor to the special purpose entity that uses them as collateral to issue the ABS.

A is incorrect because the SPE, often referred to as the issuer, is the purchaser of the collateral rather than the seller of the collateral.

C is incorrect because the underwriter neither sells nor purchases the collateral in a securitization. The underwriter performs the same functions in a securitization as it does in a standard bond offering.

7 C is correct. The first €25 (€5 + €20) million in default are absorbed by the subordinated classes (C and B). The senior Class A bonds will experience a loss only when defaults exceed €25 million.

8 A is correct. Time tranching is the process in which a set of bond classes or tranches is created that allow investors a choice in the type of prepayment risk—extension or contraction—that they prefer to bear. Senior and subordinated bond classes are used in credit tranching. Credit tranching structures allow investors to choose the amount of credit risk that they prefer to bear. Fully and partially amortizing loans are two types of amortizing loans.

9 B is correct. Credit tranching is a form of credit enhancement called subordination in which bond classes or tranches differ as to how they will share losses resulting from defaults of the borrowers whose loans are part of the collateral. This type of protection is commonly referred to as a waterfall structure because of the cascading flow of payments between bond classes in the event of default.

A is incorrect because time tranching involves the creation of bond classes that possess different expected maturities rather than bond classes that differ as to how credit losses will be shared. Time tranching involves the redistribution of prepayment risk, whereas credit tranching involves the redistribution of credit risk.

C is incorrect because although overcollateralization is a form of internal credit enhancement similar to subordination, it is the amount by which the principal amount of the pool of collateral exceeds the principal balance of the securities issued and backed by the collateral pool. Losses are absorbed first by the amount of overcollateralization and then according to the credit tranching structure.

10 A is correct. Time tranching is the creation of bond classes that possess different expected maturities so that prepayment risk can be redistributed among bond classes. When loan agreements provide borrowers the ability to alter payments, in the case of declining interest rates, this prepayment risk increases because borrowers tend to pay off part or all of their loans and refinance at lower interest rates.

B is incorrect because it is possible and quite common for a securitization to have structures with both credit tranching and time tranching.

C is incorrect because the subordinated structures of junior and senior bond classes differ as to how they will share any losses relative to defaults of the borrowers whose loans are in the collateral pool. Junior classes offer protection for senior classes, with losses first realized by the former. The classes are distinguished not by scheduled repayment terms but, rather, by a loss sharing hierarchy in the event of borrower default.

11 A is correct. The legal implication of a special purpose entity, a prerequisite for securitization, is that investors contemplating the purchase of bond classes backed by the assets of the SPE will evaluate the credit risk of those assets independently from the credit rating of the entity that sold the assets to the SPE. This separation of the seller's collateral from its credit rating provides the opportunity for the SPE to access a lower aggregate funding cost than what the seller might otherwise obtain.

B is incorrect because the absolute priority rule, under which senior creditors are paid in full before subordinated creditors, has not always been upheld in bankruptcy reorganizations. There is no assurance that if a corporate bond has collateral, the rights of the bondholders will be respected. It is this uncertainty that creates the dominant influence of credit ratings over collateral in credit spreads.

C is incorrect because corporate bond credit spreads will reflect the seller's credit rating primarily and the collateral only slightly. Securitization separates the seller's collateral from its credit rating, effectively altering the influence of collateral on the credit spread.

12 C is correct. In a partially amortizing loan, the sum of all the scheduled principal repayments is less than the amount borrowed. The last payment is for the remaining unpaid mortgage balance and is called the "balloon payment."

13 B is correct. Contraction risk is the risk that when interest rates decline, actual prepayments will be higher than forecasted. Extension risk is the risk that when interest rates rise, prepayments will be lower than forecasted. Yield maintenance results from prepayment penalties; the lender is protected from loss in yield by the imposition of prepayment penalties.

14 A is correct. In a recourse loan, the lender has a claim against the borrower for the shortfall between the amount of the mortgage balance outstanding and the proceeds received from the sale of the property. A prepayment option is a benefit to the borrower and would thus not offer protection to the lender. An interest-only mortgage requires no principal repayment for a number of years and will not protect the lender from strategic default by the borrower.

15 B is correct. Bank Nederlandse has a claim against Marolf for EUR1.5 million, the shortfall between the amount of the mortgage balance outstanding and the proceeds received from the sale of the property. This indicates that the mortgage loan is a recourse loan. The recourse/non-recourse feature indicates the rights of a lender in foreclosure. If Marolf had a non-recourse loan, the bank would have been entitled to only the proceeds from the sale of the underlying property, or EUR2.5 million. A bullet loan is a special type of interest-only mortgage for which there are no scheduled principal payments over the entire term of the loan. Since the unpaid balance is less than the original mortgage loan, it is unlikely that Marolf has an interest-only mortgage.

16 A is correct. Because the loan has a non-recourse feature, the lender can look to only the underlying property to recover the outstanding mortgage balance and has no further claim against the borrower. The lender is simply entitled to foreclose on the home and sell it.

17 A is correct. A bullet mortgage is a special type of interest-only mortgage in which there are no scheduled principal repayments over the entire life of the loan. At maturity, a balloon payment is required equal to the original loan amount.

B is incorrect because with a fully amortizing mortgage, the sum of all the scheduled principal repayments during the mortgage's life is such that when the last mortgage payment is made, the loan is fully repaid, with no balloon payment required.

C is incorrect because with a partially amortizing mortgage, the sum of all the scheduled principal repayments is less than the amount borrowed, resulting in a balloon payment equal to the unpaid mortgage balance (rather than the original loan amount).

18 A is correct. In non-recourse loan jurisdictions, the borrower may have an incentive to default on an underwater mortgage and allow the lender to foreclose on the property because the lender has no claim against the borrower for the shortfall. For this reason, such defaults, known as strategic defaults, are more likely in non-recourse jurisdictions and less likely in recourse jurisdictions, where the lender does have a claim against the borrower for the shortfall.

B is incorrect because strategic defaults in non-recourse jurisdictions do have negative consequences for the defaulting borrowers in the form of a lower credit score and a reduced ability to borrow in the future. These negative consequences can be a deterrent in the incidence of underwater mortgage defaults.

C is incorrect because when a recourse loan defaults, the lender can look to both the property and the borrower to recover the outstanding mortgage balance. In a recourse loan, the lender has a claim against the borrower for the shortfall between the amount of the outstanding mortgage balance and the proceeds received from the sale of the property.

19 A is correct. Non-agency RMBS are credit enhanced, either internally or externally, to make the securities more attractive to investors. The most common forms of internal credit enhancement are senior/subordinated structures, reserve accounts, and overcollateralization. Conforming mortgages are used as collateral for agency (not non-agency) mortgage pass-through securities. An agency RMBS, rather than a non-agency RMBS, issued by a GSE (government sponsored enterprise), is guaranteed by the respective GSE.

20 B is correct. Extension risk is the risk that when interest rate rise, fewer prepayments will occur. Homeowners will be reluctant to give up the benefit of a contractual interest rate that is lower. As a result, the mortgage pass-through security becomes longer in maturity than anticipated at the time of purchase.

21 C is correct. Using CMOs, securities can be created to closely satisfy the asset/liability needs of institutional investors. The creation of a CMO cannot eliminate prepayment risk; it can only distribute the various forms of this risk among various classes of bondholders. The collateral of CMOs is mortgage-related products, not the mortgages themselves.

22 C is correct. For a CMO with multiple sequential-pay tranches, the longest-term tranche will have the lowest contraction (prepayments greater than forecasted) risk because of the protection against this risk offered by the other tranches. The longest-term tranche is likely to have the highest average life and extension risk because it is the last tranche repaid in a sequential-pay tranche.

23 A is correct. PAC tranches have limited (but not complete) protection against both extension risk and contraction risk. This protection is provided by the support tranches. A sequential-pay tranche can protect against either extension risk or contraction risk but not both of these risks. The CMO structure with sequential-pay tranches allows investors concerned about extension risk to invest in shorter-term tranches and those concerned about contraction risk to invest in longer-term tranches.

24 C is correct. The greater predictability of cash flows provided in the planned amortization class (PAC) tranches comes at the expense of support tranches. As a result, investors in support tranches are exposed to higher extension risk and contraction risk than investors in PAC tranches. Investors will be compensated for bearing this risk because support tranches have a higher expected return than PAC tranches.

25 B is correct. CPR is an annualized rate that indicates the percentage of the outstanding mortgage pool balance at the beginning of the year that is expected to be prepaid by the end of the year.

26 C is correct. When interest rates decline, a mortgage pass-through security is subject to contraction risk. Contraction risk is the risk that when interest rates decline, actual prepayments will be higher than forecasted because borrowers will refinance at now-available lower interest rates. Thus, a security backed by mortgages will have a shorter maturity than was anticipated when the security was purchased.

27 A is correct. The coupon rate of a mortgage pass-through security is called the pass-through rate, whereas the mortgage rate on the underlying pool of mortgages is calculated as a weighted average coupon rate (WAC). The pass-through rate is lower than the WAC by an amount equal to the servicing fee and other administrative fees.

28 B is correct. The SMM is a monthly measure of the prepayment rate or prepayment speed. Contraction risk is the risk that when interest rates decline, actual prepayments will be higher than forecast. So if contraction risk falls, prepayments are likely to be lower than forecast, which would imply a decrease in the SMM.

A is incorrect because the SMM is a monthly measure of the prepayment rate or prepayment speed. Extension risk is the risk that when interest rates rise, actual prepayments will be lower than forecast. So if extension risk rises, prepayments are likely to be lower than forecast, which would imply a decrease, not an increase, in the SMM.

C is incorrect because at 100 PSA, investors can expect prepayments to follow the PSA prepayment benchmark. Based on historical patterns, the PSA standard model assumes that prepayment rates are low for newly initiated mortgages and then speed up as mortgages season. Thus, 100 PSA does not imply that the SMM remains the same but, rather, implies that it will vary over the life of the mortgage.

29 A is correct. If commercial mortgage loans are non-recourse loans, the lender can look to only the income-producing property backing the loan for interest and principal repayment. If there is a default, the lender looks to the proceeds from the sale of the property for repayment and has no recourse against the borrower for any unpaid mortgage loan balance. Call protection and prepayment penalty points protect against prepayment risk.

30 A is correct. With CMBS, investors have considerable call protection. An investor in an RMBS is exposed to considerable prepayment risk, but with CMBS, call protection is available to the investor at the structure and loan level. The call protection results in CMBS trading in the market more like a corporate bond than an RMBS. Both internal credit enhancement and the debt-service-coverage (DSC) ratio address credit risk, not prepayment risk.

31 A is correct. If specific ratios of debt to service coverage are needed and those ratios cannot be met at the loan level, subordination is used to achieve the desired credit rating. Call protection protects investors against prepayment risk. Balloon payments increase the risk of the underlying loans.

32 B is correct. In a non-recourse CMBS, the lender can look only to the income-producing property backing the loan for interest and principal repayment. If a default occurs, the lender can use only the proceeds from the sale of the property for repayment and has no recourse to the borrower for any unpaid balance.

33 B is correct. A critical feature that differentiates CMBS from RMBS is the call protection provided to investors. An investor in an RMBS is exposed to considerable prepayment risk because the borrower has the right to prepay the loan before maturity. CMBS provide investors with considerable call protection that comes either at the structure level or at the loan level.

34 A is correct. An excess spread account, sometimes called excess interest cash flow, is a form of internal credit enhancement that limits credit risk. It is an amount that can be retained and deposited into a reserve account and that can serve as a first line of protection against losses. An excess spread account does not limit prepayment risk—be it extension risk or contraction risk.

35 C is correct. During the lockout period, the cash flow that is paid out to owners of credit card receivable asset-backed securities is based only on finance charges collected and fees.

36 C is correct. Because credit card receivable ABS are backed by non-amortizing loans that do not involve scheduled principal repayments, they are not affected by prepayment risk.

A is incorrect because auto loan ABS are affected by prepayment risk since they are backed by amortizing loans involving scheduled principal repayments.

B is incorrect because residential MBS are affected by prepayment risk since they are backed by amortizing loans involving scheduled principal repayments.

37 A is correct. In addition to a senior/subordinated (sequential-pay) structure, many auto loan ABS are structured with additional credit enhancement in the form of overcollateralization and a reserve account, often an excess spread account. The excess spread is an amount that can be retained and deposited into a reserve account that can serve as a first line of protection against losses.

B is incorrect because in an auto loan ABS, losses are typically applied against the excess spread account and the amount of overcollateralization before the waterfall loss absorption of the sequential-pay structure.

C is incorrect because in auto loan ABS, proceeds from the repossession and resale of autos are prepayment cash flows rather than a form of credit enhancement for loss protection.

38 C is correct. In credit card receivable ABS, the only way the principal cash flows can be altered is by triggering the early amortization provision. Such provisions are included in the ABS structure to safeguard the credit quality of the issue.

A is incorrect because expiration of the lockout period does not result in the alteration of principal cash flows but instead defines when principal repayments are distributed to the ABS investors. During the lockout period, principal repayments by cardholders are reinvested. When the lockout period expires, principal repayments by cardholders are distributed to investors.

B is incorrect because the excess spread account is a credit enhancement for loss absorption. When the excess spread account is depleted, losses are applied against the overcollateralization amount followed by the senior/subordinated structure. The only way principal cash flows can be altered is by triggering the early amortization provision.

39 C is correct. The mezzanine tranche consists of bond classes with credit ratings between senior and subordinated bond classes.

A is incorrect because the equity tranche falls within and carries the credit rating applicable to the subordinated bond classes.

B is incorrect because the residual tranche falls within and carries the credit ratings applicable to the subordinated bond classes.

40 C is correct. The key to whether a CDO is viable is whether a structure can be created that offers a competitive return for the subordinated tranche (often referred to as the residual or equity tranche). Investors in a subordinated tranche typically use borrowed funds (the bond classes issued) to generate a return above the funding cost.

A is incorrect because the viability of a CDO depends on a structure that offers a competitive return for the subordinated tranche rather than the senior tranche.

B is incorrect because the viability of a CDO depends on a structure that offers a competitive return for the subordinated tranche rather than the mezzanine tranche.

41 A is correct. When the collateral manager fails pre-specified tests, a provision is triggered that requires the payoff of the principal to the senior class until the tests are satisfied. This reduction of the senior class effectively deleverages the CDO because the CDO's cheapest funding source is reduced.

42 B is correct. CMOs are designed to redistribute cash flows of mortgage-related products to different bond classes or tranches through securitization. Although CMOs do not eliminate prepayment risk, they distribute prepayment risk among various classes of bondholders.

43 B is correct. Covered bonds usually carry lower credit risks and offer lower yields than otherwise similar ABS. The reason is, among other factors, covered bonds provide investors with dual recourse, to the cover pool and also to the issuer. Moreover, covered bonds have a dynamic cover pool, meaning sponsors must replace any prepaid or non-performing assets.

Glossary

A priori probability A probability based on logical analysis rather than on observation or personal judgment.

Abandonment option The ability to terminate an investment at some future time if the financial results are disappointing.

Abnormal return The amount by which a security's actual return differs from its expected return, given the security's risk and the market's return.

Absolute advantage A country's ability to produce a good or service at a lower absolute cost than its trading partner.

Absolute dispersion The amount of variability present without comparison to any reference point or benchmark.

Absolute frequency The actual number of observations counted for each unique value of the variable (also called raw frequency).

Accelerated book build An offering of securities by an investment bank acting as principal that is accomplished in only one or two days.

Accelerated methods Depreciation methods that allocate a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Accounting costs Monetary value of economic resources used in performing an activity. These can be explicit, out-of-pocket, current payments, or an allocation of historical payments (depreciation) for resources. They do not include implicit opportunity costs.

Accounting profit Income as reported on the income statement, in accordance with prevailing accounting standards, before the provisions for income tax expense. Also called *income before taxes* or *pretax income*.

Accounts payable Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Accrued expenses Liabilities related to expenses that have been incurred but not yet paid as of the end of an accounting period—an example of an accrued expense is rent that has been incurred but not yet paid, resulting in a liability "rent payable." Also called *accrued liabilities*.

Accrued interest Interest earned but not yet paid.

Acquisition method A method of accounting for a business combination where the acquirer is required to measure each identifiable asset and liability at fair value. This method was the result of a joint project of the IASB and FASB aiming at convergence in standards for the accounting of business combinations.

Action lag Delay from policy decisions to implementation.

Active investment An approach to investing in which the investor seeks to outperform a given benchmark.

Active return The return on a portfolio minus the return on the portfolio's benchmark.

Activity ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory. Also called *asset utilization ratios* or *operating efficiency ratios*.

Add-on rates Bank certificates of deposit, repos, and indexes such as Libor and Euribor are quoted on an add-on rate basis (bond equivalent yield basis).

Addition rule for probabilities A principle stating that the probability that A or B occurs (both occur) equals the probability that A occurs, plus the probability that B occurs, minus the probability that both A and B occur.

Agency bond See *quasi-government bond*.

Agency costs Costs associated with the conflict of interest present between principals and agents when a company is managed by non-owners. Agency costs result from the inherent conflicts of interest between managers, bondholders, and equity owners.

Agency costs of debt Costs arising from conflicts of interest between managers and debtholders.

Agency costs of equity The smaller the stake managers have in the company, the less their share in bearing the cost of excessive perquisite consumption—consequently, the less their desire to give their best efforts in running the company.

Agency RMBS In the United States, securities backed by residential mortgage loans and guaranteed by a federal agency or guaranteed by either of the two GSEs (Fannie Mae and Freddie Mac).

Aggregate demand The quantity of goods and services that households, businesses, government, and non-domestic customers want to buy at any given level of prices.

Aggregate demand curve Inverse relationship between the price level and real output.

Aggregate income The value of all the payments earned by the suppliers of factors used in the production of goods and services.

Aggregate output The value of all the goods and services produced during a specified period.

Aggregate supply The quantity of goods and services producers are willing to supply at any given level of price.

Aggregate supply curve The level of domestic output that companies will produce at each price level.

All-or-nothing (AON) orders An order that includes the instruction to trade only if the trade fills the entire quantity (size) specified.

Allocationally efficient A characteristic of a market, a financial system, or an economy that promotes the allocation of resources to their highest value uses.

Alternative data Non-traditional data types generated by the use of electronic devices, social media, satellite and sensor networks, and company exhaust.

Alternative hypothesis The hypothesis that is accepted if the null hypothesis is rejected.

Alternative investment markets Market for investments other than traditional securities investments (i.e., traditional common and preferred shares and traditional fixed income instruments). The term usually encompasses direct and indirect investment in real estate (including timberland and farmland) and commodities (including precious metals); hedge funds, private equity, and other investments requiring specialized due diligence.

Alternative trading systems Trading venues that function like exchanges but that do not exercise regulatory authority over their subscribers except with respect to the conduct

of the subscribers' trading in their trading systems. Also called *electronic communications networks* or *multilateral trading facilities*.

American depository receipt A US dollar-denominated security that trades like a common share on US exchanges.

American depository share The underlying shares on which American depository receipts are based. They trade in the issuing company's domestic market.

American-style Type of option contract that can be exercised at any time up to the option's expiration date.

Amortisation The process of allocating the cost of intangible long-term assets having a finite useful life to accounting periods; the allocation of the amount of a bond premium or discount to the periods remaining until bond maturity.

Amortised cost The historical cost (initially recognised cost) of an asset, adjusted for amortisation and impairment.

Amortizing bond Bond with a payment schedule that calls for periodic payments of interest and repayments of principal.

Amortizing loans Loans with a payment schedule that calls for periodic payments of interest and repayments of principal.

Analysis of variance (ANOVA) The analysis that breaks the total variability of a dataset (such as observations on the dependent variable in a regression) into components representing different sources of variation. With reference to regression, ANOVA provides the inputs for an *F*-test of the significance of the regression as a whole, as well as the inputs for the coefficient of determination and the standard error of the estimate.

Analytical duration The use of mathematical formulas to estimate the impact of benchmark yield-to-maturity changes on bond prices.

Anchoring and adjustment bias An information-processing bias in which the use of a psychological heuristic influences the way people estimate probabilities.

Annual percentage rate The cost of borrowing expressed as a yearly rate.

Annuity A finite set of level sequential cash flows.

Annuity due An annuity having a first cash flow that is paid immediately.

Anomalies Apparent deviations from market efficiency.

Antidilutive With reference to a transaction or a security, one that would increase earnings per share (EPS) or result in EPS higher than the company's basic EPS—antidilutive securities are not included in the calculation of diluted EPS.

Arbitrage 1) The simultaneous purchase of an undervalued asset or portfolio and sale of an overvalued but equivalent asset or portfolio, in order to obtain a riskless profit on the price differential. Taking advantage of a market inefficiency in a risk-free manner. 2) The condition in a financial market in which equivalent assets or combinations of assets sell for two different prices, creating an opportunity to profit at no risk with no commitment of money. In a well-functioning financial market, few arbitrage opportunities are possible. 3) A risk-free operation that earns an expected positive net profit but requires no net investment of money.

Arbitrage-free pricing The overall process of pricing derivatives by arbitrage and risk neutrality. Also called the *principle of no arbitrage*.

Arbitrageurs Traders who engage in arbitrage. See *arbitrage*.

Arithmetic mean The sum of the observations divided by the number of observations.

Artificial intelligence Computer systems that exhibit cognitive and decision-making ability comparable (or superior) to that of humans.

Ask The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size). See *offer*.

Ask size The maximum quantity of an asset that pertains to a specific ask price from a trader. For example, if the ask for a share issue is $30 for a size of 1,000 shares, the trader is offering to sell at $30 up to 1,000 shares.

Asset allocation The process of determining how investment funds should be distributed among asset classes.

Asset-backed securities A type of bond issued by a legal entity called a *special purpose entity* (SPE) on a collection of assets that the SPE owns. Also, securities backed by receivables and loans other than mortgages.

Asset-based valuation models Valuation based on estimates of the market value of a company's assets.

Asset beta The unlevered beta; which reflects the business risk of the assets; the assets' systematic risk.

Asset class A group of assets that have similar characteristics, attributes, and risk–return relationships.

Asset swap Converts the periodic fixed coupon of a specific bond to an MRR plus or minus a spread.

Asset utilization ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Assets Resources controlled by an enterprise as a result of past events and from which future economic benefits to the enterprise are expected to flow.

Assignment of accounts receivable The use of accounts receivable as collateral for a loan.

Asymmetric information The differential of information between corporate insiders and outsiders regarding the company's performance and prospects. Managers typically have more information about the company's performance and prospects than owners and creditors.

At the money An option in which the underlying's price equals the exercise price.

Auction A type of bond issuing mechanism often used for sovereign bonds that involves bidding.

Autarkic price The price of a good or service in an autarkic economy.

Autarky A state in which a country does not trade with other countries.

Automatic stabilizer A countercyclical factor that automatically comes into play as an economy slows and unemployment rises.

Availability bias An information-processing bias in which people take a heuristic approach to estimating the probability of an outcome based on how easily the outcome comes to mind.

Available-for-sale Under US GAAP, debt securities not classified as either held-to-maturity or held-for-trading securities. The investor is willing to sell but not actively planning to sell. In general, available-for-sale debt securities are reported at fair value on the balance sheet, with unrealized gains included as a component of other comprehensive income.

Average fixed cost Total fixed cost divided by quantity produced.

Average life See *weighted average life*.

Average product Measures the productivity of inputs on average and is calculated by dividing total product by the total number of units for a given input that is used to generate that output.

Average revenue Total revenue divided by quantity sold.

Average total cost Total cost divided by quantity produced.

Average variable cost Total variable cost divided by quantity produced.

Back-testing With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backfill bias A bias introduced if a fund backfills its past performance data to the database when it first starts reporting actual performance.

Backtesting With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backup line of credit A type of credit enhancement provided by a bank to an issuer of commercial paper to ensure that the issuer will have access to sufficient liquidity to repay maturing commercial paper if issuing new paper is not a viable option.

Backwardation A condition in the futures markets in which the spot price exceeds the futures price, the forward curve is downward sloping, and the convenience yield is high.

Balance of payments A double-entry bookkeeping system that summarizes a country's economic transactions with the rest of the world for a particular period of time, typically a calendar quarter or year.

Balance of trade deficit When the domestic economy is spending more on non-domestic goods and services than non-domestic economies are spending on domestic goods and services.

Balance sheet The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet). Also called *statement of financial position* or *statement of financial condition.*

Balance sheet ratios Financial ratios involving balance sheet items only.

Balanced With respect to a government budget, one in which spending and revenues (taxes) are equal.

Balloon payment Large payment required at maturity to retire a bond's outstanding principal amount.

Bar chart A chart for plotting the frequency distribution of categorical data, where each bar represents a distinct category and each bar's height is proportional to the frequency of the corresponding category. In technical analysis, a bar chart that plots four bits of data for each time interval—the high, low, opening, and closing prices. A vertical line connects the high and low prices. A cross-hatch left indicates the opening price and a cross-hatch right indicates the closing price.

Barter economy An economy where economic agents as house-holds, corporations, and governments "pay" for goods and services with another good or service.

Base-rate neglect A type of representativeness bias in which the base rate or probability of the categorization is not adequately considered.

Base rates The reference rate on which a bank bases lending rates to all other customers.

Basic EPS Net earnings available to common shareholders (i.e., net income minus preferred dividends) divided by the weighted average number of common shares outstanding.

Basis point Used in stating yield spreads, one basis point equals one-hundredth of a percentage point, or 0.01%.

Basket of listed depository receipts An exchange-traded fund (ETF) that represents a portfolio of depository receipts.

Bayes' formula The rule for updating the probability of an event of interest—given a set of prior probabilities for the event, information, and information given the event—if you receive new information.

Bearer bonds Bonds for which ownership is not recorded; only the clearing system knows who the bond owner is.

Bearish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from above; this movement is considered bearish. A **death cross** is a bearish crossover based on 50-day and 200-day moving averages.

Behavioral finance A field of finance that examines the psychological variables that affect and often distort the investment decision making of investors, analysts, and portfolio managers.

Behind the market Said of prices specified in orders that are worse than the best current price; e.g., for a limit buy order, a limit price below the best bid.

Benchmark A comparison portfolio; a point of reference or comparison.

Benchmark issue The latest sovereign bond issue for a given maturity. It serves as a benchmark against which to compare bonds that have the same features but that are issued by another type of issuer.

Benchmark rate Typically the yield-to-maturity on a government bond having the same or close to the same time-to-maturity.

Benchmark spread The yield spread over a specific benchmark, usually measured in basis points.

Bernoulli random variable A random variable having the outcomes 0 and 1.

Bernoulli trial An experiment that can produce one of two outcomes.

Best bid The highest bid in the market.

Best effort offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-efforts offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-in-class An ESG implementation approach that seeks to identify the most favorable companies in an industry based on ESG considerations.

Best offer The lowest offer (ask price) in the market.

Beta A measure of the sensitivity of a given investment or portfolio to movements in the overall market.

Bid The price at which a dealer or trader is willing to buy an asset, typically qualified by a maximum quantity.

Bid–ask spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid–offer spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid size The maximum quantity of an asset that pertains to a specific bid price from a trader.

Big Data The vast amount of data being generated by industry, governments, individuals, and electronic devices that arises from both traditional and non-traditional data sources.

Bilateral loan A loan from a single lender to a single borrower.

Bimodal A distribution that has two most frequently occurring values.

Binomial model A model for pricing options in which the underlying price can move to only one of two possible new prices.

Binomial random variable The number of successes in n Bernoulli trials for which the probability of success is constant for all trials and the trials are independent.

Binomial tree The graphical representation of a model of asset price dynamics in which, at each period, the asset moves up with probability p or down with probability $(1 - p)$.

Bitcoin A cryptocurrency using blockchain technology that was created in 2009.

Bivariate correlation See *Pearson correlation*.

Block brokers A broker (agent) that provides brokerage services for large-size trades.

Blockchain A type of digital ledger in which information is recorded sequentially and then linked together and secured using cryptographic methods.

Blue chip Widely held large market capitalization companies that are considered financially sound and are leaders in their respective industry or local stock market.

Bollinger Bands A price-based technical analysis indicator consisting of a line representing the moving average, a higher line representing the moving average plus a set number of standard deviations from the average (for the same number of periods as was used to calculate the moving average), and a lower line representing the moving average minus the same number of standard deviations.

Bond Contractual agreement between the issuer and the bondholders.

Bond equivalent yield A calculation of yield that is annualized using the ratio of 365 to the number of days to maturity. Bond equivalent yield allows for the restatement and comparison of securities with different compounding periods.

Bond indenture The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *trust deed*.

Bond market vigilantes Bond market participants who might reduce their demand for long-term bonds, thus pushing up their yields.

Bond yield plus risk premium approach An estimate of the cost of common equity that is produced by summing the before-tax cost of debt and a risk premium that captures the additional yield on a company's stock relative to its bonds. The additional yield is often estimated using historical spreads between bond yields and stock yields.

Bonus issue of shares A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Book building Investment bankers' process of compiling a "book" or list of indications of interest to buy part of an offering.

Book value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. Also called *carrying value*.

Boom An expansionary phase characterized by economic growth "testing the limits" of the economy.

Bootstrap A resampling method that repeatedly draws samples with replacement of the selected elements from the original observed sample. Bootstrap is usually conducted by using computer simulation and is often used to find standard error or construct confidence intervals of population parameters.

Bottom-up analysis An investment selection approach that focuses on company-specific circumstances rather than emphasizing economic cycles or industry analysis.

Box and whisker plot A graphic for visualizing the dispersion of data across quartiles. It consists of a "box" with "whiskers" connected to the box.

Breakdown A breakdown occurs when the price of an asset moves below a support level.

Breakeven point The number of units produced and sold at which the company's net income is zero (Revenues = Total cost); in the case of perfect competition, the quantity at which price, average revenue, and marginal revenue equal average total cost.

Breakout A breakout occurs when the price of an asset moves above a resistance level.

Bridge financing Interim financing that provides funds until permanent financing can be arranged.

Broad money Encompasses narrow money plus the entire range of liquid assets that can be used to make purchases.

Broker 1) An agent who executes orders to buy or sell securities on behalf of a client in exchange for a commission. 2) See *futures commission merchants*.

Broker–dealer A financial intermediary (often a company) that may function as a principal (dealer) or as an agent (broker) depending on the type of trade.

Brokered market A market in which brokers arrange trades among their clients.

Brownfield investment Investing in *existing* infrastructure assets.

Bubble line chart A line chart that uses varying-sized bubbles to represent a third dimension of the data. The bubbles are sometimes color-coded to present additional information.

Budget surplus/deficit The difference between government revenue and expenditure for a stated fixed period of time.

Bullet bond Bond in which the principal repayment is made entirely at maturity.

Bullish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from below; this movement is considered bullish. A **golden cross** is a bullish crossover based on 50-day and 200-day moving averages.

Business cycles Are recurrent expansions and contractions in economic activity affecting broad segments of the economy.

Business risk The risk associated with operating earnings. Operating earnings are uncertain because total revenues and many of the expenditures contributed to produce those revenues are uncertain.

Buy-side firm An investment management company or other investor that uses the services of brokers or dealers (i.e., the client of the sell side firms).

Buyback A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Call An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call market A market in which trades occur only at a particular time and place (i.e., when the market is called).

Call money rate The interest rate that buyers pay for their margin loan.

Call option An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call protection The time during which the issuer of the bond is not allowed to exercise the call option.

Callable bond A bond containing an embedded call option that gives the issuer the right to buy the bond back from the investor at specified prices on pre-determined dates.

Calmar ratio A comparison of the average annual compounded return to the maximum drawdown risk.

Candlestick chart A price chart with four bits of data for each time interval. A candle indicates the opening and closing price for the interval. The body of the candle is shaded if the opening price was higher than the closing price, and the body is white (or clear) if the opening price was lower than the closing price. Vertical lines known as wicks or shadows extend from the top and bottom of the candle to indicate, respectively, the high and low prices for the interval.

Cannibalization Cannibalization occurs when an investment takes customers and sales away from another part of the company.

Cap rate A metric by which real estate managers are often judged; the annual rent actually earned (net of any vacancies) divided by the price originally paid for the property.

Capacity The ability of the borrower to make its debt payments on time.

Capital account A component of the balance of payments account that measures transfers of capital.

Capital allocation line (CAL) A graph line that describes the combinations of expected return and standard deviation of return available to an investor from combining the optimal portfolio of risky assets with the risk-free asset.

Capital allocation process The process that companies use for decision making on capital investments—those projects with a life of one year or longer.

Capital asset pricing model (CAPM) An equation describing the expected return on any asset (or portfolio) as a linear function of its beta relative to the market portfolio.

Capital consumption allowance A measure of the wear and tear (depreciation) of the capital stock that occurs in the production of goods and services.

Capital deepening investment Increases the stock of capital relative to labor.

Capital expenditure Expenditure on physical capital (fixed assets).

Capital light Capital light businesses require little incremental investment in fixed assets or working capital to enable revenue growth.

Capital loss ratio The percentage of capital in deals that have been realized below cost, net of any recovered proceeds, divided by total invested capital.

Capital market expectations An investor's expectations concerning the risk and return prospects of asset classes.

Capital market line (CML) The line with an intercept point equal to the risk-free rate that is tangent to the efficient frontier of risky assets; represents the efficient frontier when a risk-free asset is available for investment.

Capital market securities Securities with maturities at issuance longer than one year.

Capital markets Financial markets that trade securities of longer duration, such as bonds and equities.

Capital rationing A capital rationing environment assumes that the company has a fixed amount of funds to invest.

Capital restrictions Controls placed on foreigners' ability to own domestic assets and/or domestic residents' ability to own foreign assets.

Capital stock The accumulated amount of buildings, machinery, and equipment used to produce goods and services.

Capital structure The mix of debt and equity that a company uses to finance its business; a company's specific mixture of long-term financing.

Carry The net of the costs and benefits of holding, storing, or "carrying" an asset.

Carrying amount The amount at which an asset or liability is valued according to accounting principles.

Carrying value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. For a bond, the purchase price plus (or minus) the amortized amount of the discount (or premium).

Cartel Participants in collusive agreements that are made openly and formally.

Cash collateral account Form of external credit enhancement whereby the issuer immediately borrows the credit-enhancement amount and then invests that amount, usually in highly rated short-term commercial paper.

Cash conversion cycle A financial metric that measures the length of time required for a company to convert cash invested in its operations to cash received as a result of its operations; equal to days of inventory on hand + days of sales outstanding – number of days of payables. Also called *net operating cycle*.

Cash flow additivity principle The principle that dollar amounts indexed at the same point in time are additive.

Cash flow from operating activities The net amount of cash provided from operating activities.

Cash flow from operations The net amount of cash provided from operating activities.

Cash flow yield The internal rate of return on a series of cash flows.

Cash market securities Money market securities settled on a "same day" or "cash settlement" basis.

Cash markets See *spot markets*.

Cash prices See *spot prices*.

Cash-settled forwards See *non-deliverable forwards*.

Catch-up clause A clause in an agreement that favors the GP. For a GP who earns a 20% performance fee, a catch-up clause allows the GP to receive 100% of the distributions above the hurdle rate *until* she receives 20% of the profits generated, and then every excess dollar is split 80/20 between the LPs and GP.

Categorical data Values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize (also called **qualitative data**).

CBOE Volatility Index (VIX) A measure of near-term market volatility as conveyed by S&P 500 stock index option prices.

Central bank funds market The market in which deposit-taking banks that have an excess reserve with their national central bank can lend money to banks that need funds for maturities ranging from overnight to one year. Called the Federal or Fed funds market in the United States.

Central bank funds rate Interest rate at which central bank funds are bought (borrowed) and sold (lent) for maturities ranging from overnight to one year. Called Federal or Fed funds rate in the United States.

Central banks The dominant bank in a country, usually with official or semi-official governmental status.

Central limit theorem The theorem that states the sum (and the mean) of a set of independent, identically distributed random variables with finite variances is normally distributed, whatever distribution the random variables follow.

Certificate of deposit An instrument that represents a specified amount of funds on deposit with a bank for a specified maturity and interest rate. CDs are issued in various denominations and can be negotiable or non-negotiable.

Change in polarity principle A tenet of technical analysis that states that once a support level is breached, it becomes a resistance level. The same holds true for resistance levels: Once breached, they become support levels.

Change of control put A covenant giving bondholders the right to require the issuer to buy back their debt, often at par or at some small premium to par value, in the event that the borrower is acquired.

Character The quality of a debt issuer's management.

Chartist An individual who uses charts or graphs of a security's historical prices or levels to forecast its future trends.

Chi-square test of independence A statistical test for detecting a potential association between categorical variables.

Classical cycle Refers to fluctuations in the level of economic activity when measured by GDP in volume terms.

Classified balance sheet A balance sheet organized so as to group together the various assets and liabilities into subcategories (e.g., current and noncurrent).

Clawback A requirement that the general partner return any funds distributed as incentive fees until the limited partners have received back their initial investment and a percentage of the total profit.

Clearing The process by which the exchange verifies the execution of a transaction and records the participants' identities.

Clearing instructions Instructions that indicate how to arrange the final settlement ("clearing") of a trade.

Clearinghouse An entity associated with a futures market that acts as middleman between the contracting parties and guarantees to each party the performance of the other.

Closed economy An economy that does not trade with other countries; an *autarkic economy*.

Closed-end fund A mutual fund in which no new investment money is accepted. New investors invest by buying existing shares, and investors in the fund liquidate by selling their shares to other investors.

Cluster sampling A procedure that divides a population into subpopulation groups (clusters) representative of the population and then randomly draws certain clusters to form a sample.

Clustered bar chart See *grouped bar chart*.

Co-investing In co-investing, the investor invests in assets *indirectly* through the fund but also possesses rights (known as co-investment rights) to invest *directly* in the same assets. Through co-investing, an investor is able to make an investment *alongside* a fund when the fund identifies deals.

Code of ethics An established guide that communicates an organization's values and overall expectations regarding member behavior. A code of ethics serves as a general guide for how community members should act.

Coefficient of determination The percentage of the variation of the dependent variable that is explained by the independent variable. Also referred to as the "R-squared" or "R^2."

Coefficient of variation The ratio of a set of observations' standard deviation to the observations' mean value.

Cognitive cost The effort involved in processing new information and updating beliefs.

Cognitive dissonance The mental discomfort that occurs when new information conflicts with previously held beliefs or cognitions.

Cognitive errors Behavioral biases resulting from faulty reasoning; cognitive errors stem from basic statistical, information-processing, or memory errors.

Coincident economic indicators Turning points that are usually close to those of the overall economy; they are believed to have value for identifying the economy's present state.

Collateral Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Collateral manager Buys and sells debt obligations for and from the CDO's portfolio of assets (i.e., the collateral) to generate sufficient cash flows to meet the obligations to the CDO bondholders.

Collateral trust bonds Bonds secured by securities, such as common shares, other bonds, or other financial assets.

Collateralized bond obligations A structured asset-backed security that is collateralized by a pool of bonds.

Collateralized debt obligation Generic term used to describe a security backed by a diversified pool of one or more debt obligations.

Collateralized loan obligations A structured asset-backed security that is collateralized by a pool of loans.

Collateralized mortgage obligation A security created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collateralized mortgage obligations Securities created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collaterals Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Combination A listing in which the order of the listed items does not matter.

Combination formula (binomial formula) The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is ${}_nC_r = \binom{n}{r} = \dfrac{n!}{(n-r)!r!}$.

Commercial paper A short-term, negotiable, unsecured promissory note that represents a debt obligation of the issuer.

Committed (Regular) lines of credit A bank commitment to extend credit up to a pre-specified amount; the commitment is considered a short-term liability and is usually in effect for 364 days (one day short of a full year).

Committed capital The amount that the limited partners have agreed to provide to the private equity fund.

Commodity swap A swap in which the underlying is a commodity such as oil, gold, or an agricultural product.

Common market Level of economic integration that incorporates all aspects of the customs union and extends it by allowing free movement of factors of production among members.

Common shares A type of security that represent an ownership interest in a company.

Common-size analysis The restatement of financial statement items using a common denominator or reference item that allows one to identify trends and major differences; an example is an income statement in which all items are expressed as a percent of revenue.

Common stock See *common shares.*

Company analysis Analysis of an individual company.

Comparable company A company that has similar business risk, usually in the same industry and preferably with a single line of business.

Comparative advantage A country's ability to produce a good or service at a lower relative cost, or opportunity cost, than its trading partner.

Competitive strategy A company's plans for responding to the threats and opportunities presented by the external environment.

Complement The event not-S, written S^C, given the event S. Note that $P(S) + P(S^C) = 1$.

Complements Goods that tend to be used together; technically, two goods whose cross-price elasticity of demand is negative.

Complete markets Informally, markets in which the variety of distinct securities traded is so broad that any desired payoff in a future state-of-the-world is achievable.

Component cost of capital The rate of return required by suppliers of capital for an individual source of a company's funding, such as debt or equity.

Compounding The process of accumulating interest on interest.

Comprehensive income The change in equity of a business enterprise during a period from nonowner sources; includes all changes in equity during a period except those resulting from investments by owners and distributions to owners; comprehensive income equals net income plus other comprehensive income.

Conditional expected value The expected value of a stated event given that another event has occurred.

Conditional pass-through covered bonds Covered bonds that convert to pass-through securities after the original maturity date if all bond payments have not yet been made and the sponsor is in default.

Conditional probability The probability of an event given (conditioned on) another event.

Conditional variances The variance of one variable, given the outcome of another.

Confidence level The complement of the level of significance.

Confirmation bias A belief perseverance bias in which people tend to look for and notice what confirms their beliefs, to ignore or undervalue what contradicts their beliefs, and to misinterpret information as support for their beliefs.

Confusion matrix A type of contingency table used for evaluating the performance of a classification model.

Conservatism bias A belief perseverance bias in which people maintain their prior views or forecasts by inadequately incorporating new information.

Consolidation The movement of a stock's price within a well-defined range of trading levels for a period of time. The price consolidates between a support level and a resistance level.

Constant-yield price trajectory A graph that illustrates the change in the price of a fixed-income bond over time assuming no change in yield-to-maturity. The trajectory shows the "pull to par" effect on the price of a bond trading at a premium or a discount to par value.

Constituent securities With respect to an index, the individual securities within an index.

Consumer surplus The difference between the value that a consumer places on units purchased and the amount of money that was required to pay for them.

Contango A condition in the futures markets in which the spot price is lower than the futures price, the forward curve is upward sloping, and there is little or no convenience yield.

Contingency provision Clause in a legal document that allows for some action if a specific event or circumstance occurs.

Contingency table A tabular format that displays the frequency distributions of two or more categorical variables simultaneously and is used for finding patterns between the variables. A contingency table for two categorical variables is also known as a two-way table.

Contingent claim A security whose value is based on some other underlying security.

Contingent claims Derivatives in which the payoffs occur if a specific event occurs; generally referred to as options.

Contingent convertible bonds Bonds that automatically convert into equity if a specific event or circumstance occurs, such as the issuer's equity capital falling below the minimum requirement set by the regulators. Also called *CoCos.*

Continuation pattern A type of pattern used in technical analysis to predict the resumption of a market trend that was in place prior to the formation of a pattern.

Continuous data Data that can be measured and can take on any numerical value in a specified range of values.

Continuous random variable A random variable for which the range of possible outcomes is the real line (all real numbers between $-\infty$ and $+\infty$) or some subset of the real line.

Continuous trading market A market in which trades can be arranged and executed any time the market is open.

Continuously compounded return The natural logarithm of 1 plus the holding period return, or equivalently, the natural logarithm of the ending price over the beginning price.

Contra account An account that offsets another account.

Contract rate See *mortgage rate.*

Contraction The period of a business cycle after the peak and before the trough; often called a *recession* or, if exceptionally severe, called a *depression*.

Contraction risk The risk that when interest rates decline, the security will have a shorter maturity than was anticipated at the time of purchase because borrowers refinance at the new, lower interest rates.

Contractionary Tending to cause the real economy to contract.

Contractionary fiscal policy A fiscal policy that has the objective to make the real economy contract.

Contracts for differences See *non-deliverable forwards.*

Contribution margin The amount available for fixed costs and profit after paying variable costs; revenue minus variable costs.

Controlling shareholders A particular shareholder or block of shareholders holding a percentage of shares that gives them significant voting power.

Convenience sampling A procedure of selecting an element from a population on the basis of whether or not it is accessible to a researcher or how easy it is for a researcher to access the element.

Convenience yield A non-monetary advantage of holding an asset.

Conventional bond See *plain vanilla bond*.

Conventional cash flows A conventional cash flow pattern is one with an initial outflow followed by a series of inflows.

Convergence The tendency for differences in output per capita across countries to diminish over time. In technical analysis, the term describes the case when an indicator moves in the same manner as the security being analyzed.

Conversion premium The difference between the convertible bond's price and its conversion value.

Conversion price For a convertible bond, the price per share at which the bond can be converted into shares.

Conversion ratio For a convertible bond, the number of common shares that each bond can be converted into.

Conversion value For a convertible bond, the current share price multiplied by the conversion ratio.

Convertible bond Bond that gives the bondholder the right to exchange the bond for a specified number of common shares in the issuing company.

Convertible preference shares A type of equity security that entitles shareholders to convert their shares into a specified number of common shares.

Convexity adjustment For a bond, one half of the annual or approximate convexity statistic multiplied by the change in the yield-to-maturity squared.

Core inflation Refers to the inflation rate calculated based on a price index of goods and services except food and energy.

Corporate governance The system of internal controls and procedures by which individual companies are managed.

Correlation A measure of the linear relationship between two random variables.

Correlation coefficient A number between −1 and +1 that measures the consistency or tendency for two investments to act in a similar way. It is used to determine the effect on portfolio risk when two assets are combined.

Cost averaging The periodic investment of a fixed amount of money.

Cost of capital The cost of financing to a company; the rate of return that suppliers of capital require as compensation for their contribution of capital.

Cost of carry See *carry*.

Cost of debt The cost of debt financing to a company, such as when it issues a bond or takes out a bank loan; the rate of return required by a company's bondholders.

Cost of equity The rate of return required by a company's common stockholders.

Cost of preferred stock The cost to a company of issuing preferred stock; the dividend yield that a company must commit to pay preferred stockholders.

Cost structure The mix of a company's variable costs and fixed costs.

Counterparty risk The risk that the other party to a contract will fail to honor the terms of the contract.

Coupon rate The interest rate promised in a contract; this is the rate used to calculate the periodic interest payments.

Cournot assumption Assumption in which each firm determines its profit-maximizing production level assuming that the other firms' output will not change.

Covariance A measure of the co-movement (linear association) between two random variables.

Covariance matrix A matrix or square array whose entries are covariances; also known as a variance–covariance matrix.

Covenants The terms and conditions of lending agreements that the issuer must comply with; they specify the actions that an issuer is obligated to perform (affirmative covenant) or prohibited from performing (negative covenant).

Covered bond Debt obligation secured by a segregated pool of assets called the cover pool. The issuer must maintain the value of the cover pool. In the event of default, bondholders have recourse against both the issuer and the cover pool.

Covered bonds Senior debt obligations issued by a financial institution and backed by a segregated pool of assets. Covered bond investors have recourse to both the underlying assets and the financial institution.

Credit analysis The evaluation of credit risk; the evaluation of the creditworthiness of a borrower or counterparty.

Credit default swap (CDS) A type of credit derivative in which one party, the credit protection buyer who is seeking credit protection against a third party, makes a series of regularly scheduled payments to the other party, the credit protection seller. The seller makes no payments until a credit event occurs.

Credit derivatives A contract in which one party has the right to claim a payment from another party in the event that a specific credit event occurs over the life of the contract.

Credit enhancements Provisions that may be used to reduce the credit risk of a bond issue.

Credit-linked coupon bond Bond for which the coupon changes when the bond's credit rating changes.

Credit-linked note (CLN) Fixed-income security in which the holder of the security has the right to withhold payment of the full amount due at maturity if a credit event occurs.

Credit migration risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *downgrade risk*.

Credit risk The risk of loss caused by a counterparty's or debtor's failure to make a promised payment. Also called *default risk*.

Credit spread option An option on the yield spread on a bond.

Credit tranching A structure used to redistribute the credit risk associated with the collateral; a set of bond classes created to allow investors a choice in the amount of credit risk that they prefer to bear.

Creditworthiness The perceived ability of the borrower to pay its debt obligations in a timely manner; it represents the ability of a company to withstand adverse impacts on its cash flows.

Critical values Values of the test statistic at which the decision changes from fail to reject the null hypothesis to reject the null hypothesis.

Cross-default Covenant or contract clause that specifies a borrower is considered in default if they default on another debt obligation.

Cross-default provisions Provisions whereby events of default, such as non-payment of interest on one bond, trigger default on all outstanding debt; implies the same default probability for all issues.

Cross-price elasticity of demand The percentage change in quantity demanded for a given percentage change in the price of another good; the responsiveness of the demand for Product A that is associated with the change in price of Product B.

Cross-sectional analysis Analysis that involves comparisons across individuals in a group over a given time period or at a given point in time.

Cross-sectional data A list of the observations of a specific variable from multiple observational units at a given point in time. The observational units can be individuals, groups, companies, trading markets, regions, etc.

Crossing networks Trading systems that match buyers and sellers who are willing to trade at prices obtained from other markets.

Crowding out The thesis that government borrowing may divert private sector investment from taking place.

Cryptocurrency An electronic medium of exchange that lacks physical form.

Cryptography An algorithmic process to encrypt data, making the data unusable if received by unauthorized parties.

Cumulative absolute frequency Cumulates (i.e., adds up) in a frequency distribution the absolute frequencies as one moves from the first bin to the last bin.

Cumulative distribution function A function giving the probability that a random variable is less than or equal to a specified value.

Cumulative frequency distribution chart A chart that plots either the cumulative absolute frequency or the cumulative relative frequency on the y-axis against the upper limit of the interval and allows one to see the number or the percentage of the observations that lie below a certain value.

Cumulative preference shares Preference shares for which any dividends that are not paid accrue and must be paid in full before dividends on common shares can be paid.

Cumulative relative frequency A sequence of partial sums of the relative frequencies in a frequency distribution.

Cumulative voting A voting process whereby shareholders can accumulate and vote all their shares for a single candidate in an election, as opposed to having to allocate their voting rights evenly among all candidates.

Currencies Monies issued by national monetary authorities.

Currency option bonds Bonds that give bondholders the right to choose the currency in which they want to receive interest payments and principal repayments.

Currency swap A swap in which each party makes interest payments to the other in different currencies.

Current account A component of the balance of payments account that measures the flow of goods and services.

Current assets Assets that are expected to be consumed or converted into cash in the near future, typically one year or less. *Also called liquid assets.*

Current cost With reference to assets, the amount of cash or cash equivalents that would have to be paid to buy the same or an equivalent asset today; with reference to liabilities, the undiscounted amount of cash or cash equivalents that would be required to settle the obligation today.

Current government spending With respect to government expenditures, spending on goods and services that are provided on a regular, recurring basis including health, education, and defense.

Current liabilities Short-term obligations, such as accounts payable, wages payable, or accrued liabilities, that are expected to be settled in the near future, typically one year or less.

Current yield The sum of the coupon payments received over the year divided by the flat price; also called the *income* or *interest yield* or *running yield*.

Curve duration The sensitivity of the bond price (or the market value of a financial asset or liability) with respect to a benchmark yield curve.

Customs union Extends the free trade area (FTA) by not only allowing free movement of goods and services among members, but also creating a common trade policy against nonmembers.

CVaR Conditional VaR, a tail loss measure. The weighted average of all loss outcomes in the statistical distribution that exceed the VaR loss.

Cyclical A cyclical company is one whose profits are strongly correlated with the strength of the overall economy.

Cyclical companies Companies with sales and profits that regularly expand and contract with the business cycle or state of economy.

Daily settlement See *mark to market* and *marking to market.*

Dark pools Alternative trading systems that do not display the orders that their clients send to them.

Data A collection of numbers, characters, words, and text—as well as images, audio, and video—in a raw or organized format to represent facts or information.

Data mining The practice of determining a model by extensive searching through a dataset for statistically significant patterns. Also called *data snooping.*

Data science An interdisciplinary field that brings computer science, statistics, and other disciplines together to analyze and produce insights from Big Data.

Data snooping The practice of determining a model by extensive searching through a dataset for statistically significant patterns.

Data table see **two-dimensional rectangular array**.

Day order An order that is good for the day on which it is submitted. If it has not been filled by the close of business, the order expires unfilled.

Days of inventory on hand An activity ratio equal to the number of days in the period divided by inventory turnover over the period.

Dealers A financial intermediary that acts as a principal in trades.

Dealing securities Securities held by banks or other financial intermediaries for trading purposes.

Death cross See **Bearish crossover**.

Debentures Type of bond that can be secured or unsecured.

Debt rating An independent measure of the quality and safety of a company's debt based upon an analysis of the company's ability to pay the promised cash flows in full and on time. It includes an analysis of any indentures.

Debt-rating approach A method for estimating a company's before-tax cost of debt based on the yield on comparably rated bonds for maturities that closely match that of the company's existing debt.

Debt tax shield The tax benefit from interest paid on debt being tax deductible from income, equal to the marginal tax rate multiplied by the value of the debt.

Debt-to-assets ratio A solvency ratio calculated as total debt divided by total assets.

Debt-to-capital ratio A solvency ratio calculated as total debt divided by total debt plus total shareholders' equity.

Debt-to-equity ratio A solvency ratio calculated as total debt divided by total shareholders' equity.

Deciles Quantiles that divide a distribution into 10 equal parts.

Declaration date The day that the corporation issues a statement declaring a specific dividend.

Decreasing returns to scale When a production process leads to increases in output that are proportionately smaller than the increase in inputs.

Deductible temporary differences Temporary differences that result in a reduction of or deduction from taxable income in a future period when the balance sheet item is recovered or settled.

Deep learning Machine learning using neural networks with many hidden layers.

Deep learning nets Machine learning using neural networks with many hidden layers.

Default probability The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default risk*.

Default risk The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default probability*.

Default risk premium An extra return that compensates investors for the possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount.

Defensive companies Companies with sales and profits that have little sensitivity to the business cycle or state of the economy.

Defensive interval ratio A liquidity ratio that estimates the number of days that an entity could meet cash needs from liquid assets; calculated as (cash + short-term marketable investments + receivables) divided by daily cash expenditures.

Deferred coupon bond Bond that pays no coupons for its first few years but then pays a higher coupon than it otherwise normally would for the remainder of its life. Also called *split coupon bond*.

Deferred income A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred tax assets A balance sheet asset that arises when an excess amount is paid for income taxes relative to accounting profit. The taxable income is higher than accounting profit and income tax payable exceeds tax expense. The company expects to recover the difference during the course of future operations when tax expense exceeds income tax payable.

Deferred tax liabilities A balance sheet liability that arises when a deficit amount is paid for income taxes relative to accounting profit. The taxable income is less than the accounting profit and income tax payable is less than tax expense. The company expects to eliminate the liability over the course of future operations when income tax payable exceeds tax expense.

Defined benefit pension plans Plans in which the company promises to pay a certain annual amount (defined benefit) to the employee after retirement. The company bears the investment risk of the plan assets.

Defined contribution pension plans Individual accounts to which an employee and typically the employer makes contributions during their working years and expect to draw on the accumulated funds at retirement. The employee bears the investment and inflation risk of the plan assets.

Deflation Negative inflation.

Degree of confidence The probability that a confidence interval includes the unknown population parameter.

Degree of financial leverage (DFL) The ratio of the percentage change in net income to the percentage change in operating income; the sensitivity of the cash flows available to owners when operating income changes.

Degree of operating leverage (DOL) The ratio of the percentage change in operating income to the percentage change in units sold; the sensitivity of operating income to changes in units sold.

Degree of total leverage The ratio of the percentage change in net income to the percentage change in units sold; the sensitivity of the cash flows to owners to changes in the number of units produced and sold.

Degrees of freedom The number of independent variables used in defining sample statistics, such as variance, and the probability distributions they measure.

Delta The sensitivity of the derivative price to a small change in the value of the underlying asset.

Demand curve Graph of the inverse demand function. A graph showing the demand relation, either the highest quantity willingly purchased at each price or the highest price willingly paid for each quantity.

Demand function A relationship that expresses the quantity demanded of a good or service as a function of own-price and possibly other variables.

Demand shock A typically unexpected disturbance to demand, such as an unexpected interruption in trade or transportation.

Dependent With reference to events, the property that the probability of one event occurring depends on (is related to) the occurrence of another event.

Dependent variable The variable whose variation about its mean is to be explained by the regression; the left-side variable in a regression equation. Also referred to as the *explained variable*.

Depository bank A bank that raises funds from depositors and other investors and lends it to borrowers.

Depository institutions Commercial banks, savings and loan banks, credit unions, and similar institutions that raise funds from depositors and other investors and lend it to borrowers.

Depository receipt A security that trades like an ordinary share on a local exchange and represents an economic interest in a foreign company.

Depreciation The process of systematically allocating the cost of long-lived (tangible) assets to the periods during which the assets are expected to provide economic benefits.

Derivative pricing rule A pricing rule used by crossing networks in which a price is taken (derived) from the price that is current in the asset's primary market.

Derivatives A financial instrument whose value depends on the value of some underlying asset or factor (e.g., a stock price, an interest rate, or exchange rate).

Descriptive statistics Measures that summarize central tendency and spread variation in the data's distribution.

Diffuse prior The assumption of equal prior probabilities.

Diffusion index Reflects the proportion of the index's components that are moving in a pattern consistent with the overall index.

Diluted EPS The EPS that would result if all dilutive securities were converted into common shares.

Diluted shares The number of shares that would be outstanding if all potentially dilutive claims on common shares (e.g., convertible debt, convertible preferred stock, and employee stock options) were exercised.

Diminishing balance method An accelerated depreciation method, i.e., one that allocates a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Diminishing marginal productivity Describes a state in which each additional unit of input produces less output than previously.

Direct format With reference to the cash flow statement, a format for the presentation of the statement in which cash flow from operating activities is shown as operating cash receipts less operating cash disbursements. Also called *direct method*.

Direct investing Occurs when an investor makes a direct investment in an asset without the use of an intermediary.

Direct method See *direct format*.

Direct taxes Taxes levied directly on income, wealth, and corporate profits.

Direct write-off method An approach to recognizing credit losses on customer receivables in which the company waits until such time as a customer has defaulted and only then recognizes the loss.

Discount To reduce the value of a future payment in allowance for how far away it is in time; to calculate the present value of some future amount. Also, the amount by which an instrument is priced below its face (par) value.

Discount margin See *required margin*.

Discount rates In general, the interest rates used to calculate present values. In the money market, however, a discount rate is a specific type of quoted rate.

Discounted cash flow models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security.

Discouraged worker A person who has stopped looking for a job or has given up seeking employment.

Discrete data Numerical values that result from a counting process; therefore, practically speaking, the data are limited to a finite number of values.

Discrete random variable A random variable that can take on at most a countable number of possible values.

Discriminatory pricing rule A pricing rule used in continuous markets in which the limit price of the order or quote that first arrived determines the trade price.

Diseconomies of scale Increase in cost per unit resulting from increased production.

Dispersion The variability of a population or sample of observations around the central tendency.

Display size The size of an order displayed to public view.

Disposition effect As a result of loss aversion, an emotional bias whereby investors are reluctant to dispose of losers. This reluctance results in an inefficient and gradual adjustment to deterioration in fundamental value.

Distributed ledger A type of database that may be shared among entities in a network.

Distributed ledger technology Technology based on a distributed ledger.

Divergence In technical analysis, a term that describes the case when an indicator moves differently from the security being analyzed.

Diversification ratio The ratio of the standard deviation of an equally weighted portfolio to the standard deviation of a randomly selected security.

Dividend A distribution paid to shareholders based on the number of shares owned.

Dividend discount model (DDM) A present value model that estimates the intrinsic value of an equity share based on the present value of its expected future dividends.

Dividend payout ratio The ratio of cash dividends paid to earnings for a period.

Divisor A number (denominator) used to determine the value of a price return index. It is initially chosen at the inception of an index and subsequently adjusted by the index provider, as necessary, to avoid changes in the index value that are unrelated to changes in the prices of its constituent securities.

Doji In the Japanese terminology used in candlestick charting, the doji signifies that after a full day of trading, the positive price influence of buyers and the negative price influence of sellers exactly counteracted each other—with opening and closing prices that are virtually equal—which suggests that the market under analysis is in balance.

Domestic content provisions Stipulate that some percentage of the value added or components used in production should be of domestic origin.

Double bottom In technical analysis, a reversal pattern that is formed when the price reaches a low, rebounds, and then declines back to the first low level. A double bottom is used to predict a change from a downtrend to an uptrend.

Double coincidence of wants A prerequisite to barter trades, in particular that both economic agents in the transaction want what the other is selling.

Double declining balance depreciation An accelerated depreciation method that involves depreciating the asset at double the straight-line rate. This rate is multiplied by the book value of the asset at the beginning of the period (a declining balance) to calculate depreciation expense.

Double top In technical analysis, a reversal pattern that is formed when an uptrend reverses twice at roughly the same high price level. A double top is used to predict a change from an uptrend to a downtrend.

Down transition probability The probability that an asset's value moves down in a model of asset price dynamics.

Downgrade risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *credit migration risk*.

Downside risk Risk of incurring returns below a specified value.

Downtrend A pattern that occurs when the price of an asset moves lower over a period of time.

Drag on liquidity When receipts lag, creating pressure from the decreased available funds.

Drawdown A percentage peak-to-trough reduction in net asset value.

Dual-currency bonds Bonds that make coupon payments in one currency and pay the par value at maturity in another currency.

DuPont analysis An approach to decomposing return on investment, e.g., return on equity, as the product of other financial ratios.

Duration A measure of the approximate sensitivity of a security to a change in interest rates (i.e., a measure of interest rate risk).

Duration gap A bond's Macaulay duration minus the investment horizon.

Dutch Book Theorem A result in probability theory stating that inconsistent probabilities create profit opportunities.

Early repayment option See *prepayment option.*

Earnings per share The amount of income earned during a period per share of common stock.

Earnings surprise The portion of a company's earnings that is unanticipated by investors and, according to the efficient market hypothesis, merits a price adjustment.

Economic costs All the remuneration needed to keep a productive resource in its current employment or to acquire the resource for productive use; the sum of total accounting costs and implicit opportunity costs.

Economic indicator A variable that provides information on the state of the overall economy.

Economic loss The amount by which accounting profit is less than normal profit.

Economic profit Equal to accounting profit less the implicit opportunity costs not included in total accounting costs; the difference between total revenue (TR) and total cost (TC). Also called *abnormal profit* or *supernormal profit.*

Economic stabilization Reduction of the magnitude of economic fluctuations.

Economic union Incorporates all aspects of a common market and in addition requires common economic institutions and coordination of economic policies among members.

Economies of scale Reduction in cost per unit resulting from increased production.

Effective annual rate The amount by which a unit of currency will grow in a year with interest on interest included.

Effective convexity A *curve convexity* statistic that measures the secondary effect of a change in a benchmark yield curve on a bond's price.

Effective duration The sensitivity of a bond's price to a change in a benchmark yield curve.

Effective interest rate The borrowing rate or market rate that a company incurs at the time of issuance of a bond.

Efficient market A market in which asset prices reflect new information quickly and rationally.

Elastic Said of a good or service when the magnitude of elasticity is greater than one.

Elasticity The percentage change in one variable for a percentage change in another variable; a general measure of how sensitive one variable is to a change in the value of another variable.

Elasticity of demand A measure of the sensitivity of quantity demanded to a change in a product's own price: $\%\Delta Q^D/\%\Delta P$.

Elasticity of supply A measure of the sensitivity of quantity supplied to a change in price: $\%\Delta Q^S/\%\Delta P$.

Electronic communications networks See *alternative trading systems.*

Embedded option Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Embedded options Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Emotional biases Behavioral biases resulting from reasoning influenced by feelings; emotional biases stem from impulse or intuition.

Empirical duration The use of statistical methods and historical bond prices to estimate the price–yield relationship for a specific bond or portfolio of bonds.

Empirical probability The probability of an event estimated as a relative frequency of occurrence.

Employed The number of people with a job.

Endowment bias An emotional bias in which people value an asset more when they hold rights to it than when they do not.

Engagement/active ownership An ESG investment approach that uses shareholder power to influence corporate behavior through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is directed by ESG guidelines.

Enterprise risk management An overall assessment of a company's risk position. A centralized approach to risk management sometimes called firmwide risk management.

Enterprise value A measure of a company's total market value from which the value of cash and short-term investments have been subtracted.

Equal weighting An index weighting method in which an equal weight is assigned to each constituent security at inception.

Equipment trust certificates Bonds secured by specific types of equipment or physical assets.

Equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Equity risk premium The expected return on equities minus the risk-free rate; the premium that investors demand for investing in equities.

Equity swap A swap transaction in which at least one cash flow is tied to the return to an equity portfolio position, often an equity index.

Error term The difference between an observation and its expected value, where the expected value is based on the true underlying population relation between the dependent and independent variables. Also known simply as the *error*.

ESG An acronym that encompasses environmental, social, and governance.

ESG integration An ESG investment approach that focuses on systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions for the purpose of achieving the product's stated investment objectives. Used interchangeably with **ESG investing**.

ESG investing The consideration of environmental, social, and governance factors in the investment process. Used interchangeably with **ESG integration**.

Estimate The particular value calculated from sample observations using an estimator.

Estimated parameters With reference to a regression analysis, the estimated values of the population intercept and population slope coefficients in a regression.

Estimator An estimation formula; the formula used to compute the sample mean and other sample statistics are examples of estimators.

Ethical principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Ethics The study of moral principles or of making good choices. Ethics encompasses a set of moral principles and rules of conduct that provide guidance for our behavior.

Eurobonds Type of bond issued internationally, outside the jurisdiction of the country in whose currency the bond is denominated.

European-style Said of an option contract that can only be exercised on the option's expiration date.

Event Any outcome or specified set of outcomes of a random variable.

Ex-dividend date The first date that a share trades without (i.e., "ex") the dividend.

Excess kurtosis Degree of kurtosis (fatness of tails) relative to the kurtosis of the normal distribution.

Exchanges Places where traders can meet to arrange their trades.

Execution instructions Instructions that indicate how to fill an order.

Exercise The process of using an option to buy or sell the underlying.

Exercise price The fixed price at which an option holder can buy or sell the underlying. Also called *strike price*, *striking price*, or *strike*.

Exercise value The value obtained if an option is exercised based on current conditions. Also known as *intrinsic value*.

Exhaustive Covering or containing all possible outcomes.

Expansion The period of a business cycle after its lowest point and before its highest point.

Expansionary Tending to cause the real economy to grow.

Expansionary fiscal policy Fiscal policy aimed at achieving real economic growth.

Expected inflation The level of inflation that economic agents expect in the future.

Expected loss Default probability times loss severity given default.

Expected return on the portfolio ($E(R_p)$) The weighted average of the expected returns (R_1 to R_n) on the component securities using their respective weights (w_1 to w_n).

Expected value The probability-weighted average of the possible outcomes of a random variable.

Expenses Outflows of economic resources or increases in liabilities that result in decreases in equity (other than decreases because of distributions to owners); reductions in net assets associated with the creation of revenues.

Export subsidy Paid by the government to the firm when it exports a unit of a good that is being subsidized.

Exports Goods and services that an economy sells to other countries.

Extension risk The risk that when interest rates rise, fewer prepayments will occur because homeowners are reluctant to give up the benefits of a contractual interest rate that now looks low. As a result, the security becomes longer in maturity than anticipated at the time of purchase.

Externalities Spillover effects of production and consumption activities onto others who are not directly involved in a particular transaction, activity, or decision.

Externality An effect of a market transaction that is borne by parties other than those who transacted.

Extra dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Extreme value theory A branch of statistics that focuses primarily on extreme outcomes.

Face value The amount of cash payable by a company to the bondholders when the bonds mature; the promised payment at maturity separate from any coupon payment.

Factoring arrangement When a company sells its accounts receivable to a lender (known as a factor) who assumes responsibility for the credit granting and collection process.

Fair value The amount at which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's-length transaction; the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants.

False discovery approach An adjustment in the *p*-values for tests performed multiple times.

False discovery rate The rate of Type I errors in testing a null hypothesis multiple times for a given level of significance.

Fat-Tailed Describes a distribution that has fatter tails than a normal distribution (also called leptokurtic).

Fed funds rate The US interbank lending rate on overnight borrowings of reserves.

Federal funds rate The US interbank lending rate on overnight borrowings of reserves.

Fiat money Money that is not convertible into any other commodity.

Fiduciary An entity designated to represent the rights and responsibilities of a beneficiary whose assets they are managing, such as a bond trustee acting on behalf of fixed-income investors.

Fiduciary call A combination of a European call and a risk-free bond that matures on the option expiration day and has a face value equal to the exercise price of the call.

FIFO method The first in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the order in which they were placed in inventory.

Fill or kill See *immediate or cancel order*.

Finance lease From the lessee perspective, under US GAAP, a type of lease which is more akin to the purchase of an asset by the lessee. From the lessor perspective, under IFRS, a lease which "transfers substantially all the risks and rewards incidental to ownership of an underlying asset."

Financial account A component of the balance of payments account that records investment flows.

Financial distress Heightened uncertainty regarding a company's ability to meet its various obligations because of diminished earnings power or actual current losses.

Financial flexibility The ability to react and adapt to financial adversity and opportunities.

Financial leverage The extent to which a company can effect, through the use of debt, a proportional change in the return on common equity that is greater than a given proportional change in operating income; also, short for the financial leverage ratio.

Financial leverage ratio A measure of financial leverage calculated as average total assets divided by average total equity.

Financial risk The risk that environmental, social, or governance risk factors will result in significant costs or other losses to a company and its shareholders; the risk arising from a company's obligation to meet required payments under its financing agreements.

Financing activities Activities related to obtaining or repaying capital to be used in the business (e.g., equity and long-term debt).

Fintech Technological innovation in the design and delivery of financial services and products in the financial industry.

Firm commitment offering See *underwritten offering*.

First-degree price discrimination Where a monopolist is able to charge each customer the highest price the customer is willing to pay.

First lien debt Debt secured by a pledge of certain assets that could include buildings, but it may also include property and equipment, licenses, patents, brands, etc.

First mortgage debt Debt secured by a pledge of a specific property.

Fiscal multiplier The ratio of a change in national income to a change in government spending.

Fiscal policy The use of taxes and government spending to affect the level of aggregate expenditures.

Fisher effect The thesis that the real rate of interest in an economy is stable over time so that changes in nominal interest rates are the result of changes in expected inflation.

Fixed charge coverage A solvency ratio measuring the number of times interest and lease payments are covered by operating income, calculated as (EBIT + lease payments) divided by (interest payments + lease payments).

Fixed costs Costs that remain at the same level regardless of a company's level of production and sales.

Fixed-for-floating interest rate swap An interest rate swap in which one party pays a fixed rate and the other pays a floating rate, with both sets of payments in the same currency. Also called *plain vanilla swap* or *vanilla swap*.

Fixed-rate perpetual preferred stock Nonconvertible, noncallable preferred stock that has a fixed dividend rate and no maturity date.

Flag A technical analysis continuation pattern formed by parallel trendlines, typically over a short period.

Flat price The full price of a bond minus the accrued interest; also called the *quoted* or *clean* price.

Float-adjusted market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by adjusting its market capitalization for its market float.

Floaters See *floating-rate notes*.

Floating-rate notes Notes on which interest payments are not fixed but instead vary from period to period depending on the current level of a reference interest rate.

Flotation cost Fees charged to companies by investment bankers and other costs associated with raising new capital.

Foreclosure Allows the lender to take possession of a mortgaged property if the borrower defaults and then sell it to recover funds.

Foreign currency reserves Holding by the central bank of non-domestic currency deposits and non-domestic bonds.

Foreign direct investment Direct investment by a firm in one country (the source country) in productive assets in a foreign country (the host country).

Foreign exchange gains (or losses) Gains (or losses) that occur when the exchange rate changes between the investor's currency and the currency that foreign securities are denominated in.

Foreign portfolio investment Shorter-term investment by individuals, firms, and institutional investors (e.g., pension funds) in foreign financial instruments such as foreign stocks and foreign government bonds.

Forward commitments Class of derivatives that provides the ability to lock in a price to transact in the future at a previously agreed-upon price.

Forward contract An agreement between two parties in which one party, the buyer, agrees to buy from the other party, the seller, an underlying asset at a later date for a price established at the start of the contract.

Forward curve A series of forward rates, each having the same time frame.

Forward market For future delivery, beyond the usual settlement time period in the cash market.

Forward price The fixed price or rate at which the transaction scheduled to occur at the expiration of a forward contract will take place. This price is agreed on at the initiation date of the contract.

Forward rate The interest rate on a bond or money market instrument traded in a forward market. A forward rate can be interpreted as an incremental, or marginal, return for extending the time-to-maturity for an additional time period.

Forward rate agreements A forward contract calling for one party to make a fixed interest payment and the other to make an interest payment at a rate to be determined at the contract expiration.

Fractile A value at or below which a stated fraction of the data lies. Also called quantile.

Fractional reserve banking Banking in which reserves constitute a fraction of deposits.

Framing bias An information-processing bias in which a person answers a question differently based on the way in which it is asked (framed).

Free cash flow The actual cash that would be available to the company's investors after making all investments necessary to maintain the company as an ongoing enterprise (also referred to as free cash flow to the firm); the internally generated funds that can be distributed to the company's investors (e.g., shareholders and bondholders) without impairing the value of the company.

Free cash flow hypothesis The hypothesis that higher debt levels discipline managers by forcing them to make fixed debt service payments and by reducing the company's free cash flow.

Free-cash-flow-to-equity models Valuation models based on discounting expected future free cash flow to equity.

Free float The number of shares that are readily and freely tradable in the secondary market.

Free trade When there are no government restrictions on a country's ability to trade.

Free trade areas One of the most prevalent forms of regional integration, in which all barriers to the flow of goods and services among members have been eliminated.

Frequency distribution A tabular display of data constructed either by counting the observations of a variable by distinct values or groups or by tallying the values of a numerical variable into a set of numerically ordered bins (also called a one-way table).

Frequency polygon A graph of a frequency distribution obtained by drawing straight lines joining successive points representing the class frequencies.

Frequency table A representation of the frequency of occurrence of two discrete variables.

Full price The price of a security with accrued interest; also called the *invoice* or *dirty* price.

Fund investing In fund investing, the investor invests in assets indirectly by contributing capital to a fund as part of a group of investors. Fund investing is available for all major alternative investment types.

Fundamental analysis The examination of publicly available information and the formulation of forecasts to estimate the intrinsic value of assets.

Fundamental value The underlying or true value of an asset based on an analysis of its qualitative and quantitative characteristics. Also called *intrinsic value*.

Fundamental weighting An index weighting method in which the weight assigned to each constituent security is based on its underlying company's size. It attempts to address the disadvantages of market-capitalization weighting by using measures that are independent of the constituent security's price.

Funds of hedge funds Funds that hold a portfolio of hedge funds, more commonly shortened to *funds of funds*.

Fungible Freely exchangeable, interchangeable, or substitutable with other things of the same type. Money and commodities are the most common examples.

Future value (FV) The amount to which a payment or series of payments will grow by a stated future date.

Futures contract A variation of a forward contract that has essentially the same basic definition but with some additional features, such as a clearinghouse guarantee against credit losses, a daily settlement of gains and losses, and an organized electronic or floor trading facility.

Futures price The agreed-upon price of a futures contract.

FX swap The combination of a spot and a forward FX transaction.

G-spread The yield spread in basis points over an actual or interpolated government bond.

Gains Asset inflows not directly related to the ordinary activities of the business.

Game theory The set of tools decision makers use to incorporate responses by rival decision makers into their strategies.

Gamma A numerical measure of how sensitive an option's delta (the sensitivity of the derivative's price) is to a change in the value of the underlying.

Gap opening A gap is an area of a chart where a security's price either rises or falls from the previous day's close with no trading occurring in between. A gap opening is the start of a new trading session with a gap.

GDP deflator A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

General partner (GP) The partner that runs the business and theoretically bears unlimited liability for the business's debts and obligations.

Geometric mean A measure of central tendency computed by taking the nth root of the product of n non-negative values.

Giffen goods Goods that are consumed more as the price of the good rises because it is a very inferior good whose income effect overwhelms its substitution effect when price changes.

Gilts Bonds issued by the UK government.

Global depository receipt A depository receipt that is issued outside of the company's home country and outside of the United States.

Global minimum-variance portfolio The portfolio on the minimum-variance frontier with the smallest variance of return.

Global registered share A common share that is traded on different stock exchanges around the world in different currencies.

Gold standard With respect to a currency, if a currency is on the gold standard a given amount can be converted into a prespecified amount of gold.

Golden Cross See **Bullish crossover**.

Good-on-close An execution instruction specifying that an order can only be filled at the close of trading. Also called *market on close*.

Good-on-open An execution instruction specifying that an order can only be filled at the opening of trading.

Good-till-cancelled order An order specifying that it is valid until the entity placing the order has cancelled it (or, commonly, until some specified amount of time such as 60 days has elapsed, whichever comes sooner).

Goodwill An intangible asset that represents the excess of the purchase price of an acquired company over the value of the net assets acquired.

Government equivalent yield A yield that restates a yield-to-maturity based on a 30/360 day count to one based on actual/actual.

Green bonds Bonds used in green finance whereby the proceeds are earmarked toward environmental-related products.

Green finance A type of finance that addresses environmental concerns while achieving economic growth.

Green loans Any loan instruments made available exclusively to finance or re-finance, in whole or in part, new and/or existing eligible green projects. Green loans are commonly aligned in the market with the Green Loan Principles.

Greenfield investment Investing in infrastructure assets *that are to be constructed*.

Grey market The forward market for bonds about to be issued. Also called "when issued" market.

Gross domestic product The market value of all final goods and services produced within the economy during a given period (output definition) or, equivalently, the aggregate income earned by all households, all companies, and the government within the economy during a given period (income definition).

Gross margin Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit margin The ratio of gross profit to revenues.

Grouped bar chart A bar chart for showing joint frequencies for two categorical variables (also known as a **clustered bar chart**).

Grouping by function With reference to the presentation of expenses in an income statement, the grouping together of expenses serving the same function, e.g. all items that are costs of goods sold.

Grouping by nature With reference to the presentation of expenses in an income statement, the grouping together of expenses by similar nature, e.g., all depreciation expenses.

Growth cycle Refers to fluctuations in economic activity around the long-term potential trend growth level, focusing on how much actual economic activity is below or above trend growth in economic activity.

Growth cyclical A term sometimes used to describe companies that are growing rapidly on a long-term basis but that still experience above-average fluctuation in their revenues and profits over the course of a business cycle.

Growth investors With reference to equity investors, investors who seek to invest in high-earnings-growth companies.

Growth option The ability to make additional investments in a project at some future time if the financial results are strong. Also called *expansion option*.

Growth rate cycle Refers to fluctuations in the growth rate of economic activity.

Guarantee certificate A type of structured financial instrument that provides investors with capital protection. It combines a zero-coupon bond and a call option on some underlying asset.

Haircut See *repo margin*.

Halo effect An emotional bias that extends a favorable evaluation of some characteristics to other characteristics.

Hard-bullet covered bonds Covered bonds for which a bond default is triggered and bond payments are accelerated in the event of sponsor default if payments do not occur according to the original maturity schedule.

Harmonic mean A type of weighted mean computed as the reciprocal of the arithmetic average of the reciprocals.

Head and shoulders pattern In technical analysis, a reversal pattern that is formed in three parts: a left shoulder, a head, and a right shoulder. A head and shoulders pattern is used to predict a change from an uptrend to a downtrend.

Headline inflation Refers to the inflation rate calculated based on the price index that includes all goods and services in an economy.

Heat map A type of graphic that organizes and summarizes data in a tabular format and represents it using a color spectrum.

Hedge funds Private investment vehicles that typically use leverage, derivatives, and long and short investment strategies.

Hedge portfolio A hypothetical combination of the derivative and its underlying that eliminates risk.

Held-to-maturity Debt (fixed-income) securities that a company intends to hold to maturity; these are presented at their original cost, updated for any amortisation of discounts or premiums.

Herding Clustered trading that may or may not be based on information.

Heteroskedasticity A violation of the simple linear regression assumption of homoskedasticity; with heteroskedasticity, the variability of the residuals is not equal across all observations.

Hidden order An order that is exposed not to the public but only to the brokers or exchanges that receive it.

High-frequency trading A form of algorithmic trading that makes use of vast quantities of data to execute trades on ultra-high-speed networks in fractions of a second.

High-water mark The highest value, net of fees, which a fund has reached in history. It reflects the highest cumulative return used to calculate an incentive fee.

Hindsight bias A bias with selective perception and retention aspects in which people may see past events as having been predictable and reasonable to expect.

Histogram A chart that presents the distribution of numerical data by using the height of a bar or column to represent the absolute frequency of each bin or interval in the distribution.

Historical cost In reference to assets, the amount paid to purchase an asset, including any costs of acquisition and/or preparation; with reference to liabilities, the amount of proceeds received in exchange in issuing the liability.

Historical equity risk premium approach An estimate of a country's equity risk premium that is based on the historical averages of the risk-free rate and the rate of return on the market portfolio.

Holder-of-record date The date that a shareholder listed on the corporation's books will be deemed to have ownership of the shares for purposes of receiving an upcoming dividend.

Holding period return The return that an investor earns during a specified holding period; a synonym for total return.

Home bias A preference for securities listed on the exchanges of one's home country.

Homogeneity of expectations The assumption that all investors have the same economic expectations and thus have the same expectations of prices, cash flows, and other investment characteristics.

Homoskedasticity An assumption of simple linear regression such that the variability of the residuals is equal across all observations.

Horizon yield The internal rate of return between the total return (the sum of reinvested coupon payments and the sale price or redemption amount) and the purchase price of the bond.

Horizontal analysis Common-size analysis that involves comparing a specific financial statement with that statement in prior or future time periods; also, cross-sectional analysis of one company with another.

Horizontal demand schedule Implies that at a given price, the response in the quantity demanded is infinite.

Hostile takeover An attempt by one entity to acquire a company without the consent of the company's management.

Household A person or a group of people living in the same residence, taken as a basic unit in economic analysis.

Human capital The accumulated knowledge and skill that workers acquire from education, training, or life experience and the corresponding present value of future earnings to be generated by said skilled individual.

Hurdle rate The minimum rate of return on investment that a fund must reach before a GP receives carried interest.

Hypothesis A proposed explanation or theory that can be tested.

Hypothesis testing The process of testing of hypotheses about one or more populations using statistical inference.

I-spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Iceberg order An order in which the display size is less than the order's full size.

If-converted method A method for accounting for the effect of convertible securities on earnings per share (EPS) that specifies what EPS would have been if the convertible securities had been converted at the beginning of the period, taking account of the effects of conversion on net income and the weighted average number of shares outstanding.

Illusion of control bias A bias in which people tend to believe that they can control or influence outcomes when, in fact, they cannot.

Immediate or cancel order An order that is valid only upon receipt by the broker or exchange. If such an order cannot be filled in part or in whole upon receipt, it cancels immediately. Also called *fill or kill*.

Impact investing Investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.

Impact lag The lag associated with the result of actions affecting the economy with delay.

Implicit price deflator for GDP A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

Implicit selection bias One type of selection bias introduced through the presence of a threshold that filters out some unqualified members.

Implied forward rates Calculated from spot rates, an implied forward rate is a breakeven reinvestment rate that links the return on an investment in a shorter-term zero-coupon bond to the return on an investment in a longer-term zero-coupon bond.

Implied volatility The volatility that option traders use to price an option, implied by the price of the option and a particular option-pricing model.

Import license Specifies the quantity of a good that can be imported into a country.

Imports Goods and services that a domestic economy (i.e., house-holds, firms, and government) purchases from other countries.

In the money Options that, if exercised, would result in the value received being worth more than the payment required to exercise.

Income Increases in economic benefits in the form of inflows or enhancements of assets, or decreases of liabilities that result in an increase in equity (other than increases resulting from contributions by owners).

Income elasticity of demand A measure of the responsiveness of demand to changes in income, defined as the percentage change in quantity demanded divided by the percentage change in income.

Income tax paid The actual amount paid for income taxes in the period; not a provision, but the actual cash outflow.

Income tax payable The income tax owed by the company on the basis of taxable income.

Increasing marginal returns When the marginal product of a resource increases as additional units of that input are employed.

Increasing returns to scale When a production process leads to increases in output that are proportionately larger than the increase in inputs.

Incremental cash flow The cash flow that is realized because of a decision; the changes or increments to cash flows resulting from a decision or action.

Indenture Legal contract that describes the form of a bond, the obligations of the issuer, and the rights of the bondholders. Also called the *trust deed*.

Independent With reference to events, the property that the occurrence of one event does not affect the probability of another event occurring. With reference to two random variables X and Y, they are independent if and only if $P(X,Y) = P(X)P(Y)$.

Independent projects Independent projects are capital investments whose cash flows are independent of each other.

Independent variable A variable used to explain the dependent variable in a regression; a right-side variable in a regression equation. Also referred to as the *explanatory variable*.

Independently and identically distributed With respect to random variables, the property of random variables that are independent of each other but follow the identical probability distribution.

Index-linked bond Bond for which coupon payments and/or principal repayment are linked to a specified index.

Indexing An investment strategy in which an investor constructs a portfolio to mirror the performance of a specified index.

Indicator variable A variable that takes on only one of two values, 0 or 1, based on a condition. In simple linear regression, the slope is the difference in the dependent variable for the two conditions. Also referred to as a *dummy variable*.

Indifference curve A curve representing all the combinations of two goods or attributes such that the consumer is entirely indifferent among them.

Indirect format With reference to cash flow statements, a format for the presentation of the statement which, in the operating cash flow section, begins with net income then shows additions and subtractions to arrive at operating cash flow. Also called *indirect method*.

Indirect method See *indirect format*.

Indirect taxes Taxes such as taxes on spending, as opposed to direct taxes.

Industry A group of companies offering similar products and/or services.

Industry analysis The analysis of a specific branch of manufacturing, service, or trade.

Inelastic Said of a good or service when the magnitude of elasticity is less than one. Insensitive to price changes.

Inferior goods A good whose consumption decreases as income increases.

Inflation The percentage increase in the general price level from one period to the next; a sustained rise in the overall level of prices in an economy.

Inflation-linked bond Type of index-linked bond that offers investors protection against inflation by linking the bond's coupon payments and/or the principal repayment to an index of consumer prices. Also called *linkers*.

Inflation premium An extra return that compensates investors for expected inflation.

Inflation reports A type of economic publication put out by many central banks.

Inflation uncertainty The degree to which economic agents view future rates of inflation as difficult to forecast.

Information cascade The transmission of information from those participants who act first and whose decisions influence the decisions of others.

Information-motivated traders Traders that trade to profit from information that they believe allows them to predict future prices.

Informationally efficient market A market in which asset prices reflect new information quickly and rationally.

Initial coin offering An unregulated process whereby companies raise capital by selling crypto tokens to investors in exchange for fiat money or another agreed-upon cryptocurrency.

Initial margin The amount that must be deposited in a clearinghouse account when entering into a futures contract.

Initial margin requirement The margin requirement on the first day of a transaction as well as on any day in which additional margin funds must be deposited.

Initial public offering (IPO) The first issuance of common shares to the public by a formerly private corporation.

Input productivity The amount of output produced by workers in a given period of time—for example, output per hour worked; measures the efficiency of labor.

Intangible assets Assets lacking physical substance, such as patents and trademarks.

Interbank market The market of loans and deposits between banks for maturities ranging from overnight to one year.

Intercept The expected value of the dependent variable when the independent variable in a simple linear regression is equal to zero.

Interest Payment for lending funds.

Interest coverage A solvency ratio calculated as EBIT divided by interest payments.

Interest-only mortgage A loan in which no scheduled principal repayment is specified for a certain number of years.

Interest rate A rate of return that reflects the relationship between differently dated cash flows; a discount rate.

Interest rate effect The effect through which price level changes, through demand for money, impact interest rate, which in turn impacts investment and consumption.

Interest rate swap A swap in which the underlying is an interest rate. Can be viewed as a currency swap in which both currencies are the same and can be created as a combination of currency swaps.

Intermarket analysis A field within technical analysis that combines analysis of the major categories of securities—namely, equities, bonds, currencies, and commodities—to identify market trends and possible inflections in trends.

Internal rate of return The discount rate that makes net present value equal 0; the discount rate that makes the present value of an investment's costs (outflows) equal to the present value of the investment's benefits (inflows).

Internet of Things A network arrangement of structures and devices whereby the objects on the network are able to interact and share information.

Interpolated spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Interquartile range The difference between the third and first quartiles of a dataset.

Interval With reference to grouped data, a set of values within which an observation falls.

Intrinsic value See *exercise value*.

Inventory investment Net change in business inventory.

Inventory turnover An activity ratio calculated as cost of goods sold divided by average inventory.

Inverse demand function A restatement of the demand function in which price is stated as a function of quantity.

Inverse floater A type of leveraged structured financial instrument. The cash flows are adjusted periodically and move in the opposite direction of changes in the reference rate.

Inverse transformation method A method using randomly generated numbers from the continuous uniform distribution to generate random observations from any distribution.

Investing activities Activities associated with the acquisition and disposal of property, plant, and equipment; intangible assets; other long-term assets; and both long-term and short-term investments in the equity and debt (bonds and loans) issued by other companies.

Investment banks Financial intermediaries that provide advice to their mostly corporate clients and help them arrange transactions such as initial and seasoned securities offerings.

Investment policy statement (IPS) A written planning document that describes a client's investment objectives and risk tolerance over a relevant time horizon, along with constraints that apply to the client's portfolio.

Investment property Property used to earn rental income or capital appreciation (or both).

Jackknife A resampling method that repeatedly draws samples by taking the original observed data sample and leaving out one observation at a time (without replacement) from the set.

January effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January. Also called *turn-of-the-year effect.*

Joint frequencies The entry in the cells of the contingency table that represent the joining of one variable from a row and the other variable from a column to count observations.

Joint probability The probability of the joint occurrence of stated events.

Joint probability function A function giving the probability of joint occurrences of values of stated random variables.

Judgmental Sampling A procedure of selectively handpicking elements from the population based on a researcher's knowledge and professional judgment.

Key rate duration A method of measuring the interest rate sensitivities of a fixed-income instrument or portfolio to shifts in key points along the yield curve.

Keynesians Economists who believe that fiscal policy can have powerful effects on aggregate demand, output, and employment when there is substantial spare capacity in an economy.

Kurtosis The statistical measure that indicates the combined weight of the tails of a distribution relative to the rest of the distribution.

Labor force The portion of the working age population (over the age of 16) that is employed or is available for work but not working (unemployed).

Labor productivity The quantity of goods and services (real GDP) that a worker can produce in one hour of work.

Lagging economic indicators Turning points that take place later than those of the overall economy; they are believed to have value in identifying the economy's past condition.

Law of demand The principle that as the price of a good rises, buyers will choose to buy less of it, and as its price falls, they will buy more.

Law of diminishing marginal returns The observation that a variable factor's marginal product must eventually fall as more of it is added to a fixed amount of the other factors.

Law of diminishing returns The smallest output that a firm can produce such that its long run average costs are minimized.

Law of one price The condition in a financial market in which two equivalent financial instruments or combinations of financial instruments can sell for only one price. Equivalent to the principle that no arbitrage opportunities are possible.

Lead underwriter The lead investment bank in a syndicate of investment banks and broker–dealers involved in a securities underwriting.

Leading economic indicators Turning points that usually precede those of the overall economy; they are believed to have value for predicting the economy's future state, usually near-term.

Lease A contract that conveys the right to use an asset for a period of time in exchange for consideration.

Legal tender Something that must be accepted when offered in exchange for goods and services.

Lender of last resort An entity willing to lend money when no other entity is ready to do so.

Leptokurtic Describes a distribution that has fatter tails than a normal distribution (also called fat-tailed).

Lessee The party obtaining the use of an asset through a lease.

Lessor The owner of an asset that grants the right to use the asset to another party.

Letter of credit Form of external credit enhancement whereby a financial institution provides the issuer with a credit line to reimburse any cash flow shortfalls from the assets backing the issue.

Level of significance The probability of a Type I error in testing a hypothesis.

Leverage In the context of corporate finance, leverage refers to the use of fixed costs within a company's cost structure. Fixed costs that are operating costs (such as depreciation or rent) create operating leverage. Fixed costs that are financial costs (such as interest expense) create financial leverage.

Leveraged buyout A transaction whereby the target company's management team converts the target to a privately held company by using heavy borrowing to finance the purchase of the target company's outstanding shares.

Liabilities Present obligations of an enterprise arising from past events, the settlement of which is expected to result in an outflow of resources embodying economic benefits; creditors' claims on the resources of a company.

LIFO layer liquidation With respect to the application of the LIFO inventory method, the liquidation of old, relatively low-priced inventory; happens when the volume of sales rises above the volume of recent purchases so that some sales are made from relatively old, low-priced inventory. Also called *LIFO liquidation*.

LIFO method The last in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the reverse order the items were placed in inventory (i.e., inventory produced or acquired last are assumed to be sold first).

LIFO reserve The difference between the reported LIFO inventory carrying amount and the inventory amount that would have been reported if the FIFO method had been used (in other words, the FIFO inventory value less the LIFO inventory value).

Likelihood The probability of an observation, given a particular set of conditions.

Limit down A limit move in the futures market in which the price at which a transaction would be made is at or below the lower limit.

Limit order Instructions to a broker or exchange to obtain the best price immediately available when filling an order, but in no event accept a price higher than a specified (limit) price when buying or accept a price lower than a specified (limit) price when selling.

Limit order book The book or list of limit orders to buy and sell that pertains to a security.

Limit up A limit move in the futures market in which the price at which a transaction would be made is at or above the upper limit.

Limitations on liens Meant to put limits on how much secured debt an issuer can have.

Limited partners Partners with limited liability. Limited partnerships in hedge and private equity funds are typically restricted to investors who are expected to understand and to be able to assume the risks associated with the investments.

Limited partnership agreement (LPA) A legal document that outlines the rules of the partnership and establishes the framework that ultimately guides the fund's operations throughout its life.

Lin-log model A regression model in which the independent variable is in logarithmic form.

Line chart A type of graph used to visualize ordered observations. In technical analysis, a plot of price data, typically closing prices, with a line connecting the points.

Linear interpolation The estimation of an unknown value on the basis of two known values that bracket it, using a straight line between the two known values.

Linear regression Regression that models the straight-line relationship between the dependent and independent variables. Also known as *least squares regression* and *ordinary least squares regression*.

Linear scale A scale in which equal distances correspond to equal absolute amounts. Also called an *arithmetic scale*.

Linker See *inflation-linked bond*.

Liquid market Said of a market in which traders can buy or sell with low total transaction costs when they want to trade.

Liquidation To sell the assets of a company, division, or subsidiary piecemeal, typically because of bankruptcy; the form of bankruptcy that allows for the orderly satisfaction of creditors' claims after which the company ceases to exist.

Liquidity The extent to which a company is able to meet its short-term obligations using cash flows and those assets that can be readily transformed into cash.

Liquidity premium An extra return that compensates investors for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

Liquidity ratios Financial ratios measuring the company's ability to meet its short-term obligations to creditors.

Liquidity risk The risk that a financial instrument cannot be purchased or sold without a significant concession in price due to the size of the market.

Liquidity trap A condition in which the demand for money becomes infinitely elastic (horizontal demand curve) so that injections of money into the economy will not lower interest rates or affect real activity.

Load fund A mutual fund in which, in addition to the annual fee, a percentage fee is charged to invest in the fund and/or for redemptions from the fund.

Loan-to-value ratio The amount of the property's mortgage divided by the purchase price

Locked limit A condition in the futures markets in which a transaction cannot take place because the price would be beyond the limits.

Lockup period The minimum holding period before investors are allowed to make withdrawals or redeem shares from a fund.

Log-lin model A regression model in which the dependent variable is in logarithmic form.

Log-log model A regression model in which both the dependent and independent variables are in logarithmic form. Also known as the *double-log model*.

Logarithmic scale A scale in which equal distances represent equal proportional changes in the underlying quantity.

London interbank offered rate (Libor) Collective name for multiple rates at which a select set of banks believes they could borrow unsecured funds from other banks in the London interbank market for different currencies and different borrowing periods ranging from overnight to one year.

Long The buyer of a derivative contract. Also refers to the position of owning a derivative.

Long-lived assets Assets that are expected to provide economic benefits over a future period of time, typically greater than one year. Also called *long-term assets*.

Long position A position in an asset or contract in which one owns the asset or has an exercisable right under the contract.

Long-run average total cost The curve describing average total cost when no costs are considered fixed.

Look-ahead bias A bias caused by using information that was unavailable on the test date.

Loss aversion The tendency of people to dislike losses more than they like comparable gains.

Loss-aversion bias A bias in which people tend to strongly prefer avoiding losses as opposed to achieving gains.

Loss severity Portion of a bond's value (including unpaid interest) an investor loses in the event of default.

Losses Asset outflows not directly related to the ordinary activities of the business.

Lower bound The lowest possible value of an option.

M^2 A measure of what a portfolio would have returned if it had taken on the same total risk as the market index.

M^2 alpha Difference between the risk-adjusted performance of the portfolio and the performance of the benchmark.

Macaulay duration The approximate amount of time a bond would have to be held for the market discount rate at purchase to be realized if there is a single change in interest rate. It indicates the point in time when the coupon reinvestment and price effects of a change in yield-to-maturity offset each other.

Machine learning Computer based techniques that seek to extract knowledge from large amounts of data by "learning" from known examples and then generating structure or predictions. ML algorithms aim to "find the pattern, apply the pattern."

Macroeconomics The branch of economics that deals with aggregate economic quantities, such as national output and national income.

Maintenance covenants Covenants in bank loan agreements that require the borrower to satisfy certain financial ratio tests while the loan is outstanding.

Maintenance margin The minimum amount that is required by a futures clearinghouse to maintain a margin account and to protect against default. Participants whose margin balances drop below the required maintenance margin must replenish their accounts.

Maintenance margin requirement The margin requirement on any day other than the first day of a transaction.

Management buy-in Leveraged buyout in which the current management team is being replaced and the acquiring team will be involved in managing the company.

Management buyout A leveraged buyout event in which a group of investors consisting primarily of the company's existing management purchases at least controlling interest in its outstanding shares. At the extreme, they may purchase all shares and take the company private.

Management fee A fee based on assets under management or committed capital, as applicable—also called a *base fee*.

MAR ratio A variation of the Calmar ratio that uses a full investment history and the average drawdown.

Margin The amount of money that a trader deposits in a margin account. The term is derived from the stock market practice in which an investor borrows a portion of the money required to purchase a certain amount of stock. In futures markets, there is no borrowing so the margin is more of a down payment or performance bond.

Margin bond A cash deposit required by the clearinghouse from the participants to a contract to provide a credit guarantee. Also called a *performance bond*.

Margin call A request for the short to deposit additional funds to bring their balance up to the initial margin.

Margin loan Money borrowed from a broker to purchase securities.

Marginal cost The cost of producing an additional unit of a good.

Marginal frequencies The sums determined by adding joint frequencies across rows or across columns in a contingency table.

Marginal product Measures the productivity of each unit of input and is calculated by taking the difference in total product from adding another unit of input (assuming other resource quantities are held constant).

Marginal propensity to consume The proportion of an additional unit of disposable income that is consumed or spent; the change in consumption for a small change in income.

Marginal propensity to save The proportion of an additional unit of disposable income that is saved (not spent).

Marginal revenue The change in total revenue divided by the change in quantity sold; simply, the additional revenue from selling one more unit.

Marginal value curve A curve describing the highest price consumers are willing to pay for each additional unit of a good.

Mark to market The revaluation of a financial asset or liability to its current market value or fair value.

Mark-to-market Refers to the current expected fair market value for which a given security would likely be available for purchase or sale if traded in current market conditions.

Market anomaly Change in the price or return of a security that cannot directly be linked to current relevant information known in the market or to the release of new information into the market.

Market bid–ask spread The difference between the best bid and the best offer.

Market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its market capitalization by the total market capitalization (sum of the market capitalization) of all securities in the index. Also called *value weighting*.

Market discount rate The rate of return required by investors given the risk of the investment in a bond; also called the *required yield* or the *required rate of return*.

Market float The number of shares that are available to the investing public.

Market liquidity risk The risk that the price at which investors can actually transact—buying or selling—may differ from the price indicated in the market.

Market model A regression equation that specifies a linear relationship between the return on a security (or portfolio) and the return on a broad market index.

Market multiple models Valuation models based on share price multiples or enterprise value multiples.

Market-on-close An execution instruction specifying that an order can only be filled at the close of trading.

Market order Instructions to a broker or exchange to obtain the best price immediately available when filling an order.

Market-oriented investors With reference to equity investors, investors whose investment disciplines cannot be clearly categorized as value or growth.

Market rate of interest The rate demanded by purchasers of bonds, given the risks associated with future cash payment obligations of the particular bond issue.

Market reference rate (MRR) Collective name for a set of rates covering different currencies for different maturities, ranging from overnight to one year.

Market risk The risk that arises from movements in interest rates, stock prices, exchange rates, and commodity prices.

Market value The price at which an asset or security can currently be bought or sold in an open market.

Marketable limit order A buy limit order in which the limit price is placed above the best offer, or a sell limit order in which the limit price is placed below the best bid. Such orders generally will partially or completely fill right away.

Markowitz efficient frontier The graph of the set of portfolios offering the maximum expected return for their level of risk (standard deviation of return).

Matching principle The accounting principle that expenses should be recognized in the same period in which the associated revenue is recognized.

Matrix pricing Process of estimating the market discount rate and price of a bond based on the quoted or flat prices of more frequently traded comparable bonds.

Maturity premium An extra return that compensates investors for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended.

Maturity structure A factor explaining the differences in yields on similar bonds; also called *term structure*.

Mean absolute deviation With reference to a sample, the mean of the absolute values of deviations from the sample mean.

Mean square error (MSE) The sum of squares error divided by the degrees of freedom, $n - k - 1$; in a simple linear regression, $n - k - 1 = n - 2$.

Mean square regression (MSR) The sum of squares regression divided by the number of independent variables k; in a simple linear regression, $k = 1$.

Mean–variance analysis An approach to portfolio analysis using expected means, variances, and covariances of asset returns.

Measure of central tendency A quantitative measure that specifies where data are centered.

Measure of value A standard for measuring value; a function of money.

Measures of location Quantitative measures that describe the location or distribution of data. They include not only measures of central tendency but also other measures, such as percentiles.

Median The value of the middle item of a set of items that has been sorted into ascending or descending order (i.e., the 50th percentile).

Medium of exchange Any asset that can be used to purchase goods and services or to repay debts; a function of money.

Medium-term note A corporate bond offered continuously to investors by an agent of the issuer, designed to fill the funding gap between commercial paper and long-term bonds.

Mental accounting bias An information-processing bias in which people treat one sum of money differently from another equal-sized sum based on which mental account the money is assigned to.

Menu costs A cost of inflation in which businesses constantly have to incur the costs of changing the advertised prices of their goods and services.

Mesokurtic Describes a distribution with kurtosis equal to that of the normal distribution, namely, kurtosis equal to three.

Microeconomics The branch of economics that deals with markets and decision making of individual economic units, including consumers and businesses.

Minimum efficient scale The smallest output that a firm can produce such that its long-run average total cost is minimized.

Minimum-variance portfolio The portfolio with the minimum variance for each given level of expected return.

Minority shareholders Particular shareholders or a block of shareholders holding a small proportion of a company's outstanding shares, resulting in a limited ability to exercise control in voting activities.

Minsky moment Named for Hyman Minksy. A point in a business cycle when, after individuals become overextended in borrowing to finance speculative investments, people start realizing that something is likely to go wrong and a panic ensues, leading to asset sell-offs.

Modal interval With reference to grouped data, the interval containing the greatest number of observations (i.e., highest frequency).

Mode The most frequently occurring value in a distribution.

Modern portfolio theory (MPT) The analysis of rational portfolio choices based on the efficient use of risk.

Modified duration A measure of the percentage price change of a bond given a change in its yield-to-maturity.

Momentum oscillator A graphical representation of market sentiment that is constructed from price data and calculated so that it oscillates either between a low and a high or around some number.

Monetarists Economists who believe that the rate of growth of the money supply is the primary determinant of the rate of inflation.

Monetary policy Actions taken by a nation's central bank to affect aggregate output and prices through changes in bank reserves, reserve requirements, or its target interest rate.

Monetary transmission mechanism The process whereby a central bank's interest rate gets transmitted through the economy and ultimately affects the rate of increase of prices.

Monetary union An economic union in which the members adopt a common currency.

Money A generally accepted medium of exchange and unit of account.

Money convexity For a bond, the annual or approximate convexity multiplied by the full price.

Money creation The process by which changes in bank reserves translate into changes in the money supply.

Money duration A measure of the price change in units of the currency in which the bond is denominated given a change in its yield-to-maturity.

Money market The market for short-term debt instruments (one-year maturity or less).

Money market securities Fixed-income securities with maturities at issuance of one year or less.

Money multiplier Describes how a change in reserves is expected to affect the money supply; in its simplest form, 1 divided by the reserve requirement.

Money neutrality The thesis that an increase in the money supply leads in the long-run to an increase in the price level, while leaving real variables like output and employment unaffected.

Money-weighted return The internal rate of return on a portfolio, taking account of all cash flows.

Moneyness The relationship between the price of the underlying and an option's exercise price.

Monopolistic competition Highly competitive form of imperfect competition; the competitive characteristic is a notably large number of firms, while the monopoly aspect is the result of product differentiation.

Monopoly In pure monopoly markets, there are no substitutes for the given product or service. There is a single seller, which exercises considerable power over pricing and output decisions.

Monte Carlo simulation A technique involving the use of computer software to represent the operation of a complex financial system. A characteristic feature is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Moral principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Mortgage-backed securities Debt obligations that represent claims to the cash flows from pools of mortgage loans, most commonly on residential property.

Mortgage loan A loan secured by the collateral of some specified real estate property that obliges the borrower to make a predetermined series of payments to the lender.

Mortgage pass-through security A security created when one or more holders of mortgages form a pool of mortgages and sell shares or participation certificates in the pool.

Mortgage rate The interest rate on a mortgage loan; also called *contract rate* or *note rate*.

Moving average The average of the closing price of a security over a specified number of periods. With each new period, the average is recalculated.

Moving-average convergence/divergence oscillator A momentum oscillator that is based on the difference between short-term and long-term moving averages of a security's price.

Multi-factor model A model that explains a variable in terms of the values of a set of factors.

Multi-market indexes Comprised of indexes from different countries, designed to represent multiple security markets.

Multi-step format With respect to the format of the income statement, a format that presents a subtotal for gross profit (revenue minus cost of goods sold).

Multilateral trading facilities See *alternative trading systems*.

Multinational corporation A company operating in more than one country or having subsidiary firms in more than one country.

Multinomial formula (general formula for labeling problems) The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by $\frac{n!}{n_1!n_2!\ldots n_k!}$.

Multiple of invested capital (MOIC) A simplified calculation that measures the total value of all distributions and residual asset values relative to an initial total investment; also known as a *money multiple*.

Multiple testing problem The risk of getting statistically significant test results when performing a test multiple times.

Multiplication rule for counting If one task can be done in n_1 ways, and a second task, given the first, can be done in n_2 ways, and a third task, given the first two tasks, can be done in n_3 ways, and so on for k tasks, then the number of ways the k tasks can be done is $(n_1)(n_2)(n_3) \ldots (n_k)$.

Multiplication rule for expected value of the product of uncorrelated random variables This rule is $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated.

Multiplication rule for independent events The rule that when two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B.

Multiplication rule for probability The rule that the joint probability of events A and B equals the probability of A given B times the probability of B.

Multiplier models Valuation models based on share price multiples or enterprise value multiples.

Multivariate distribution A probability distribution that specifies the probabilities for a group of related random variables.

Multivariate normal distribution A probability distribution for a group of random variables that is completely defined by the means and variances of the variables plus all the correlations between pairs of the variables.

Muni See *municipal bond*.

Municipal bond A type of non-sovereign bond issued by a state or local government in the United States. It very often (but not always) offers income tax exemptions.

Mutual fund A comingled investment pool in which investors in the fund each have a pro-rata claim on the income and value of the fund.

Mutually exclusive Indicates that only one event can occur at a time.

Mutually exclusive projects Mutually exclusive projects compete directly with each other. For example, if Projects A and B are mutually exclusive, you can choose A or B but you cannot choose both.

***n* Factorial** For a positive integer n, the product of the first n positive integers; 0 factorial equals 1 by definition. n factorial is written as $n!$.

Narrow money The notes and coins in circulation in an economy, plus other very highly liquid deposits.

Nash equilibrium When two or more participants in a non-coop-erative game have no incentive to deviate from their respective equilibrium strategies given their opponent's strategies.

National income The income received by all factors of production used in the generation of final output. National income equals gross domestic product (or, in some countries, gross national product) minus the capital consumption allowance and a statistical discrepancy.

Natural language processing Computer programs developed to analyze and interpret human language.

Natural rate of unemployment Effective unemployment rate, below which pressure emerges in labor markets.

Negative screening An ESG investment style that focuses on the exclusion of certain sectors, companies, or practices in a fund or portfolio on the basis of specific ESG criteria.

Net book value The remaining (undepreciated) balance of an asset's purchase cost. For liabilities, the face value of a bond minus any unamortized discount, or plus any unamortized premium.

Net exports The difference between the value of a country's exports and the value of its imports (i.e., value of exports minus imports).

Net income The difference between revenue and expenses; what remains after subtracting all expenses (including depreciation, interest, and taxes) from revenue.

Net present value The present value of an investment's cash inflows (benefits) minus the present value of its cash outflows (costs).

Net profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also called *profit margin* or *return on sales*.

Net realisable value Estimated selling price in the ordinary course of business less the estimated costs necessary to make the sale.

Net revenue Revenue after adjustments (e.g., for estimated returns or for amounts unlikely to be collected).

Net tax rate The tax rate net of transfer payments.

Neural networks Computer programs based on how our own brains learn and process information.

Neutral rate of interest The rate of interest that neither spurs on nor slows down the underlying economy.

No-load fund A mutual fund in which there is no fee for investing in the fund or for redeeming fund shares, although there is an annual fee based on a percentage of the fund's net asset value.

Node Each value on a binomial tree from which successive moves or outcomes branch.

Nominal data Categorical values that are not amenable to being organized in a logical order. An example of nominal data is the classification of publicly listed stocks into sectors.

Nominal GDP The value of goods and services measured at current prices.

Nominal risk-free interest rate The sum of the real risk-free interest rate and the inflation premium.

Non-accelerating inflation rate of unemployment Effective unemployment rate below which pressure emerges in labor markets.

Non-agency RMBS In the United States, securities issued by private entities that are not guaranteed by a federal agency or a GSE.

Non-Bank lender Unlike typical banks, which make loans and take deposits, these lenders only make loans. Although they lend money to companies, they usually specialize in issuing mortgages to first-time homebuyers and parties looking to refinance an existing mortgage.

Non-cumulative preference shares Preference shares for which dividends that are not paid in the current or subsequent periods are forfeited permanently (instead of being accrued and paid at a later date).

Non-current assets Assets that are expected to benefit the company over an extended period of time (usually more than one year).

Non-current liabilities Obligations that broadly represent a probable sacrifice of economic benefits in periods generally greater than one year in the future.

Non-cyclical A company whose performance is largely independent of the business cycle. Also known as defensive.

Non-deliverable forwards Cash-settled forward contracts, used predominately with respect to foreign exchange forwards. Also called *contracts for differences*.

Non-financial risks Risks that arise from sources other than changes in the external financial markets, such as changes in accounting rules, legal environment, or tax rates.

Non-participating preference shares Preference shares that do not entitle shareholders to share in the profits of the company. Instead, shareholders are only entitled to receive a fixed dividend payment and the par value of the shares in the event of liquidation.

Non-probability sampling A sampling plan dependent on factors other than probability considerations, such as a sampler's judgment or the convenience to access data.

Non-recourse loan A loan in which the lender does not have a shortfall claim against the borrower, so the lender can look only to the property to recover the outstanding mortgage balance.

Non-renewable resources Finite resources that are depleted once they are consumed, such as oil and coal.

Non-sovereign bond A bond issued by a government below the national level, such as a province, region, state, or city.

Nonconventional cash flows In a nonconventional cash flow pattern, the initial outflow is not followed by inflows only, but the cash flows can flip from positive (inflows) to negative (outflows) again or even change signs several times.

Nonparametric test A test that is not concerned with a parameter or that makes minimal assumptions about the population from which a sample comes.

Nonsystematic risk Unique risk that is local or limited to a particular asset or industry that need not affect assets outside of that asset class.

Normal distribution A continuous, symmetric probability distribution that is completely described by its mean and its variance.

Normal goods Goods that are consumed in greater quantities as income increases.

Normal profit The level of accounting profit needed to just cover the implicit opportunity costs ignored in accounting costs.

Notching Ratings adjustment methodology where specific issues from the same borrower may be assigned different credit ratings.

Note rate See *mortgage rate*.

Notice period The length of time (typically 30 to 90 days) in advance that investors may be required to notify a fund of their intent to redeem some or all of their investment.

Notional principal An imputed principal amount.

Null hypothesis The hypothesis that is tested.

Numerical data Values that represent measured or counted quantities as a number. Also called **quantitative data**.

Objective probabilities Probabilities that generally do not vary from person to person; includes a priori and empirical probabilities.

Observation The value of a specific variable collected at a point in time or over a specified period of time.

Odds against *E* The reciprocal of odds for *E*.

Odds for *E* The probability of *E* divided by 1 minus the probability of *E*.

Off-the-run Seasoned government bonds are off-the-run securities; they are not the most recently issued or the most actively traded.

Offer The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size).

Official interest rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks. Also called *official policy rate* or *policy rate*.

Official policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Oligopoly Market structure with a relatively small number of firms supplying the market.

On-the-run The most recently issued and most actively traded sovereign securities.

One-dimensional array The simplest format for representing a collection of data of the same data type.

One-sided hypothesis test A test in which the null hypothesis is rejected only if the evidence indicates that the population parameter is greater than or less than the hypothesized parameter; occurs when the alternative hypothesis is stated either as greater than or less than the hypothesized population parameter.

Open economy An economy that trades with other countries.

Open-end fund A mutual fund that accepts new investment money and issues additional shares at a value equal to the net asset value of the fund at the time of investment.

Open interest The number of outstanding contracts in a clearinghouse at any given time. The open interest figure changes daily as some parties open up new positions, while other parties offset their old positions.

Open market operations The purchase or sale of bonds by the national central bank to implement monetary policy. The bonds traded are usually sovereign bonds issued by the national government.

Operating activities Activities that are part of the day-to-day business functioning of an entity, such as selling inventory and providing services.

Operating breakeven The number of units produced and sold at which the company's operating profit is zero (revenues = operating costs).

Operating cash flow The net amount of cash provided from operating activities.

Operating efficiency ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Operating lease A type of lease which is more akin to the rental of the underlying asset.

Operating leverage The use of fixed costs in operations.

Operating profit A company's profits on its usual business activities before deducting taxes. Also called *operating income*.

Operating profit margin A profitability ratio calculated as operating income (i.e., income before interest and taxes) divided by revenue. Also called *operating margin*.

Operating risk The risk attributed to the operating cost structure, in particular the use of fixed costs in operations; the risk arising from the mix of fixed and variable costs; the risk that a company's operations may be severely affected by environmental, social, and governance risk factors.

Operational independence A bank's ability to execute monetary policy and set interest rates in the way it thought would best meet the inflation target.

Operational risk The risk that arises from inadequate or failed people, systems, and internal policies, procedures, and processes, as well as from external events that are beyond the control of the organization but that affect its operations.

Operationally efficient Said of a market, a financial system, or an economy that has relatively low transaction costs.

Opportunity cost The value that investors forgo by choosing a particular course of action; the value of something in its best alternative use.

Optimal capital structure The capital structure at which the value of the company is maximized.

Option A financial instrument that gives one party the right, but not the obligation, to buy or sell an underlying asset from or to another party at a fixed price over a specific period of time. Also referred to as *contingent claim* or *option contract.*

Option-adjusted price The value of the embedded option plus the flat price of the bond.

Option-adjusted spread OAS = Z-spread – Option value (in basis points per year).

Option-adjusted yield The required market discount rate whereby the price is adjusted for the value of the embedded option.

Option contract See *option*.

Option premium The amount of money a buyer pays and seller receives to engage in an option transaction.

Order A specification of what instrument to trade, how much to trade, and whether to buy or sell.

Order-driven markets A market (generally an auction market) that uses rules to arrange trades based on the orders that traders submit; in their pure form, such markets do not make use of dealers.

Order precedence hierarchy With respect to the execution of orders to trade, a set of rules that determines which orders execute before other orders.

Ordinal data Categorical values that can be logically ordered or ranked.

Ordinary annuity An annuity with a first cash flow that is paid one period from the present.

Ordinary shares Equity shares that are subordinate to all other types of equity (e.g., preferred equity). Also called *common stock* or *common shares.*

Organized exchange A securities marketplace where buyers and seller can meet to arrange their trades.

Other comprehensive income Items of comprehensive income that are not reported on the income statement; comprehensive income minus net income.

Out-of-sample test A test of a strategy or model using a sample outside the period on which the strategy or model was developed.

Out of the money Options that, if exercised, would require the payment of more money than the value received and therefore would not be currently exercised.

Outcome A possible value of a random variable.

Over-the-counter (OTC) market A decentralized market where buy and sell orders initiated from various locations are matched through a communications network.

Overbought A market condition in which market sentiment is thought to be unsustainably bullish.

Overcollateralization Form of internal credit enhancement that refers to the process of posting more collateral than needed to obtain or secure financing.

Overconfidence bias A bias in which people demonstrate unwarranted faith in their own intuitive reasoning, judgments, and/or cognitive abilities.

Overfitting An undesirable result from fitting a model so closely to a dataset that it does not perform well on new data.

Oversold A market condition in which market sentiment is thought to be unsustainably bearish.

Own price The price of a good or service itself (as opposed to the price of something else).

Own-price elasticity of demand The percentage change in quantity demanded for a percentage change in good's own price, holding all other things constant.

Owners' equity The excess of assets over liabilities; the residual interest of shareholders in the assets of an entity after deducting the entity's liabilities. Also called *shareholders' equity* or *shareholders' funds*.

***p*-Value** The smallest level of significance at which the null is rejected.

Paired comparisons test See *test of the mean of the differences.*

Panel data A mix of time-series and cross-sectional data that contains observations through time on characteristics of across multiple observational units.

Par curve A sequence of yields-to-maturity such that each bond is priced at par value. The bonds are assumed to have the same currency, credit risk, liquidity, tax status, and annual yields stated for the same periodicity.

Par value The amount of principal on a bond.

Parallel shift A parallel yield curve shift implies that all rates change by the same amount in the same direction.

Parameter A descriptive measure computed from or used to describe a population of data, conventionally represented by Greek letters.

Parametric test Any test (or procedure) concerned with parameters or whose validity depends on assumptions concerning the population generating the sample.

Pari passu Covenant or contract clause that ensures a debt obligation is treated the same as the borrower's other senior debt instruments and is not subordinated to similar obligations.

Partial duration See *key rate duration.*

Participating preference shares Preference shares that entitle shareholders to receive the standard preferred dividend plus the opportunity to receive an additional dividend if the company's profits exceed a pre-specified level.

Pass-through rate The coupon rate of a mortgage pass-through security.

Passive investment A buy and hold approach in which an investor does not make portfolio changes based on short-term expectations of changing market or security performance.

Payable date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payment date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payments system The system for the transfer of money.

Pearson correlation A parametric measure of the relationship between two variables.

Pecking order theory The theory that managers consider how their actions might be interpreted by outsiders and thus order their preferences for various forms of corporate financing. Forms of financing that are least visible to outsiders (e.g., internally generated funds) are most preferable to managers, and those that are most visible (e.g., equity) are least preferable.

Peer company See *comparable company*.

Peer group A group of companies engaged in similar business activities whose economics and valuation are influenced by closely related factors.

Pennant A technical analysis continuation pattern formed by trendlines that converge to form a triangle, typically over a short period.

Per capita real GDP Real GDP divided by the size of the population, often used as a measure of a country's average standard of living.

Per unit contribution margin The amount that each unit sold contributes to covering fixed costs—that is, the difference between the price per unit and the variable cost per unit.

Percentiles Quantiles that divide a distribution into 100 equal parts that sum to 100.

Perfect capital markets Markets in which, by assumption, there are no taxes, transaction costs, or bankruptcy costs and in which all investors have equal ("symmetric") information.

Perfect competition A market structure in which the individual firm has virtually no impact on market price, because it is assumed to be a very small seller among a very large number of firms selling essentially identical products.

Perfectly elastic When the quantity demanded or supplied of a given good is infinitely sensitive to a change in the value of a specified variable (e.g., price).

Perfectly inelastic When the quantity demanded or supplied of a given good is completely insensitive to a change in the value of a specified variable (e.g., price).

Performance bond See *margin bond.*

Performance evaluation The measurement and assessment of the outcomes of investment management decisions.

Performance fee Fees paid to the general partner from the limited partner(s) based on realized net profits.

Period costs Costs (e.g., executives' salaries) that cannot be directly matched with the timing of revenues and which are thus expensed immediately.

Periodicity The assumed number of periods in the year; typically matches the frequency of coupon payments.

Permanent differences Differences between tax and financial reporting of revenue (expenses) that will not be reversed at some future date. These result in a difference between the company's effective tax rate and statutory tax rate and do not result in a deferred tax item.

Permissioned networks Networks that are fully open only to select participants on a DLT network.

Permissionless networks Networks that are fully open to any user on a DLT network.

Permutation An ordered listing.

Permutation formula The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is ${}_nP_r = \dfrac{n!}{(n-r)!}$.

Perpetual bonds Bonds with no stated maturity date.

Perpetuity A perpetual annuity, or a set of never-ending level sequential cash flows, with the first cash flow occurring one period from now. A bond that does not mature.

Personal income A broad measure of household income that includes all income received by households, whether earned or unearned; measures the ability of consumers to make purchases.

Pet projects Investments in which influential managers want the corporation to invest. Often, unfortunately, pet projects are selected without undergoing normal capital allocation analysis.

Plain vanilla bond Bond that makes periodic, fixed coupon payments during the bond's life and a lump-sum payment of principal at maturity. Also called *conventional bond*.

Platykurtic Describes a distribution that has relatively less weight in the tails than the normal distribution (also called thin-tailed).

Point estimate A single numerical estimate of an unknown quantity, such as a population parameter.

Policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Population All members of a specified group.

Portfolio company In private equity, the company in which the private equity fund is investing.

Portfolio demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments.

Portfolio planning The process of creating a plan for building a portfolio that is expected to satisfy a client's investment objectives.

Position The quantity of an asset that an entity owns or owes.

Positive screening The process of including sectors or companies based on specific ESG criteria, typically ESG performance relative to industry peers

Posterior probability An updated probability that reflects or comes after new information.

Potential GDP The level of real GDP that can be produced at full employment; measures the productive capacity of the economy.

Power of a test The probability of correctly rejecting the null—that is, rejecting the null hypothesis when it is false.

Precautionary money balances Money held to provide a buffer against unforeseen events that might require money.

Preference shares A type of equity interest which ranks above common shares with respect to the payment of dividends and the distribution of the company's net assets upon liquidation. They have characteristics of both debt and equity securities. Also called *preferred stock*.

Preferred stock See *preference shares*.

Premium In the case of bonds, premium refers to the amount by which a bond is priced above its face (par) value. In the case of an option, the amount paid for the option contract.

Prepaid expense A normal operating expense that has been paid in advance of when it is due.

Prepayment option Contractual provision that entitles the borrower to prepay all or part of the outstanding mortgage principal prior to the scheduled due date when the principal must be repaid. Also called *early repayment option*.

Prepayment penalty mortgages Mortgages that stipulate a monetary penalty if a borrower prepays within a certain time period after the mortgage is originated.

Prepayment risk The uncertainty that the timing of the actual cash flows will be different from the scheduled cash flows as set forth in the loan agreement due to the borrowers' ability to alter payments, usually to take advantage of interest rate movements.

Present value (PV) The present discounted value of future cash flows: For assets, the present discounted value of the future net cash inflows that the asset is expected to generate; for liabilities, the present discounted value of the future net cash outflows that are expected to be required to settle the liabilities.

Present value models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security. Also called *discounted cash flow models*.

Pretax margin A profitability ratio calculated as earnings before taxes divided by revenue.

Price elasticity of demand Measures the percentage change in the quantity demanded, given a percentage change in the price of a given product.

Price index Represents the average prices of a basket of goods and services.

Price limits Limits imposed by a futures exchange on the price change that can occur from one day to the next.

Price multiple A ratio that compares the share price with some sort of monetary flow or value to allow evaluation of the relative worth of a company's stock.

Price priority The principle that the highest priced buy orders and the lowest priced sell orders execute first.

Price relative A ratio of an ending price over a beginning price; it is equal to 1 plus the holding period return on the asset.

Price return Measures *only* the price appreciation or percentage change in price of the securities in an index or portfolio.

Price return index An index that reflects *only* the price appreciation or percentage change in price of the constituent securities. Also called *price index*.

Price-setting option The option to adjust prices when demand varies from what is forecast.

Price stability In economics, refers to an inflation rate that is low on average and not subject to wide fluctuation.

Price takers Producers that must accept whatever price the market dictates.

Price to book value A valuation ratio calculated as price per share divided by book value per share.

Price to cash flow A valuation ratio calculated as price per share divided by cash flow per share.

Price to earnings ratio (P/E ratio or P/E) The ratio of share price to earnings per share.

Price to sales A valuation ratio calculated as price per share divided by sales per share.

Price value of a basis point A version of money duration, it is an estimate of the change in the full price of a bond given a 1 basis point change in the yield-to-maturity.

Price weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its price by the sum of all the prices of the constituent securities.

Priced risk Risk for which investors demand compensation for bearing (e.g., equity risk, company-specific factors, macroeconomic factors).

Primary bond market A market in which issuers first sell bonds to investors to raise capital.

Primary capital markets (primary markets) The market where securities are first sold and the issuers receive the proceeds.

Primary dealer Financial institution that is authorized to deal in new issues of sovereign bonds and that serves primarily as a trading counterparty of the office responsible for issuing sovereign bonds.

Primary market The market where securities are first sold and the issuers receive the proceeds.

Prime brokers Brokers that provide services that commonly include custody, administration, lending, short borrowing, and trading.

Principal The amount of funds originally invested in a project or instrument; the face value to be paid at maturity.

Principal–agent relationship A relationship in which a principal hires an agent to perform a particular task or service; also known as an *agency relationship*.

Principal amount Amount that an issuer agrees to repay the debtholders on the maturity date.

Principal business activity The business activity from which a company derives a majority of its revenues and/or earnings.

Principal value Amount that an issuer agrees to repay the debtholders on the maturity date.

Principle of no arbitrage See *arbitrage-free pricing*.

Prior probabilities Probabilities reflecting beliefs prior to the arrival of new information.

Priority of claims Priority of payment, with the most senior or highest ranking debt having the first claim on the cash flows and assets of the issuer.

Private equity fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *venture capital fund*.

Private equity funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *venture capital funds*.

Private equity securities Securities that are not listed on public exchanges and have no active secondary market. They are issued primarily to institutional investors via non-public offerings, such as private placements.

Private investment in public equity (PIPE) An investment in the equity of a publicly traded firm that is made at a discount to the market value of the firm's shares.

Private placement Typically, a non-underwritten, unregistered offering of securities that are sold only to an investor or a small group of investors. It can be accomplished directly between the issuer and the investor(s) or through an investment bank.

Probability A number between 0 and 1 describing the chance that a stated event will occur.

Probability density function A function with non-negative values such that probability can be described by areas under the curve graphing the function.

Probability distribution A distribution that specifies the probabilities of a random variable's possible outcomes.

Probability function A function that specifies the probability that the random variable takes on a specific value.

Probability sampling A sampling plan that allows every member of the population to have an equal chance of being selected.

Probability tree diagram A diagram with branches emanating from nodes representing either mutually exclusive chance events or mutually exclusive decisions.

Production-flexibility options The options to alter production when demand varies from what is forecast.

Production function Provides the quantitative link between the levels of output that the economy can produce and the inputs used in the production process.

Productivity The amount of output produced by workers during a given period—for example, output per hour worked measures the efficiency of labor.

Profession An occupational group that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition.

Profit The return that owners of a company receive for the use of their capital and the assumption of financial risk when making their investments.

Profit and loss (P&L) statement A financial statement that provides information about a company's profitability over a stated period of time. Also called the *income statement*.

Profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses.

Profitability ratios Ratios that measure a company's ability to generate profitable sales from its resources (assets).

Project sequencing To defer the decision to invest in a future project until the outcome of some or all of a current investment is known. Investments are sequenced over time, so that making an investment creates the option to invest in future projects.

Promissory note A written promise to pay a certain amount of money on demand.

Property, plant, and equipment Tangible assets that are expected to be used for more than one period in either the production or supply of goods or services, or for administrative purposes.

Prospectus The document that describes the terms of a new bond issue and helps investors perform their analysis on the issue.

Protective put An option strategy in which a long position in an asset is combined with a long position in a put.

Proxy contest Corporate takeover mechanism in which shareholders are persuaded to vote for a group seeking a controlling position on a company's board of directors.

Proxy voting A process that enables shareholders who are unable to attend a meeting to authorize another individual to vote on their behalf.

Public offer See *public offering*.

Public offering An offering of securities in which any member of the public may buy the securities. Also called *public offer*.

Public–private partnership (PPP) An agreement between the public sector and the private sector to finance, build, and operate public infrastructure, such as hospitals and toll roads.

Pull on liquidity When disbursements are paid too quickly or trade credit availability is limited, requiring companies to expend funds before they receive funds from sales that could cover the liability.

Pure discount bonds See *zero-coupon bond*.

Put An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Put–call–forward parity The relationship among puts, calls, and forward contracts.

Put–call parity An equation expressing the equivalence (parity) of a portfolio of a call and a bond with a portfolio of a put and the underlying, which leads to the relationship between put and call prices.

Put/call ratio A technical analysis indicator that evaluates market sentiment based on the volume of put options traded divided by the volume of call options traded for a particular financial instrument.

Put option An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Putable bonds Bonds that give the bondholder the right to sell the bond back to the issuer at a predetermined price on specified dates.

Qualitative data see **categorical data**.

Quantile A value at or below which a stated fraction of the data lies. Also referred to as a fractile.

Quantitative data see **numerical data**.

Quantitative easing An expansionary monetary policy based on aggressive open market purchase operations.

Quantity equation of exchange An expression that over a given period, the amount of money used to purchase all goods and services in an economy, $M \times V$, is equal to monetary value of this output, $P \times Y$.

Quantity theory of money Asserts that total spending (in money terms) is proportional to the quantity of money.

Quartiles Quantiles that divide a distribution into four equal parts.

Quasi-fixed cost A cost that stays the same over a range of production but can change to another constant level when production moves outside of that range.

Quasi-government bond A bond issued by an entity that is either owned or sponsored by a national government. Also called *agency bond*.

Quintiles Quantiles that divide a distribution into five equal parts.

Quota rents Profits that foreign producers can earn by raising the price of their goods higher than they would without a quota.

Quotas Government policies that restrict the quantity of a good that can be imported into a country, generally for a specified period of time.

Quote-driven market A market in which dealers acting as principals facilitate trading.

Quoted interest rate A quoted interest rate that does not account for compounding within the year. Also called *stated annual interest rate*.

Quoted margin The specified yield spread over the reference rate, used to compensate an investor for the difference in the credit risk of the issuer and that implied by the reference rate.

Random number An observation drawn from a uniform distribution.

Random number generator An algorithm that produces uniformly distributed random numbers between 0 and 1.

Random variable A quantity whose future outcomes are uncertain.

Range The difference between the maximum and minimum values in a dataset.

Raw data Data available in their original form as collected.

Real exchange rate effect The effect through which changing price level impacts real exchange rate which in turn impacts net exports and aggregate demand.

Real GDP The value of goods and services produced, measured at base year prices.

Real income Income adjusted for the effect of inflation on the purchasing power of money. Also known as the *purchasing power of income*. If income remains constant and a good's price falls, real income is said to rise, even though the number of monetary units (e.g., dollars) remains unchanged.

Real interest rate Nominal interest rate minus the expected rate of inflation.

Real risk-free interest rate The single-period interest rate for a completely risk-free security if no inflation were expected.

Realizable (settlement) value With reference to assets, the amount of cash or cash equivalents that could currently be obtained by selling the asset in an orderly disposal; with reference to liabilities, the undiscounted amount of cash or cash equivalents expected to be paid to satisfy the liabilities in the normal course of business.

Rebalancing Adjusting the weights of the constituent securities in an index.

Rebalancing policy The set of rules that guide the process of restoring a portfolio's asset class weights to those specified in the strategic asset allocation.

Recession A period during which real GDP decreases (i.e., negative growth) for at least two successive quarters, or a period of significant decline in total output, income, employment, and sales usually lasting from six months to a year.

Recognition lag The lag in government response to an economic problem resulting from the delay in confirming a change in the state of the economy.

Recourse loan A loan in which the lender has a claim against the borrower for any shortfall between the outstanding mortgage balance and the proceeds received from the sale of the property.

Redemption yield See *yield-to-maturity*.

Redemptions Withdrawals of funds by investors, as allowed by the notice period and other terms in the partnership agreement.

Refinancing rate A type of central bank policy rate.

Registered bonds Bonds for which ownership is recorded by either name or serial number.

Regression analysis A tool for examining whether a variable is useful for explaining another variable.

Regression coefficients The intercept and slope coefficient of a regression.

Regret The feeling that an opportunity has been missed; typically, an expression of *hindsight bias*.

Regret-aversion bias An emotional bias in which people tend to avoid making decisions that will result in action out of fear that the decision will turn out poorly.

Relative dispersion The amount of dispersion relative to a reference value or benchmark.

Relative frequency The absolute frequency of each unique value of the variable divided by the total number of observations of the variable.

Relative price The price of a specific good or service in comparison with those of other goods and services.

Relative strength analysis A comparison of the performance of one asset with the performance of another asset or a benchmark, based on changes in the ratio of the two assets' prices over time.

Relative strength index (RSI) A technical analysis momentum oscillator that compares a security's gains with its losses over a set period.

Renewable resources Resources that can be replenished, such as a forest.

Rent Payment for the use of property.

Reorganization Agreements made by a company in bankruptcy under which a company's capital structure is altered and/or alternative arrangements are made for debt repayment; US Chapter 11 bankruptcy. The company emerges from bankruptcy as a going concern.

Replication The creation of an asset or portfolio from another asset, portfolio, and/or derivative.

Repo A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy back the same security from the purchaser at an agreed-on price and future date. The party who sells the

security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repo margin The difference between the market value of the security used as collateral and the value of the loan. Also called *haircut*.

Repo rate The interest rate on a repurchase agreement.

Representativeness bias A belief perseverance bias in which people tend to classify new information based on past experiences and classifications.

Repurchase agreement A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy the same security back from the purchaser at an agreed-on price and future date. The party who sells the security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repurchase date The date when the party who sold the security at the inception of a repurchase agreement buys back the security from the cash lending counterparty.

Repurchase price The price at which the party who sold the security at the inception of the repurchase agreement buys back the security from the cash lending counterparty.

Required margin The yield spread over or under the reference rate such that an FRN is priced at par value on a rate reset date.

Required rate of return See *market discount rate*.

Required yield See *market discount rate*.

Required yield spread The difference between the yield-to-maturity on a new bond and the benchmark rate; additional compensation required by investors for the difference in risk and tax status of a bond relative to a government bond. Sometimes called the *spread over the benchmark*.

Resampling A statistical method that repeatedly draws samples from the original observed data sample for the statistical inference of population parameters.

Reserve accounts Form of internal credit enhancement that relies on creating accounts and depositing in these accounts cash that can be used to absorb losses. Also called *reserve funds*.

Reserve funds See *reserve accounts*.

Reserve requirement The requirement for banks to hold reserves in proportion to the size of deposits.

Residual The difference between an observation and its predicted value, where the predicted value is based on the estimated linear relation between the dependent and independent variables using sample data.

Resistance In technical analysis, a price range in which selling activity is sufficient to stop the rise in the price of a security.

Responsible investing An overall (umbrella) term for several investment strategies—ESG investing, SRI investing, thematic investing, and impact investing—that incorporate ESG analysis in their investment processes.

Restricted payments A bond covenant meant to protect creditors by limiting how much cash can be paid out to shareholders over time.

Retracement In technical analysis, a reversal in the movement of a security's price such that it is counter to the prevailing longer-term price trend.

Return-generating model A model that can provide an estimate of the expected return of a security given certain parameters and estimates of the values of the independent variables in the model.

Return on assets (ROA) A profitability ratio calculated as net income divided by average total assets; indicates a company's net profit generated per dollar invested in total assets.

Return on equity (ROE) A profitability ratio calculated as net income divided by average shareholders' equity.

Return on invested capital A measure of the profitability of a company relative to the amount of capital invested by the equity- and debtholders.

Return on sales An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also referred to as *net profit margin*.

Return on total capital A profitability ratio calculated as EBIT divided by the sum of short- and long-term debt and equity.

Revaluation model Under IFRS, the process of valuing long-lived assets at fair value, rather than at cost less accumulated depreciation. Any resulting profit or loss is either reported on the income statement and/or through equity under revaluation surplus.

Revenue The amount charged for the delivery of goods or services in the ordinary activities of a business over a stated period; the inflows of economic resources to a company over a stated period.

Reversal pattern A type of pattern used in technical analysis to predict the end of a trend and a change in the direction of a security's price.

Reverse repo A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse repurchase agreement A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse stock split A reduction in the number of shares outstanding with a corresponding increase in share price, but no change to the company's underlying fundamentals.

Revolving credit agreements (also known as revolvers) The most reliable form of short-term bank borrowing facilities; they are in effect for multiple years (e.g., three to five years) and can have optional medium-term loan features.

Rho The sensitivity of the option price to the risk-free rate.

Ricardian equivalence An economic theory that implies that it makes no difference whether a government finances a deficit by increasing taxes or issuing debt.

Risk Exposure to uncertainty. The chance of a loss or adverse outcome as a result of an action, inaction, or external event.

Risk averse The assumption that an investor will choose the least risky alternative.

Risk aversion The degree of an investor's inability and unwillingness to take risk.

Risk budgeting The establishment of objectives for individuals, groups, or divisions of an organization that takes into account the allocation of an acceptable level of risk.

Risk exposure The state of being exposed or vulnerable to a risk. The extent to which an organization is sensitive to underlying risks.

Risk governance The top-down process and guidance that directs risk management activities to align with and support the overall enterprise.

Risk management The process of identifying the level of risk an organization wants, measuring the level of risk the organization currently has, taking actions that bring the actual

level of risk to the desired level of risk, and monitoring the new actual level of risk so that it continues to be aligned with the desired level of risk.

Risk management framework The infrastructure, process, and analytics needed to support effective risk management in an organization.

Risk-neutral pricing Sometimes said of derivatives pricing, uses the fact that arbitrage opportunities guarantee that a risk-free portfolio consisting of the underlying and the derivative must earn the risk-free rate.

Risk-neutral probabilities Weights that are used to compute a binomial option price. They are the probabilities that would apply if a risk-neutral investor valued an option.

Risk premium An extra return expected by investors for bearing some specified risk.

Risk shifting Actions to change the distribution of risk outcomes.

Risk tolerance The amount of risk an investor is willing and able to bear to achieve an investment goal.

Risk transfer Actions to pass on a risk to another party, often, but not always, in the form of an insurance policy.

Robo-adviser A machine-based analytical tool or service that provides technology-driven investment solutions through online platforms.

Rule of 72 The principle that the approximate number of years necessary for an investment to double is 72 divided by the stated interest rate.

Running yield See *current yield*.

Safety-first rules Rules for portfolio selection that focus on the risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Sales Generally, a synonym for revenue; "sales" is generally understood to refer to the sale of goods, whereas "revenue" is understood to include the sale of goods or services.

Sales risk Uncertainty with respect to the quantity of goods and services that a company is able to sell and the price it is able to achieve; the risk related to the uncertainty of revenues.

Sample A subset of a population.

Sample correlation coefficient A standardized measure of how two variables in a sample move together. It is the ratio of the sample covariance to the product of the two variables' standard deviations.

Sample covariance A measure of how two variables in a sample move together.

Sample excess kurtosis A sample measure of the degree of a distribution's kurtosis in excess of the normal distribution's kurtosis.

Sample mean The sum of the sample observations divided by the sample size.

Sample selection bias Bias introduced by systematically excluding some members of the population according to a particular attribute—for example, the bias introduced when data availability leads to certain observations being excluded from the analysis.

Sample-size neglect A type of representativeness bias in which financial market participants incorrectly assume that small sample sizes are representative of populations (or "real" data).

Sample skewness A sample measure of the degree of asymmetry of a distribution.

Sample standard deviation The positive square root of the sample variance.

Sample statistic A quantity computed from or used to describe a sample.

Sample variance The sum of squared deviations around the mean divided by the degrees of freedom.

Sampling The process of obtaining a sample.

Sampling distribution The distribution of all distinct possible values that a statistic can assume when computed from samples of the same size randomly drawn from the same population.

Sampling error The difference between the observed value of a statistic and the estimate resulting from using subsets of the population.

Sampling plan The set of rules used to select a sample.

Say on pay A process whereby shareholders may vote on executive remuneration (compensation) matters.

Scatter plot A chart in which two variables are plotted along the axis and points on the chart represent pairs of the two variables. In regression, the dependent variable is plotted on the vertical axis and the independent variable is plotted along the horizontal axis. Also known as a scattergram and a *scatter diagram*.

Scatter plot matrix A tool for organizing scatter plots between pairs of variables, making it easy to inspect all pairwise relationships in one combined visual.

Scenario analysis Analysis that shows the changes in key financial quantities that result from given (economic) events, such as the loss of customers, the loss of a supply source, or a catastrophic event; a risk management technique involving examination of the performance of a portfolio under specified situations. Closely related to stress testing.

Screening The application of a set of criteria to reduce a set of potential investments to a smaller set having certain desired characteristics.

Seasoned offering An offering in which an issuer sells additional units of a previously issued security.

Second-degree price discrimination When the monopolist charges different per-unit prices using the quantity purchased as an indicator of how highly the customer values the product.

Second lien A secured interest in the pledged assets that ranks below first lien debt in both collateral protection and priority of payment.

Secondary bond markets Markets in which existing bonds are traded among investors.

Secondary market The market where securities are traded among investors.

Secondary precedence rules Rules that determine how to rank orders placed at the same time.

Sector A group of related industries (GICS definition).

Sector indexes Indexes that represent and track different economic sectors—such as consumer goods, energy, finance, health care, and technology—on either a national, regional, or global basis.

Secured ("asset-based") loan A loan that is backed by specific, secured company assets.

Secured bonds Bonds secured by assets or financial guarantees pledged to ensure debt repayment in case of default.

Secured debt Debt in which the debtholder has a direct claim—a pledge from the issuer—on certain assets and their associated cash flows.

Securitization A process that involves moving assets into a special legal entity, which then uses the assets as guarantees to secure a bond issue.

Securitized assets Assets that are typically used to create asset-backed bonds; for example, when a bank securitizes a pool of loans, the loans are said to be securitized.

Security characteristic line A plot of the excess return of a security on the excess return of the market.

Security market index A portfolio of securities representing a given security market, market segment, or asset class.

Security market line (SML) The graph of the capital asset pricing model.

Security selection The process of selecting individual securities; typically, security selection has the objective of generating superior risk-adjusted returns relative to a portfolio's benchmark.

Self-attribution bias A bias in which people take too much credit for successes (*self-enhancing*) and assign responsibility to others for failures (*self-protecting*).

Self-control bias A bias in which people fail to act in pursuit of their long-term, overarching goals because of a lack of self-discipline.

Self-investment limits With respect to investment limitations applying to pension plans, restrictions on the percentage of assets that can be invested in securities issued by the pension plan sponsor.

Sell-side firm A broker/dealer that sells securities and provides independent investment research and recommendations to their clients (i.e., buy-side firms).

Semi-strong-form efficient market A market in which security prices reflect all publicly known and available information.

Semiannual bond basis yield An annual rate having a periodicity of two; also known as a *semiannual bond equivalent yield*.

Semiannual bond equivalent yield See *semiannual bond basis yield*.

Seniority ranking Priority of payment of various debt obligations.

Sensitivity analysis Analysis that shows the range of possible outcomes as specific assumptions are changed.

Separately managed account (SMA) An investment portfolio managed exclusively for the benefit of an individual or institution.

Serial maturity structure Structure for a bond issue in which the maturity dates are spread out during the bond's life; a stated number of bonds mature and are paid off each year before final maturity.

Settlement The process that occurs after a trade is completed, the securities are passed to the buyer, and payment is received by the seller.

Settlement date Date when the buyer makes cash payment and the seller delivers the security.

Settlement price The official price, designated by the clearinghouse, from which daily gains and losses will be determined and marked to market.

Share repurchase A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Shareholder activism Strategies used by shareholders to attempt to compel a company to act in a desired manner.

Shareholder engagement The process whereby companies engage with their shareholders.

Shareholders' equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Sharpe ratio The average return in excess of the risk-free rate divided by the standard deviation of return; a measure of the average excess return earned per unit of standard deviation of return.

Shelf registration A type of public offering that allows the issuer to file a single, all-encompassing offering circular that covers a series of bond issues.

Short The seller of an asset or derivative contract. Also refers to the position of being short an asset or derivative contract.

Short position A position in an asset or contract in which one has sold an asset one does not own, or in which a right under a contract can be exercised against oneself.

Short-run average total cost The curve describing average total cost when some costs are considered fixed.

Short selling A transaction in which borrowed securities are sold with the intention to repurchase them at a lower price at a later date and return them to the lender.

Shortfall risk The risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Shutdown point The point at which average revenue is equal to the firm's average variable cost.

Side letters Side agreements created between the GP and specific LPs. These agreements exist *outside* the LPA. These agreements provide additional terms and conditions related to the investment agreement.

Simple interest The interest earned each period on the original investment; interest calculated on the principal only.

Simple linear regression (SLR) A regression that summarizes the relation between the dependent variable and a single independent variable.

Simple random sample A subset of a larger population created in such a way that each element of the population has an equal probability of being selected to the subset.

Simple random sampling The procedure of drawing a sample to satisfy the definition of a simple random sample.

Simple yield The sum of the coupon payments plus the straight-line amortized share of the gain or loss, divided by the flat price.

Simulation Computer-generated sensitivity or scenario analysis that is based on probability models for the factors that drive outcomes.

Simulation trial A complete pass through the steps of a simulation.

Single-step format With respect to the format of the income statement, a format that does not subtotal for gross profit (revenue minus cost of goods sold).

Sinking fund arrangement Provision that reduces the credit risk of a bond issue by requiring the issuer to retire a portion of the bond's principal outstanding each year.

Situational influences External factors, such as environmental or cultural elements, that shape our behavior.

Skewed Not symmetrical.

Skewness A quantitative measure of skew (lack of symmetry); a synonym of skew. It is computed as the average cubed deviation from the mean standardized by dividing by the standard deviation cubed.

Slope coefficient The coefficient of an independent variable that represents the average change in the dependent variable for a one-unit change in the independent variable.

Small country A country that is a price taker in the world market for a product and cannot influence the world market price.

Smart beta Involves the use of simple, transparent, rules-based strategies as a basis for investment decisions.

Smart contract A computer program that is designed to self-execute on the basis of pre-specified terms and conditions agreed to by parties to a contract.

Socially responsible investing An investment approach that excludes investments in companies or industries that deviate from an organization's beliefs and sometimes includes investments with favorable environmental or social profiles.

Soft-bullet covered bonds Covered bonds for which bond default and payment acceleration of bond cash flows may be delayed upon sponsor default until a new final maturity date is reached.

Solvency With respect to financial statement analysis, the ability of a company to fulfill its long-term obligations.

Solvency ratios Ratios that measure a company's ability to meet its long-term obligations.

Solvency risk The risk that an organization does not survive or succeed because it runs out of cash, even though it might otherwise be solvent.

Sovereign bond A bond issued by a national government. Also called "Sovereign."

Spearman rank correlation coefficient A measure of correlation applied to ranked data.

Special dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Special purpose entity A non-operating entity created to carry out a specified purpose, such as leasing assets or securitizing receivables; can be a corporation, partnership, trust, or limited liability partnership formed to facilitate a specific type of business activity. Also called *special purpose vehicle*, *special purpose company*, or *variable interest entity*.

Special purpose vehicle See *special purpose entity*.

Specific identification method An inventory accounting method that identifies which specific inventory items were sold and which remained in inventory to be carried over to later periods.

Speculative demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments. Also called *portfolio demand for money*.

Speculative money balances Monies held in anticipation that other assets will decline in value.

Split coupon bond See *deferred coupon bond*.

Sponsored A type of depository receipt in which the foreign company whose shares are held by the depository has a direct involvement in the issuance of the receipts.

Spot curve A sequence of yields-to-maturity on zero-coupon bonds. Sometimes called *zero* or *strip curve* (because coupon payments are "stripped" off the bonds).

Spot markets Markets in which assets are traded for immediate delivery.

Spot prices The price of an asset for immediately delivery.

Spot rates A sequence of market discount rates that correspond to the cash flow dates; yields-to-maturity on zero-coupon bonds maturing at the date of each cash flow.

Spread In general, the difference in yield between different fixed-income securities. Often used to refer to the difference between the yield-to-maturity and the benchmark.

Spread over the benchmark See *required yield spread*.

Spread risk Bond price risk arising from changes in the yield spread on credit-risky bonds; reflects changes in the market's assessment and/or pricing of credit migration (or downgrade) risk and market liquidity risk.

Spurious correlation Refers to: 1) correlation between two variables that reflects chance relationships in a particular dataset; 2) correlation induced by a calculation that mixes each of two variables with a third variable; and 3) correlation between two variables arising not from a direct relation between them but from their relation to a third variable.

Stacked bar chart An alternative form for presenting the frequency distribution of two categorical variables, where bars representing the sub-groups are placed on top of each other to form a single bar. Each sub-section is shown in a different color to represent the contribution of each sub-group, and the overall height of the stacked bar represents the marginal frequency for the category.

Stackelberg model A prominent model of strategic decision making in which firms are assumed to make their decisions sequentially.

Stagflation The combination of a high inflation rate with a high level of unemployment and a slowdown of the economy.

Staggered boards Election process whereby directors are typically divided into multiple classes that are elected separately in consecutive years—that is, one class every year.

Stakeholder management The identification, prioritization, and understanding of the interests of stakeholder groups and managing the company's relationships with these groups.

Stakeholders Individuals or groups of individuals who may be affected either directly or indirectly by a decision and thus have an interest, or stake, in the decision.

Standard deviation The positive square root of the variance; a measure of dispersion in the same units as the original data.

Standard error of the estimate A measure of the fit of a regression line, calculated as the square root of the mean square error. Also known as the *standard error of the regression* and the *root mean square error*.

Standard error of the forecast A measure of the uncertainty associated with a forecasted value of the dependent variable that depends on the standard error of the estimate, the variability of the independent variable, the deviation of the forecasted independent variable from the mean in the regression, and the number of observations.

Standard error of the slope coefficient The standard error of the slope, which in a simple linear regression is the ratio of the model's standard error of the estimate (s_e) to the square root of the variation of the independent variable.

Standard normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Standardizing A transformation that involves subtracting the mean and dividing the result by the standard deviation.

Standards of conduct Behaviors required by a group; established benchmarks that clarify or enhance a group's code of ethics.

Standing limit orders A limit order at a price below market and which therefore is waiting to trade.

Stated annual interest rate A quoted interest rate that does not account for compounding within the year. Also called *quoted interest rate*.

Statement of changes in equity (statement of owners' equity) A financial statement that reconciles the beginning-of-period and end-of-period balance sheet values of

shareholders' equity; provides information about all factors affecting shareholders' equity. Also called *statement of owners' equity*.

Statement of financial condition The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of financial position The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of operations A financial statement that provides information about a company's profitability over a stated period of time.

Static trade-off theory of capital structure A theory pertaining to a company's optimal capital structure. The optimal level of debt is found at the point where additional debt would cause the costs of financial distress to increase by a greater amount than the benefit of the additional tax shield.

Statistic A summary measure of a sample of observations.

Statistically significant A result indicating that the null hypothesis can be rejected; with reference to an estimated regression coefficient, frequently understood to mean a result indicating that the corresponding population regression coefficient is different from zero.

Status quo bias An emotional bias in which people do nothing (i.e., maintain the status quo) instead of making a change.

Statutory voting A common method of voting where each share represents one vote.

Step-up coupon bond Bond for which the coupon, fixed or floating, increases by specified margins at specified dates.

Stochastic oscillator A momentum indicator that compares a particular closing price of a security to a range of the security's prices over a certain period of time.

Stock dividend A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Stock split An increase in the number of shares outstanding with a consequent decrease in share price, but no change to the company's underlying fundamentals.

Stop-loss order See *stop order*.

Stop order An order in which a trader has specified a stop price condition. Also called *stop-loss order*.

Store of value The quality of tending to preserve value.

Store of wealth Goods that depend on the fact that they do not perish physically over time, and on the belief that others would always value the good.

Straight-line method A depreciation method that allocates evenly the cost of a long-lived asset less its estimated residual value over the estimated useful life of the asset.

Straight voting A shareholder voting process in which shareholders receive one vote for each share owned.

Strategic analysis Analysis of the competitive environment with an emphasis on the implications of the environment for corporate strategy.

Strategic asset allocation The set of exposures to IPS-permissible asset classes that is expected to achieve the client's long-term objectives given the client's investment constraints.

Stratified random sampling A procedure that first divides a population into subpopulations (strata) based on classification criteria and then randomly draws samples from each stratum in sizes proportional to that of each stratum in the population.

Street convention A yield measure that neglects weekends and holidays; the internal rate of return on cash flows assuming payments are made on the scheduled dates, even when the scheduled date falls on a weekend or holiday.

Stress testing A specific type of scenario analysis that estimates losses in rare and extremely unfavorable combinations of events or scenarios.

Strong-form efficient market A market in which security prices reflect all public and private information.

Structural (or cyclically adjusted) budget deficit The deficit that would exist if the economy was at full employment (or full potential output).

Structural subordination Arises in a holding company structure when the debt of operating subsidiaries is serviced by the cash flow and assets of the subsidiaries before funds can be passed to the holding company to service debt at the parent level.

Structured data Data that are highly organized in a predefined manner, usually with repeating patterns.

Structured financial instrument A financial instrument that shares the common attribute of repackaging risks. Structured financial instruments include asset-backed securities, collateralized debt obligations, and other structured financial instruments such as capital protected, yield enhancement, participation, and leveraged instruments.

Subjective probability A probability drawing on personal or subjective judgment.

Subordinated debt A class of unsecured debt that ranks below a firm's senior unsecured obligations.

Subordination Form of internal credit enhancement that relies on creating more than one bond tranche and ordering the claim priorities for ownership or interest in an asset between the tranches. The ordering of the claim priorities is called a senior/subordinated structure, where the tranches of highest seniority are called senior followed by subordinated or junior tranches. Also called *credit tranching*.

Substitutes Said of two goods or services such that if the price of one increases the demand for the other tends to increase, holding all other things equal (e.g., butter and margarine).

Sum of squares error (SSE) The sum of the squared deviations of (1) the value of the dependent variable and (2) the value of the dependent variable based on the estimated regression line. Also referred to as the *residual sum of squares*.

Sum of squares regression (SSR) The sum of the squared deviations of (1) the value of the dependent variable based on the estimated regression line and (2) the mean of the dependent variable.

Sum of squares total (SST) The sum of the squared deviations of the dependent variable from its mean; the variation of the dependent variable. Also referred to as the *total sum of squares*.

Sunk cost A cost that has already been incurred.

Supervised learning A machine learning approach that makes use of labeled training data.

Supply shock A typically unexpected disturbance to supply.

Support In technical analysis, a price range in which buying activity is sufficient to stop the decline in the price of a security.

Support tranches Classes or tranches in CMOs that protect PAC tranches from prepayment risk.

Supranational bond A bond issued by a supranational agency such as the World Bank.

Surety bond Form of external credit enhancement whereby a rated and regulated insurance company guarantees to reimburse bondholders for any losses incurred up to a maximum amount if the issuer defaults.

Survey approach An estimate of the equity risk premium that is based on estimates provided by a panel of finance experts.

Survivorship bias Relates to the inclusion of only current investment funds in a database. As such, the returns of funds that are no longer available in the marketplace (have been liquidated) are excluded from the database. In addition, *backfill bias* is another problem, whereby certain surviving hedge funds may be added to databases and various hedge fund indexes only after they are initially successful and start to report their returns.

Sustainability linked loans These are any types of loan instruments and/or contingent facilities (such as bonding lines, guarantee lines, or letters of credit) that incentivize the borrower's achievement of ambitious, pre-determined sustainability performance objectives.

Sustainable growth rate The rate of dividend (and earnings) growth that can be sustained over time for a given level of return on equity, keeping the capital structure constant and without issuing additional common stock.

Sustainable investing A term used in a similar context to responsible investing, but its key focus is on factoring in sustainability issues while investing.

Sustainable rate of economic growth The rate of increase in the economy's productive capacity or potential GDP.

Swap contract An agreement between two parties to exchange a series of future cash flows.

Syndicated loan A loan from a group of lenders to a single borrower.

Syndicated offering A bond issue underwritten by a group of investment banks.

Systematic risk Risk that affects the entire market or economy; it cannot be avoided and is inherent in the overall market. Systematic risk is also known as non-diversifiable or market risk.

Systematic sampling A procedure of selecting every *k*th member until reaching a sample of the desired size. The sample that results from this procedure should be approximately random.

Tactical asset allocation The decision to deliberately deviate from the strategic asset allocation in an attempt to add value based on forecasts of the near-term relative performance of asset classes.

Tactical asset allocation (TAA) A portfolio strategy that shifts the percentages of assets held in various asset classes (or categories) to take advantage of market opportunities. Allocation shifts can occur within an asset class or across asset classes.

Tag cloud see **word cloud**.

Target capital structure A company's chosen proportions of debt and equity.

Target independent A bank's ability to determine the definition of inflation that they target, the rate of inflation that they target, and the horizon over which the target is to be achieved.

Target semideviation A measure of downside risk, calculated as the square root of the average of the squared deviations of observations below the target (also called target downside deviation).

Tariffs Taxes that a government levies on imported goods.

Tax base The amount at which an asset or liability is valued for tax purposes.

Tax expense An aggregate of an entity's income tax payable (or recoverable in the case of a tax benefit) and any changes in deferred tax assets and liabilities. It is essentially the income tax payable or recoverable if these had been determined based on accounting profit rather than taxable income.

Tax loss carry forward A taxable loss in the current period that may be used to reduce future taxable income.

Taxable income The portion of an entity's income that is subject to income taxes under the tax laws of its jurisdiction.

Taxable temporary differences Temporary differences that result in a taxable amount in a future period when determining the taxable profit as the balance sheet item is recovered or settled.

Technical analysis A form of security analysis that uses price and volume data, often displayed graphically, in decision making.

Technology The process a company uses to transform inputs into outputs.

Tender offer Corporate takeover mechanism that involves shareholders selling their interests directly to the group seeking to gain control.

Tenor The time-to-maturity for a bond or derivative contract. Also called *term to maturity*.

Term maturity structure Structure for a bond issue in which the bond's notional principal is paid off in a lump sum at maturity.

Term structure See *maturity structure*.

Term structure of credit spreads The relationship between the spreads over the "risk-free" (or benchmark) rates and times-to-maturity.

Term structure of yield volatility The relationship between the volatility of bond yields-to-maturity and times-to-maturity.

Terminal stock value The expected value of a share at the end of the investment horizon—in effect, the expected selling price. Also called *terminal value*.

Terminal value The expected value of a share at the end of the investment horizon—in effect, the expected selling price.

Terms of trade The ratio of the price of exports to the price of imports, representing those prices by export and import price indexes, respectively.

Test of the mean of the differences A statistical test for differences based on paired observations drawn from samples that are dependent on each other.

Text analytics The use of computer programs to analyze and derive meaning from typically large, unstructured text- or voice-based datasets.

Thematic investing An ESG investment approach that focuses on investing in themes or assets specifically relating to ESG factors, such as clean energy, green technology, or sustainable agriculture.

Thin-Tailed Describes a distribution that has relatively less weight in the tails than the normal distribution (also called platykurtic).

Third-degree price discrimination When the monopolist segregates customers into groups based on demographic or other characteristics and offers different pricing to each group.

Time-period bias The possibility that when we use a time-series sample, our statistical conclusion may be sensitive to the starting and ending dates of the sample.

Time-series data A sequence of observations for a single observational unit of a specific variable collected over time and at discrete and typically equally spaced intervals of time (such as daily, weekly, monthly, annually, or quarterly).

Time tranching The creation of classes or tranches in an ABS/MBS that possess different (expected) maturities.

Time value The difference between the market price of the option and its intrinsic value.

Time value decay Said of an option when, at expiration, no time value remains and the option is worth only its exercise value.

Time value of money The principles governing equivalence relationships between cash flows with different dates.

Time-weighted rate of return The compound rate of growth of one unit of currency invested in a portfolio during a stated measurement period; a measure of investment performance that is not sensitive to the timing and amount of withdrawals or additions to the portfolio.

Tokenization The process of representing ownership rights to physical assets on a blockchain or distributed ledger.

Top-down analysis An investment selection approach that begins with consideration of macroeconomic conditions and then evaluates markets and industries based upon such conditions.

Total comprehensive income The change in equity during a period resulting from transaction and other events, other than those changes resulting from transactions with owners in their capacity as owners.

Total cost The summation of all costs, for which costs are classified as fixed or variable.

Total factor productivity A scale factor that reflects the portion of growth unaccounted for by explicit factor inputs (e.g., capital and labor).

Total fixed cost The summation of all expenses that do not change as the level of production varies.

Total invested capital The sum of market value of common equity, book value of preferred equity, and face value of debt.

Total probability rule A rule explaining the unconditional probability of an event in terms of probabilities of the event conditional on mutually exclusive and exhaustive scenarios.

Total probability rule for expected value A rule explaining the expected value of a random variable in terms of expected values of the random variable conditional on mutually exclusive and exhaustive scenarios.

Total return Measures the price appreciation, or percentage change in price of the securities in an index or portfolio, plus any income received over the period.

Total return index An index that reflects the price appreciation or percentage change in price of the constituent securities plus any income received since inception.

Total return swap A swap in which one party agrees to pay the total return on a security. Often used as a credit derivative, in which the underlying is a bond.

Total variable cost The summation of all variable expenses.

Tracking error The standard deviation of the differences between a portfolio's returns and its benchmark's returns; a synonym of active risk.

Tracking risk The standard deviation of the differences between a portfolio's returns and its benchmarks returns. Also called *tracking error*.

Trade creation When regional integration results in the replacement of higher cost domestic production by lower cost imports from other members.

Trade credit A spontaneous form of credit in which a purchaser of the goods or service is financing its purchase by delaying the date on which payment is made.

Trade diversion When regional integration results in lower-cost imports from non-member countries being replaced with higher-cost imports from members.

Trade payables Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Trade protection Government policies that impose restrictions on trade, such as tariffs and quotas.

Trade surplus (deficit) When the value of exports is greater (less) than the value of imports.

Trading securities Under US GAAP, a category of debt securities held by a company with the intent to trade them. Also called *held-for-trading securities*.

Traditional investment markets Markets for traditional investments, which include all publicly traded debts and equities and shares in pooled investment vehicles that hold publicly traded debts and/or equities.

Transactions money balances Money balances that are held to finance transactions.

Transfer payments Welfare payments made through the social security system that exist to provide a basic minimum level of income for low-income households.

Transparency Said of something (e.g., a market) in which information is fully disclosed to the public and/or regulators.

Treasury stock method A method for accounting for the effect of options (and warrants) on earnings per share (EPS) that specifies what EPS would have been if the options and warrants had been exercised and the company had used the proceeds to repurchase common stock.

Tree-Map Another graphical tool for displaying categorical data. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group.

Trend A long-term pattern of movement in a particular direction.

Treynor ratio A measure of risk-adjusted performance that relates a portfolio's excess returns to the portfolio's beta.

Triangle pattern In technical analysis, a continuation chart pattern that forms as the range between high and low prices narrows, visually forming a triangle.

Trimmed mean A mean computed after excluding a stated small percentage of the lowest and highest observations.

Trimodal A distribution that has the three most frequently occurring values.

Triple bottom In technical analysis, a reversal pattern that results when the price forms three troughs at roughly the same price level. A triple bottom is used to predict a change from a downtrend to an uptrend.

Triple top In technical analysis, a reversal pattern that results when the price forms three peaks at roughly the same price level. A triple top is used to predict a change from an uptrend to a downtrend.

True yield The internal rate of return on cash flows using the actual calendar, including weekends and bank holidays.

Trust deed The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *bond indenture*.

Turn-of-the-year effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January.

Two-dimensional rectangular array A popular form for organizing data for processing by computers or for presenting data visually. It is comprised of columns and rows to hold multiple variables and multiple observations, respectively (also called a **data table**).

Two-fund separation theorem The theory that all investors regardless of taste, risk preferences, and initial wealth will hold a combination of two portfolios or funds: a risk-free asset and an optimal portfolio of risky assets.

Two-sided hypothesis test A test in which the null hypothesis is rejected in favor of the alternative hypothesis if the evidence indicates that the population parameter is either smaller or larger than a hypothesized value; occurs when the alternative hypothesis is stated as not equal to the hypothesized population parameters.

Two-way table See *contingency table*.

Two-week repo rate The interest rate on a two-week repurchase agreement; may be used as a policy rate by a central bank.

Type I error The error of rejecting a true null hypothesis; a false positive.

Type II error The error of not rejecting a false null hypothesis; false negative.

Unanticipated (unexpected) inflation The component of inflation that is a surprise.

Uncommitted lines of credit The least reliable form of bank borrowing in which a bank offers, without formal commitment, a line of credit for an extended period of time but reserves the right to refuse any request for its use.

Unconditional probability The probability of an event *not* conditioned on another event.

Underemployed A person who has a job but has the qualifications to work a significantly higher-paying job.

Underlying An asset that trades in a market in which buyers and sellers meet, decide on a price, and the seller then delivers the asset to the buyer and receives payment. The underlying is the asset or other derivative on which a particular derivative is based. The market for the underlying is also referred to as the *spot market*.

Underwriter A firm, usually an investment bank, that takes the risk of buying the newly issued securities from the issuer and then reselling them to investors or to dealers, thus guaranteeing the sale of the securities at the offering price negotiated with the issuer.

Underwritten offering A type of securities issue mechanism in which the investment bank guarantees the sale of the securities at an offering price that is negotiated with the issuer. Also known as *firm commitment offering*.

Unearned revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service. Also called *deferred revenue* or *deferred income*.

Unemployed People who are actively seeking employment but are currently without a job.

Unemployment rate The ratio of unemployed to the labor force.

Unexpected inflation The component of inflation that is a surprise.

Unimodal A distribution with a single value that is most frequently occurring.

Unit elastic An elasticity with a magnitude of negative one. Also called *unitary elastic*.

Unit labor cost The average labor cost to produce one unit of output.

Unit normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Unitranche debt Consists of a hybrid or blended loan structure that combines different tranches of secured and unsecured debt into a single loan with a single, blended interest rate.

Units-of-production method A depreciation method that allocates the cost of a long-lived asset based on actual usage during the period.

Univariate distribution A distribution that specifies the probabilities for a single random variable.

Universal owners Long-term investors, such as pension funds, that have significant assets invested in globally diversified portfolios.

Unlimited funds An unlimited funds environment assumes that the company can raise the funds it wants for all profitable investments simply by paying the required rate of return.

Unsecured debt Debt that gives the debtholder only a general claim on an issuer's assets and cash flow.

Unsponsored A type of depository receipt in which the foreign company whose shares are held by the depository has no involvement in the issuance of the receipts.

Unstructured data Data that do not follow any conventionally organized forms.

Unsupervised learning A machine learning approach that does not make use of labeled training data.

Up transition probability The probability that an asset's value moves up.

Uptrend A pattern that occurs when the price of an asset moves higher over a period of time.

Validity instructions Instructions which indicate when the order may be filled.

Valuation allowance A reserve created against deferred tax assets, based on the likelihood of realizing the deferred tax assets in future accounting periods.

Valuation ratios Ratios that measure the quantity of an asset or flow (e.g., earnings) in relation to the price associated with a specified claim (e.g., a share or ownership of the enterprise).

Value at risk A money measure of the minimum value of losses expected during a specified time period at a given level of probability.

Value investors With reference to equity investors, investors who are focused on paying a relatively low share price in relation to earnings or assets per share.

VaR See *value at risk*.

Variable A characteristic or quantity that can be measured, counted, or categorized and that is subject to change (also called a field, an attribute, or a feature).

Variable costs Costs that fluctuate with the level of production and sales.

Variance The expected value (the probability-weighted average) of squared deviations from a random variable's expected value.

Variation margin Additional margin that must be deposited in an amount sufficient to bring the balance up to the initial margin requirement.

Veblen goods Goods that increase in desirability with increasing price.

Vega A measure of the sensitivity of an option's price to changes in the underlying's volatility.

Venture capital Investments that provide "seed" or startup capital, early-stage financing, or later-stage financing (including mezzanine-stage financing) to companies that are in early development stages and require additional capital for expansion or preparation for an initial public offering.

Venture capital fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *private equity fund*.

Venture capital funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *private equity funds.*

Vertical analysis Common-size analysis using only one reporting period or one base financial statement; for example, an income statement in which all items are stated as percentages of sales.

Vertical demand schedule Implies that some fixed quantity is demanded, regardless of price.

Visual technique The most common and readily available method of initial data assessment. Experts in pattern recognition maintain that the visual (or "eyeball") technique is still the most effective way of searching for recognizable patterns.

Visualization The presentation of data in a pictorial or graphical format for the purpose of increasing understanding and for gaining insights into the data.

Volatility The standard deviation of the continuously compounded returns on the underlying asset.

Voluntarily unemployed A person voluntarily outside the labor force, such as a jobless worker refusing an available vacancy.

Voluntary export restraint A trade barrier under which the exporting country agrees to limit its exports of the good to its trading partners to a specific number of units.

Vote by proxy A mechanism that allows a designated party—such as another shareholder, a shareholder representative, or management—to vote on the shareholder's behalf.

Warrant Attached option that gives its holder the right to buy the underlying stock of the issuing company at a fixed exercise price until the expiration date.

Waterfall Represents the distribution method that defines the order in which allocations are made to LPs and GPs. There are two major types of waterfall: *deal by deal* (or *American*) and *whole of fund* (or *European*).

Weak-form efficient market hypothesis The belief that security prices fully reflect all past market data, which refers to all historical price and volume trading information.

Wealth effect An increase (decrease) in household wealth increases (decreases) consumer spending out of a given level of current income.

Web-based lender A lender that operates primarily on the internet, offering loans in relatively small amounts, typically to small businesses in need of cash.

Weighted average cost method An inventory accounting method that averages the total cost of available inventory items over the total units available for sale.

Weighted average cost of capital A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average cost of capital (WACC) A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average coupon rate Weighting the mortgage rate of each mortgage loan in the pool by the percentage of the mortgage outstanding relative to the outstanding amount of all the mortgages in the pool.

Weighted average life A measure that gives investors an indication of how long they can expect to hold the MBS before it is paid off; the convention-based average time to receipt of all principal repayments. Also called *average life*.

Weighted average maturity Weighting the remaining number of months to maturity for each mortgage loan in the pool by the amount of the outstanding mortgage balance.

Weighted mean An average in which each observation is weighted by an index of its relative importance.

Winsorized mean A mean computed after assigning a stated percentage of the lowest values equal to one specified low value and a stated percentage of the highest values equal to one specified high value.

Word cloud A visual device for representing textual data, which consists of words extracted from a source of textual data. The size of each distinct word is proportional to the frequency with which it appears in the given text (also known as **tag cloud**).

Working capital The difference between current assets and current liabilities.

Working capital management The management of a company's short-term assets (such as inventory) and short-term liabilities (such as money owed to suppliers).

World price The price prevailing in the world market.

Yield duration The sensitivity of the bond price with respect to the bond's own yield-to-maturity.

Yield to maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield to redemption* or *redemption yield*.

Yield-to-maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield-to-redemption* or *redemption yield*.

Yield-to-redemption See *yield-to-maturity*.

Yield-to-worst The lowest of the sequence of yields-to-call and the yield-to-maturity.

Zero-coupon bond A bond that does not pay interest during its life. It is issued at a discount to par value and redeemed at par. Also called *pure discount bond*.

Zero-volatility spread (Z-spread) Calculates a constant yield spread over a government (or interest rate swap) spot curve.

CFA Institute®
CFA Program

FIXED INCOME, DERIVATIVES, ALTERNATIVE INVESTMENTS, AND PORTFOLIO MANAGEMENT

CFA® Program Curriculum
2022 • LEVEL I • VOLUME 5

WILEY

CONTENTS

How to Use the CFA Program Curriculum ix
Background on the CBOK ix
Organization of the Curriculum x
Features of the Curriculum x
Designing Your Personal Study Program xi
CFA Institute Learning Ecosystem (LES) xii
Prep Providers xiii
Feedback xiv

Fixed Income

Study Session 14 **Fixed Income (2)** 3

Reading 43 **Understanding Fixed-Income Risk and Return** 5
Introduction 6
Sources of Return 7
Macaulay and Modified Duration 14
    Macaulay, Modified, and Approximate Duration 14
Approximate Modified and Macaulay Duration 19
Effective and Key Rate Duration 22
    Key Rate Duration 25
Properties of Bond Duration 26
Duration of a Bond Portfolio 32
Money Duration and the Price Value of a Basis Point 34
Bond Convexity 36
Investment Horizon, Macaulay Duration and Interest Rate Risk 45
    Yield Volatility 45
    Investment Horizon, Macaulay Duration, and Interest Rate Risk 47
Credit and Liquidity Risk 51
Empirical Duration 52
Summary 53
Practice Problems 58
Solutions 63

Reading 44 **Fundamentals of Credit Analysis** 69
Introduction 70
Credit Risk 70
Capital Structure, Seniority Ranking, and Recovery Rates 72
    Capital Structure 73
    Seniority Ranking 73
    Recovery Rates 74
Rating Agencies, Credit Ratings, and Their Role in the Debt Markets 78
    Credit Ratings 79
    Issuer vs. Issue Ratings 80
    ESG Ratings 81

Risks in Relying on Agency Ratings 82
Traditional Credit Analysis: Corporate Debt Securities 87
Credit Analysis vs. Equity Analysis: Similarities and Differences 87
The Four Cs of Credit Analysis: A Useful Framework 88
Credit Risk vs. Return: Yields and Spreads 105
Credit Risk vs. Return: The Price Impact of Spread Changes 109
High-Yield, Sovereign, and Non-Sovereign Credit Analysis 112
High Yield 113
Sovereign Debt 120
Non-Sovereign Government Debt 124
Summary 126
Practice Problems 131
Solutions 140

Derivatives

Study Session 15 **Derivatives** 147

Reading 45 **Derivative Markets and Instruments** 149
Derivatives: Introduction, Definitions, and Uses 149
Derivatives: Definitions and Uses 150
The Structure of Derivative Markets 153
Exchange-Traded Derivatives Markets 154
Over-the-Counter Derivatives Markets 155
Types of Derivatives: Introduction, Forward Contracts 158
Forward Commitments 158
Types of Derivatives: Futures 162
Types of Derivatives: Swaps 166
Contingent Claims: Options 170
Options 170
Contingent Claims: Credit Derivatives 178
Types of Derivatives: Asset-Backed Securities and Hybrids 182
Hybrids 183
Derivatives Underlyings 184
Equities 185
Fixed-Income Instruments and Interest Rates 185
Currencies 185
Commodities 186
Credit 186
Other 186
The Purposes and Benefits of Derivatives 188
Risk Allocation, Transfer, and Management 189
Information Discovery 189
Operational Advantages 190
Market Efficiency 191
Criticisms and Misuses of Derivatives 191
Speculation and Gambling 192
Destabilization and Systemic Risk 192

Elementary Principles of Derivative Pricing 194
    Storage 195
    Arbitrage 196
Summary 201
Practice Problems 204
Solutions 211

Reading 46 **Basics of Derivative Pricing and Valuation** 221
Introduction 222
Basic Derivative Concepts, Pricing the Underlying 222
    Basic Derivative Concepts 222
    Pricing the Underlying 224
The Principle of Arbitrage 228
    The (In)Frequency of Arbitrage Opportunities 229
    Arbitrage and Derivatives 229
    Arbitrage and Replication 230
    Risk Aversion, Risk Neutrality, and Arbitrage-Free Pricing 231
    Limits to Arbitrage 232
Pricing and Valuation of Forward Contracts: Pricing vs. Valuation; Expiration; Initiation 234
    Pricing and Valuation of Forward Commitments 235
Pricing and Valuation of Forward Contracts: Between Initiation and Expiration; Forward Rate Agreements 239
    A Word about Forward Contracts on Interest Rates 240
Pricing and Valuation of Futures Contracts 242
Pricing and Valuation of Swap Contracts 245
Pricing and Valuation of Options 248
    European Option Pricing 249
Lower Limits for Prices of European Options 254
Put-Call Parity, Put-Call-Forward Parity 258
    Put–Call–Forward Parity 261
Binomial Valuation of Options 263
American Option Pricing 268
Summary 271
Practice Problems 273
Solutions 279

Alternative Investments

Study Session 16 **Alternative Investments** 287

Reading 47 **Introduction to Alternative Investments** 289
Introduction 289
    Why Investors Consider Alternative Investments 290
    Categories of Alternative Investments 291
Investment Methods 295
    Methods of Investing in Alternative Investments 295
    Advantages and Disadvantages of Direct Investing, Co-Investing, and Fund Investing 296

Due Diligence for Fund Investing, Direct Investing, and Co-Investing **299**
Investment and Compensation Structures **303**
Partnership Structures **303**
Compensation Structures **305**
Common Investment Clauses, Provisions, and Contingencies **306**
Hedge Funds **310**
Characteristics of Hedge Funds **311**
Hedge Fund Strategies **314**
Hedge Funds and Diversification Benefits **319**
Private Capital **321**
Overview of Private Capital **321**
Description: Private Equity **322**
Description: Private Debt **326**
Risk/Return of Private Equity **328**
Risk/Return of Private Debt **329**
Diversification Benefits of Investing in Private Capital **331**
Natural Resources **334**
Overview of Natural Resources **334**
Characteristics of Natural Resources **335**
Risk/Return of Natural Resources **339**
Diversification Benefits of Natural Resources **342**
Instruments **347**
Real Estate **350**
Overview of the Real Estate Market **350**
Characteristics: Forms of Real Estate Ownership **353**
Characteristics: Real Estate Investment Categories **356**
Risk and Return Characteristics **359**
Diversification Benefits **362**
Infrastructure **366**
Introduction and Overview **366**
Description **368**
Risk and Return Characteristics **370**
Diversification Benefits **372**
Issues in Performance Appraisal **375**
Overview of Performance Appraisal for Alternative Investments **376**
Common Approaches to Performance Appraisal and Application Challenges **376**
Private Equity and Real Estate Performance Evaluation **379**
Hedge Funds: Leverage, Illiquidity, and Redemption Terms **381**
Calculating Fees and Returns **387**
Alternative Asset Fee Structures and Terms **387**
Custom Fee Arrangements **388**
Alignment of Interests and Survivorship Bias **392**
Summary **396**

Portfolio Management

Study Session 17 **Portfolio Management (1)** **403**

Reading 48 **Portfolio Management: An Overview** **405**
Introduction **405**
Portfolio Perspective: Diversification and Risk Reduction **406**
  Historical Example of Portfolio Diversification: Avoiding Disaster **406**
  Portfolios: Reduce Risk **408**
Portfolio Perspective: Risk-Return Trade-off, Downside Protection, Modern Portfolio Theory **411**
  Historical Portfolio Example: Not Necessarily Downside Protection **412**
  Portfolios: Modern Portfolio Theory **414**
Steps in the Portfolio Management Process **415**
  Step One: The Planning Step **415**
  Step Two: The Execution Step **416**
  Step Three: The Feedback Step **418**
Types of Investors **419**
  Individual Investors **419**
  Institutional Investors **420**
The Asset Management Industry **425**
  Active versus Passive Management **426**
  Traditional versus Alternative Asset Managers **427**
  Ownership Structure **427**
  Asset Management Industry Trends **428**
Pooled Interest - Mutual Funds **430**
  Mutual Funds **430**
Pooled Interest - Type of Mutual Funds **432**
  Money Market Funds **432**
  Bond Mutual Funds **433**
  Stock Mutual Funds **433**
  Hybrid/Balanced Funds **434**
Pooled Interest - Other Investment Products **434**
  Exchange-Traded Funds **434**
  Hedge Funds **435**
  Private Equity and Venture Capital Funds **435**
Summary **436**
Practice Problems **438**
Solutions **440**

Reading 49 **Portfolio Risk and Return: Part I** **441**
Introduction **441**
Investment Characteristics of Assets: Return **442**
  Return **442**
Money-Weighted Return or Internal Rate of Return **445**
Time-Weighted Rate of Return **449**
Annualized Return **454**
Other Major Return Measures and their Applications **456**
  Gross and Net Return **456**

Pre-tax and After-tax Nominal Return 457
Real Returns 457
Leveraged Return 458
Historical Return and Risk 459
Historical Mean Return and Expected Return 459
Nominal Returns of Major US Asset Classes 460
Real Returns of Major US Asset Classes 461
Nominal and Real Returns of Asset Classes in Major Countries 462
Risk of Major Asset Classes 462
Risk–Return Trade-off 463
Other Investment Characteristics 463
Distributional Characteristics 464
Market Characteristics 466
Risk Aversion and Portfolio Selection & The Concept of Risk Aversion 467
The Concept of Risk Aversion 467
Utility Theory and Indifference Curves 468
Indifference Curves 469
Application of Utility Theory to Portfolio Selection 473
Portfolio Risk & Portfolio of Two Risky Assets 476
Portfolio of Two Risky Assets 476
Portfolio of Many Risky Assets 484
Importance of Correlation in a Portfolio of Many Assets 485
The Power of Diversification 485
Correlation and Risk Diversification 487
Historical Risk and Correlation 487
Historical Correlation among Asset Classes 488
Avenues for Diversification 489
Efficient Frontier: Investment Opportunity Set & Minimum Variance Portfolios 491
Investment Opportunity Set 491
Minimum-Variance Portfolios 492
Efficient Frontier: A Risk-Free Asset and Many Risky Assets 494
Capital Allocation Line and Optimal Risky Portfolio 494
The Two-Fund Separation Theorem 495
Efficient Frontier: Optimal Investor Portfolio 497
Investor Preferences and Optimal Portfolios 502
Summary 503
Practice Problems 505
Solutions 513

Reading 50 **Portfolio Risk and Return: Part II** 519
Introduction 519
Capital Market Theory: Risk-Free and Risky Assets 520
Portfolio of Risk-Free and Risky Assets 520
Capital Market Theory: The Capital Market Line 524
Passive and Active Portfolios 524
What Is the “Market”? 525
The Capital Market Line (CML) 525

Capital Market Theory: CML - Leveraged Portfolios 528
Leveraged Portfolios with Different Lending and Borrowing Rates 530
Systematic and Nonsystematic Risk 532
Systematic Risk and Nonsystematic Risk 532
Return Generating Models 534
Return-Generating Models 534
Decomposition of Total Risk for a Single-Index Model 536
Return-Generating Models: The Market Model 537
Calculation and Interpretation of Beta 537
Estimation of Beta 539
Beta and Expected Return 540
Capital Asset Pricing Model: Assumptions and the Security Market Line 541
Assumptions of the CAPM 542
The Security Market Line 543
Capital Asset Pricing Model: Applications 546
Estimate of Expected Return 547
Beyond CAPM: Limitations and Extensions of CAPM 548
Limitations of the CAPM 548
Extensions to the CAPM 549
Portfolio Performance Appraisal Measures 551
The Sharpe Ratio 552
The Treynor Ratio 552
M^2: Risk-Adjusted Performance (RAP) 553
Jensen's Alpha 554
Applications of the CAPM in Portfolio Construction 557
Security Characteristic Line 558
Security Selection 558
Implications of the CAPM for Portfolio Construction 560
Summary 563
Practice Problems 565
Solutions 571

Glossary G-1

How to Use the CFA Program Curriculum

Congratulations on your decision to enter the Chartered Financial Analyst (CFA®) Program. This exciting and rewarding program of study reflects your desire to become a serious investment professional. You are embarking on a program noted for its high ethical standards and the breadth of knowledge, skills, and abilities (competencies) it develops. Your commitment should be educationally and professionally rewarding.

The credential you seek is respected around the world as a mark of accomplishment and dedication. Each level of the program represents a distinct achievement in professional development. Successful completion of the program is rewarded with membership in a prestigious global community of investment professionals. CFA charterholders are dedicated to life-long learning and maintaining currency with the ever-changing dynamics of a challenging profession. CFA Program enrollment represents the first step toward a career-long commitment to professional education.

The CFA exam measures your mastery of the core knowledge, skills, and abilities required to succeed as an investment professional. These core competencies are the basis for the Candidate Body of Knowledge (CBOK™). The CBOK consists of four components:

- A broad outline that lists the major CFA Program topic areas (www.cfainstitute.org/programs/cfa/curriculum/cbok);
- Topic area weights that indicate the relative exam weightings of the top-level topic areas (www.cfainstitute.org/programs/cfa/curriculum);
- Learning outcome statements (LOS) that advise candidates about the specific knowledge, skills, and abilities they should acquire from readings covering a topic area (LOS are provided in candidate study sessions and at the beginning of each reading); and
- CFA Program curriculum that candidates receive upon exam registration.

Therefore, the key to your success on the CFA exams is studying and understanding the CBOK. The following sections provide background on the CBOK, the organization of the curriculum, features of the curriculum, and tips for designing an effective personal study program.

BACKGROUND ON THE CBOK

CFA Program is grounded in the practice of the investment profession. CFA Institute performs a continuous practice analysis with investment professionals around the world to determine the competencies that are relevant to the profession, beginning with the Global Body of Investment Knowledge (GBIK®). Regional expert panels and targeted surveys are conducted annually to verify and reinforce the continuous feedback about the GBIK. The practice analysis process ultimately defines the CBOK. The CBOK reflects the competencies that are generally accepted and applied by investment professionals. These competencies are used in practice in a generalist context and are expected to be demonstrated by a recently qualified CFA charterholder.

© 2021 CFA Institute. All rights reserved.

The CFA Institute staff—in conjunction with the Education Advisory Committee and Curriculum Level Advisors, who consist of practicing CFA charterholders—designs the CFA Program curriculum in order to deliver the CBOK to candidates. The exams, also written by CFA charterholders, are designed to allow you to demonstrate your mastery of the CBOK as set forth in the CFA Program curriculum. As you structure your personal study program, you should emphasize mastery of the CBOK and the practical application of that knowledge. For more information on the practice analysis, CBOK, and development of the CFA Program curriculum, please visit www.cfainstitute.org.

ORGANIZATION OF THE CURRICULUM

The Level I CFA Program curriculum is organized into 10 topic areas. Each topic area begins with a brief statement of the material and the depth of knowledge expected. It is then divided into one or more study sessions. These study sessions should form the basic structure of your reading and preparation. Each study session includes a statement of its structure and objective and is further divided into assigned readings. An outline illustrating the organization of these study sessions can be found at the front of each volume of the curriculum.

The readings are commissioned by CFA Institute and written by content experts, including investment professionals and university professors. Each reading includes LOS and the core material to be studied, often a combination of text, exhibits, and in-text examples and questions. End of Reading Questions (EORQs) followed by solutions help you understand and master the material. The LOS indicate what you should be able to accomplish after studying the material. The LOS, the core material, and the EORQs are dependent on each other, with the core material and EORQs providing context for understanding the scope of the LOS and enabling you to apply a principle or concept in a variety of scenarios.

The entire readings, including the EORQs, are the basis for all exam questions and are selected or developed specifically to teach the knowledge, skills, and abilities reflected in the CBOK.

You should use the LOS to guide and focus your study because each exam question is based on one or more LOS and the core material and practice problems associated with the LOS. As a candidate, you are responsible for the entirety of the required material in a study session.

We encourage you to review the information about the LOS on our website (www.cfainstitute.org/programs/cfa/curriculum/study-sessions), including the descriptions of LOS "command words" on the candidate resources page at www.cfainstitute.org.

FEATURES OF THE CURRICULUM

End of Reading Questions/Solutions *All End of Reading Questions (EORQs) as well as their solutions are part of the curriculum and are required material for the exam.* In addition to the in-text examples and questions, these EORQs help demonstrate practical applications and reinforce your understanding of the concepts presented. Some of these EORQs are adapted from past CFA exams and/or may serve as a basis for exam questions.

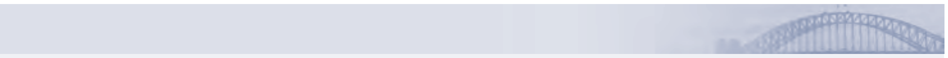
Glossary For your convenience, each volume includes a comprehensive Glossary. Throughout the curriculum, a **bolded** word in a reading denotes a term defined in the Glossary.

Note that the digital curriculum that is included in your exam registration fee is searchable for key words, including Glossary terms.

LOS Self-Check We have inserted checkboxes next to each LOS that you can use to track your progress in mastering the concepts in each reading.

Source Material The CFA Institute curriculum cites textbooks, journal articles, and other publications that provide additional context or information about topics covered in the readings. As a candidate, you are not responsible for familiarity with the original source materials cited in the curriculum.

Note that some readings may contain a web address or URL. The referenced sites were live at the time the reading was written or updated but may have been deactivated since then.



Some readings in the curriculum cite articles published in the *Financial Analysts Journal*®, which is the flagship publication of CFA Institute. Since its launch in 1945, the *Financial Analysts Journal* has established itself as the leading practitioner-oriented journal in the investment management community. Over the years, it has advanced the knowledge and understanding of the practice of investment management through the publication of peer-reviewed practitioner-relevant research from leading academics and practitioners. It has also featured thought-provoking opinion pieces that advance the common level of discourse within the investment management profession. Some of the most influential research in the area of investment management has appeared in the pages of the *Financial Analysts Journal*, and several Nobel laureates have contributed articles.

Candidates are not responsible for familiarity with *Financial Analysts Journal* articles that are cited in the curriculum. But, as your time and studies allow, we strongly encourage you to begin supplementing your understanding of key investment management issues by reading this, and other, CFA Institute practice-oriented publications through the Research & Analysis webpage (www.cfainstitute.org/en/research).

Errata The curriculum development process is rigorous and includes multiple rounds of reviews by content experts. Despite our efforts to produce a curriculum that is free of errors, there are times when we must make corrections. Curriculum errata are periodically updated and posted by exam level and test date online (www.cfainstitute.org/en/programs/submit-errata). If you believe you have found an error in the curriculum, you can submit your concerns through our curriculum errata reporting process found at the bottom of the Curriculum Errata webpage.

DESIGNING YOUR PERSONAL STUDY PROGRAM

Create a Schedule An orderly, systematic approach to exam preparation is critical. You should dedicate a consistent block of time every week to reading and studying. Complete all assigned readings and the associated problems and solutions in each study session. Review the LOS both before and after you study each reading to ensure that

you have mastered the applicable content and can demonstrate the knowledge, skills, and abilities described by the LOS and the assigned reading. Use the LOS self-check to track your progress and highlight areas of weakness for later review.

Successful candidates report an average of more than 300 hours preparing for each exam. Your preparation time will vary based on your prior education and experience, and you will probably spend more time on some study sessions than on others.

You should allow ample time for both in-depth study of all topic areas and additional concentration on those topic areas for which you feel the least prepared.

CFA INSTITUTE LEARNING ECOSYSTEM (LES)

As you prepare for your exam, we will email you important exam updates, testing policies, and study tips. Be sure to read these carefully.

Your exam registration fee includes access to the CFA Program Learning Ecosystem (LES). This digital learning platform provides access, even offline, to all of the readings and End of Reading Questions found in the print curriculum organized as a series of shorter online lessons with associated EORQs. This tool is your one-stop location for all study materials, including practice questions and mock exams.

The LES provides the following supplemental study tools:

Structured and Adaptive Study Plans The LES offers two ways to plan your study through the curriculum. The first is a structured plan that allows you to move through the material in the way that you feel best suits your learning. The second is an adaptive study plan based on the results of an assessment test that uses actual practice questions.

Regardless of your chosen study path, the LES tracks your level of proficiency in each topic area and presents you with a dashboard of where you stand in terms of proficiency so that you can allocate your study time efficiently.

Flashcards and Game Center The LES offers all the Glossary terms as Flashcards and tracks correct and incorrect answers. Flashcards can be filtered both by curriculum topic area and by action taken—for example, answered correctly, unanswered, and so on. These Flashcards provide a flexible way to study Glossary item definitions.

The Game Center provides several engaging ways to interact with the Flashcards in a game context. Each game tests your knowledge of the Glossary terms a in different way. Your results are scored and presented, along with a summary of candidates with high scores on the game, on your Dashboard.

Discussion Board The Discussion Board within the LES provides a way for you to interact with other candidates as you pursue your study plan. Discussions can happen at the level of individual lessons to raise questions about material in those lessons that you or other candidates can clarify or comment on. Discussions can also be posted at the level of topics or in the initial Welcome section to connect with other candidates in your area.

Practice Question Bank The LES offers access to a question bank of hundreds of practice questions that are in addition to the End of Reading Questions. These practice questions, only available on the LES, are intended to help you assess your mastery of individual topic areas as you progress through your studies. After each practice question, you will receive immediate feedback noting the correct response and indicating the relevant assigned reading so you can identify areas of weakness for further study.

Mock Exams The LES also includes access to three-hour Mock Exams that simulate the morning and afternoon sessions of the actual CFA exam. These Mock Exams are intended to be taken after you complete your study of the full curriculum and take practice questions so you can test your understanding of the curriculum and your readiness for the exam. If you take these Mock Exams within the LES, you will receive feedback afterward that notes the correct responses and indicates the relevant assigned readings so you can assess areas of weakness for further study. We recommend that you take Mock Exams during the final stages of your preparation for the actual CFA exam. For more information on the Mock Exams, please visit www.cfainstitute.org.

PREP PROVIDERS

You may choose to seek study support outside CFA Institute in the form of exam prep providers. After your CFA Program enrollment, you may receive numerous solicitations for exam prep courses and review materials. When considering a prep course, make sure the provider is committed to following the CFA Institute guidelines and high standards in its offerings.

Remember, however, that there are no shortcuts to success on the CFA exams; reading and studying the CFA Program curriculum *is* the key to success on the exam. The CFA Program exams reference only the CFA Institute assigned curriculum; no prep course or review course materials are consulted or referenced.

SUMMARY

Every question on the CFA exam is based on the content contained in the required readings and on one or more LOS. Frequently, an exam question is based on a specific example highlighted within a reading or on a specific practice problem and its solution. To make effective use of the CFA Program curriculum, please remember these key points:

1. All pages of the curriculum are required reading for the exam.
2. All questions, problems, and their solutions are part of the curriculum and are required study material for the exam. These questions are found at the end of the readings in the print versions of the curriculum. In the LES, these questions appear directly after the lesson with which they are associated. The LES provides immediate feedback on your answers and tracks your performance on these questions throughout your study.
3. We strongly encourage you to use the CFA Program Learning Ecosystem. In addition to providing access to all the curriculum material, including EORQs, in the form of shorter, focused lessons, the LES offers structured and adaptive study planning, a Discussion Board to communicate with other candidates, Flashcards, a Game Center for study activities, a test bank of practice questions, and online Mock Exams. Other supplemental study tools, such as eBook and PDF versions of the print curriculum, and additional candidate resources are available at www.cfainstitute.org.
4. Using the study planner, create a schedule and commit sufficient study time to cover the study sessions. You should also plan to review the materials, answer practice questions, and take Mock Exams.
5. Some of the concepts in the study sessions may be superseded by updated rulings and/or pronouncements issued after a reading was published. Candidates are expected to be familiar with the overall analytical framework contained in the assigned readings. Candidates are not responsible for changes that occur after the material was written.

FEEDBACK

At CFA Institute, we are committed to delivering a comprehensive and rigorous curriculum for the development of competent, ethically grounded investment professionals. We rely on candidate and investment professional comments and feedback as we work to improve the curriculum, supplemental study tools, and candidate resources.

Please send any comments or feedback to info@cfainstitute.org. You can be assured that we will review your suggestions carefully. Ongoing improvements in the curriculum will help you prepare for success on the upcoming exams and for a lifetime of learning as a serious investment professional.

Fixed Income

STUDY SESSIONS

Study Session 13	Fixed Income (1)
Study Session 14	Fixed Income (2)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to describe fixed-income securities and their markets, yield measures, risk factors, and valuation measures and drivers. The candidate should also be able to calculate yields and values of fixed-income securities.

Fixed-income securities continue to represent the largest capital market segment in the financial ecosystem and the primary means in which institutions, governments, and other issuers raise capital globally. Institutions and individuals use fixed-income investments in a wide range of applications including asset liability management, income generation, and principal preservation. Since the global financial crisis of 2008, evaluating risk—in particular, credit risk—for fixed-income securities has become an increasingly important aspect for this asset class.

© 2021 CFA Institute. All rights reserved.

FIXED INCOME

STUDY SESSION

14

Fixed Income (2)

This study session examines the fundamental elements underlying bond returns and risks with a specific focus on interest rate and credit risk. Duration, convexity, and other key measures for assessing a bond's sensitivity to interest rate risk are introduced. An explanation of credit risk and the use of credit analysis for risky bonds concludes the session.

READING ASSIGNMENTS

Reading 43	Understanding Fixed-Income Risk and Return by James F. Adams, PhD, CFA, and Donald J. Smith, PhD
Reading 44	Fundamentals of Credit Analysis by Christopher L. Gootkind, CFA

© 2021 CFA Institute. All rights reserved.

READING

43

Understanding Fixed-Income Risk and Return

by James F. Adams, PhD, CFA, and Donald J. Smith, PhD

James F. Adams, PhD, CFA, is at New York University (USA). Donald J. Smith, PhD, is at Boston University Questrom School of Business (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** calculate and interpret the sources of return from investing in a fixed-rate bond;
☐	**b.** define, calculate, and interpret Macaulay, modified, and effective durations;
☐	**c.** explain why effective duration is the most appropriate measure of interest rate risk for bonds with embedded options;
☐	**d.** define key rate duration and describe the use of key rate durations in measuring the sensitivity of bonds to changes in the shape of the benchmark yield curve;
☐	**e.** explain how a bond's maturity, coupon, and yield level affect its interest rate risk;
☐	**f.** calculate the duration of a portfolio and explain the limitations of portfolio duration;
☐	**g.** calculate and interpret the money duration of a bond and price value of a basis point (PVBP);
☐	**h.** calculate and interpret approximate convexity and compare approximate and effective convexity;
☐	**i.** calculate the percentage price change of a bond for a specified change in yield, given the bond's approximate duration and convexity;
☐	**j.** describe how the term structure of yield volatility affects the interest rate risk of a bond;

(continued)

© 2021 CFA Institute. All rights reserved.

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**k.** describe the relationships among a bond's holding period return, its duration, and the investment horizon;
☐	**l.** explain how changes in credit spread and liquidity affect yield-to-maturity of a bond and how duration and convexity can be used to estimate the price effect of the changes.
☐	**m.** describe the difference between empirical duration and analytical duration.

1

INTRODUCTION

Successful analysts must develop a solid understanding of the risk and return characteristics of fixed-income investments. Beyond the vast global market for public and private fixed-rate bonds, many financial assets and liabilities with known future cash flows you will encounter throughout your career are evaluated using similar principles. This analysis starts with the yield-to-maturity, or internal rate of return on future cash flows, introduced in the fixed-income valuation reading. Fixed-rate bond returns are affected by many factors, the most important of which is the full receipt of all interest and principal payments on scheduled dates. Assuming no default, return is also affected by interest rate changes that affect coupon reinvestment and the bond price if it is sold prior to maturity. Price change measures may be derived from the mathematical relationship used to calculate a bond's price. Specifically, duration estimates the price change for a given change in interest rates, and convexity improves on the duration estimate by considering that the price and yield-to-maturity relationship of a fixed-rate bond is non-linear.

Sources of return on a fixed-rate bond investment include the receipt and reinvestment of coupon payments and either the redemption of principal if the bond is held to maturity or capital gains (or losses) if the bond is sold earlier. Fixed-income investors holding the same bond may have different interest rate risk exposures if their investment horizons differ.

We introduce bond duration and convexity, showing how these statistics are calculated and used as interest rate risk measures. Although procedures and formulas exist to calculate duration and convexity, these statistics can be approximated using basic bond-pricing techniques and a financial calculator. Commonly used versions of the statistics are covered, including Macaulay, modified, effective, and key rate durations, and we distinguish between risk measures based on changes in the bond's yield-to-maturity (i.e., *yield* duration and convexity) and on benchmark yield curve changes (i.e., *curve* duration and convexity).

We then return to the investment time horizon. When an investor has a short-term horizon, duration and convexity are used to estimate the change in the bond price. Note that yield volatility matters, because bonds with varying times-to-maturity have different degrees of yield volatility. When an investor has a long-term horizon, the interaction between coupon reinvestment risk and market price risk matters. The relationship among interest rate risk, bond duration, and the investment horizon is explored.

Finally, we discuss how duration and convexity may be extended to credit and liquidity risks and highlight how these factors can affect a bond's return and risk. In addition, we highlight the use of statistical methods and historical data to establish empirical as opposed to analytical duration estimates.

SOURCES OF RETURN

2

a calculate and interpret the sources of return from investing in a fixed-rate bond

Fixed-rate bond investors have three sources of return: (1) receipt of promised coupon and principal payments on the scheduled dates, (2) reinvestment of coupon payments, and (3) potential capital gains or losses on the sale of the bond prior to maturity. In this section, it is assumed that the issuer makes the coupon and principal payments as scheduled. Here, the focus is primarily on how interest rate changes affect the reinvestment of coupon payments and a bond's market price if sold prior to maturity. Credit risk is considered later and is also the primary subject of a subsequent reading.

When a bond is purchased at a premium or a discount, it adds another aspect to the rate of return. Recall from the fixed-income valuation reading that a discount bond offers the investor a "deficient" coupon rate below the market discount rate. The amortization of this discount in each period brings the return in line with the market discount rate as the bond's carrying value is "pulled to par." For a premium bond, the coupon rate exceeds the market discount rate and the amortization of the premium adjusts the return to match the market discount rate. Through amortization, the bond's carrying value reaches par value at maturity.

A series of examples will demonstrate the effect of a change in interest rates on two investors' realized rate of returns. Interest rates are the rates at which coupon payments are reinvested and the market discount rates at the time of purchase and at the time of sale if the bond is not held to maturity. In Examples 1 and 2, interest rates are unchanged. The two investors, however, have different time horizons for holding the bond. Examples 3 and 4 show the impact of higher interest rates on the two investors' total return. Examples 5 and 6 show the impact of lower interest rates. In each of the six examples, an investor initially buys a 10-year, 8% annual coupon payment bond at a price of 85.503075 per 100 of par value. The bond's yield-to-maturity is 10.40%.

$$85.503075 = \frac{8}{(1+r)^1} + \frac{8}{(1+r)^2} + \frac{8}{(1+r)^3} + \frac{8}{(1+r)^4} + \frac{8}{(1+r)^5} + \frac{8}{(1+r)^6} + \frac{8}{(1+r)^7} + \frac{8}{(1+r)^8} + \frac{8}{(1+r)^9} + \frac{108}{(1+r)^{10}},$$

$$r = 0.1040$$

EXAMPLE 1

A "buy-and-hold" investor purchases a 10-year, 8% annual coupon payment bond at 85.503075 per 100 of par value and holds it until maturity. The investor receives the series of 10 coupon payments of 8 (per 100 of par value) for a total of 80, plus the redemption of principal (100) at maturity. In addition to collecting the coupon interest and the principal, the investor may reinvest the cash flows. If the coupon payments are reinvested at 10.40%, the future value of the coupons on the bond's maturity date is 129.970678 per 100 of par value.

$$\left[8 \times (1.1040)^9\right] + \left[8 \times (1.1040)^8\right] + \left[8 \times (1.1040)^7\right] + \left[8 \times (1.1040)^6\right] + \left[8 \times (1.1040)^5\right] + \left[8 \times (1.1040)^4\right] + \left[8 \times (1.1040)^3\right] + \left[8 \times (1.1040)^2\right] + \left[8 \times (1.1040)^1\right] + 8 = 129.970678$$

The first coupon payment of 8 is reinvested at 10.40% for nine years until maturity, the second is reinvested for eight years, and so forth. The future value of the annuity is obtained easily on a financial calculator, using 8 for the payment that is received at the end of each of the 10 periods. The amount in excess of the coupons, 49.970678 (= 129.970678 – 80), is the "interest-on-interest" gain from compounding.

The investor's total return is 229.970678, the sum of the reinvested coupons (129.970678) and the redemption of principal at maturity (100). The realized rate of return is 10.40%.

$$85.503075 = \frac{229.970678}{(1 + r)^{10}}, \ r = 0.1040$$

Example 1 demonstrates that the yield-to-maturity at the time of purchase measures the investor's rate of return under three assumptions: (1) The investor holds the bond to maturity, (2) there is no default by the issuer, and (3) the coupon interest payments are reinvested at that same rate of interest.

Example 2 considers another investor who buys the 10-year, 8% annual coupon payment bond and pays the same price. This investor, however, has a four-year investment horizon. Therefore, coupon interest is only reinvested for four years, and the bond is sold immediately after receiving the fourth coupon payment.

EXAMPLE 2

A second investor buys the 10-year, 8% annual coupon payment bond and sells the bond after four years. Assuming that the coupon payments are reinvested at 10.40% for four years, the future value of the reinvested coupons is 37.347111 per 100 of par value.

$$\left[8 \times (1.1040)^3\right] + \left[8 \times (1.1040)^2\right] + \left[8 \times (1.1040)^1\right] + 8 = 37.347111$$

The interest-on-interest gain from compounding is 5.347111 (= 37.347111 – 32). After four years, when the bond is sold, it has six years remaining until maturity. If the yield-to-maturity remains 10.40%, the sale price of the bond is 89.668770.

$$\frac{8}{(1.1040)^1} + \frac{8}{(1.1040)^2} + \frac{8}{(1.1040)^3} + \frac{8}{(1.1040)^4} + \frac{8}{(1.1040)^5} + \frac{8}{(1.1040)^6} = 89.668770$$

The total return is 127.015881 (= 37.347111 + 89.668770), and the realized rate of return is 10.40%.

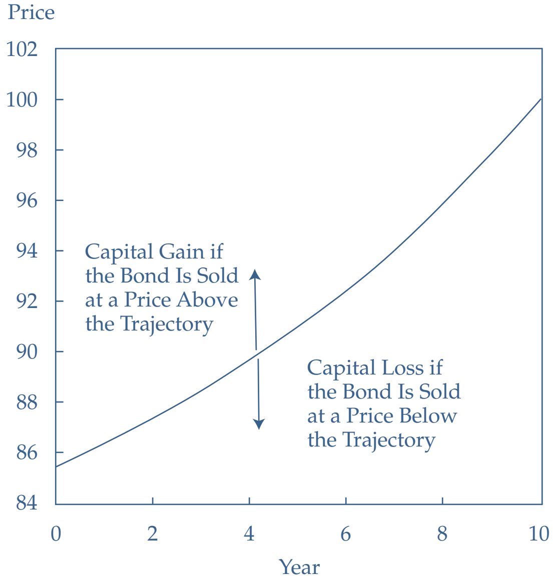
$$85.503075 = \frac{127.015881}{(1 + r)^4}, r = 0.1040$$

In Example 2, the investor's **horizon yield** is 10.40%. A horizon yield is the internal rate of return between the total return (the sum of reinvested coupon payments and the sale price or redemption amount) and the purchase price of the bond. The horizon yield on a bond investment is the annualized holding-period rate of return.

Example 2 demonstrates that the realized horizon yield matches the original yield-to-maturity if: (1) coupon payments are reinvested at the same interest rate as the original yield-to-maturity, and (2) the bond is sold at a price on the constant-yield price trajectory, which implies that the investor does not have any capital gains or losses when the bond is sold.

Capital gains arise if a bond is sold at a price above its constant-yield price trajectory and capital losses occur if a bond is sold at a price below its constant-yield price trajectory. This trajectory is based on the yield-to-maturity when the bond is purchased. The trajectory is shown in Exhibit 1 for a 10-year, 8% annual payment bond purchased at a price of 85.503075 per 100 of par value.

Exhibit 1 Constant-Yield Price Trajectory for a 10-Year, 8% Annual Payment Bond



Note: Price is price per 100 of par value.

A point on the trajectory represents the **carrying value** of the bond at that time. The carrying value is the purchase price plus the amortized amount of the discount if the bond is purchased at a price below par value. If the bond is purchased at a price above par value, the carrying value is the purchase price minus the amortized amount of the premium.

The amortized amount for each year is the change in the price between two points on the trajectory. The initial price of the bond is 85.503075 per 100 of par value. Its price (the carrying value) after one year is 86.395394, calculated using the original yield-to-maturity of 10.40%. Therefore, the amortized amount for the first year is 0.892320 (= 86.395394 − 85.503075). The bond price in Example 2 increases from 85.503075 to 89.668770, and that increase over the four years is movement *along* the constant-yield price trajectory. At the time the bond is sold, its carrying value is also 89.668770, so there is no capital gain or loss.

Examples 3 and 4 demonstrate the impact on investors' realized horizon yields if interest rates go up by 100 basis points (bps). The market discount rate on the bond increases from 10.40% to 11.40%. Coupon reinvestment rates go up by 100 bps as well.

EXAMPLE 3

The buy-and-hold investor purchases the 10-year, 8% annual payment bond at 85.503075. After the bond is purchased and before the first coupon is received, interest rates go up to 11.40%. The future value of the reinvested coupons at 11.40% for 10 years is 136.380195 per 100 of par value.

$$\left[8\times(1.1140)^9\right]+\left[8\times(1.1140)^8\right]+\left[8\times(1.1140)^7\right]+\left[8\times(1.1140)^6\right]+$$
$$\left[8\times(1.1140)^5\right]+\left[8\times(1.1140)^4\right]+\left[8\times(1.1140)^3\right]+\left[8\times(1.1140)^2\right]+$$
$$\left[8\times(1.1140)^1\right]+8=136.380195$$

The total return is 236.380195 (= 136.380195 + 100). The investor's realized rate of return is 10.70%.

$$85.503075=\frac{236.380195}{(1+r)^{10}},\ r=0.1070$$

In Example 3, the buy-and-hold investor benefits from the higher coupon reinvestment rate. The realized horizon yield is 10.70%, 30 bps higher than the outcome in Example 1, when interest rates are unchanged. There is no capital gain or loss because the bond is held until maturity. The carrying value at the maturity date is par value, the same as the redemption amount.

EXAMPLE 4

The second investor buys the 10-year, 8% annual payment bond at 85.503075 and sells it in four years. After the bond is purchased, interest rates go up to 11.40%. The future value of the reinvested coupons at 11.40% after four years is 37.899724 per 100 of par value.

$$\left[8\times(1.1140)^3\right]+\left[8\times(1.1140)^2\right]+\left[8\times(1.1140)^1\right]+8=37.899724$$

The sale price of the bond after four years is 85.780408.

$$\frac{8}{(1.1140)^1}+\frac{8}{(1.1140)^2}+\frac{8}{(1.1140)^3}+\frac{8}{(1.1140)^4}+$$
$$\frac{8}{(1.1140)^5}+\frac{108}{(1.1140)^6}=85.780408$$

The total return is 123.680132 (= 37.899724 + 85.780408), resulting in a realized four-year horizon yield of 9.67%.

$$85.503075 = \frac{123.680132}{(1+r)^4},\ r = 0.0967$$

In Example 4, the second investor has a lower realized rate of return compared with the investor in Example 2, in which interest rates are unchanged. The future value of reinvested coupon payments goes up by 0.552613 (= 37.899724 – 37.347111) per 100 of par value because of the higher interest rates. But there is a *capital loss* of 3.888362 (= 89.668770 – 85.780408) per 100 of par value. Notice that the capital loss is measured from the bond's carrying value, the point on the constant-yield price trajectory, and not from the original purchase price. The bond is now sold at a price below the constant-yield price trajectory. The reduction in the realized four-year horizon yield from 10.40% to 9.67% is a result of the capital loss being greater than the gain from reinvesting coupons at a higher rate, which reduces the investor's total return.

Examples 5 and 6 complete the series of rate-of-return calculations for the two investors. Interest rates decline by 100 bps. The required yield on the bond falls from 10.40% to 9.40% after the purchase of the bond. The interest rates at which the coupon payments are reinvested fall as well.

EXAMPLE 5

The buy-and-hold investor purchases the 10-year bond at 85.503075 and holds the security until it matures. After the bond is purchased and before the first coupon is received, interest rates go down to 9.40%. The future value of reinvesting the coupon payments at 9.40% for 10 years is 123.888356 per 100 of par value.

$$\left[8\times(1.0940)^9\right] + \left[8\times(1.0940)^8\right] + \left[8\times(1.0940)^7\right] + \left[8\times(1.0940)^6\right] +$$
$$\left[8\times(1.0940)^5\right] + \left[8\times(1.0940)^4\right] + \left[8\times(1.0940)^3\right] + \left[8\times(1.0940)^2\right] +$$
$$\left[8\times(1.0940)^1\right] + 8 = 123.888356$$

The total return is 223.888356, the sum of the future value of reinvested coupons and the redemption of par value. The investor's realized rate of return is 10.10%.

$$85.503075 = \frac{223.888365}{(1+r)^{10}},\ r = 0.1010$$

In Example 5, the buy-and-hold investor suffers from the lower coupon reinvestment rates. The realized horizon yield is 10.10%, 30 bps lower than the result in Example 1, when interest rates are unchanged. There is no capital gain or loss because the bond is held until maturity. Examples 1, 3, and 5 indicate that the interest rate risk for a buy-and-hold investor arises entirely from changes in coupon reinvestment rates.

EXAMPLE 6

The second investor buys the 10-year bond at 85.503075 and sells it in four years. After the bond is purchased, interest rates go down to 9.40%. The future value of the reinvested coupons at 9.40% is 36.801397 per 100 of par value.

$$\left[8 \times (1.0940)^3\right] + \left[8 \times (1.0940)^2\right] + \left[8 \times (1.0940)^1\right] + 8 = 36.801397$$

This reduction in future value is offset by the higher sale price of the bond, which is 93.793912 per 100 of par value.

$$\frac{8}{(1.0940)^1} + \frac{8}{(1.0940)^2} + \frac{8}{(1.0940)^3} + \frac{8}{(1.0940)^4} +$$
$$\frac{8}{(1.0940)^5} + \frac{108}{(1.0940)^6} = 93.793912$$

The total return is 130.595309 (= 36.801397 + 93.793912), and the realized yield is 11.17%.

$$85.503075 = \frac{130.595309}{(1+r)^4}, \quad r = 0.1117$$

The investor in Example 6 has a capital gain of 4.125142 (= 93.793912 – 89.668770). The capital gain is measured from the carrying value, the point on the constant-yield price trajectory. That gain offsets the reduction in the future value of reinvested coupons of 0.545714 (= 37.347111 – 36.801397). The total return is higher than that in Example 2, in which the interest rate remains at 10.40%.

In these examples, interest income for the investor is the return associated with the *passage of time*. Therefore, interest income includes the receipt of coupon interest, the reinvestment of those cash flows, and the amortization of the discount from purchase at a price below par value (or the premium from purchase at a price above par value) to bring the return back in line with the market discount rate. A capital gain or loss is the return to the investor associated with the *change in the value* of the security. On the fixed-rate bond, a change in value arises from a change in the yield-to-maturity, which is the implied market discount rate. In practice, the way interest income and capital gains and losses are calculated and reported on financial statements depends on financial and tax accounting rules.

This series of examples illustrates an important point about fixed-rate bonds: The *investment horizon* is at the heart of understanding interest rate risk and return. There are two offsetting types of interest rate risk that affect the bond investor: coupon reinvestment risk and market price risk. The future value of reinvested coupon payments (and, in a portfolio, the principal on bonds that mature before the horizon date) *increases* when interest rates rise and *decreases* when rates fall. The sale price on a bond that matures after the horizon date (and thus needs to be sold) *decreases* when interest rates rise and *increases* when rates fall. Coupon reinvestment risk matters more when the investor has a long-term horizon relative to the time-to-maturity of the bond. For instance, a buy-and-hold investor only has coupon reinvestment risk. Market price risk matters more when the investor has a short-term horizon relative to the time-to-maturity. For example, an investor who sells the bond before the first coupon is received has only market price risk. Therefore, two investors holding the same bond (or bond portfolio) can have different exposures to interest rate risk if they have different investment horizons.

EXAMPLE 7

An investor buys a four-year, 10% annual coupon payment bond priced to yield 5.00%. The investor plans to sell the bond in two years once the second coupon payment is received. Calculate the purchase price for the bond and the horizon yield assuming that the coupon reinvestment rate after the bond purchase and the yield-to-maturity at the time of sale are (1) 3.00%, (2) 5.00%, and (3) 7.00%.

Solution:

The purchase price is 117.729753.

$$\frac{10}{(1.0500)^1} + \frac{10}{(1.0500)^2} + \frac{10}{(1.0500)^3} + \frac{110}{(1.0500)^4} = 117.729753$$

1 3.00%: The future value of reinvested coupons is 20.300.

$$(10 \times 1.0300) + 10 = 20.300$$

The sale price of the bond is 113.394288.

$$\frac{10}{(1.0300)^1} + \frac{110}{(1.0300)^2} = 113.394288$$

Total return: 20.300 + 113.394288 = 133.694288.

If interest rates go down from 5.00% to 3.00%, the realized rate of return over the two-year investment horizon is 6.5647%, higher than the original yield-to-maturity of 5.00%.

$$117.729753 = \frac{133.694288}{(1+r)^2}, \quad r = 0.065647$$

2 5.00%: The future value of reinvested coupons is 20.500.

$$(10 \times 1.0500) + 10 = 20.500$$

The sale price of the bond is 109.297052.

$$\frac{10}{(1.0500)^1} + \frac{110}{(1.0500)^2} = 109.297052$$

Total return: 20.500 + 109.297052 = 129.797052.

If interest rates remain 5.00% for reinvested coupons and for the required yield on the bond, the realized rate of return over the two-year investment horizon is equal to the yield-to-maturity of 5.00%.

$$117.729753 = \frac{129.797052}{(1+r)^2}, \quad r = 0.050000$$

3 7.00%: The future value of reinvested coupons is 20.700.

$$(10 \times 1.0700) + 10 = 20.700$$

The bond is sold at 105.424055.

$$\frac{10}{(1.0700)^1} + \frac{110}{(1.0700)^2} = 105.424055$$

Total return: 20.700 + 105.424055 = 126.124055.

$$117.729753 = \frac{126.124055}{(1 + r)^2}, \quad r = 0.035037$$

If interest rates go up from 5.00% to 7.00%, the realized rate of return over the two-year investment horizon is 3.5037%, lower than the yield-to-maturity of 5.00%.

3 MACAULAY AND MODIFIED DURATION

b define, calculate, and interpret Macaulay, modified, and effective durations

This section covers two commonly used measures of interest rate risk: duration and convexity. It distinguishes between risk measures based on changes in a bond's own yield-to-maturity (yield duration and convexity) and those that affect the bond based on changes in a benchmark yield curve (curve duration and convexity).

3.1 Macaulay, Modified, and Approximate Duration

The duration of a bond measures the sensitivity of the bond's full price (including accrued interest) to changes in the bond's yield-to-maturity or, more generally, to changes in benchmark interest rates. Duration estimates changes in the bond price assuming that variables other than the yield-to-maturity or benchmark rates are held constant. Most importantly, the time-to-maturity is unchanged. Therefore, duration measures the *instantaneous* (or, at least, same-day) change in the bond price. The accrued interest is the same, so it is the flat price that goes up or down when the full price changes. Duration is a useful measure because it represents the approximate amount of time a bond would have to be held for the market discount rate at purchase to be realized if there is a single change in interest rate. If the bond is held for the duration period, an increase from reinvesting coupons is offset by a decrease in price if interest rates increase and a decrease from reinvesting coupons is offset by an increase in price if interest rates decrease.

There are several types of bond duration. In general, these can be divided into **yield duration** and **curve duration**. Yield duration is the sensitivity of the bond price with respect to the bond's own yield-to-maturity. Curve duration is the sensitivity of the bond price (or more generally, the market value of a financial asset or liability) with respect to a benchmark yield curve. The benchmark yield curve could be the government yield curve on coupon bonds, the spot curve, or the forward curve, but in practice, the government par curve is often used. Yield duration statistics used in fixed-income analysis include Macaulay duration, modified duration, money duration, and the price value of a basis point (PVBP). A curve duration statistic often used is effective duration. Effective duration is covered later in this reading.

Macaulay duration is named after Frederick Macaulay, the Canadian economist who first wrote about the statistic in 1938. Equation 1 is a general formula to calculate the Macaulay duration (MacDur) of a traditional fixed-rate bond.

$$\text{MacDur} = \left[\frac{\frac{(1-t/T)\times PMT}{(1+r)^{1-t/T}} + \frac{(2-t/T)\times PMT}{(1+r)^{2-t/T}} + \cdots + \frac{(N-t/T)\times(PMT+FV)}{(1+r)^{N-t/T}}}{\frac{PMT}{(1+r)^{1-t/T}} + \frac{PMT}{(1+r)^{2-t/T}} + \cdots + \frac{PMT+FV}{(1+r)^{N-t/T}}}\right] \quad (1)$$

where

t = the number of days from the last coupon payment to the settlement date

T = the number of days in the coupon period

t/T = the fraction of the coupon period that has gone by since the last payment

PMT = the coupon payment per period

FV = the future value paid at maturity, or the par value of the bond

r = the yield-to-maturity, or the market discount rate, per period

N = the number of evenly spaced periods to maturity as of the beginning of the current period

The denominator in Equation 1 is the full price (PV^{Full}) of the bond including accrued interest. It is the present value of the coupon interest and principal payments, with each cash flow discounted by the same market discount rate, r.

$$PV^{Full} = \frac{PMT}{(1+r)^{1-t/T}} + \frac{PMT}{(1+r)^{2-t/T}} + \cdots + \frac{PMT+FV}{(1+r)^{N-t/T}} \quad (2)$$

Equation 3 combines Equations 1 and 2 to reveal an important aspect of the Macaulay duration: Macaulay duration is a weighted average of the time to receipt of the bond's promised payments, where the weights are the shares of the full price that correspond to each of the bond's promised future payments.

$$\text{MacDur} = \left\{\begin{array}{l} (1-t/T)\left[\dfrac{\frac{PMT}{(1+r)^{1-t/T}}}{PV^{Full}}\right] + (2-t/T)\left[\dfrac{\frac{PMT}{(1+r)^{2-t/T}}}{PV^{Full}}\right] + \cdots + \\ (N-t/T)\left[\dfrac{\frac{PMT+FV}{(1+r)^{N-t/T}}}{PV^{Full}}\right] \end{array}\right\} \quad (3)$$

The times to receipt of cash flow measured in terms of time periods are $1 - t/T$, $2 - t/T, \ldots, N - t/T$. The weights are the present values of the cash flows divided by the full price. Therefore, Macaulay duration is measured in terms of time periods. A couple of examples will clarify this calculation.

Consider first the 10-year, 8% annual coupon payment bond used in Examples 1–6. The bond's yield-to-maturity is 10.40%, and its price is 85.503075 per 100 of par value. This bond has 10 evenly spaced periods to maturity. Settlement is on a coupon payment date so that $t/T = 0$. Exhibit 2 illustrates the calculation of the bond's Macaulay duration.

Exhibit 2 Macaulay Duration of a 10-Year, 8% Annual Payment Bond

Period	Cash Flow	Present Value	Weight	Period × Weight
1	8	7.246377	0.08475	0.0847
2	8	6.563747	0.07677	0.1535
3	8	5.945423	0.06953	0.2086
4	8	5.385347	0.06298	0.2519
5	8	4.878032	0.05705	0.2853
6	8	4.418507	0.05168	0.3101
7	8	4.002271	0.04681	0.3277
8	8	3.625245	0.04240	0.3392
9	8	3.283737	0.03840	0.3456
10	108	40.154389	0.46963	4.6963
		85.503075	1.00000	7.0029

The first two columns of Exhibit 2 show the number of periods to the receipt of the cash flow and the amount of the payment per 100 of par value. The third column is the present value of the cash flow. For example, the final payment is 108 (the last coupon payment plus the redemption of principal) and its present value is 40.154389.

$$\frac{108}{(1.1040)^{10}} = 40.154389$$

The sum of the present values is the full price of the bond. The fourth column is the weight, the share of total market value corresponding to each cash flow. The final payment of 108 per 100 of par value is 46.963% of the bond's market value.

$$\frac{40.154389}{85.503075} = 0.46963$$

The sum of the weights is 1.00000. The fifth column is the number of periods to the receipt of the cash flow (the first column) multiplied by the weight (the fourth column). The sum of that column is 7.0029, which is the Macaulay duration of this 10-year, 8% annual coupon payment bond. This statistic is sometimes reported as 7.0029 *years,* although the time frame is not needed in most applications.

Now consider an example *between* coupon payment dates. A 6% semiannual payment corporate bond that matures on 14 February 2027 is purchased for settlement on 11 April 2019. The coupon payments are 3 per 100 of par value, paid on 14 February and 14 August of each year. The yield-to-maturity is 6.00% quoted on a street-convention semiannual bond basis. The full price of this bond comprises the flat price plus accrued interest. The flat price for the bond is 99.990423 per 100 of par value. The accrued interest is calculated using the 30/360 method to count days. This settlement date is 57 days into the 180-day semiannual period, so $t/T = 57/180$. The accrued interest is 0.950000 (= 57/180 × 3) per 100 of par value. The full price for the bond is 100.940423 (= 99.990423 + 0.950000). Exhibit 3 shows the calculation of the bond's Macaulay duration.

Exhibit 3 Macaulay Duration of an Eight-Year, 6% Semiannual Payment Bond Priced to Yield 6.00%

Period	Time to Receipt	Cash Flow	Present Value	Weight	Time × Weight
1	0.6833	3	2.940012	0.02913	0.019903
2	1.6833	3	2.854381	0.02828	0.047601
3	2.6833	3	2.771244	0.02745	0.073669
4	3.6833	3	2.690528	0.02665	0.098178
5	4.6833	3	2.612163	0.02588	0.121197
6	5.6833	3	2.536080	0.02512	0.142791
7	6.6833	3	2.462214	0.02439	0.163025
8	7.6833	3	2.390499	0.02368	0.181959
9	8.6833	3	2.320873	0.02299	0.199652
10	9.6833	3	2.253275	0.02232	0.216159
11	10.6833	3	2.187645	0.02167	0.231536
12	11.6833	3	2.123927	0.02104	0.245834
13	12.6833	3	2.062065	0.02043	0.259102
14	13.6833	3	2.002005	0.01983	0.271389
15	14.6833	3	1.943694	0.01926	0.282740
16	15.6833	103	64.789817	0.64186	10.066535
			100.940423	1.00000	12.621268

There are 16 semiannual periods to maturity between the last coupon payment date of 14 February 2019 and maturity on 14 February 2027. The time to receipt of cash flow in semiannual periods is in the second column: 0.6833 = 1 – 57/180, 1.6833 = 2 – 57/180, etc. The cash flow for each period is in the third column. The annual yield-to-maturity is 6.00%, so the yield per semiannual period is 3.00%. When that yield is used to get the present value of each cash flow, the full price of the bond is 100.940423, the sum of the fourth column. The weights, which are the shares of the full price corresponding to each cash flow, are in the fifth column. The Macaulay duration is the sum of the items in the sixth column, which is the weight multiplied by the time to receipt of each cash flow. The result, 12.621268, is the Macaulay duration on an eight-year, 6% semiannual payment bond for settlement on 11 April 2019 measured in *semiannual periods*. Similar to coupon rates and yields-to-maturity, duration statistics invariably are annualized in practice. Therefore, the Macaulay duration typically is reported as 6.310634 *years* (= 12.621268/2). (Such precision for the duration statistic is not needed in practice. Typically, "6.31 years" is enough. The full precision is shown here to illustrate calculations.) Microsoft Excel users can obtain the Macaulay duration using the DURATION financial function—DURATION(DATE(2019,4,11),DATE(2027,2,14),0.06,0.06,2,0)—and inputs that include the settlement date, maturity date, annual coupon rate as a decimal, annual yield-to-maturity as a decimal, periodicity, and day count code (0 for 30/360, 1 for actual/actual).

Another approach to calculating the Macaulay duration is to use a closed-form equation derived using calculus and algebra (see Smith 2014). Equation 4 is a general closed-form formula for determining the Macaulay duration of a fixed-rate bond, where c is the coupon rate per period (PMT/FV).

$$\text{MacDur} = \left\{\frac{1+r}{r} - \frac{1+r+\left[N\times(c-r)\right]}{c\times\left[(1+r)^N-1\right]+r}\right\} - (t/T) \tag{4}$$

The Macaulay duration of the 10-year, 8% annual payment bond is calculated by entering $r = 0.1040$, $c = 0.0800$, $N = 10$, and $t/T = 0$ into Equation 4.

$$\text{MacDur} = \frac{1+0.1040}{0.1040} - \frac{1+0.1040+\left[10\times(0.0800-0.1040)\right]}{0.0800\times\left[(1+0.1040)^{10}-1\right]+0.1040} = 7.0029$$

Therefore, the weighted average time to receipt of the interest and principal payments that will result in realization of the initial market discount rate on this 10-year bond is 7.00 years.

The Macaulay duration of the 6% semiannual payment bond maturing on 14 February 2027 is obtained by entering $r = 0.0300$, $c = 0.0300$, $N = 16$, and $t/T = 57/180$ into Equation 4.

$$\text{MacDur} = \left[\frac{1+0.0300}{0.0300} - \frac{1+0.0300+\left[16\times(0.0300-0.0300)\right]}{0.0300\times\left[(1+0.0300)^{16}-1\right]+0.0300}\right] - (57/180)$$
$$= 12.621268$$

Equation 4 uses the yield-to-maturity *per period*, the coupon rate *per period*, the number of *periods* to maturity, and the fraction of the current *period* that has gone by. Its output is the Macaulay duration in terms of *periods*. It is converted to annual duration by dividing by the number of periods in the year.

The calculation of the **modified duration** (ModDur) statistic of a bond requires a simple adjustment to Macaulay duration. It is the Macaulay duration statistic divided by one plus the yield per period.

$$\text{ModDur} = \frac{\text{MacDur}}{1+r} \tag{5}$$

For example, the modified duration of the 10-year, 8% annual payment bond is 6.3432.

$$\text{ModDur} = \frac{7.0029}{1.1040} = 6.3432$$

The modified duration of the 6% semiannual payment bond maturing on 14 February 2027 is 12.253658 semiannual periods.

$$\text{ModDur} = \frac{12.621268}{1.0300} = 12.253658$$

The annualized modified duration of the bond is 6.126829 (= 12.253658/2).

Microsoft Excel users can obtain the modified duration using the MDURATION financial function using the same inputs as for the Macaulay duration: MDURATION(DATE(2019,4,11),DATE(2027,2,14),0.06,0.06,2,0). Although modified duration might seem to be just a Macaulay duration with minor adjustments, it has an important application in risk measurement: Modified duration provides an estimate of the percentage price change for a bond given a change in its yield-to-maturity.

$$\%\Delta PV^{Full} \approx -\text{AnnModDur} \times \Delta\text{Yield} \tag{6}$$

The percentage price change refers to the full price, including accrued interest. The AnnModDur term in Equation 6 is the *annual* modified duration, and the ΔYield term is the change in the *annual* yield-to-maturity. The ≈ sign indicates that this calculation is an estimation. The minus sign indicates that bond prices and yields-to-maturity move inversely.

If the annual yield on the 6% semiannual payment bond that matures on 14 February 2027 jumps by 100 bps, from 6.00% to 7.00%, the estimated loss in value for the bond is 6.1268%.

$$\%\Delta PV^{Full} \approx -6.126829 \times 0.0100 = -0.061268$$

If the yield-to-maturity were to drop by 100 bps to 5.00%, the estimated gain in value is also 6.1268%.

$$\%\Delta PV^{Full} \approx -6.126829 \times -0.0100 = 0.061268$$

Modified duration provides a *linear* estimate of the percentage price change. In terms of absolute value, the change is the same for either an increase or a decrease in the yield-to-maturity. Recall that for a given coupon rate and time-to-maturity, the percentage price change is greater (in absolute value) when the market discount rate goes down than when it goes up. Later in this reading, a "convexity adjustment" to duration is introduced. It improves the accuracy of this estimate, especially when a large change in yield-to-maturity (such as 100 bps) is considered.

APPROXIMATE MODIFIED AND MACAULAY DURATION

4

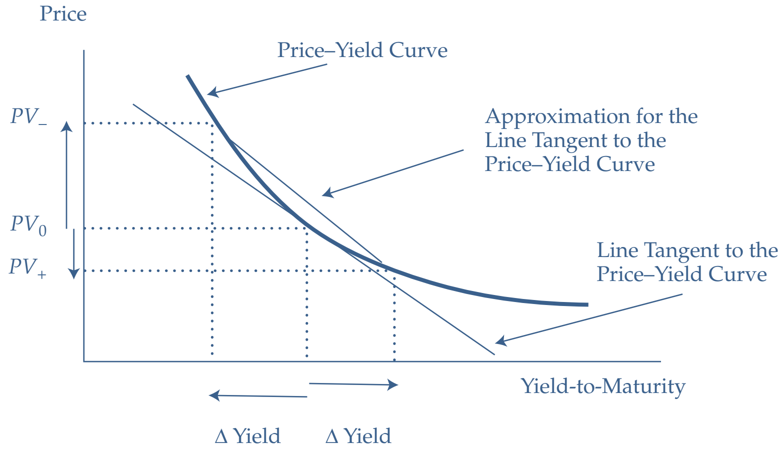
b define, calculate, and interpret Macaulay, modified, and effective durations

The modified duration statistic for a fixed-rate bond is easily obtained if the Macaulay duration is already known. An alternative approach is to *approximate* modified duration directly. Equation 7 is the approximation formula for annual modified duration.

$$\text{ApproxModDur} = \frac{(PV_{-}) - (PV_{+})}{2 \times (\Delta\text{Yield}) \times (PV_0)} \tag{7}$$

The objective of the approximation is to estimate the slope of the line tangent to the price–yield curve. The slope of the tangent and the approximated slope are shown in Exhibit 4.

Exhibit 4 Approximate Modified Duration


Price
Price–Yield Curve
PV_-
PV_0
PV_+
Approximation for the Line Tangent to the Price–Yield Curve
Line Tangent to the Price–Yield Curve
Yield-to-Maturity
Δ Yield
Δ Yield

To estimate the slope, the yield-to-maturity is changed up and down by the same amount—the ΔYield. Then the bond prices given the new yields-to-maturity are calculated. The price when the yield is increased is denoted PV_+. The price when the yield-to-maturity is reduced is denoted PV_-. The original price is PV_0. These prices are the full prices, including accrued interest. The slope of the line based on PV_+ and PV_- is the approximation for the slope of the line tangent to the price–yield curve. The following example illustrates the remarkable accuracy of this approximation. In fact, as ΔYield approaches zero, the approximation approaches AnnModDur.

Consider the 6% semiannual coupon payment corporate bond maturing on 14 February 2027. For settlement on 11 April 2019, the full price (PV_0) is 100.940423 given that the yield-to-maturity is 6.00%.

$$PV_0 = \left[\frac{3}{(1.03)^1} + \frac{3}{(1.03)^2} + \cdots + \frac{103}{(1.03)^{16}}\right] \times (1.03)^{57/180} = 100.940423$$

Raise the annual yield-to-maturity by five bps, from 6.00% to 6.05%. This increase corresponds to an increase in the yield-to-maturity per semiannual period of 2.5 bps, from 3.00% to 3.025% per period. The new full price (PV_+) is 100.631781.

$$PV_+ = \left[\frac{3}{(1.03025)^1} + \frac{3}{(1.03025)^2} + \cdots + \frac{103}{(1.03025)^{16}}\right] \times (1.03025)^{57/180} = 100.631781$$

Lower the annual yield-to-maturity by five bps, from 6.00% to 5.95%. This decrease corresponds to a decrease in the yield-to-maturity per semiannual period of 2.5 bps, from 3.00% to 2.975% per period. The new full price (PV_-) is 101.250227.

$$PV_- = \left[\frac{3}{(1.02975)^1} + \frac{3}{(1.02975)^2} + \cdots + \frac{103}{(1.02975)^{16}}\right] \times (1.02975)^{57/180} = 101.250227$$

Enter these results into Equation 7 for the 5 bp change in the annual yield-to-maturity, or ΔYield = 0.0005:

$$\text{ApproxModDur} = \frac{101.250227 - 100.631781}{2 \times 0.0005 \times 100.940423} = 6.126842$$

The "exact" annual modified duration for this bond is 6.126829 and the "approximation" is 6.126842—virtually identical results. Therefore, although duration can be calculated using the approach in Exhibits 2 and 3—basing the calculation on the weighted average time to receipt of each cash flow—or using the closed-form formula as in Equation 4, it can also be estimated quite accurately using the basic bond-pricing equation and a financial calculator. The Macaulay duration can be approximated as well—the approximate modified duration multiplied by one plus the yield per period.

$$\text{ApproxMacDur} = \text{ApproxModDur} \times (1 + r) \tag{8}$$

The approximation formulas produce results for *annualized* modified and Macaulay durations. The frequency of coupon payments and the periodicity of the yield-to-maturity are included in the bond price calculations.

EXAMPLE 8

Assume that the 3.75% US Treasury bond that matures on 15 August 2041 is priced to yield 5.14% for settlement on 15 October 2020. Coupons are paid semiannually on 15 February and 15 August. The yield-to-maturity is stated on a street-convention semiannual bond basis. This settlement date is 61 days into a 184-day coupon period, using the actual/actual day-count convention. Compute the approximate modified duration and the approximate Macaulay duration for this Treasury bond assuming a 5 bp change in the yield-to-maturity.

Solution:

The yield-to-maturity per semiannual period is 0.0257 (= 0.0514/2). The coupon payment per period is 1.875 (= 3.75/2). At the beginning of the period, there are 21 years (42 semiannual periods) to maturity. The fraction of the period that has passed is 61/184. The full price at that yield-to-maturity is 82.967530 per 100 of par value.

$$PV_0 = \left[\frac{1.875}{(1.0257)^1} + \frac{1.875}{(1.0257)^2} + \cdots + \frac{101.875}{(1.0257)^{42}}\right] \times (1.0257)^{61/184} = 82.96753$$

Raise the yield-to-maturity from 5.14% to 5.19%—therefore, from 2.57% to 2.595% per semiannual period—and the price becomes 82.411395 per 100 of par value.

$$PV_+ = \left[\frac{1.875}{(1.02595)^1} + \frac{1.875}{(1.02595)^2} + \cdots + \frac{101.875}{(1.02595)^{42}}\right] \times (1.02595)^{61/184}$$
$$= 82.411395$$

Lower the yield-to-maturity from 5.14% to 5.09%—therefore, from 2.57% to 2.545% per semiannual period—and the price becomes 83.528661 per 100 of par value.

$$PV_- = \left[\frac{1.875}{(1.02545)^1} + \frac{1.875}{(1.02545)^2} + \cdots + \frac{101.875}{(1.02545)^{42}}\right] \times (1.02545)^{61/184}$$
$$= 83.528661$$

The approximate annualized modified duration for the Treasury bond is 13.466.

$$\text{ApproxModDur} = \frac{83.528661 - 82.411395}{2 \times 0.0005 \times 82.967530} = 13.466$$

The approximate annualized Macaulay duration is 13.812.

$$\text{ApproxMacDur} = 13.466 \times 1.0257 = 13.812$$

Therefore, from these statistics, the investor knows that the weighted average time to receipt of interest and principal payments is 13.812 years (the Macaulay duration) and that the estimated loss in the bond's market value is 13.466% (the modified duration) if the market discount rate were to suddenly go up by 1% from 5.14% to 6.14%.

5 EFFECTIVE AND KEY RATE DURATION

b define, calculate, and interpret Macaulay, modified, and effective durations

c explain why effective duration is the most appropriate measure of interest rate risk for bonds with embedded options

Another approach to assess the interest rate risk of a bond is to estimate the percentage change in price given a change in a benchmark yield curve—for example, the government par curve. This estimate, which is very similar to the formula for approximate modified duration, is called the **effective duration**. The effective duration of a bond is the sensitivity of the bond's price to a change in a benchmark yield curve. The formula to calculate effective duration (EffDur) is Equation 9.

$$\text{EffDur} = \frac{(PV_-) - (PV_+)}{2 \times (\Delta\text{Curve}) \times (PV_0)} \tag{9}$$

The difference between approximate modified duration and effective duration is in the denominator. Modified duration is a *yield duration* statistic in that it measures interest rate risk in terms of a change in the bond's own yield-to-maturity (ΔYield). Effective duration is a *curve duration* statistic in that it measures interest rate risk in terms of a parallel shift in the benchmark yield curve (ΔCurve).

Effective duration is essential to the measurement of the interest rate risk of a complex bond, such as a bond that contains an embedded call option. The duration of a callable bond is *not* the sensitivity of the bond price to a change in the yield-to-worst (i.e., the lowest of the yield-to-maturity, yield-to-first-call, yield-to-second-call, and so forth). The problem is that future cash flows are uncertain because they are contingent on future interest rates. The issuer's decision to call the bond depends on the ability to refinance the debt at a lower cost of funds. In brief, a callable bond does not have a well-defined internal rate of return (yield-to-maturity). Therefore, yield duration statistics, such as modified and Macaulay durations, do not apply; effective duration is the appropriate duration measure.

The specific option-pricing models that are used to produce the inputs to effective duration for a callable bond are covered in later readings. However, as an example, suppose that the full price of a callable bond is 101.060489 per 100 of par value. The option-pricing model inputs include (1) the length of the call protection period, (2) the schedule of call prices and call dates, (3) an assumption about credit spreads over benchmark yields (which includes any liquidity spread as well), (4) an assumption about future interest rate volatility, and (5) the level of market interest rates (e.g., the government par curve). The analyst then holds the first four inputs constant and raises and lowers the fifth input. Suppose that when the government par curve is raised and lowered by 25 bps, the new full prices for the callable bond from the model are 99.050120 and 102.890738, respectively. Therefore, PV_0 = 101.060489, PV_+ = 99.050120, PV_- = 102.890738, and ΔCurve = 0.0025. The effective duration for the callable bond is 7.6006.

$$\text{EffDur} = \frac{102.890738 - 99.050120}{2 \times 0.0025 \times 101.060489} = 7.6006$$

This curve duration measure indicates the bond's sensitivity to the benchmark yield curve—in particular, the government par curve—assuming no change in the credit spread. In practice, a callable bond issuer might be able to exercise the call option and obtain a lower cost of funds if (1) benchmark yields fall and the credit spread over the benchmark is unchanged or (2) benchmark yields are unchanged and the credit spread is reduced (e.g., because of an upgrade in the issuer's rating). A pricing model can be used to determine a "credit duration" statistic—that is, the sensitivity of the bond price to a change in the credit spread. On a traditional fixed-rate bond, modified duration estimates the percentage price change for a change in the benchmark yield and/or the credit spread. For bonds that do not have a well-defined internal rate of return because the future cash flows are not fixed—for instance, callable bonds and floating-rate notes—pricing models are used to produce different statistics for changes in benchmark interest rates and for changes in credit risk.

Another fixed-income security for which yield duration statistics, such as modified and Macaulay durations, are not relevant is a mortgage-backed bond. These securities arise from a residential (or commercial) loan portfolio securitization. The key point for measuring interest rate risk on a mortgage-backed bond is that the cash flows are contingent on homeowners' ability to refinance their debt at a lower rate. In effect, the homeowners have call options on their mortgage loans.

A practical consideration in using effective duration is in setting the change in the benchmark yield curve. With approximate modified duration, accuracy is improved by choosing a smaller yield-to-maturity change. But the pricing models for more-complex securities, such as callable and mortgage-backed bonds, include assumptions about the behavior of the corporate issuers, businesses, or homeowners. Rates typically need to change by a minimum amount to affect the decision to call a bond or refinance a mortgage loan because issuing new debt involves transaction costs. Therefore, estimates of interest rate risk using effective duration are not necessarily improved by choosing a smaller change in benchmark rates. Effective duration has become an important tool in the financial analysis of not only traditional bonds but also financial liabilities. Example 9 demonstrates such an application of effective duration.

EXAMPLE 9

Defined-benefit pension schemes typically pay retirees a monthly amount based on their wage level at the time of retirement. The amount could be fixed in nominal terms or indexed to inflation. These programs are referred to as "defined-benefit pension plans" when US GAAP or IFRS accounting standards are used. In Australia, they are called "superannuation funds."

A British defined-benefit pension scheme seeks to measure the sensitivity of its retirement obligations to market interest rate changes. The pension scheme manager hires an actuarial consultancy to model the present value of its liabilities under three interest rate scenarios: (1) a base rate of 5%, (2) a 100 bp increase in rates, up to 6%, and (3) a 100 bp drop in rates, down to 4%.

The actuarial consultancy uses a complex valuation model that includes assumptions about employee retention, early retirement, wage growth, mortality, and longevity. The following chart shows the results of the analysis.

Interest Rate Assumption	Present Value of Liabilities
4%	GBP973.5 million
5%	GBP926.1 million
6%	GBP871.8 million

Compute the effective duration of the pension scheme's liabilities.

Solution:

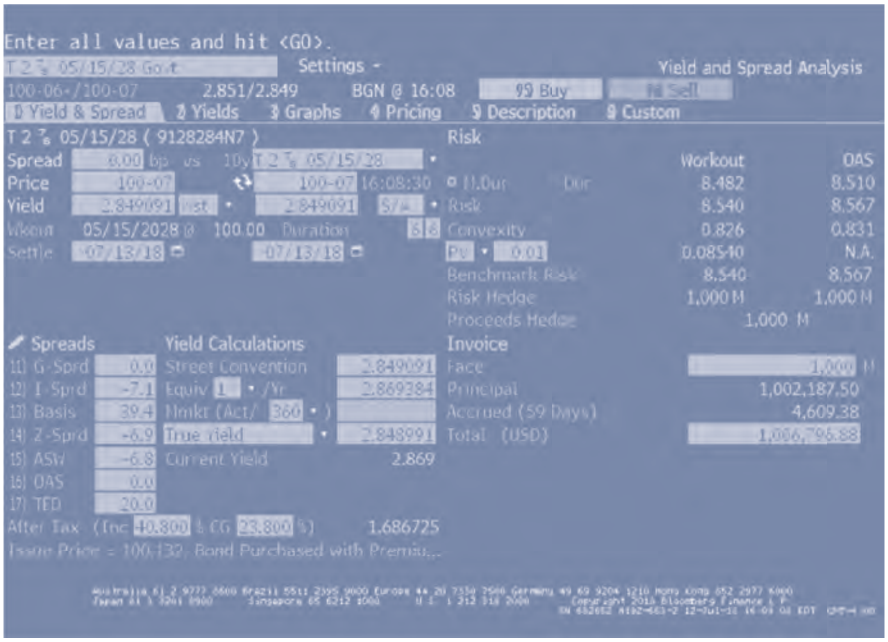
$PV_0 = 926.1$, $PV_+ = 871.8$, $PV_- = 973.5$, and $\Delta\text{Curve} = 0.0100$. The effective duration of the pension scheme's liabilities is 5.49.

$$\text{EffDur} = \frac{973.5 - 871.8}{2 \times 0.0100 \times 926.1} = 5.49$$

This effective duration statistic for the pension scheme's liabilities might be used in asset allocation decisions to decide the mix of equity, fixed income, and alternative assets.

Although effective duration is the most appropriate interest rate risk measure for bonds with embedded options, it also is useful with traditional bonds to supplement the information provided by the Macaulay and modified yield durations. Exhibit 5 displays the Bloomberg Yield and Spread (YAS) Analysis page for the 2.875% US Treasury note that matures on 15 May 2028.

Exhibit 5 Bloomberg YAS Page for the 2.875% US Treasury Note



```
Enter all values and hit <GO>.
T 2 7/8 05/15/28 Govt              Settings -                              Yield and Spread Analysis
100-06+/100-07        2.851/2.849      BGN @ 16:08     99 Buy     98 Sell
1) Yield & Spread   2) Yields   3) Graphs   4) Pricing   5) Description   6) Custom
T 2 7/8 05/15/28 ( 9128284N7 )                          Risk
Spread      0.00 bp  vs   10y T 2 7/8 05/15/28                          Workout        OAS
Price     100-07                100-07  16:08:30        M.Dur     Dur    8.482         8.510
Yield   2.849091  Wst       2.849091    S/A             Risk             8.540         8.567
Wkout   05/15/2028 @   100.00   Duration         6      Convexity        0.826         0.831
Settle   07/13/18                07/13/18               PV    0.01       0.08540       N.A.
                                                        Benchmark Risk   8.540         8.567
                                                        Risk Hedge       1,000 M       1,000 M
                                                        Proceeds Hedge          1,000 M
Spreads             Yield Calculations                  Invoice
11) G-Sprd    0.0   Street Convention      2.849091     Face                       1,000 M
12) I-Sprd   -7.1   Equiv  1  /Yr          2.869384     Principal           1,002,187.50
13) Basis    39.4   Mmkt (Act/ 360 )                    Accrued (59 Days)       4,609.38
14) Z-Sprd   -6.9   True Yield             2.848991     Total   (USD)       1,006,796.88
15) ASW      -6.8   Current Yield          2.869
16) OAS       0.0
17) TED      20.0
After Tax  (Inc 40.800 % CG 23.800 %)      1.686725
Issue Price = 100.132. Bond Purchased with Premiu...
```

© 2019 Bloomberg L.P. All rights reserved. Reproduced with permission.

In Exhibit 5, the quoted (flat) asked price for the bond is 100-07, which is equal to 100 and 7/32nds per 100 of par value for settlement on 13 July 2018. Most bond prices are stated in decimals, but US Treasuries are usually quoted in fractions. As a decimal, the flat price is 100.21875. The accrued interest uses the actual/actual day-count method. That settlement date is 59 days into a 184-day semiannual coupon payment period. The accrued interest is 0.4609375 per 100 of par value (= 59/184 × 0.02875/2 × 100). The full price of the bond is 100.679688. The yield-to-maturity of the bond is 2.849091%, stated on a street-convention semiannual bond basis.

The modified duration for the bond is shown in Exhibit 5 to be 8.482, which is the conventional *yield* duration statistic. Its *curve* duration, however, is 8.510, which is the price sensitivity with respect to changes in the US Treasury par curve. On Bloomberg, the effective duration is called the "OAS duration" because it is based on the option-pricing model that is also used to calculate the option-adjusted spread. The small difference arises because the government yield curve is not flat. When the par curve is shifted in the model, the government spot curve is also shifted, although not in the same "parallel" manner. Therefore, the change in the bond price is not the same as it would be if its own yield-to-maturity changed by the same amount as the change in the par curve. In general, the modified duration and effective duration on a traditional option-free bond are not identical. The difference narrows when the yield curve is flatter, the time-to-maturity is shorter, and the bond is priced closer to par value (so that the difference between the coupon rate and the yield-to-maturity is smaller). The modified duration and effective duration on an option-free bond are identical only in the rare circumstance of a flat yield curve.

5.1 Key Rate Duration

d define key rate duration and describe the use of key rate durations in measuring the sensitivity of bonds to changes in the shape of the benchmark yield curve

The effective duration for a sample callable bond was calculated previously as

$$\text{EffDur} = \frac{102.890738 - 99.050120}{2 \times 0.0025 \times 101.060489} = 7.6006$$

This duration measure indicates the bond's sensitivity to the benchmark yield curve if all yields change by the same amount. "Key rate" duration provides further insight into a bond's sensitivity to non-parallel benchmark yield curve changes. A **key rate duration** (or **partial duration**) is a measure of a bond's sensitivity to a change in the benchmark yield at a specific maturity. Key rate durations define a security's price sensitivity over a set of maturities along the yield curve, with the sum of key rate durations being identical to the effective duration:

$$\text{KeyRateDur}^k = -\frac{1}{PV} \times \frac{\Delta PV}{\Delta r^k} \tag{10}$$

$$\sum_{k=1}^{n} \text{KeyRateDur}^k = \text{EffDur}$$

where r^k represents the *k*th key rate. In contrast to effective duration, key rate durations help identify "shaping risk" for a bond—that is, a bond's sensitivity to changes in the shape of the benchmark yield curve (e.g., the yield curve becoming steeper or flatter).

The previous illustration of effective duration assumed a parallel shift of 25 bps at all maturities. However, the analyst may want to know how the price of the callable bond is expected to change if short-term benchmark rates (say, for a current two-year Treasury note with modified duration of 1.9) rise by 25 bps but longer-maturity benchmark rates remain unchanged. This scenario would represent a flattening of the yield curve, given that the yield curve is upward sloping. Using key rate durations, the expected price change would be approximately equal to minus the key rate duration for the short-maturity segment (–1.9) times the 0.0025 interest rate shift at that segment, or –0.475% based on the following formula:

$$\frac{\Delta PV}{PV} = -\text{KeyRateDur}^{k} \times \Delta r^{k}$$

Of course, for parallel shifts in the benchmark yield curve, key rate durations will indicate the same interest rate sensitivity as effective duration.

6 PROPERTIES OF BOND DURATION

e explain how a bond's maturity, coupon, and yield level affect its interest rate risk

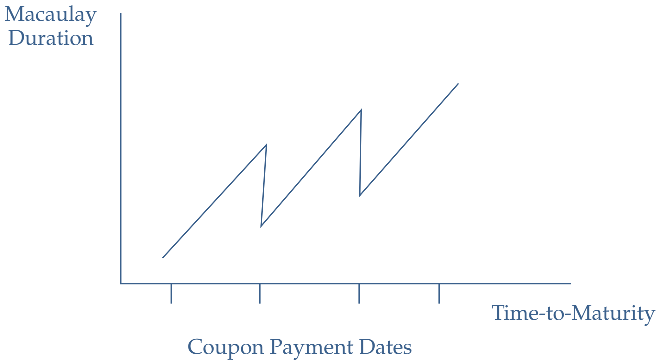
The Macaulay and modified yield duration statistics for a traditional fixed-rate bond are functions of the input variables: the coupon rate or payment per period, the yield-to-maturity per period, the number of periods to maturity (as of the beginning of the period), and the fraction of the period that has gone by. The properties of bond duration are obtained by changing one of these variables while holding the others constant. Because duration is the basic measure of interest rate risk on a fixed-rate bond, these properties are important to understand.

The closed-form formula for Macaulay duration, presented as Equation 4 and again here, is useful in demonstrating the characteristics of the bond duration statistic.

$$\text{MacDur} = \left\{\frac{1+r}{r} - \frac{1+r+\left[N \times (c-r)\right]}{c \times \left[(1+r)^{N} - 1\right] + r}\right\} - (t/T)$$

The same characteristics hold for modified duration. Consider first the fraction of the period that has gone by (t/T). Macaulay and modified durations depend on the day-count basis used to obtain the yield-to-maturity. The duration of a bond that uses the actual/actual method to count days is slightly different from that of an otherwise comparable bond that uses the 30/360 method. The key point is that for a constant yield-to-maturity (r), the expression in braces is unchanged as time passes during the period. Therefore, the Macaulay duration decreases smoothly as t goes from $t = 0$ to $t = T$, which creates a “saw-tooth” pattern. This pattern for a typical fixed-rate bond is illustrated in Exhibit 6.

Exhibit 6 Macaulay Duration between Coupon Payments with a Constant Yield-to-Maturity



As times passes during the coupon period (moving from right to left in the diagram), the Macaulay duration declines smoothly and then jumps upward after the coupon is paid.

The characteristics of bond duration related to changes in the coupon rate, the yield-to-maturity, and the time-to-maturity are illustrated in Exhibit 7.

Exhibit 7 Properties of the Macaulay Yield Duration

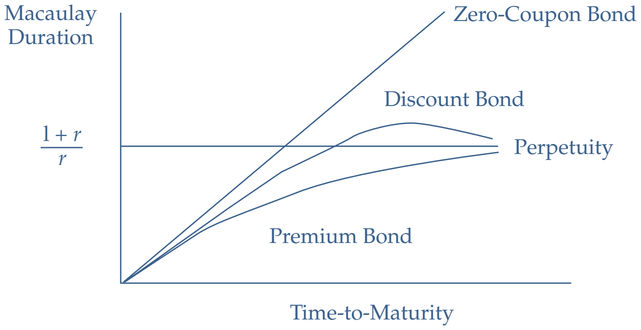


Exhibit 7 shows the graph for coupon payment dates when $t/T = 0$, thus not displaying the saw-tooth pattern between coupon payments. The relationship between the Macaulay duration and the time-to-maturity for a zero-coupon bond is the 45-degree line: MacDur = N when $c = 0$ (and $t/T = 0$). Therefore, the Macaulay duration of a zero-coupon bond is its time-to-maturity.

A **perpetuity** or perpetual bond, which also is called a consol, is a bond that does not mature. There is no principal to redeem. The investor receives a fixed coupon payment forever unless the bond is callable. Non-callable perpetuities are rare, but they have an interesting Macaulay duration: MacDur = $(1 + r)/r$ as N approaches infinity. In effect, the second expression within the braces approaches zero as the number of periods to maturity increases because N in the numerator is a coefficient but N in the denominator is an exponent and the denominator increases faster than the numerator as N grows larger.

Typical fixed-rate coupon bonds with a stated maturity date are portrayed in Exhibit 7 as the premium and discount bonds. The usual pattern is that longer times-to-maturity correspond to higher Macaulay duration statistics. This pattern always holds for bonds trading at par value or at a premium above par. In Equation 4, the second expression within the braces is a positive number for premium and par bonds.

The numerator is positive because the coupon rate (c) is greater than or equal to the yield-to-maturity (r), whereas the denominator is always positive. Therefore, the Macaulay duration is always less than $(1 + r)/r$, and it approaches that threshold from below as the time-to-maturity increases.

The curious result displayed in Exhibit 7 is in the pattern for discount bonds. Generally, the Macaulay duration increases for a longer time-to-maturity. But at some point when the time-to-maturity is high enough, the Macaulay duration exceeds $(1 + r)/r$, reaches a maximum, and then approaches the threshold from above. In Equation 4, such a pattern develops when the number of periods (N) is large and the coupon rate (c) is below the yield-to-maturity (r). Then the numerator of the second expression within the braces can become negative. The implication is that on long-term discount bonds, the interest rate risk can actually be less than on a shorter-term bond, which explains why the word "generally" is needed in describing the maturity effect for the relationship between bond prices and yields-to-maturity. Generally, for the same coupon rate, a longer-term bond has a greater percentage price change than a shorter-term bond when their yields-to-maturity change by the same amount. The exception is when the longer-term bond has a lower duration statistic.

Coupon rates and yields-to-maturity are both inversely related to the Macaulay duration. In Exhibit 7, for the same time-to-maturity and yield-to-maturity, the Macaulay duration is higher for a zero-coupon bond than for a low-coupon bond trading at a discount. Also, the low-coupon bond trading at a discount has a higher duration than a high-coupon bond trading at a premium. Therefore, all else being equal, a lower-coupon bond has a higher duration and more interest rate risk than a higher-coupon bond. The same pattern holds for the yield-to-maturity. A higher yield-to-maturity reduces the weighted average of the time to receipt of cash flow. More weight is on the cash flows received in the near term, and less weight is on the cash flows received in the more-distant future periods if those cash flows are discounted at a higher rate.

In summary, the Macaulay and modified duration statistics for a fixed-rate bond depend primarily on the coupon rate, yield-to-maturity, and time-to-maturity. A higher coupon rate or a higher yield-to-maturity reduces the duration measures. A longer time-to-maturity *usually* leads to a higher duration. It *always* does so for a bond priced at a premium or at par value. But if the bond is priced at a discount, a longer time-to-maturity *might* lead to a lower duration. This situation only occurs if the coupon rate is low (but not zero) relative to the yield and the time-to-maturity is long.

EXAMPLE 10

A hedge fund specializes in investments in emerging market sovereign debt. The fund manager believes that the implied default probabilities are too high, which means that the bonds are viewed as "cheap" and the credit spreads are too high. The hedge fund plans to take a position on one of these available bonds.

Bond	Time-to-Maturity	Coupon Rate	Price	Yield-to-Maturity
(A)	10 years	10%	58.075279	20%
(B)	20 years	10%	51.304203	20%
(C)	30 years	10%	50.210636	20%

The coupon payments are annual. The yields-to-maturity are effective annual rates. The prices are per 100 of par value.

1 Compute the approximate modified duration of each of the three bonds using a 1 bp change in the yield-to-maturity and keeping precision to six decimals (because approximate duration statistics are very sensitive to rounding).

2 Which of the three bonds is expected to have the highest percentage price increase if the yield-to-maturity on each decreases by the same amount—for instance, by 10 bps from 20% to 19.90%?

Solution to 1:

Bond A:

$PV_0 = 58.075279$

$PV_+ = 58.047598$

$$\frac{10}{(1.2001)^1} + \frac{10}{(1.2001)^2} + \cdots + \frac{110}{(1.2001)^{10}} = 58.047598$$

$PV_- = 58.102981$

$$\frac{10}{(1.1999)^1} + \frac{10}{(1.1999)^2} + \cdots + \frac{110}{(1.1999)^{10}} = 58.102981$$

The approximate modified duration of Bond A is 4.768.

$$\text{ApproxModDur} = \frac{58.102981 - 58.047598}{2 \times 0.0001 \times 58.075279} = 4.768$$

Bond B:

$PV_0 = 51.304203$

$PV_+ = 51.277694$

$$\frac{10}{(1.2001)^1} + \frac{10}{(1.2001)^2} + \cdots + \frac{110}{(1.2001)^{20}} = 51.277694$$

$PV_- = 51.330737$

$$\frac{10}{(1.1999)^1} + \frac{10}{(1.1999)^2} + \cdots + \frac{110}{(1.1999)^{20}} = 51.330737$$

The approximate modified duration of Bond B is 5.169.

$$\text{ApproxModDur} = \frac{51.330737 - 51.277694}{2 \times 0.0001 \times 51.304203} = 5.169$$

Bond C:

$PV_0 = 50.210636$

$PV_+ = 50.185228$

$$\frac{10}{(1.2001)^1} + \frac{10}{(1.2001)^2} + \cdots + \frac{110}{(1.2001)^{30}} = 50.185228$$

$PV_- = 50.236070$

$$\frac{10}{(1.1999)^1} + \frac{10}{(1.1999)^2} + \cdots + \frac{110}{(1.1999)^{30}} = 50.236070$$

The approximate modified duration of Bond C is 5.063.

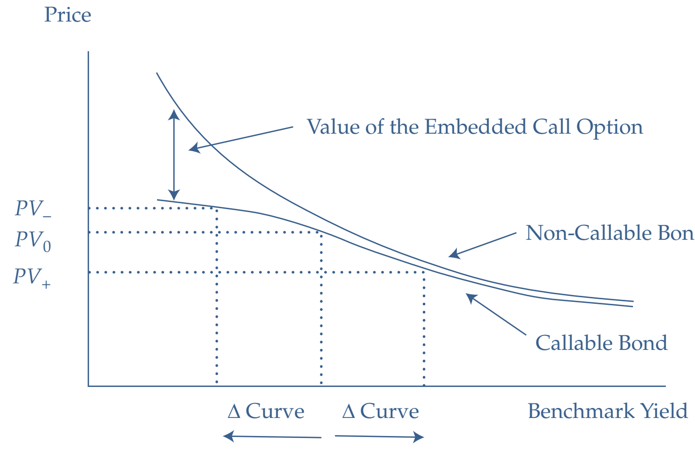
$$\text{ApproxModDur} = \frac{50.236070 - 50.185228}{2 \times 0.0001 \times 50.210636} = 5.063$$

Solution to 2:

Despite the significant differences in times-to-maturity (10, 20, and 30 years), the approximate modified durations on the three bonds are fairly similar (4.768, 5.169, and 5.063). Because the yields-to-maturity are so high, the additional time to receipt of interest and principal payments on the 20- and 30-year bonds have low weight. Nevertheless, Bond B, with 20 years to maturity, has the highest modified duration. If the yield-to-maturity on each is decreased by the same amount—for instance, by 10 bps, from 20% to 19.90%—Bond B would be expected to have the highest percentage price increase because it has the highest modified duration. This example illustrates the relationship between the Macaulay duration and the time-to-maturity on discount bonds in Exhibit 7. The 20-year bond has a higher duration than the 30-year bond.

Callable bonds require the use of effective duration because Macaulay and modified yield duration statistics are not relevant. The yield-to-maturity for callable bonds is not well-defined because future cash flows are uncertain. Exhibit 8 illustrates the impact of the change in the benchmark yield curve (ΔCurve) on the price of a callable bond price compared with that on a comparable non-callable bond. The two bonds have the same credit risk, coupon rate, payment frequency, and time-to-maturity. The vertical axis is the bond price. The horizontal axis is a benchmark yield—for instance, a point on the par curve for government bonds.

Exhibit 8 Interest Rate Risk Characteristics of a Callable Bond

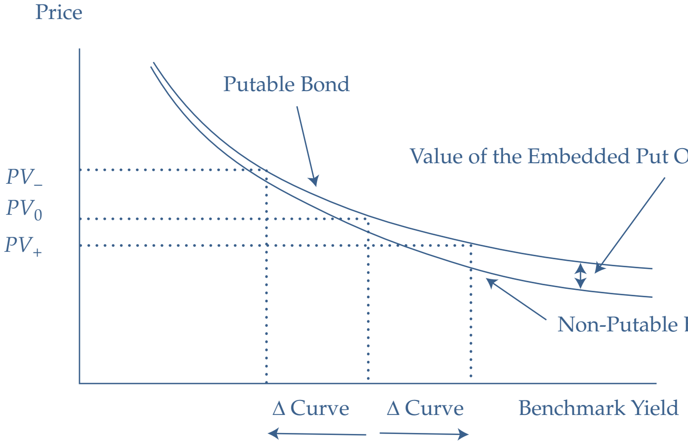


As shown in Exhibit 8, the price of the non-callable bond is always greater than that of the callable bond with otherwise identical features. The difference is the value of the embedded call option. Recall that the call option is an option to the issuer and not the holder of the bond. When interest rates are high compared with the coupon rate, the value of the call option is low. When rates are low, the value of the call option is much greater because the issuer is more likely to exercise the option to refinance the debt at a lower cost of funds. The investor bears the "call risk" because if the bond is called, the investor must reinvest the proceeds at a lower interest rate.

Exhibit 8 shows the inputs for calculating the effective duration of the callable bond. The entire benchmark curve is raised and lowered by the same amount, ΔCurve. The key point is that when benchmark yields are high, the effective durations of the callable and non-callable bonds are very similar. Although the exhibit does not illustrate it, the slopes of the lines tangent to the price–yield curve are about the same in such a situation. But when interest rates are low, the effective duration of the callable bond is lower than that of the otherwise comparable non-callable bond. That is because the callable bond price does not increase as much when benchmark yields fall. The slope of the line tangent to the price–yield curve would be flatter. The presence of the call option limits price appreciation. Therefore, an embedded call option reduces the effective duration of the bond, especially when interest rates are falling and the bond is more likely to be called. The lower effective duration can also be interpreted as a shorter expected life—the weighted average of time to receipt of cash flow is reduced.

Exhibit 9 considers another embedded option—a put option.

Exhibit 9 Interest Rate Risk Characteristics of a Putable Bond



A putable bond allows the investor to sell the bond back to the issuer prior to maturity, usually at par value, which protects the investor from higher benchmark yields or credit spreads that otherwise would drive the bond to a discounted price. Therefore, the price of a putable bond is always higher than that of an otherwise comparable non-putable bond. The price difference is the value of the embedded put option.

An embedded put option reduces the effective duration of the bond, especially when rates are rising. If interest rates are low compared with the coupon rate, the value of the put option is low and the impact of a change in the benchmark yield on the bond's price is very similar to the impact on the price of a non-putable bond. But when benchmark interest rates rise, the put option becomes more valuable to the investor. The ability to sell the bond at par value limits the price depreciation as rates rise. In summary, the presence of an embedded option reduces the sensitivity of the bond price to changes in the benchmark yield curve, assuming no change in credit risk.

7

DURATION OF A BOND PORTFOLIO

f calculate the duration of a portfolio and explain the limitations of portfolio duration

Similar to equities, bonds are typically held in a portfolio. There are two ways to calculate the duration of a bond portfolio: (1) the weighted average of time to receipt of the *aggregate* cash flows, and (2) the weighted average of the individual bond durations that comprise the portfolio. The first method is the theoretically correct approach, but it is difficult to use in practice. The second method is commonly used by fixed-income portfolio managers, but it has its own limitations. The differences in these two methods to compute portfolio duration can be examined with a numerical example.

Suppose an investor holds the following portfolio of two *zero-coupon* bonds:

Bond	Maturity	Price	Yield	Macaulay Duration	Modified Duration	Par Value	Market Value	Weight
(X)	1 year	98.00	2.0408%	1	0.980	10,000,000	9,800,000	0.50
(Y)	30 years	9.80	8.0503%	30	27.765	100,000,000	9,800,000	0.50

The prices are per 100 of par value. The yields-to-maturity are effective annual rates. The total market value for the portfolio is 19,600,000. The portfolio is evenly weighted in terms of market value between the two bonds.

The first approach views the portfolio as a series of aggregated cash flows. Its **cash flow yield** is 7.8611%. A cash flow yield is the internal rate of return on a series of cash flows, usually used on a complex security such as a mortgage-backed bond (using projected cash flows based on a model of prepayments as a result of refinancing) or a portfolio of fixed-rate bonds. It is the solution for r in the following equation.

$$19{,}600{,}000 = \frac{10{,}000{,}000}{(1+r)^1} + \frac{0}{(1+r)^2} + \cdots + \frac{0}{(1+r)^{29}} + \frac{100{,}000{,}000}{(1+r)^{30}}, \quad r = 0.078611$$

The Macaulay duration of the portfolio in this approach is the weighted average of the times to receipt of aggregated cash flow. The cash flow yield is used to obtain the weights. This calculation is similar to Equation 1, and the portfolio duration is 16.2825.

$$\text{MacDur} = \left[\frac{\dfrac{1 \times 10{,}000{,}000}{(1.078611)^1} + \dfrac{30 \times 100{,}000{,}000}{(1.078611)^{30}}}{\dfrac{10{,}000{,}000}{(1.078611)^1} + \dfrac{100{,}000{,}000}{(1.078611)^{30}}}\right] = 16.2825$$

There are just two future cash flows in the portfolio—the redemption of principal on the two zero-coupon bonds. In more complex portfolios, a series of coupon and principal payments may occur on some dates, with an aggregated cash flow composed of coupon interest on some bonds and principal on those that mature.

The modified duration of the portfolio is the Macaulay duration divided by one plus the cash flow yield per period (here, the periodicity is 1).

$$\text{ModDur} = \frac{16.2825}{1.078611} = 15.0958$$

The modified duration for the portfolio is 15.0958. That statistic indicates the percentage change in the market value given a change in the cash flow yield. If the cash flow yield increases or decreases by 100 bps, the market value of the portfolio is expected to decrease or increase by about 15.0958%.

Although this approach is theoretically correct, it is difficult to use in practice. First, the cash flow yield is not commonly calculated for bond portfolios. Second, the amount and timing of future coupon and principal payments are uncertain if the portfolio contains callable or putable bonds or floating-rate notes. Third, interest rate risk is usually expressed as a change in benchmark interest rates, not as a change in the cash flow yield. Fourth, the change in the cash flow yield is not necessarily the same amount as the change in the yields-to-maturity on the individual bonds. For instance, if the yields-to-maturity on the two zero-coupon bonds in this portfolio both increase or decrease by 10 bps, the cash flow yield increases or decreases by only 9.52 bps.

In practice, the second approach to portfolio duration is commonly used. The Macaulay and modified durations for the portfolio are calculated as the weighted average of the statistics for the individual bonds. The shares of overall portfolio market value are the weights. This weighted average approximates the theoretically correct portfolio duration, which is obtained using the first approach. This approximation becomes more accurate when the differences in the yields-to-maturity on the bonds in the portfolio are smaller. When the yield curve is flat, the two approaches produce the same portfolio duration.

Given the equal "50/50" weights in this simple numerical example, this version of portfolio duration is easily computed.

$$\text{Average Macaulay duration} = (1 \times 0.50) + (30 \times 0.50) = 15.50$$

$$\text{Average modified duration} = (0.980 \times 0.50) + (27.765 \times 0.50) = 14.3725$$

Note that 0.980 = 1/1.020404 and 27.765 = 30/1.080503. An advantage of the second approach is that callable bonds, putable bonds, and floating-rate notes can be included in the weighted average using the effective durations for these securities.

The main advantage to the second approach is that it is easily used as a measure of interest rate risk. For instance, if the yields-to-maturity on the bonds in the portfolio increase by 100 bps, the estimated drop in the portfolio value is 14.3725%. However, this advantage also indicates a limitation: This measure of portfolio duration implicitly assumes a **parallel shift** in the yield curve. A parallel yield curve shift implies that all rates change by the same amount in the same direction. In reality, interest rate changes frequently result in a steeper or flatter yield curve. Yield volatility is discussed later in this reading.

EXAMPLE 11

An investment fund owns the following portfolio of three fixed-rate government bonds:

	Bond A	Bond B	Bond C
Par value	EUR25,000,000	EUR25,000,000	EUR50,000,000
Coupon rate	9%	11%	8%
Time-to-maturity	6 years	8 years	12 years
Yield-to-maturity	9.10%	9.38%	9.62%
Market value	EUR24,886,343	EUR27,243,887	EUR44,306,787
Macaulay duration	4.761	5.633	7.652

The total market value of the portfolio is EUR96,437,017. Each bond is on a coupon date so that there is no accrued interest. The market values are the full prices given the par value. Coupons are paid semiannually. The yields-to-maturity are stated on a semiannual bond basis, meaning an annual rate for a periodicity of 2. The Macaulay durations are annualized.

1 Calculate the average (annual) modified duration for the portfolio using the shares of market value as the weights.

2 Estimate the percentage loss in the portfolio's market value if the (annual) yield-to-maturity on each bond goes up by 20 bps.

Solution to 1:

The average (annual) modified duration for the portfolio is 6.0495.

$$\left(\frac{4.761}{1+\frac{0.0910}{2}} \times \frac{24{,}886{,}343}{96{,}437{,}017}\right) + \left(\frac{5.633}{1+\frac{0.0938}{2}} \times \frac{27{,}243{,}887}{96{,}437{,}017}\right) +$$

$$\left(\frac{7.652}{1+\frac{0.0962}{2}} \times \frac{44{,}306{,}787}{96{,}437{,}017}\right) = 6.0495$$

Note that the annual modified duration for each bond is the annual Macaulay duration, which is given, divided by one plus the yield-to-maturity per semiannual period.

Solution to 2:

The estimated decline in market value if each yield rises by 20 bps is 1.21%: $-6.0495 \times 0.0020 = -0.0121$.

8 MONEY DURATION AND THE PRICE VALUE OF A BASIS POINT

g calculate and interpret the money duration of a bond and price value of a basis point (PVBP)

Modified duration is a measure of the *percentage price change* of a bond given a change in its yield-to-maturity. A related statistic is **money duration**. The money duration of a bond is a measure of the *price change* in units of the currency in which the bond is denominated. The money duration can be stated per 100 of par value or in terms of the actual position size of the bond in the portfolio. In the United States, money duration is commonly called "dollar duration."

Money duration (MoneyDur) is calculated as the annual modified duration times the full price (PV^{Full}) of the bond, including accrued interest.

$$\text{MoneyDur} = \text{AnnModDur} \times PV^{Full} \tag{11}$$

The estimated change in the bond price in currency units is calculated using Equation 12, which is very similar to Equation 6. The difference is that for a given change in the annual yield-to-maturity (ΔYield), modified duration estimates the percentage price change and money duration estimates the change in currency units.

$$\Delta PV^{Full} \approx -\text{MoneyDur} \times \Delta\text{Yield} \tag{12}$$

For a theoretical example of money duration, consider the 6% semiannual coupon payment bond that matures on 14 February 2027 and is priced to yield 6.00% for settlement on 11 April 2019. The full price of the bond is 100.940423 per 100 of par value, and the annual modified duration is 6.1268. Suppose that a Nairobi based life insurance company has a position in the bond for a par value of KES100,000,000. The market value of the investment is KES 100,940,423. The money duration of this bond is KES 618,441,784 (= 6.1268 × KES 100,940,423). Therefore, if the yield-to-maturity rises by 100 bps—from 6.00% to 7.00%—the expected loss is approximately –KES 6,184,418 (= –KES 618,441,784 × 0.0100). On a percentage basis, that expected loss is approximately 6.1268%. The "convexity adjustment" introduced in the next section makes these estimates more accurate.

Another version of money duration is the value of one basis point in price terms. The **price value of a basis point** (or PVBP) is an estimate of the change in the full bond price given a 1 bp change in the yield-to-maturity. The PVBP can be calculated using a formula similar to that for the approximate modified duration. Equation 13 is the formula for the PVBP.

$$\text{PVBP} = \frac{(PV_-) - (PV_+)}{2} \tag{13}$$

PV_- and PV_+ are the full prices calculated by decreasing and increasing the yield-to-maturity by 1 bp. The PVBP is also called the "PV01," standing for the "price value of an 01" or "present value of an 01," where "01" means 1 bp. In the United States, it is commonly called the "DV01," or the "dollar value of a 01." The PVBP is particularly useful for bonds where future cash flows are uncertain, such as callable bonds. A related statistic called a "basis point value" (BPV) is simply the money duration times 0.0001 (1 bp).

For a numerical example of the PVBP calculation, consider the 2.875% semiannual coupon payment US Treasury note that matures on 15 May 2028. In Exhibit 5, the PVBP for the Treasury note is shown to be 0.08540. Its yield-to-maturity is 2.849091%, and the settlement date is 59 days into a 184-day period. To confirm this result, calculate the new prices by increasing and decreasing the yield-to-maturity. First, increase the yield by 1 bp (0.01%), from 2.849091% to 2.859091%, to solve for a PV_+ of 100.594327.

$$PV_+ = \left[\frac{1.4375}{\left(1+\frac{0.02859091}{2}\right)^1} + \cdots + \frac{101.4375}{\left(1+\frac{0.02859091}{2}\right)^{20}}\right] \times \left(1+\frac{0.02859091}{2}\right)^{59/184}$$
$$= 101.594327$$

Then, decrease the yield-to-maturity by 1 bp, from 2.849091% to 2.839091%, to solve for a PV_- of 100.765123.

$$PV_- = \left[\frac{1.4375}{\left(1+\frac{0.02839091}{2}\right)^1} + \cdots + \frac{101.4375}{\left(1+\frac{0.02839091}{2}\right)^{20}}\right] \times \left(1+\frac{0.02839091}{2}\right)^{59/184}$$
$$= 100.765123$$

The PVBP is obtained by substituting these results into Equation 13.

$$\text{PVBP} = \frac{100.765123 - 100.594327}{2} = 0.08540$$

Another money duration statistic reported on the Bloomberg YAS page is "risk." It is shown to be 8.540. Bloomberg's risk statistic is simply the PVBP (or PV01) times 100.

EXAMPLE 12

A life insurance company holds a USD10 million (par value) position in a 5.95% Dominican Republic bond that matures on 25 January 2027. The bond is priced (flat) at 101.996 per 100 of par value to yield 5.6511% on a street-convention semiannual bond basis for settlement on 24 July 2018. The total market value of the position, including accrued interest, is USD10,495,447, or 101.95447 per 100 of par value. The bond's (annual) Macaulay duration is 6.622.

1 Calculate the money duration per 100 in par value for the sovereign bond.
2 Using the money duration, estimate the loss on the position for each 1 bp increase in the yield-to-maturity for that settlement date.

Solution to 1:

The money duration is the annual modified duration times the full price of the bond per 100 of par value.

$$\left(\frac{6.622}{1+\frac{0.056511}{2}}\right) \times \text{USD}104.95447 = \text{USD}675.91$$

Solution to 2:

For each 1 bp increase in the yield-to-maturity, the loss is estimated to be USD 0.067591 per 100 of par value: USD 675.91 × 0.0001 = USD 0.067591.

Given a position size of USD 10 million in par value, the estimated loss per basis-point increase in the yield is USD 6,759.10. The money duration is per 100 of par value, so the position size of USD10 million is divided by USD 100.

$$\text{USD}0.067591 \times \frac{\text{USD}10{,}000{,}000}{\text{USD}100} = \text{USD}6{,}759.10$$

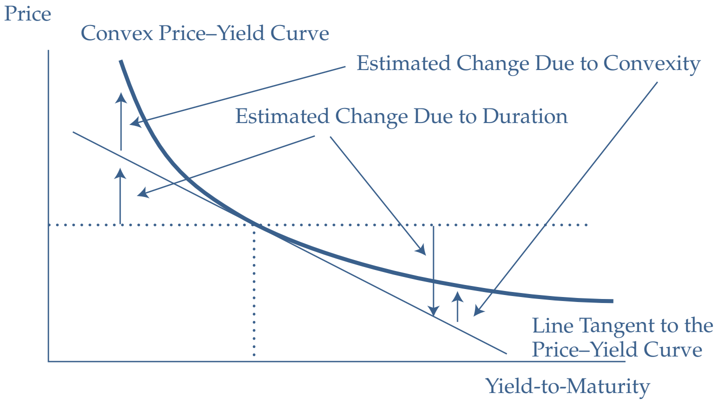
9 BOND CONVEXITY

h calculate and interpret approximate convexity and compare approximate and effective convexity

i calculate the percentage price change of a bond for a specified change in yield, given the bond's approximate duration and convexity

Modified duration measures the primary effect on a bond's percentage price change given a change in the yield-to-maturity. A secondary effect is measured by the convexity statistic, which is illustrated in Exhibit 10 for a traditional (option-free) fixed-rate bond.

Exhibit 10 Convexity of a Traditional (Option-Free) Fixed-Rate Bond


Price
Convex Price–Yield Curve
Estimated Change Due to Convexity
Estimated Change Due to Duration
Line Tangent to the Price–Yield Curve
Yield-to-Maturity

The true relationship between the bond price and the yield-to-maturity is the curved (convex) line shown in Exhibit 10. This curved line shows the actual bond price given its market discount rate. Duration (in particular, money duration) estimates the change in the bond price along the straight line that is tangent to the curved line. For small yield-to-maturity changes, there is little difference between the lines. But for larger changes, the difference becomes significant.

The convexity statistic for the bond is used to improve the estimate of the percentage price change provided by modified duration alone. Equation 14 is the convexity-adjusted estimate of the percentage change in the bond's full price.

$$\%\Delta PV^{Full} \approx$$

$$(-\text{AnnModDur} \times \Delta\text{Yield}) + \left[\frac{1}{2} \times \text{AnnConvexity} \times (\Delta\text{Yield})^2\right] \quad (14)$$

The first bracketed expression, the "first-order" effect, is the same as Equation 6. The (annual) modified duration, AnnModDur, is multiplied by the change in the (annual) yield-to-maturity, ΔYield. The second bracketed expression, the "second-order" effect, is the **convexity adjustment**. The convexity adjustment is the annual convexity statistic, AnnConvexity, times one-half, multiplied by the change in the yield-to-maturity *squared*. This additional term is a positive amount on a traditional (option-free) fixed-rate bond for either an increase or decrease in the yield. In Exhibit 10, this amount adds to the linear estimate provided by the duration alone, which brings the adjusted estimate very close to the actual price on the curved line. But it still is an estimate, so the ≈ sign is used.

Similar to the Macaulay and modified durations, the annual convexity statistic can be calculated in several ways. It can be calculated using tables, such as Exhibits 2 and 3. It also is possible to derive a closed-form equation for the convexity of a fixed-rate bond on and between coupon payment dates using calculus and algebra (see D. Smith, 2014). But like modified duration, convexity can be approximated with accuracy. Equation 15 is the formula for the approximate convexity statistic, ApproxCon.

$$\text{ApproxCon} = \frac{(PV_-) + (PV_+) - [2 \times (PV_0)]}{(\Delta\text{Yield})^2 \times (PV_0)} \quad (15)$$

This equation uses the same inputs as Equation 7 for ApproxModDur. The new price when the yield-to-maturity is increased is PV_+. The new price when the yield is decreased by the same amount is PV_-. The original price is PV_0. These are the full prices, including accrued interest, for the bond.

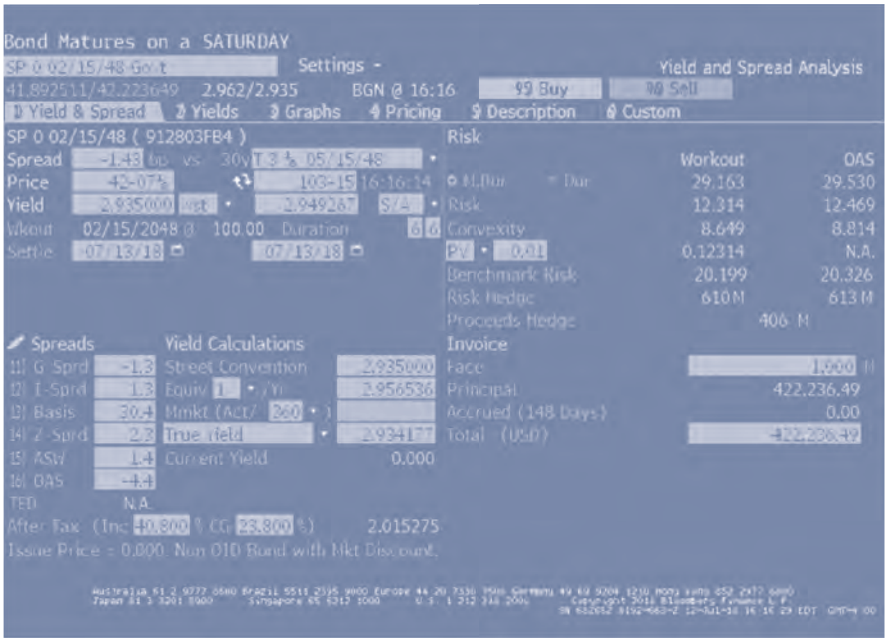
The accuracy of this approximation can be demonstrated with the special case of a zero-coupon bond. The absence of coupon payments simplifies the interest rate risk measures. The Macaulay duration of a zero-coupon bond is $N - t/T$ in terms of periods to maturity. The exact convexity statistic of a zero-coupon bond, also in terms of periods, is calculated with Equation 16.

$$\text{Convexity (of a zero-coupon bond)} = \frac{\left[N - (t/T)\right] \times \left[N + 1 - (t/T)\right]}{(1 + r)^2} \qquad \textbf{(16)}$$

N is the number of periods to maturity as of the beginning of the current period, t/T is the fraction of the period that has gone by, and r is the yield-to-maturity per period.

For an example of this calculation, consider a long-term, zero-coupon US Treasury bond. The bond's Bloomberg YAS page is shown in Exhibit 11.

Exhibit 11 Bloomberg YAS Page for the Zero-Coupon US Treasury Bond



Bond Matures on a SATURDAY
SP 0 02/15/48 Govt
Settings -
Yield and Spread Analysis
41.892511/42.223649 2.962/2.935 BGN @ 16:16
99 Buy
98 Sell
1) Yield & Spread 2) Yields 3) Graphs 4) Pricing 5) Description 6) Custom

SP 0 02/15/48 (912803FB4)
Spread -1.43 bp vs 30y T 3 ⅛ 05/15/48
Price 42-07⅛ 103-15 16:16:14
Yield 2.935000 Wst 2.949267 S/A
Wkout 02/15/2048 @ 100.00 Duration 6 6
Settle 07/13/18 07/13/18

Risk
	Workout	OAS
M.Dur Dur	29.163	29.530
Risk	12.314	12.469
Convexity	8.649	8.814
PV 0.01	0.12314	N.A.
Benchmark Risk	20.199	20.326
Risk Hedge	610 M	613 M
Proceeds Hedge	406 M	

Spreads
11) G Sprd	-1.3
12) I-Sprd	1.3
13) Basis	30.4
14) Z-Sprd	2.3
15) ASW	1.4
16) OAS	-4.4
TED	N.A.

Yield Calculations
Street Convention	2.935000
Equiv 1 /Yr	2.956536
Mmkt (Act/ 360)	
True Yield	2.934177
Current Yield	0.000

After Tax (Inc 40.800 % CG 23.800 %) 2.015275
Issue Price = 0.000. Non OID Bond with Mkt Discount.

Invoice
Face	1,000 M
Principal	422,236.49
Accrued (148 Days)	0.00
Total (USD)	422,236.49

© 2019 Bloomberg L.P. All rights reserved. Reproduced with permission.

The bond matures on 15 February 2048 and its asked price was 42.223649 per 100 of par value for settlement on 13 July 2018. Its yield-to-maturity was 2.935% stated on a street-convention semiannual bond basis. Even though it is a zero-coupon bond, its yield-to-maturity is based on the actual/actual day-count convention. That settlement date was 148 days into a 181-day period. The annual modified duration was 29.163.

For this bond, $N = 60$, $t/T = 148/181$, and $r = 0.02935/2$. Entering these variables into Equation 16 produces a convexity of 3,459.45 in terms of semiannual periods.

$$\frac{\left[60 - (148/181)\right] \times \left[60 + 1 - (148/181)\right]}{\left(1 + \dfrac{0.02935}{2}\right)^2} = 3{,}459.45$$

As with the other statistics, convexity is annualized in practice and for use in the convexity adjustment in Equation 14. It is divided by the periodicity *squared*. The yield-to-maturity on this zero-coupon bond is stated on a semiannual bond basis, meaning a periodicity of 2. Therefore, the annualized convexity statistic is 864.9.

$$\frac{3{,}459.45}{4} = 864.9$$

For example, suppose that the yield-to-maturity is expected to fall by 10 bps, from 2.935% to 2.835%. Given the (annual) modified duration of 29.163 and (annual) convexity of 864.9, the expected percentage price gain is 2.9595%.

$$\begin{aligned}\%\Delta PV^{Full} &\approx \left[-29.163 \times -0.0010\right] + \left[\frac{1}{2} \times 864.9 \times \left(-0.0010\right)^2\right] \\ &= 0.029163 + 0.000432 \\ &= 0.029595\end{aligned}$$

Modified duration alone (under)estimates the gain to be 2.9163%. The convexity adjustment adds 4.32 bps.

The long-term, zero-coupon bond of Exhibit 11 demonstrates the difference between *yield* duration and convexity and *curve* duration and convexity, even on an option-free bond. Its modified duration is 29.163, whereas its effective duration is 29.530. Its yield convexity is reported on the Bloomberg page to be 8.649, and its effective convexity is 8.814. (Note that although Bloomberg scales the convexity statistics by dividing by 100, either raw or scaled convexity figures are acceptable in practice.) In general, the differences are heightened when the benchmark yield curve is not flat, when the bond has a long time-to-maturity, and the bond is priced at a significant discount or premium.

To obtain the ApproxCon for this long-term, zero-coupon bond, calculate PV_0, PV_+, and PV_- for yields-to-maturity of 2.935%, 2.945%, and 2.925%, respectively. For this exercise, ΔYield = 0.0001.

$$PV_0 = \frac{100}{\left(1 + \frac{0.02935}{2}\right)^{60}} \times \left(1 + \frac{0.02935}{2}\right)^{148/181} = 42.223649$$

$$PV_+ = \frac{100}{\left(1 + \frac{0.02945}{2}\right)^{60}} \times \left(1 + \frac{0.02945}{2}\right)^{148/181} = 42.100694$$

$$PV_- = \frac{100}{\left(1 + \frac{0.02925}{2}\right)^{60}} \times \left(1 + \frac{0.02925}{2}\right)^{148/181} = 42.346969$$

Using these results, first calculate ApproxModDur using Equation 7 to confirm that these inputs are correct. In Exhibit 11, modified duration is stated to be 29.163.

$$\text{ApproxModDur} = \frac{42.346969 - 42.100694}{2 \times 0.0001 \times 42.223649} = 29.163$$

Using Equation 15, ApproxCon is 864.9.

$$\text{ApproxCon} = \frac{42.346969 + 42.100694 - \left(2 \times 42.223649\right)}{\left(0.0001\right)^2 \times 42.223649} = 864.9$$

This result, 864.9, is an approximation for *annualized* convexity. The number of periods in the year is included in the price calculations. This approximation in this example is the same as the "exact" result using the closed-form equation for the special case of the zero-coupon bond. Any small difference is not likely to be meaningful for practical applications.

Because this is an individual zero-coupon bond, it is easy to calculate the new price if the yield-to-maturity does go down by 50 bps, to 2.435%.

$$\frac{100}{\left(1+\dfrac{0.02435}{2}\right)^{60}} \times \left(1+\frac{0.02435}{2}\right)^{148/181} = 48.860850$$

Therefore, the actual percentage price increase is 15.7192%.

$$\frac{48.860850 - 42.223649}{42.223649} = 0.157192$$

The convexity-adjusted estimate is 15.6626%.

$$\begin{aligned}\%\Delta PV^{Full} &\approx (-29.163 \times -0.0050) + \left[\frac{1}{2} \times 864.9 \times (-0.0050)^2\right] \\ &= 0.145815 + 0.010811 \\ &= 0.156626\end{aligned}$$

EXAMPLE 13

A Dutch bank holds a large position in a zero-coupon Federal Republic of Germany government bond maturing on 11 April 2025. The yield-to-maturity is –0.72% for settlement on 11 May 2020, stated as an effective annual rate on an Actual/Actual basis. That settlement date is 30 days into the 365-day year using this day count method.

1. Calculate the full price of the bond per 100 of par value.
2. Calculate the approximate modified duration and approximate convexity using a 1 bp increase and decrease in the yield-to-maturity.
3. Calculate the estimated convexity-adjusted percentage price change resulting from a 100 bp increase in the yield-to-maturity.
4. Compare the estimated percentage price change with the actual change, assuming the yield-to-maturity jumps 100 bps to 0.28% on that settlement date.

Solution to 1:

There are five years from the beginning of the current period on 11 April 2020 to maturity on 11 April 2025.

The full price of the bond is 103.617526 per 100 of par value. Note that

$$1 + r = 1 + (-0.0072) = 0.9928$$

$$PV_0 = \left[\frac{100}{(0.9928)^5}\right] \times (0.9928)^{30/365} = 103.617526$$

Solution to 2:

$PV_+ = 103.566215$, and $PV_- = 103.668868$.

$$PV_+ = \left[\frac{100}{(0.9929)^5}\right] \times (0.9929)^{30/365} = 103.566215$$

$$PV_- = \left[\frac{100}{(0.9927)^5}\right] \times (0.9927)^{30/365} = 103.668868$$

The approximate modified duration is 4.9535.

$$\text{ApproxModDur} = \frac{103.668868 - 103.566215}{2 \times 0.0001 \times 103.617526} = 4.9535$$

The approximate convexity is 29.918.

$$\text{ApproxCon} = \frac{103.668868 + 103.566215 - (2 \times 103.617526)}{(0.0001)^2 \times 103.617526} = 29.918$$

Solution to 3:

The convexity-adjusted percentage price drop resulting from a 100 bp increase in the yield-to-maturity is estimated to be 4.80391%. Modified duration alone estimates the percentage drop to be 4.9535%. The convexity adjustment adds 14.96 bps.

$$\begin{aligned}\%\Delta PV^{Full} &\approx -(4.9535 \times 0.0100) + \left[\frac{1}{2} \times 29.918 \times (-0.0100)^2\right]\\ &= -0.049535 + 0.001496\\ &= -0.0480391\end{aligned}$$

Solution to 4:

The new full price if the yield-to-maturity goes from –0.72% to 0.28% on that settlement date is 98.634349.

$$PF^{Full} = \left[\frac{100}{(1.0028)^5}\right] \times (1.0028)^{30/365} = 98.634349$$

$$\%\Delta PV^{Full} = \frac{98.634349 - 103.617526}{103.617526} = -0.04809203$$

The actual percentage change in the bond price is –4.809203%. The convexity-adjusted estimate is –4.80391%, whereas the estimated change using modified duration alone is –4.9535%.

The money duration of a bond indicates the first-order effect on the full price of a bond in units of currency given a change in the yield-to-maturity. The **money convexity** statistic (MoneyCon) is the second-order effect. The money convexity of the bond is the annual convexity multiplied by the full price, such that

$$\Delta PV^{Full} \approx -(\text{MoneyDur} \times \Delta\text{Yield}) + \left[\frac{1}{2} \times \text{MoneyCon} \times (\Delta\text{Yield})^2\right] \quad (17)$$

For a money convexity example, consider again the Nairobi-based life insurance company that has a KES100 million position in the 6.00% bond that matures on 14 February 2027. Previously, using the money duration alone, the estimated loss was KES6,184,418 if the yield-to-maturity increased by 100 bps. The money duration for the position is KES618,441,784. That estimation is improved by including the convexity adjustment. Given the approximate modified duration of 6.1268 for a 5 bp change in the yield-to-maturity (ΔYield = 0.0005) and given that PV_0 = 100.940423, PV_+ = 100.631781, and PV_- = 101.250227, we use Equation 15 to calculate the approximate convexity:

$$\text{ApproxCon} = \frac{101.250227 + 100.631781 - (2 \times 100.940423)}{(0.0005)^2 \times 100.940423} = 46.047$$

The money convexity is 46.047 times the market value of the position, KES100,940,423. The convexity-adjusted loss given a 100 bp jump in the yield-to-maturity is KES5,952,018:

$$-\left[(6.1268 \times \text{KES}100{,}940{,}423) \times 0.0100\right] +$$

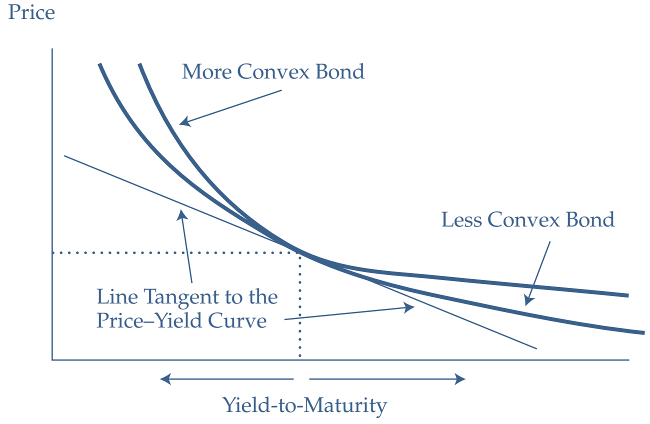
$$\left[\frac{1}{2} \times (46.047 \times \text{KES}100{,}940{,}423) \times (0.0100)^2\right]$$

$$= -\text{KES}6{,}184{,}418 + \text{KES}232{,}400$$

$$= -\text{KES}5{,}952{,}018$$

The factors that lead to greater convexity are the same as for duration. A fixed-rate bond with a longer time-to-maturity, a lower coupon rate, and a lower yield-to-maturity has greater convexity than a bond with a shorter time-to-maturity, a higher coupon rate, and a higher yield-to-maturity. Another factor is the dispersion of cash flows, meaning the degree to which payments are spread out over time. If two bonds have the same duration, the one that has the greater dispersion of cash flows has the greater convexity. The positive attributes of greater convexity for an investor are shown in Exhibit 12.

Exhibit 12 The Positive Attributes of Greater Bond Convexity on a Traditional (Option-Free) Bond



The two bonds in Exhibit 12 are assumed to have the same price, yield-to-maturity, and modified duration. Therefore, they share the same line tangent to their price–yield curves. The benefit of greater convexity occurs when their yields-to-maturity change. For the same decrease in yield-to-maturity, the more convex bond *appreciates more* in price. And for the same increase in yield-to-maturity, the more convex bond *depreciates less* in price. The conclusion is that the more convex bond outperforms the less convex bond in both bull (rising price) and bear (falling price) markets. This conclusion assumes, however, that this positive attribute is not "priced into" the bond. To the extent that it is included, the more convex bond would have a higher price (and lower yield-to-maturity). That does not diminish the value of convexity. It only suggests that the investor has to pay for it. As economists say, "There is no such thing as a free lunch."

EXAMPLE 14

The investment manager for a UK defined-benefit pension scheme is considering two bonds about to be issued by a large life insurance company. The first is a 30-year, 4% semiannual coupon payment bond. The second is a 100-year, 4% semiannual coupon payment "century" bond. Both bonds are expected to trade at par value at issuance.

Calculate the approximate modified duration and approximate convexity for each bond using a 5 bp increase and decrease in the annual yield-to-maturity. Retain accuracy to six decimals per 100 of par value.

Solution:

In the calculations, the yield per semiannual period goes up by 2.5 bps to 2.025% and down by 2.5 bps to 1.975%. The 30-year bond has an approximate modified duration of 17.381 and an approximate convexity of 420.80.

$$PV_+ = \frac{2}{(1.02025)^1} + \cdots + \frac{102}{(1.02025)^{60}} = 99.136214$$

$$PV_- = \frac{2}{(1.01975)^1} + \cdots + \frac{102}{(1.01975)^{60}} = 100.874306$$

$$\text{ApproxModDur} = \frac{100.874306 - 99.136214}{2 \times 0.0005 \times 100} = 17.381$$

$$\text{ApproxCon} = \frac{100.874306 + 99.136214 - (2 \times 100)}{(0.0005)^2 \times 100} = 420.80$$

The 100-year century bond has an approximate modified duration of 24.527 and an approximate convexity of 1,132.88.

$$PV_+ = \frac{2}{(1.02025)^1} + \cdots + \frac{102}{(1.02025)^{200}} = 98.787829$$

$$PV_- = \frac{2}{(1.01975)^1} + \cdots + \frac{102}{(1.01975)^{200}} = 101.240493$$

$$\text{ApproxModDur} = \frac{101.240493 - 98.787829}{2 \times 0.0005 \times 100} = 24.527$$

$$\text{ApproxCon} = \frac{101.240493 + 98.787829 - (2 \times 100)}{(0.0005)^2 \times 100} = 1{,}132.88$$

The century bond offers a higher modified duration, 24.527 compared with 17.381, and a much greater degree of convexity, 1,132.88 compared with 420.80.

In the same manner that the primary, or first-order, effect of a shift in the benchmark yield curve is measured by effective duration, the secondary, or second-order, effect is measured by **effective convexity**. The effective convexity of a bond is a *curve convexity* statistic that measures the secondary effect of a change in a benchmark yield curve. A pricing model is used to determine the new prices when the benchmark curve is shifted upward (PV_+) and downward (PV_-) by the same amount (ΔCurve). These changes are made holding other factors constant—for example, the credit spread. Then, Equation 18 is used to calculate the effective convexity (EffCon) given the initial price (PV_0).

$$\text{EffCon} = \frac{\left[(PV_-) + (PV_+)\right] - \left[2 \times (PV_0)\right]}{(\Delta\text{Curve})^2 \times (PV_0)} \tag{18}$$

This equation is very similar to Equation 15, for approximate *yield* convexity. The difference is that in Equation 15, the denominator includes the change in the yield-to-maturity squared, $(\Delta\text{Yield})^2$. Here, the denominator includes the change in the benchmark yield curve squared, $(\Delta\text{Curve})^2$.

Consider again the callable bond example in our initial discussion of effective duration. It is assumed that an option-pricing model is used to generate these callable bond prices: PV_0 = 101.060489, PV_+ = 99.050120, PV_- = 102.890738, and ΔCurve = 0.0025. The effective duration for the callable bond is 7.6006.

$$\text{EffDur} = \frac{102.890738 - 99.050120}{2 \times 0.0025 \times 101.060489} = 7.6006$$

Using these inputs in Equation 18, the effective convexity is –285.17.

$$\text{EffCon} = \frac{102.890738 + 99.050120 - (2 \times 101.060489)}{(0.0025)^2 \times 101.060489} = -285.17$$

Negative convexity, which could be called "concavity," is an important feature of callable bonds. Putable bonds, on the other hand, always have positive convexity. As a second-order effect, effective convexity indicates the change in the first-order effect (i.e., effective duration) as the benchmark yield curve is changed. In Exhibit 8, as the benchmark yield goes down, the slope of the line tangent to the curve for the non-callable bond steepens, which indicates positive convexity. But the slope of the line tangent to the callable bond flattens as the benchmark yield goes down. Technically, it reaches an inflection point, which is when the effective convexity shifts from positive to negative.

In summary, when the benchmark yield is high and the value of the embedded call option is low, the callable and the non-callable bonds experience very similar effects from interest rate changes. They both have positive convexity. But as the benchmark yield is reduced, the curves diverge. At some point, the callable bond moves into the range of negative convexity, which indicates that the embedded call option has more value to the issuer and is more likely to be exercised. This situation limits the potential price appreciation of the bond arising from lower interest rates, whether because of a lower benchmark yield or a lower credit spread.

Another way to understand why a callable bond can have negative convexity is to rearrange Equation 18.

$$\text{EffCon} = \frac{\left[(PV_-) - (PV_0)\right] - \left[(PV_0) - (PV_+)\right]}{(\Delta\text{Curve})^2 \times (PV_0)}$$

In the numerator, the first bracketed expression is the increase in price when the benchmark yield curve is lowered. The second expression is the decrease in price when the benchmark yield curve is raised. On a non-callable bond, the increase is always larger than the decrease (in absolute value). This result is the "convexity effect" for the relationship between bond prices and yields-to-maturity. On a callable bond, the increase can be smaller than the decrease (in absolute value). That creates negative convexity, as illustrated in Exhibit 8.

INVESTMENT HORIZON, MACAULAY DURATION AND INTEREST RATE RISK

10

j describe how the term structure of yield volatility affects the interest rate risk of a bond

k describe the relationships among a bond's holding period return, its duration, and the investment horizon

This section explores the effect of yield volatility on the investment horizon, and on the interaction between the investment horizon, market price risk, and coupon reinvestment risk.

10.1 Yield Volatility

An important aspect in understanding the interest rate risk and return characteristics of an investment in a fixed-rate bond is the time horizon. This section considers a short-term horizon. A primary concern for the investor is the change in the price of the bond given a sudden (i.e., same-day) change in its yield-to-maturity. The accrued interest does not change, so the impact of the change in the yield is on the flat price of the bond. Next, we consider a long-term horizon. The reinvestment of coupon interest then becomes a key factor in the investor's horizon yield.

Bond duration is the primary measure of risk arising from a change in the yield-to-maturity. Convexity is the secondary risk measure. In the discussion of the impact on the bond price, the phrase "for a *given* change in the yield-to-maturity" is used repeatedly. For instance, the given change in the yield-to-maturity could be 1 bp, 25 bps, or 100 bps. In comparing two bonds, it is assumed that the "given change" is the same for both securities. When the government bond par curve is shifted up or down by the same amount to calculate effective duration and effective convexity, the events are described as "parallel" yield curve shifts. Because yield curves are rarely (if ever) straight lines, this shift may also be described as a "shape-preserving" shift to the yield curve. The key assumption is that all yields-to-maturity under consideration rise or fall by the same amount across the curve.

Although the assumption of a parallel shift in the yield curve is common in fixed-income analysis, it is not always realistic. The shape of the yield curve changes based on factors affecting the supply and demand of shorter-term versus longer-term securities. In fact, the term structure of bond yields (also called the "term structure of interest rates") is typically upward sloping. However, the **term structure of yield volatility** may have a different shape depending on a number of factors. The term structure of yield volatility is the relationship between the volatility of bond yields-to-maturity and times-to-maturity.

For example, a central bank engaging in expansionary monetary policy might cause the yield curve to steepen by reducing short-term interest rates. But this policy might cause greater *volatility* in short-term bond yields-to-maturity than in longer-term bonds, resulting in a downward-sloping term structure of yield volatility. Longer-term bond yields are mostly determined by future inflation and economic growth expectations. Those expectations often tend to be less volatile.

The importance of yield volatility in measuring interest rate risk is that bond price changes are products of two factors: (1) the impact *per* basis-point change in the yield-to-maturity and (2) the *number* of basis points in the yield-to-maturity change. The first factor is duration or the combination of duration and convexity, and the second factor is the yield volatility. For example, consider a 5-year bond with a modified duration of 4.5 and a 30-year bond with a modified duration of 18.0. Clearly, for a *given* change in yield-to-maturity, the 30-year bond represents much more interest rate risk to an investor who has a short-term horizon. In fact, the 30-year bond appears to have *four times* the risk given the ratio of the modified durations. But that assumption neglects the possibility that the 30-year bond might have half the yield volatility of the 5-year bond.

Equation 14, restated here, summarizes the two factors.

$$\%\Delta PV^{Full} \approx (-\text{AnnModDur} \times \Delta\text{Yield}) + \left[\frac{1}{2} \times \text{AnnConvexity} \times (\Delta\text{Yield})^2\right]$$

The estimated percentage change in the bond price depends on the modified duration and convexity as well as on the yield-to-maturity change. Parallel shifts between two bond yields and along a benchmark yield curve are common assumptions in fixed-income analysis. However, an analyst must be aware that non-parallel shifts frequently occur in practice.

EXAMPLE 15

A fixed-income analyst is asked to rank three bonds in terms of interest rate risk. Interest rate risk here means the potential price decrease on a percentage basis given a sudden change in financial market conditions. The increases in the yields-to-maturity represent the "worst case" for the scenario being considered.

Bond	Modified Duration	Convexity	ΔYield
A	3.72	12.1	25 bps
B	5.81	40.7	15 bps
C	12.39	158.0	10 bps

The modified duration and convexity statistics are annualized. ΔYield is the increase in the annual yield-to-maturity. Rank the bonds in terms of interest rate risk.

Solution:

Calculate the estimated percentage price change for each bond:

Bond A:

$$(-3.72 \times 0.0025) + \left[\frac{1}{2} \times 12.1 \times (0.0025)^2\right] = -0.009262$$

Bond B:

$$(-5.81 \times 0.0015) + \left[\frac{1}{2} \times 40.7 \times (0.0015)^2\right] = -0.008669$$

Bond C:

$$(-12.39 \times 0.0010) + \left[\frac{1}{2} \times 158.0 \times (0.0010)^2\right] = -0.012311$$

Based on these assumed changes in the yield-to-maturity and the modified duration and convexity risk measures, Bond C has the highest degree of interest rate risk (a potential loss of 1.2311%), followed by Bond A (a potential loss of 0.9262%) and Bond B (a potential loss of 0.8669%).

10.2 Investment Horizon, Macaulay Duration, and Interest Rate Risk

Although short-term interest rate risk is a concern to some investors, other investors have a long-term horizon. Day-to-day changes in bond prices cause *unrealized* capital gains and losses. Those unrealized gains and losses might need to be accounted for in financial statements. This section considers a long-term investor concerned only with the total return over the investment horizon. Therefore, interest rate risk is important to this investor. The investor faces coupon reinvestment risk as well as market price risk if the bond needs to be sold prior to maturity.

Earlier, we discussed examples of interest rate risk using a 10-year, 8% annual coupon payment bond that is priced at 85.503075 per 100 of par value. The bond's yield-to-maturity is 10.40%. A key result in Example 3 is that an investor with a 10-year time horizon is concerned only with coupon reinvestment risk. This situation assumes, of course, that the issuer makes all of the coupon and principal payments as scheduled. The buy-and-hold investor has a higher total return if interest rates rise (see Example 3) and a lower total return if rates fall (see Example 5). The investor in Examples 4 and 6 has a four-year horizon. This investor faces market price risk in addition to coupon reinvestment risk. In fact, the market price risk dominates because this investor has a higher total return if interest rates fall (see Example 6) and a lower return if rates rise (see Example 4).

Now, consider a third investor who has a seven-year time horizon. If interest rates remain at 10.40%, the future value of reinvested coupon interest is 76.835787 per 100 of par value.

$$\left[8 \times (1.1040)^6\right] + \left[8 \times (1.1040)^5\right] + \left[8 \times (1.1040)^4\right] + \left[8 \times (1.1040)^3\right] +$$
$$\left[8 \times (1.1040)^2\right] + \left[8 \times (1.1040)^1\right] + 8 = 76.835787$$

The bond is sold for a price of 94.073336, assuming that the bond stays on the constant-yield price trajectory and continues to be "pulled to par."

$$\frac{8}{(1.1040)^1} + \frac{8}{(1.1040)^2} + \frac{108}{(1.1040)^3} = 94.073336$$

The total return is 170.909123 (= 76.835787 + 94.073336) per 100 of par value, and the horizon yield, as expected, is 10.40%.

$$85.503075 = \frac{170.909123}{(1+r)^7}, \quad r = 0.1040$$

Following Examples 3 and 4, assume that the yield-to-maturity on the bond rises to 11.40%. Also, coupon interest is now reinvested each year at 11.40%. The future value of reinvested coupons becomes 79.235183 per 100 of par value.

$$\left[8 \times (1.1140)^6\right] + \left[8 \times (1.1140)^5\right] + \left[8 \times (1.1140)^4\right] + \left[8 \times (1.1140)^3\right] +$$
$$\left[8 \times (1.1140)^2\right] + \left[8 \times (1.1140)^1\right] + 8 = 79.235183$$

After receiving the seventh coupon payment, the bond is sold. There is a capital loss because the price, although much higher than at purchase, is below the constant-yield price trajectory.

$$\frac{8}{(1.1140)^1} + \frac{8}{(1.1140)^2} + \frac{108}{(1.1140)^3} = 91.748833$$

The total return is 170.984016 (= 79.235183 + 91.748833) per 100 of par value and the holding-period rate of return is 10.407%.

$$85.503075 = \frac{170.984016}{(1 + r)^7}, \quad r = 0.10407$$

Following Examples 5 and 6, assume that the coupon reinvestment rates and the bond yield-to-maturity fall to 9.40%. The future value of reinvested coupons is 74.512177.

$$\left[8 + (1.0940)^6\right] + \left[8 + (1.0940)^5\right] + \left[8 + (1.0940)^4\right] + \left[8 + (1.0940)^3\right] +$$
$$\left[8 + (1.0940)^2\right] + \left[8 + (1.0940)^1\right] + 8 = 74.52177$$

The bond is sold at a capital gain because the price is above the constant-yield price trajectory.

$$\frac{8}{(1.0940)^1} + \frac{8}{(1.0940)^2} + \frac{108}{(1.0940)^3} = 96.481299$$

The total return is 170.993476 (= 74.512177 + 96.481299) per 100 of par value, and the horizon yield is 10.408%.

$$85.503075 = \frac{170.993476}{(1 + r)^7}, \quad r = 0.10408$$

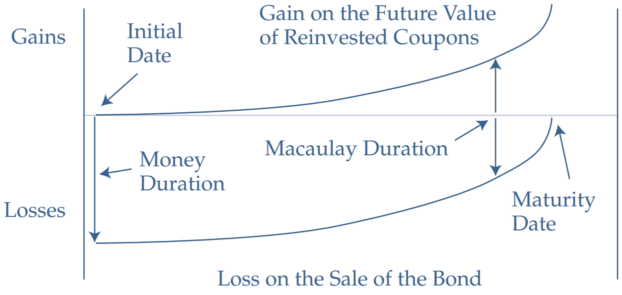
These results are summarized in the following table to reveal the remarkable outcome: The total returns and horizon yields are virtually the same. The investor with the 7-year horizon, unlike those having a 4- or 10-year horizon, achieves the same holding-period rate of return whether interest rates rise, fall, or remain the same. Note that the terms "horizon yield" and "holding-period rate of return" are used interchangeably in this reading. Sometimes "horizon yield" refers to yields on bonds that need to be sold at the end of the investor's holding period.

Interest Rate	Future Value of Reinvested Coupon	Sale Price	Total Return	Horizon Yield
9.40%	74.512177	96.481299	170.993476	10.408%
10.40%	76.835787	94.073336	170.909123	10.400%
11.40%	79.235183	91.748833	170.984016	10.407%

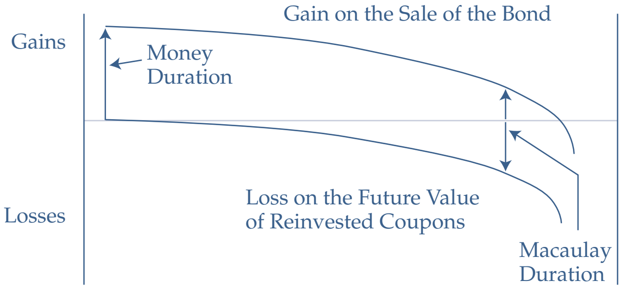
This particular bond was chosen as an example to demonstrate an important property of Macaulay duration: For a particular assumption about yield volatility, Macaulay duration indicates the investment horizon for which coupon reinvestment risk and market price risk offset each other. In Exhibit 2, the Macaulay duration of this 10-year, 8% annual payment bond is calculated to be 7.0029 years. This is one of the applications for duration in which "years" is meaningful and in which Macaulay duration is used rather than modified duration. The particular assumption about yield volatility is that there is a one-time "parallel" shift in the yield curve that occurs before the next coupon payment date. Exhibit 13 illustrates this property of bond duration, assuming that the bond is initially priced at par value.

Exhibit 13 Interest Rate Risk, Macaulay Duration, and the Investment Horizon

A. Interest Rates Rise


Gains
Initial Date
Gain on the Future Value of Reinvested Coupons
Money Duration
Macaulay Duration
Maturity Date
Losses
Loss on the Sale of the Bond

B. Interest Rates Fall


Gains
Gain on the Sale of the Bond
Money Duration
Losses
Loss on the Future Value of Reinvested Coupons
Macaulay Duration

As demonstrated in Panel A of Exhibit 13, when interest rates rise, duration measures the immediate drop in value. In particular, the money duration indicates the change in price. Then as time passes, the bond price is "pulled to par." The gain in the future value of reinvested coupons starts small but builds over time as more coupons are received. The curve indicates the additional future value of reinvested coupons because of the higher interest rate. At some point in the lifetime of the bond, those two effects offset each other and the gain on reinvested coupons is equal to the loss on the sale of the bond. That point in time is the Macaulay duration statistic.

The same pattern is displayed in the Panel B when interest rates fall, which leads to a reduction in the bond yield and the coupon reinvestment rate. There is an immediate jump in the bond price, as measured by the money duration, but then the "pull to par" effect brings the price down as time passes. The impact from reinvesting at a lower rate starts small but then becomes more significant over time. The loss on reinvested coupons is with respect to the future value if interest rates had not fallen. Once again, the bond's Macaulay duration indicates the point in time when the two effects offset each other and the gain on the sale of the bond matches the loss on coupon reinvestment.

The earlier numerical example and Exhibit 13 allow for a statement of the general relationships among interest rate risk, the Macaulay duration, and the investment horizon.

1 When the investment horizon is greater than the Macaulay duration of a bond, coupon reinvestment risk dominates market price risk. The investor's risk is to lower interest rates.

2 When the investment horizon is equal to the Macaulay duration of a bond, coupon reinvestment risk offsets market price risk.

3 When the investment horizon is less than the Macaulay duration of the bond, market price risk dominates coupon reinvestment risk. The investor's risk is to higher interest rates.

In the numerical example, the Macaulay duration of the bond is 7.0 years. Statement 1 reflects the investor with the 10-year horizon; Statement 2, the investor with the 7-year horizon; and Statement 3, the investor with the 4-year horizon.

The difference between the Macaulay duration of a bond and the investment horizon is called the **duration gap**. The duration gap is a bond's Macaulay duration minus the investment horizon. The investor with the 10-year horizon has a negative duration gap and currently is at risk of lower rates. The investor with the 7-year horizon has a duration gap of zero and currently is hedged against interest rate risk. The investor with the 4-year horizon has a positive duration gap and currently is at risk of higher rates. The word "currently" is important because interest rate risk is connected to an *immediate* change in the bond's yield-to-maturity and the coupon reinvestment rates. As time passes, the investment horizon is reduced and the Macaulay duration of the bond also changes. Therefore, the duration gap changes as well.

EXAMPLE 16

An investor plans to retire in 10 years. As part of the retirement portfolio, the investor buys a newly issued, 12-year, 8% annual coupon payment bond. The bond is purchased at par value, so its yield-to-maturity is 8.00% stated as an effective annual rate.

1 Calculate the approximate Macaulay duration for the bond, using a 1 bp increase and decrease in the yield-to-maturity and calculating the new prices per 100 of par value to six decimal places.

2 Calculate the duration gap at the time of purchase.

3 Does this bond at purchase entail the risk of higher or lower interest rates? Interest rate risk here means an immediate, one-time, parallel yield curve shift.

Solution to 1:

The approximate modified duration of the bond is 7.5361. $PV_0 = 100$, $PV_+ = 99.924678$, and $PV_- = 100.075400$.

$$PV_+ = \frac{8}{(1.0801)^1} + \cdots + \frac{108}{(1.0801)^{12}} = 99.924678$$

$$PV_- = \frac{8}{(1.0799)^1} + \cdots + \frac{108}{(1.0799)^{12}} = 100.075400$$

$$\text{ApproxModDur} = \frac{100.075400 - 99.924678}{2 \times 0.0001 \times 100} = 7.5361$$

The approximate Macaulay duration is 8.1390 (= 7.5361 × 1.08).

Solution to 2:

Given an investment horizon of 10 years, the duration gap for this bond at purchase is negative: 8.1390 – 10 = –1.8610.

Solution to 3:

A negative duration gap entails the risk of lower interest rates. To be precise, the risk is an immediate, one-time, parallel, downward yield curve shift because coupon reinvestment risk dominates market price risk. The loss from reinvesting coupons at a rate lower than 8% is larger than the gain from selling the bond at a price above the constant-yield price trajectory.

CREDIT AND LIQUIDITY RISK

11

l. explain how changes in credit spread and liquidity affect yield-to-maturity of a bond and how duration and convexity can be used to estimate the price effect of the changes

The focus of this reading is to demonstrate how bond duration and convexity estimate the bond price change, either in percentage terms or in units of currency, given an assumed yield-to-maturity change. This section addresses the *source* of the change in the yield-to-maturity. In general, the yield-to-maturity on a corporate bond is composed of a government *benchmark* yield and a *spread* over that benchmark. A change in the bond's yield-to-maturity can originate in either component or a combination of the two.

The key point is that for a traditional (option-free) fixed-rate bond, the same duration and convexity statistics apply for a change in the benchmark yield as for a change in the spread. The "building blocks" approach covered in an earlier reading shows that these yield-to-maturity changes can be broken down further. A change in the benchmark yield can arise from a change in either the expected inflation rate or the expected real rate of interest. A change in the spread can arise from a change in the credit risk of the issuer or in the liquidity of the bond. Therefore, for a fixed-rate bond, the "inflation duration," the "real rate duration," the "credit duration," and the "liquidity duration" are all the same number. The inflation duration would indicate the change in the bond price if expected inflation were to change by a certain amount. In the same manner, the real rate duration would indicate the bond price change if the real rate were to go up or down. The credit duration and liquidity duration would indicate the price sensitivity that would arise from changes in those building blocks in the yield-to-maturity. A bond with a modified duration of 5.00 and a convexity of 32.00 will appreciate in value by about 1.26% if its yield-to-maturity goes down by 25 bps: $(–5.00 \times –0.0025) + [1/2 \times 32.00 \times (–0.0025)^2] = +0.0126$, regardless of the *source* of the yield-to-maturity change.

Suppose that the yield-to-maturity on a corporate bond is 6.00%. If the benchmark yield is 4.25%, the spread is 1.75%. An analyst believes that credit risk makes up 1.25% of the spread and liquidity risk, the remaining 0.50%. Credit risk includes the probability of default as well as the recovery of assets if default does occur. A credit rating downgrade or an adverse change in the ratings outlook for a borrower reflects a higher risk of default. Liquidity risk refers to the transaction costs associated with selling a bond. In general, a bond with greater frequency of trading and a higher volume of trading provides fixed-income investors with more opportunity to purchase or sell the security and thus has less liquidity risk. In practice, there is a difference between the *bid* (or purchase) and the *offer* (or sale) price. This difference depends on the type of bond, the size of the transaction, and the time of execution, among other factors. For instance, government bonds often trade at just a few basis points between the purchase and sale prices. More thinly traded corporate bonds can have a much wider difference between the bid and offer prices.

The problem for a fixed-income analyst is that it is rare for the changes in the components of the overall yield-to-maturity to occur in isolation. In practice, the analyst is concerned with the *interaction* between changes in benchmark yields and spreads, between changes in expected inflation and the expected real rate, and between changes in credit and liquidity risk. For example, during a financial crisis, a "flight to quality" can cause government benchmark yields to fall as credit spreads widen. An unexpected credit downgrade on a corporate bond can result in greater credit as well as liquidity risk.

EXAMPLE 17

The (flat) price on a fixed-rate corporate bond falls one day from 92.25 to 91.25 per 100 of par value because of poor earnings and an unexpected ratings downgrade of the issuer. The (annual) modified duration for the bond is 7.24. Which of the following is *closest* to the estimated change in the credit spread on the corporate bond, assuming benchmark yields are unchanged?

A 15 bps

B 100 bps

C 108 bps

Solution:

Given that the price falls from 92.25 to 91.25, the percentage price decrease is 1.084%.

$$\frac{91.25 - 92.25}{92.25} = -0.01084$$

Given an annual modified duration of 7.24, the change in the yield-to-maturity is 14.97 bps.

$$-0.01084 \approx -7.24 \times \Delta\text{Yield},\ \Delta\text{Yield} = 0.001497$$

Therefore, the answer is A. The change in price reflects a credit spread increase on the bond of about 15 bps.

12 EMPIRICAL DURATION

m describe the difference between empirical duration and analytical duration

The approach taken in this reading to estimate duration and convexity statistics using mathematical formulas is often referred to as **analytical duration**. These estimates of the impact of benchmark yield changes on bond prices implicitly assume that government bond yields and spreads are independent variables that are uncorrelated with one another. Analytical duration offers a reasonable approximation of the price–yield relationship in many situations, but fixed-income professionals often use historical data in statistical models that incorporate various factors affecting bond prices to calculate **empirical duration** estimates. These estimates calculated over time and in different interest rate environments inform the fixed-income portfolio decision-making process and will be addressed in detail in later readings.

For instance, in the "flight to quality" example cited earlier in which investors sell risky assets during market turmoil and purchase default-risk-free government bonds, we might expect analytical and empirical duration estimates to differ among bond types. For example, on the one hand, for a government bond with little or no credit risk, we would expect analytical and empirical duration to be similar because benchmark yield changes largely drive bond prices. On the other hand, the same macroeconomic factors driving government bond yields lower in a market stress scenario will cause high-yield bond credit spreads to widen because of an increase in expected default risk. Since credit spreads and benchmark yields are negatively correlated under this scenario, wider credit spreads will partially or fully offset the decline in government benchmark yields, resulting in lower empirical duration estimates than analytical duration estimates. Importantly, analysts must consider the correlation between benchmark yields and credit spreads when deciding whether to use empirical or analytical duration estimates.

EXAMPLE 18

AFC Investment Ltd. is a fixed-income investment firm that actively manages a government bond fund and a corporate bond fund. Holdings of the government bond fund are mainly medium-term US Treasury securities but also include debt of highly rated developed-market sovereign issuers. About half of the corporate bond fund is invested in investment-grade issues, and the other half consists of high-yield issues, all with a mix of maturities and from a mix of North American, European, and Asian companies.

Explain why empirical duration is likely to be a more accurate risk measure for AFC's corporate bond fund than for its government bond fund.

Solution:

The government bond fund includes debt securities of the US government and other highly rated developed-market sovereign issuers. Since benchmark yields are the primary driver of changes in overall bond yields in this fund, the results of analytical duration and empirical duration should be broadly similar.

The corporate bond fund includes a wide variety of debt securities with varying levels of credit quality and liquidity and, therefore, different credit and liquidity spreads. Interactions between benchmark yield changes and credit and liquidity spreads would tend to offset each other, particularly during stressed market conditions, making empirical duration significantly lower than analytical duration. As a result, empirical duration may be the more accurate risk measure for the corporate bond fund.

SUMMARY

This reading covers the risk and return characteristics of fixed-rate bonds. The focus is on the widely used measures of interest rate risk—duration and convexity. These statistics are used extensively in fixed-income analysis. The following are the main points made in the reading:

- The three sources of return on a fixed-rate bond purchased at par value are: (1) receipt of the promised coupon and principal payments on the scheduled dates, (2) reinvestment of coupon payments, and (3) potential capital gains, as well as losses, on the sale of the bond prior to maturity.

- For a bond purchased at a discount or premium, the rate of return also includes the effect of the price being "pulled to par" as maturity nears, assuming no default.
- The total return is the future value of reinvested coupon interest payments and the sale price (or redemption of principal if the bond is held to maturity).
- The horizon yield (or holding period rate of return) is the internal rate of return between the total return and purchase price of the bond.
- Coupon reinvestment risk increases with a higher coupon rate and a longer reinvestment time period.
- Capital gains and losses are measured from the carrying value of the bond and not from the purchase price. The carrying value includes the amortization of the discount or premium if the bond is purchased at a price below or above par value. The carrying value is any point on the constant-yield price trajectory.
- Interest income on a bond is the return associated with the passage of time. Capital gains and losses are the returns associated with a change in the value of a bond as indicated by a change in the yield-to-maturity.
- The two types of interest rate risk on a fixed-rate bond are coupon reinvestment risk and market price risk. These risks offset each other to a certain extent. An investor gains from higher rates on reinvested coupons but loses if the bond is sold at a capital loss because the price is below the constant-yield price trajectory. An investor loses from lower rates on reinvested coupon but gains if the bond is sold at a capital gain because the price is above the constant-yield price trajectory.
- Market price risk dominates coupon reinvestment risk when the investor has a short-term horizon (relative to the time-to-maturity on the bond).
- Coupon reinvestment risk dominates market price risk when the investor has a long-term horizon (relative to the time-to-maturity)—for instance, a buy-and-hold investor.
- Bond duration, in general, measures the sensitivity of the full price (including accrued interest) to a change in interest rates.
- Yield duration statistics measuring the sensitivity of a bond's full price to the bond's own yield-to-maturity include the Macaulay duration, modified duration, money duration, and price value of a basis point.
- Curve duration statistics measuring the sensitivity of a bond's full price to the benchmark yield curve are usually called "effective durations."
- Macaulay duration is the weighted average of the time to receipt of coupon interest and principal payments, in which the weights are the shares of the full price corresponding to each payment. This statistic is annualized by dividing by the periodicity (number of coupon payments or compounding periods in a year).
- Modified duration provides a linear estimate of the percentage price change for a bond given a change in its yield-to-maturity.
- Approximate modified duration approaches modified duration as the change in the yield-to-maturity approaches zero.
- Effective duration is very similar to approximate modified duration. The difference is that approximate modified duration is a yield duration statistic that measures interest rate risk in terms of a change in the bond's own yield-to-maturity, whereas effective duration is a curve duration statistic that measures interest rate risk assuming a parallel shift in the benchmark yield curve.

- Key rate duration is a measure of a bond's sensitivity to a change in the benchmark yield curve at specific maturity segments. Key rate durations can be used to measure a bond's sensitivity to changes in the shape of the yield curve.
- Bonds with an embedded option do not have a meaningful internal rate of return because future cash flows are contingent on interest rates. Therefore, effective duration is the appropriate interest rate risk measure, not modified duration.
- The effective duration of a traditional (option-free) fixed-rate bond is its sensitivity to the benchmark yield curve, which can differ from its sensitivity to its own yield-to-maturity. Therefore, modified duration and effective duration on a traditional (option-free) fixed-rate bond are not necessarily equal.
- During a coupon period, Macaulay and modified durations decline smoothly in a "saw-tooth" pattern, assuming the yield-to-maturity is constant. When the coupon payment is made, the durations jump upward.
- Macaulay and modified durations are inversely related to the coupon rate and the yield-to-maturity.
- Time-to-maturity and Macaulay and modified durations are *usually* positively related. They are *always* positively related on bonds priced at par or at a premium above par value. They are *usually* positively related on bonds priced at a discount below par value. The exception is on long-term, low-coupon bonds, on which it is possible to have a lower duration than on an otherwise comparable shorter-term bond.
- The presence of an embedded call option reduces a bond's effective duration compared with that of an otherwise comparable non-callable bond. The reduction in the effective duration is greater when interest rates are low and the issuer is more likely to exercise the call option.
- The presence of an embedded put option reduces a bond's effective duration compared with that of an otherwise comparable non-putable bond. The reduction in the effective duration is greater when interest rates are high and the investor is more likely to exercise the put option.
- The duration of a bond portfolio can be calculated in two ways: (1) the weighted average of the time to receipt of *aggregate* cash flows and (2) the weighted average of the durations of individual bonds that compose the portfolio.
- The first method to calculate portfolio duration is based on the cash flow yield, which is the internal rate of return on the aggregate cash flows. It cannot be used for bonds with embedded options or for floating-rate notes.
- The second method is simpler to use and quite accurate when the yield curve is relatively flat. Its main limitation is that it assumes a parallel shift in the yield curve in that the yields on all bonds in the portfolio change by the same amount.
- Money duration is a measure of the price change in terms of units of the currency in which the bond is denominated.
- The price value of a basis point (PVBP) is an estimate of the change in the full price of a bond given a 1 bp change in the yield-to-maturity.
- Modified duration is the primary, or first-order, effect on a bond's percentage price change given a change in the yield-to-maturity. Convexity is the secondary, or second-order, effect. It indicates the change in the modified duration as the yield-to-maturity changes.
- Money convexity is convexity times the full price of the bond. Combined with money duration, money convexity estimates the change in the full price of a bond in units of currency given a change in the yield-to-maturity.

- Convexity is a positive attribute for a bond. Other things being equal, a more convex bond appreciates in price more than a less convex bond when yields fall and depreciates less when yields rise.
- Effective convexity is the second-order effect on a bond price given a change in the benchmark yield curve. It is similar to approximate convexity. The difference is that approximate convexity is based on a yield-to-maturity change and effective convexity is based on a benchmark yield curve change.
- Callable bonds have negative effective convexity when interest rates are low. The increase in price when the benchmark yield is reduced is less in absolute value than the decrease in price when the benchmark yield is raised.
- The change in a bond price is the product of: (1) the impact per basis-point change in the yield-to-maturity and (2) the number of basis points in the yield change. The first factor is estimated by duration and convexity. The second factor depends on yield volatility.
- The investment horizon is essential in measuring the interest rate risk on a fixed-rate bond.
- For a particular assumption about yield volatility, the Macaulay duration indicates the investment horizon for which coupon reinvestment risk and market price risk offset each other. The assumption is a one-time parallel shift to the yield curve in which the yield-to-maturity and coupon reinvestment rates change by the same amount in the same direction.
- When the investment horizon is greater than the Macaulay duration of the bond, coupon reinvestment risk dominates price risk. The investor's risk is to lower interest rates. The duration gap is negative.
- When the investment horizon is equal to the Macaulay duration of the bond, coupon reinvestment risk offsets price risk. The duration gap is zero.
- When the investment horizon is less than the Macaulay duration of the bond, price risk dominates coupon reinvestment risk. The investor's risk is to higher interest rates. The duration gap is positive.
- Credit risk involves the probability of default and degree of recovery if default occurs, whereas liquidity risk refers to the transaction costs associated with selling a bond.
- For a traditional (option-free) fixed-rate bond, the same duration and convexity statistics apply if a change occurs in the benchmark yield or a change occurs in the spread. The change in the spread can result from a change in credit risk or liquidity risk.
- In practice, there often is interaction between changes in benchmark yields and in the spread over the benchmark.
- Empirical duration uses statistical methods and historical bond prices to derive the price–yield relationship for specific bonds or bond portfolios.

REFERENCES

Smith, Donald J. 2014. *Bond Math: The Theory behind the Formulas*. 2nd ed. Hoboken, NJ: John Wiley & Sons.

PRACTICE PROBLEMS

1 A "buy-and-hold" investor purchases a fixed-rate bond at a discount and holds the security until it matures. Which of the following sources of return is *least likely* to contribute to the investor's total return over the investment horizon, assuming all payments are made as scheduled?

A Capital gain

B Principal payment

C Reinvestment of coupon payments

2 Which of the following sources of return is *most likely* exposed to interest rate risk for an investor of a fixed-rate bond who holds the bond until maturity?

A Capital gain or loss

B Redemption of principal

C Reinvestment of coupon payments

3 An investor purchases a bond at a price above par value. Two years later, the investor sells the bond. The resulting capital gain or loss is measured by comparing the price at which the bond is sold to the:

A carrying value.

B original purchase price.

C original purchase price value plus the amortized amount of the premium.

The following information relates to Questions 4–6

An investor purchases a nine-year, 7% annual coupon payment bond at a price equal to par value. After the bond is purchased and before the first coupon is received, interest rates increase to 8%. The investor sells the bond after five years. Assume that interest rates remain unchanged at 8% over the five-year holding period.

4 Per 100 of par value, the future value of the reinvested coupon payments at the end of the holding period is *closest* to:

A 35.00.

B 40.26.

C 41.07.

5 The capital gain/loss per 100 of par value resulting from the sale of the bond at the end of the five-year holding period is *closest* to a:

A loss of 8.45.

B loss of 3.31.

C gain of 2.75.

6 Assuming that all coupons are reinvested over the holding period, the investor's five-year horizon yield is *closest* to:

A 5.66%.

© 2021 CFA Institute. All rights reserved.

B 6.62%.

C 7.12%.

---

7 An investor buys a three-year bond with a 5% coupon rate paid annually. The bond, with a yield-to-maturity of 3%, is purchased at a price of 105.657223 per 100 of par value. Assuming a 5-basis point change in yield-to-maturity, the bond's approximate modified duration is *closest* to:

A 2.78.

B 2.86.

C 5.56.

8 Which of the following statements about duration is correct? A bond's:

A effective duration is a measure of yield duration.

B modified duration is a measure of curve duration.

C modified duration cannot be larger than its Macaulay duration (assuming a positive yield-to-maturity).

9 An investor buys a 6% annual payment bond with three years to maturity. The bond has a yield-to-maturity of 8% and is currently priced at 94.845806 per 100 of par. The bond's Macaulay duration is *closest* to:

A 2.62.

B 2.78.

C 2.83.

10 The interest rate risk of a fixed-rate bond with an embedded call option is *best* measured by:

A effective duration.

B modified duration.

C Macaulay duration.

11 Which of the following is *most* appropriate for measuring a bond's sensitivity to shaping risk?

A Key rate duration

B Effective duration

C Modified duration

12 A Canadian pension fund manager seeks to measure the sensitivity of her pension liabilities to market interest rate changes. The manager determines the present value of the liabilities under three interest rate scenarios: a base rate of 7%, a 100 basis point increase in rates up to 8%, and a 100 basis point drop in rates down to 6%. The results of the manager's analysis are presented below:

Interest Rate Assumption	Present Value of Liabilities
6%	CAD510.1 million
7%	CAD455.4 million
8%	CAD373.6 million

The effective duration of the pension fund's liabilities is *closest* to:

A 1.49.

B 14.99.

C 29.97.

13 Which of the following statements about Macaulay duration is correct?

A A bond's coupon rate and Macaulay duration are positively related.

B A bond's Macaulay duration is inversely related to its yield-to-maturity.

C The Macaulay duration of a zero-coupon bond is less than its time-to-maturity.

14 Assuming no change in the credit risk of a bond, the presence of an embedded put option:

A reduces the effective duration of the bond.

B increases the effective duration of the bond.

C does not change the effective duration of the bond.

15 A bond portfolio consists of the following three fixed-rate bonds. Assume annual coupon payments and no accrued interest on the bonds. Prices are per 100 of par value.

Bond	Maturity	Market Value	Price	Coupon	Yield-to-Maturity	Modified Duration
A	6 years	170,000	85.0000	2.00%	4.95%	5.42
B	10 years	120,000	80.0000	2.40%	4.99%	8.44
C	15 years	100,000	100.0000	5.00%	5.00%	10.38

The bond portfolio's modified duration is *closest* to:

A 7.62.

B 8.08.

C 8.20.

16 A limitation of calculating a bond portfolio's duration as the weighted average of the yield durations of the individual bonds that compose the portfolio is that it:

A assumes a parallel shift to the yield curve.

B is less accurate when the yield curve is less steeply sloped.

C is not applicable to portfolios that have bonds with embedded options.

17 Using the information below, which bond has the *greatest* money duration per 100 of par value assuming annual coupon payments and no accrued interest?

Bond	Time-to-Maturity	Price Per 100 of Par Value	Coupon Rate	Yield-to-Maturity	Modified Duration
A	6 years	85.00	2.00%	4.95%	5.42
B	10 years	80.00	2.40%	4.99%	8.44
C	9 years	85.78	3.00%	5.00%	7.54

A Bond A

B Bond B

C Bond C

18 A bond with exactly nine years remaining until maturity offers a 3% coupon rate with annual coupons. The bond, with a yield-to-maturity of 5%, is priced at 85.784357 per 100 of par value. The estimated price value of a basis point for the bond is *closest* to:

A 0.0086.

B 0.0648.

C 0.1295.

19 The “second-order” effect on a bond’s percentage price change given a change in yield-to-maturity can be *best* described as:

A duration.

B convexity.

C yield volatility.

20 A bond is currently trading for 98.722 per 100 of par value. If the bond’s yield-to-maturity (YTM) rises by 10 basis points, the bond’s full price is expected to fall to 98.669. If the bond’s YTM decreases by 10 basis points, the bond’s full price is expected to increase to 98.782. The bond’s approximate convexity is *closest* to:

A 0.071.

B 70.906.

C 1,144.628.

21 A bond has an annual modified duration of 7.020 and annual convexity of 65.180. If the bond’s yield-to-maturity decreases by 25 basis points, the expected percentage price change is *closest* to:

A 1.73%.

B 1.76%.

C 1.78%.

22 A bond has an annual modified duration of 7.140 and annual convexity of 66.200. The bond’s yield-to-maturity is expected to increase by 50 basis points. The expected percentage price change is *closest* to:

A –3.40%.

B –3.49%.

C –3.57%.

23 Which of the following statements relating to yield volatility is *most* accurate? If the term structure of yield volatility is downward sloping, then:

A short-term rates are higher than long-term rates.

B long-term yields are more stable than short-term yields.

C short-term bonds will always experience greater price fluctuation than long-term bonds.

24 The holding period for a bond at which the coupon reinvestment risk offsets the market price risk is *best* approximated by:

A duration gap.

B modified duration.

C Macaulay duration.

25 When the investor’s investment horizon is less than the Macaulay duration of the bond she owns:

A the investor is hedged against interest rate risk.

B reinvestment risk dominates, and the investor is at risk of lower rates.

C market price risk dominates, and the investor is at risk of higher rates.

26 An investor purchases an annual coupon bond with a 6% coupon rate and exactly 20 years remaining until maturity at a price equal to par value. The investor’s investment horizon is eight years. The approximate modified duration of the bond is 11.470 years. The duration gap at the time of purchase is *closest* to:

A –7.842.

B 3.470.

C 4.158.

27 A manufacturing company receives a ratings upgrade and the price increases on its fixed-rate bond. The price increase was *most likely* caused by a(n):

A decrease in the bond's credit spread.

B increase in the bond's liquidity spread.

C increase of the bond's underlying benchmark rate.

28 Empirical duration is likely the best measure of the impact of yield changes on portfolio value, especially under stressed market conditions, for a portfolio consisting of:

A 100% sovereign bonds of several AAA rated euro area issuers.

B 100% covered bonds of several AAA rated euro area corporate issuers.

C 25% AAA rated sovereign bonds, 25% AAA rated corporate bonds, and 50% high-yield (i.e., speculative-grade) corporate bonds, all from various euro area sovereign and corporate issuers.

SOLUTIONS

1 A is correct. A capital gain is least likely to contribute to the investor's total return. There is no capital gain (or loss) because the bond is held to maturity. The carrying value of the bond at maturity is par value, the same as the redemption amount. When a fixed-rate bond is held to its maturity, the investor receives the principal payment at maturity. This principal payment is a source of return for the investor. A fixed-rate bond pays periodic coupon payments, and the reinvestment of these coupon payments is a source of return for the investor. The investor's total return is the redemption of principal at maturity and the sum of the reinvested coupons.

2 C is correct. Because the fixed-rate bond is held to maturity (a "buy-and-hold" investor), interest rate risk arises entirely from changes in coupon reinvestment rates. Higher interest rates increase income from reinvestment of coupon payments, and lower rates decrease income from coupon reinvestment. There will not be a capital gain or loss because the bond is held until maturity. The carrying value at the maturity date is par value, the same as the redemption amount. The redemption of principal does not expose the investor to interest rate risk. The risk to a bond's principal is credit risk.

3 A is correct. Capital gains (losses) arise if a bond is sold at a price above (below) its constant-yield price trajectory. A point on the trajectory represents the carrying value of the bond at that time. That is, the capital gain/loss is measured from the bond's carrying value, the point on the constant-yield price trajectory, and not from the original purchase price. The carrying value is the original purchase price plus the amortized amount of the discount if the bond is purchased at a price below par value. If the bond is purchased at a price above par value, the carrying value is the original purchase price minus (not plus) the amortized amount of the premium. The amortized amount for each year is the change in the price between two points on the trajectory.

4 C is correct. The future value of reinvested cash flows at 8% after five years is closest to 41.07 per 100 of par value.

$$\left[7 \times (1.08)^4\right] + \left[7 \times (1.08)^3\right] + \left[7 \times (1.08)^2\right] + \left[7 \times (1.08)^1\right] + 7 = 41.0662$$

The 6.07 difference between the sum of the coupon payments over the five-year holding period (35) and the future value of the reinvested coupons (41.07) represents the "interest-on-interest" gain from compounding.

5 B is correct. The capital loss is closest to 3.31 per 100 of par value. After five years, the bond has four years remaining until maturity and the sale price of the bond is 96.69, calculated as

$$\frac{7}{(1.08)^1} + \frac{7}{(1.08)^2} + \frac{7}{(1.08)^3} + \frac{107}{(1.08)^4} = 96.69$$

The investor purchased the bond at a price equal to par value (100). Because the bond was purchased at a price equal to its par value, the carrying value is par value. Therefore, the investor experienced a capital loss of 96.69 – 100 = –3.31.

6 B is correct. The investor's five-year horizon yield is closest to 6.62%. After five years, the sale price of the bond is 96.69 (from problem 5) and the future value of reinvested cash flows at 8% is 41.0662 (from problem 4) per 100 of par value. The total return is 137.76 (= 41.07 + 96.69), resulting in a realized five-year horizon yield of 6.62%:

$$100.00 = \frac{137.76}{(1+r)^5}, \quad r = 0.0662$$

7 A is correct. The bond's approximate modified duration is closest to 2.78. Approximate modified duration is calculated as

$$\text{ApproxModDur} = \frac{(PV_-) - (PV_+)}{2 \times (\Delta\text{Yield}) \times (PV_0)}$$

Lower yield-to-maturity by 5 bps to 2.95%:

$$PV_- = \frac{5}{(1+0.0295)^1} + \frac{5}{(1+0.0295)^2} + \frac{5+100}{(1+0.0295)^3} = 105.804232$$

Increase yield-to-maturity by 5 bps to 3.05%:

$$PV_+ = \frac{5}{(1+0.0305)^1} + \frac{5}{(1+0.0305)^2} + \frac{5+100}{(1+0.0305)^3} = 105.510494$$

$$PV_0 = 105.657223,\ \Delta\text{Yield} = 0.0005$$

$$\text{ApproxModDur} = \frac{105.804232 - 105.510494}{2 \times 0.0005 \times 105.657223} = 2.78$$

8 C is correct. A bond's modified duration cannot be larger than its Macaulay duration assuming a positive yield-to-maturity. The formula for modified duration is

$$\text{ModDur} = \frac{\text{MacDur}}{1+r}$$

where r is the bond's yield-to-maturity per period. Therefore, ModDur will typically be less than MacDur.

Effective duration is a measure of curve duration. Modified duration is a measure of yield duration.

9 C is correct. The bond's Macaulay duration is closest to 2.83. Macaulay duration (MacDur) is a weighted average of the times to the receipt of cash flow. The weights are the shares of the full price corresponding to each coupon and principal payment.

Period	Cash Flow	Present Value	Weight	Period × Weight
1	6	5.555556	0.058575	0.058575
2	6	5.144033	0.054236	0.108472
3	106	84.146218	0.887190	2.661570
		94.845806	1.000000	2.828617

Thus, the bond's Macaulay duration (MacDur) is 2.83.

Alternatively, Macaulay duration can be calculated using the following closed-form formula:

$$\text{MacDur} = \left\{\frac{1+r}{r} - \frac{1+r+\left[N \times (c-r)\right]}{c \times \left[(1+r)^N - 1\right] + r}\right\} - (t/T)$$

$$\text{MacDur} = \left\{\frac{1.08}{0.08} - \frac{1.08+\left[3 \times (0.06-0.08)\right]}{0.06 \times \left[(1.08)^3 - 1\right] + 0.08}\right\} - 0$$

MacDur = 13.50 − 10.67 = 2.83

10 A is correct. The interest rate risk of a fixed-rate bond with an embedded call option is best measured by effective duration. A callable bond's future cash flows are uncertain because they are contingent on future interest rates. The issuer's decision to call the bond depends on future interest rates. Therefore, the yield-to-maturity on a callable bond is not well defined. Only effective duration, which takes into consideration the value of the call option, is the appropriate interest rate risk measure. Yield durations like Macaulay and modified durations are not relevant for a callable bond because they assume no changes in cash flows when interest rates change.

11 A is correct. Key rate duration is used to measure a bond's sensitivity to a shift at one or more maturity segments of the yield curve which result in a change to yield curve shape. Modified and effective duration measure a bond's sensitivity to parallel shifts in the entire curve.

12 B is correct. The effective duration of the pension fund's liabilities is closest to 14.99. The effective duration is calculated as follows:

$$\text{EffDur} = \frac{(PV_-) - (PV_+)}{2 \times (\Delta\text{Curve}) \times (PV_0)}$$

$PV_0 = 455.4$, $PV_+ = 373.6$, $PV_- = 510.1$, and $\Delta\text{Curve} = 0.0100$.

$$\text{EffDur} = \frac{510.1 - 373.6}{2 \times 0.0100 \times 455.4} = 14.99$$

13 B is correct. A bond's yield-to-maturity is inversely related to its Macaulay duration: The higher the yield-to-maturity, the lower its Macaulay duration and the lower the interest rate risk. A higher yield-to-maturity decreases the weighted average of the times to the receipt of cash flow, and thus decreases the Macaulay duration.

A bond's coupon rate is inversely related to its Macaulay duration: The lower the coupon, the greater the weight of the payment of principal at maturity. This results in a higher Macaulay duration. Zero-coupon bonds do not pay periodic coupon payments; therefore, the Macaulay duration of a zero-coupon bond is its time-to-maturity.

14 A is correct. The presence of an embedded put option reduces the effective duration of the bond, especially when rates are rising. If interest rates are low compared with the coupon rate, the value of the put option is low and the impact of the change in the benchmark yield on the bond's price is very similar to the impact on the price of a non-putable bond. But when benchmark interest rates rise, the put option becomes more valuable to the investor. The ability to sell the bond at par value limits the price depreciation as rates rise. The presence of an embedded put option reduces the sensitivity of the bond price to changes in the benchmark yield, assuming no change in credit risk.

15 A is correct. The portfolio's modified duration is closest to 7.62. Portfolio duration is commonly estimated as the market-value-weighted average of the yield durations of the individual bonds that compose the portfolio.

The total market value of the bond portfolio is 170,000 + 120,000 + 100,000 = 390,000.

The portfolio duration is 5.42 × (170,000/390,000) + 8.44 × (120,000/390,000) + 10.38 × (100,000/390,000) = 7.62.

16 A is correct. A limitation of calculating a bond portfolio's duration as the weighted average of the yield durations of the individual bonds is that this measure implicitly assumes a parallel shift to the yield curve (all rates change by the same amount in the same direction). In reality, interest rate changes frequently result in a steeper or flatter yield curve. This approximation of the "theoretically correct" portfolio duration is *more* accurate when the yield curve is flatter (less steeply sloped). An advantage of this approach is that it can be used with portfolios that include bonds with embedded options. Bonds with embedded options can be included in the weighted average using the effective durations for these securities.

17 B is correct. Bond B has the greatest money duration per 100 of par value. Money duration (MoneyDur) is calculated as the annual modified duration (AnnModDur) times the full price (PV^{Full}) of the bond including accrued interest. Bond B has the highest money duration per 100 of par value.

$$\text{MoneyDur} = \text{AnnModDur} \times PV^{Full}$$

MoneyDur of Bond A = 5.42 × 85.00 = 460.70

MoneyDur of Bond B = 8.44 × 80.00 = 675.20

MoneyDur of Bond C = 7.54 × 85.78 = 646.78

18 B is correct. The PVBP is closest to 0.0648. The formula for the price value of a basis point is

$$\text{PVBP} = \frac{(PV_-) - (PV_+)}{2}$$

where

PVBP = price value of a basis point
PV_- = full price calculated by lowering the yield-to-maturity by 1 bp
PV_+ = full price calculated by raising the yield-to-maturity by 1 bp

Lowering the yield-to-maturity by 1 bp to 4.99% results in a bond price of 85.849134:

$$PV_- = \frac{3}{(1+0.0499)^1} + \cdots + \frac{3+100}{(1+0.0499)^9} = 85.849134$$

Increasing the yield-to-maturity by 1 bp to 5.01% results in a bond price of 85.719638:

$$PV_+ = \frac{3}{(1+0.0501)^1} + \cdots + \frac{3+100}{(1+0.0501)^9} = 85.719638$$

$$\text{PVBP} = \frac{85.849134 - 85.719638}{2} = 0.06475$$

Alternatively, the PVBP can be derived using modified duration:

$$\text{ApproxModDur} = \frac{(PV_{-}) - (PV_{+})}{2 \times (\Delta\text{Yield}) \times (PV_0)}$$

$$\text{ApproxModDur} = \frac{85.849134 - 85.719638}{2 \times 0.0001 \times 85.784357} = 7.548$$

$$\text{PVBP} = 7.548 \times 85.784357 \times 0.0001 = 0.06475$$

19 B is correct. Convexity measures the “second order” effect on a bond’s percentage price change given a change in yield-to-maturity. Convexity adjusts the percentage price change estimate provided by modified duration to better approximate the true relationship between a bond’s price and its yield-to-maturity which is a curved line (convex).

Duration estimates the change in the bond’s price along the straight line that is tangent to this curved line (“first order” effect). Yield volatility measures the magnitude of changes in the yields along the yield curve.

20 B is correct. The bond’s approximate convexity is closest to 70.906. Approximate convexity (ApproxCon) is calculated using the following formula:

$$\text{ApproxCon} = [PV_{-} + PV_{+} - (2 \times PV_0)]/(\Delta\text{Yield}^2 \times PV_0)$$

where

PV_{-} = new price when the yield-to-maturity is decreased
PV_{+} = new price when the yield-to-maturity is increased
PV_0 = original price
ΔYield = change in yield-to-maturity

$$\text{ApproxCon} = [98.782 + 98.669 - (2 \times 98.722)]/(0.001^2 \times 98.722) = 70.906$$

21 C is correct. The expected percentage price change is closest to 1.78%. The convexity-adjusted percentage price change for a bond given a change in the yield-to-maturity is estimated by

$$\%\Delta PV^{Full} \approx [-\text{AnnModDur} \times \Delta\text{Yield}] + [0.5 \times \text{AnnConvexity} \times (\Delta\text{Yield})^2]$$

$$\%\Delta PV^{Full} \approx [-7.020 \times (-0.0025)] + [0.5 \times 65.180 \times (-0.0025)^2] = 0.017754,$$ or 1.78%

22 B is correct. The expected percentage price change is closest to −3.49%. The convexity-adjusted percentage price change for a bond given a change in the yield-to-maturity is estimated by

$$\%\Delta PV^{Full} \approx [-\text{AnnModDur} \times \Delta\text{Yield}] + [0.5 \times \text{AnnConvexity} \times (\Delta\text{Yield})^2]$$

$$\%\Delta PV^{Full} \approx [-7.140 \times 0.005] + [0.5 \times 66.200 \times (0.005)^2] = -0.034873,$$ or −3.49%

23 B is correct. If the term structure of yield volatility is downward-sloping, then short-term bond yields-to-maturity have greater volatility than for long-term bonds. Therefore, long-term yields are more stable than short-term yields. Higher volatility in short-term rates does not necessarily mean that the level of short-term rates is higher than long-term rates. With a downward-sloping term structure of yield volatility, short-term bonds will not always experience greater price fluctuation than long-term bonds. The estimated percentage change in a bond price depends on the modified duration and convexity as well as on the yield-to-maturity change.

24 C is correct. When the holder of a bond experiences a one-time parallel shift in the yield curve, the Macaulay duration statistic identifies the number of years necessary to hold the bond so that the losses (or gains) from coupon reinvestment offset the gains (or losses) from market price changes. The duration gap is the difference between the Macaulay duration and the investment horizon. Modified duration approximates the percentage price change of a bond given a change in its yield-to-maturity.

25 C is correct. The duration gap is equal to the bond's Macaulay duration minus the investment horizon. In this case, the duration gap is positive, and price risk dominates coupon reinvestment risk. The investor risk is to higher rates.

The investor is hedged against interest rate risk if the duration gap is zero; that is, the investor's investment horizon is equal to the bond's Macaulay duration. The investor is at risk of lower rates only if the duration gap is negative; that is, the investor's investment horizon is greater than the bond's Macaulay duration. In this case, coupon reinvestment risk dominates market price risk.

26 C is correct. The duration gap is closest to 4.158. The duration gap is a bond's Macaulay duration minus the investment horizon. The approximate Macaulay duration is the approximate modified duration times one plus the yield-to-maturity. It is 12.158 (= 11.470 × 1.06).

Given an investment horizon of eight years, the duration gap for this bond at purchase is positive: 12.158 – 8 = 4.158. When the investment horizon is less than the Macaulay duration of the bond, the duration gap is positive, and price risk dominates coupon reinvestment risk.

27 A is correct. The price increase was most likely caused by a decrease in the bond's credit spread. The ratings upgrade most likely reflects a lower expected probability of default and/or a greater level of recovery of assets if default occurs. The decrease in credit risk results in a smaller credit spread. The increase in the bond price reflects a decrease in the yield-to-maturity due to a smaller credit spread. The change in the bond price was not due to a change in liquidity risk or an increase in the benchmark rate.

28 C is correct. Empirical duration is the best measure—better than analytical duration—of the impact of yield changes on portfolio value, especially under stressed market conditions, for a portfolio consisting of a variety of different bonds from different issuers, such as the portfolio described in Answer C. In this portfolio, credit spread changes on the high-yield bonds may partly or fully offset yield changes on the AAA rated sovereign bonds and spread changes on the AAA rated corporate bonds; this interaction is best captured using empirical duration. The portfolios described in Answers A and B consist of the same types of bonds from similar issuers—sovereign bonds from similar-rated sovereign issuers (A) and covered bonds from similar-rated corporate issuers (B)—so empirical and analytical durations should be roughly similar in each of these portfolios.

READING

44

Fundamentals of Credit Analysis

by Christopher L. Gootkind, CFA

Christopher L. Gootkind, CFA, is at Loomis, Sayles & Company, LP (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe credit risk and credit-related risks affecting corporate bonds;
☐	**b.** describe default probability and loss severity as components of credit risk;
☐	**c.** describe seniority rankings of corporate debt and explain the potential violation of the priority of claims in a bankruptcy proceeding;
☐	**d.** compare and contrast corporate issuer credit ratings and issue credit ratings and describe the rating agency practice of "notching";
☐	**e.** explain risks in relying on ratings from credit rating agencies;
☐	**f.** explain the four Cs (Capacity, Collateral, Covenants, and Character) of traditional credit analysis;
☐	**g.** calculate and interpret financial ratios used in credit analysis;
☐	**h.** evaluate the credit quality of a corporate bond issuer and a bond of that issuer, given key financial ratios of the issuer and the industry;
☐	**i.** describe macroeconomic, market, and issuer-specific factors that influence the level and volatility of yield spreads;
☐	**j.** explain special considerations when evaluating the credit of high-yield, sovereign, and non-sovereign government debt issuers and issues.

The author would like to thank several of his Fixed Income Research colleagues at Loomis, Sayles & Company for their assistance with this reading: Paul Batterton, CFA, Diana Moskowitz, CFA, Diana Monteith, Shannon O'Mara, CFA, and Laura Sarlo, CFA.

© 2021 CFA Institute. All rights reserved.

1 INTRODUCTION

a describe credit risk and credit-related risks affecting corporate bonds

b describe default probability and loss severity as components of credit risk

With bonds outstanding worth many trillions of US dollars, the debt markets play a critical role in the global economy. Companies and governments raise capital in the debt market to fund current operations; buy equipment; build factories, roads, bridges, airports, and hospitals; acquire assets; and so on. By channeling savings into productive investments, the debt markets facilitate economic growth. Credit analysis has a crucial function in the debt capital markets—efficiently allocating capital by properly assessing credit risk, pricing it accordingly, and repricing it as risks change. How do fixed-income investors determine the riskiness of that debt, and how do they decide what they need to earn as compensation for that risk?

In the sections that follow, we cover basic principles of credit analysis, which may be broadly defined as the process by which credit risk is evaluated. Readers will be introduced to the definition of credit risk, the interpretation of credit ratings, the four Cs of traditional credit analysis, and key financial measures and ratios used in credit analysis. We explain, among other things, how to compare bond issuer creditworthiness within a given industry as well as across industries and how credit risk is priced in the bond market.

Our coverage focuses primarily on analysis of corporate debt; however, credit analysis of sovereign and nonsovereign, particularly municipal, government bonds will also be addressed. Structured finance, a segment of the debt markets that includes securities backed by such pools of assets as residential and commercial mortgages as well as other consumer loans, will not be covered here.

We first introduce the key components of credit risk—default probability and loss severity— along with such credit-related risks as spread risk, credit migration risk, and liquidity risk. We then discuss the relationship between credit risk and the capital structure of the firm before turning attention to the role of credit rating agencies. We also explore the process of analyzing the credit risk of corporations and examine the impact of credit spreads on risk and return. Finally, we look at special considerations applicable to the analysis of (i) high-yield (low-quality) corporate bonds and (ii) government bonds.

2 CREDIT RISK

Credit risk is the risk of loss resulting from the borrower (issuer of debt) failing to make full and timely payments of interest and/or principal. Credit risk has two components. The first is known as **default risk**, or **default probability**, which is the probability that a borrower defaults—that is, fails to meet its obligation to make full and timely payments of principal and interest according to the terms of the debt security. The second component is **loss severity** (also known as “loss given default”) in the event of default—that is, the portion of a bond’s value (including unpaid interest) an investor loses. A default can lead to losses of various magnitudes. In most instances, in the event of default, bondholders will recover some value, so there will not be a total loss on the investment. Thus, credit risk is reflected in the distribution of potential losses that may arise if the investor is not paid in full and on time. Although it is sometimes important to consider the entire distribution of potential losses and their respective probabilities—for instance, when losses have a disproportionate impact on the various tranches of a securitized pool of loans—it is often convenient to summarize the risk with a single default probability and loss severity and focus on the **expected loss**:

Expected loss = Default probability × Loss severity given default.

The loss severity, and hence the expected loss, can be expressed as either a monetary amount (e.g., €450,000) or as a percentage of the principal amount (e.g., 45%). The latter form of expression is generally more useful for analysis because it is independent of the amount of investment. Loss severity is often expressed as (1 – Recovery rate), where the recovery rate is the percentage of the principal amount recovered in the event of default.

Because default risk (default probability) is quite low for most high-quality debt issuers, bond investors tend to focus primarily on assessing this probability and devote less effort to assessing the potential loss severity arising from default. However, as an issuer's default risk rises, investors will focus more on what the recovery rate might be in the event of default. This issue will be discussed in more detail later. Important credit-related risks include the following:

- **Spread risk**. Corporate bonds and other "credit-risky" debt instruments typically trade at a yield premium, or spread, to bonds that have been considered "default-risk free," such as US Treasury bonds or German government bonds. Yield spreads, expressed in basis points, widen based on factors specific to the issuer, such as a decline in creditworthiness, sometimes referred to as credit migration or downgrade risk, or factors associated with the market as a whole, such as an increase in **market liquidity risk** or a general aversion to risk during periods of financial distress.
- **Credit migration risk** or **downgrade risk**. This is the risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher and thus causing the yield spreads on the issuer's bonds to widen and the price of its bonds to fall. The term "downgrade" refers to action by the major bond rating agencies, whose role will be covered in more detail later.
- **Market liquidity risk**. This is the risk that the price at which investors can actually transact—buying or selling—may differ from the price indicated in the market. To compensate investors for the risk that there may not be sufficient market liquidity for them to buy or sell bonds in the quantity they desire, the spread or yield premium on corporate bonds includes a market liquidity component, in addition to a credit risk component. Unlike stocks, which trade on exchanges, most markets bonds trade primarily over the counter through broker–dealers trading for their own accounts. Their ability and willingness to make markets, as reflected in the bid–ask spread, is an important determinant of market liquidity risk. The two main issuer-specific factors that affect market liquidity risk are (1) the size of the issuer (that is, the amount of publicly traded debt an issuer has outstanding) and (2) the credit quality of the issuer. In general, the less debt an issuer has outstanding, the less frequently its debt trades and thus the higher the market liquidity risk. And the lower the quality of the issuer, the higher the market liquidity risk.

During times of financial stress or crisis, such as in late 2008, market liquidity can decline sharply, causing yield spreads on corporate bonds and other credit-risky debt to widen and their prices to drop.

EXAMPLE 1

Defining Credit Risk

1 Which of the following *best* defines credit risk?
   - A The probability of default times the severity of loss given default
   - B The loss of principal and interest payments in the event of bankruptcy
   - C The risk of not receiving full interest and principal payments on a timely basis

2 Which of the following is the *best* measure of credit risk?
   - A The expected loss
   - B The severity of loss
   - C The probability of default

3 Which of the following is NOT credit or credit-related risk?
   - A Default risk
   - B Interest rate risk
   - C Downgrade or credit migration risk

Solution to 1:

C is correct. Credit risk is the risk that the borrower will not make full and timely payments.

Solution to 2:

A is correct. The expected loss captures both of the key components of credit risk: (the product of) the probability of default and the loss severity in the event of default. Neither component alone fully reflects the risk.

Solution to 3:

B is correct. Bond price changes due to general interest rate movements are not considered credit risk.

3 CAPITAL STRUCTURE, SENIORITY RANKING, AND RECOVERY RATES

c describe seniority rankings of corporate debt and explain the potential violation of the priority of claims in a bankruptcy proceeding

The various debt obligations of a given borrower will not necessarily all have the same **seniority ranking**, or priority of payment. In this section, we will introduce the topic of an issuer's capital structure and discuss the various types of debt claims that may arise from that structure, as well as their ranking and how those rankings can influence recovery rates in the event of default. The term "creditors" is used throughout our coverage to mean holders of debt instruments, such as bonds and bank loans. Unless specifically stated, it does not include such obligations as trade credit, tax liens, or employment-related obligations.

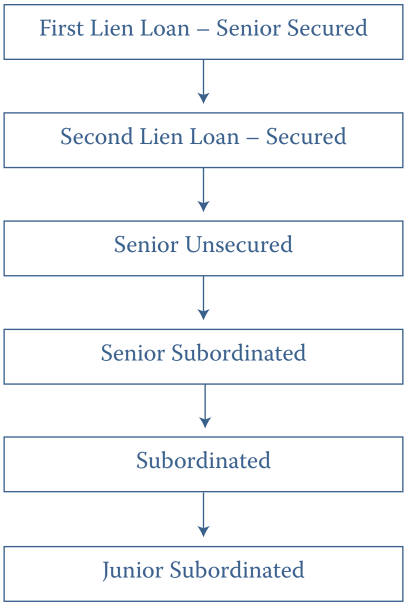
3.1 Capital Structure

The composition and distribution across operating units of a company's debt and equity—including bank debt, bonds of all seniority rankings, preferred stock, and common equity—is referred to as its **capital structure**. Some companies and industries have straightforward capital structures, with all the debt equally ranked and issued by one main operating entity. Other companies and industries, due to their frequent acquisitions and divestitures (e.g., media companies or conglomerates) or high levels of regulation (e.g., banks and utilities), tend to have more complicated capital structures. Companies in these industries often have many different subsidiaries, or operating companies, that have their own debt outstanding and parent holding companies that also issue debt, with different levels or rankings of seniority. Similarly, the cross-border operations of multinational corporations tend to increase the complexity of their capital structures.

3.2 Seniority Ranking

Just as borrowers can issue debt with many different maturity dates and coupons, they can also have many different rankings in terms of seniority. The ranking refers to the priority of payment, with the most senior or highest-ranking debt having the first claim on the cash flows and assets of the issuer. This level of seniority can affect the value of an investor's claim in the event of default and restructuring. Broadly, there is **secured debt** and **unsecured debt**. Secured debt means the debtholder has a direct claim—a pledge from the issuer—on certain assets and their associated cash flows. Unsecured bondholders have only a general claim on an issuer's assets and cash flows. In the event of default, unsecured debtholders' claims rank below (i.e., get paid after) those of secured creditors under what's known as the **priority of claims**.

Exhibit 1 Seniority Ranking



Within each category of debt are finer gradations of types and rankings. Within secured debt, there is first mortgage and first lien debt, which are the highest-ranked debt in terms of priority of repayment. **First mortgage debt** or loan refers to the pledge of a specific property (e.g., a power plant for a utility or a specific casino for a gaming company). **First lien debt** or loan refers to a pledge of certain assets that could include buildings but might also include property and equipment, licenses, patents, brands, and so on. There can also be **second lien**, or even third lien, secured debt, which, as the name implies, has a secured interest in the pledged assets but ranks below first lien debt in both collateral protection and priority of payment.

Within unsecured debt, there can also be finer gradations and seniority rankings. The highest-ranked unsecured debt is senior unsecured debt. It is the most common type of all corporate bonds outstanding. Other, lower-ranked debt includes **subordinated debt** and junior subordinated debt. Among the various creditor classes, these obligations have among the lowest priority of claims and frequently have little or no recovery in the event of default. That is, their loss severity can be as high as 100%. (See Exhibit 1 for a sample seniority ranking.) For regulatory and capital purposes, banks in Europe and the United States have issued debt and debt-like securities that rank even lower than subordinated debt, typically referred to as hybrids or trust preferred, and are intended to provide a capital cushion in times of financial distress. Many of them did not work as intended during the global financial crisis that began in 2008, and most were phased out, potentially to be replaced by more effective instruments that automatically convert to equity in certain circumstances.

Companies issue—and investors buy—debt with different seniority rankings for many reasons. Issuers are interested in optimizing their cost of capital—finding the right mix of the various types of both debt and equity—for their industry and type of business. Issuers may offer secured debt because that is what the market (i.e., investors) may require—given a company's perceived riskiness or because secured debt is generally lower cost due to the reduced credit risk inherent in its higher priority of claims. Or, issuers may offer subordinated debt because (1) they believe it is less expensive than issuing equity, as debtholders require a lower rate of return due to their superior place in line in the event of default; (2) doing so prevents dilution to existing shareholders; (3) it is typically less restrictive than issuing senior debt; and (4) investors are willing to buy it because they believe the yield being offered is adequate compensation for the risk they perceive.

EXAMPLE 2

Seniority Ranking

The Acme Company has senior unsecured bonds as well as both first and second lien debt in its capital structure. Which ranks higher with respect to priority of claims: senior unsecured bonds or second lien debt?

Solution:

Second lien debt ranks higher than senior unsecured bonds because of its secured position.

3.3 Recovery Rates

All creditors at the same level of the capital structure are treated as one class; thus, a senior unsecured bondholder whose debt is due in 30 years has the same pro rata claim in bankruptcy as one whose debt matures in six months. This provision is referred to as bonds ranking **pari passu** ("on an equal footing") in right of payment.

Defaulted debt will often continue to be traded by investors and broker–dealers based on their assessment that either in liquidation of the bankrupt company's assets or in reorganization, the bonds will have some recovery value. In the case of reorganization or restructuring (whether through formal bankruptcy or on a voluntary basis), new debt, equity, cash, or some combination thereof could be issued in exchange for the original defaulted debt.

As discussed, recovery rates vary by seniority of ranking in a company's capital structure, under the priority of claims treatment in bankruptcy. Over many decades, there have been enough defaults to generate statistically meaningful historical data on recovery rates by seniority ranking. Exhibit 2 provides recovery rates by seniority ranking for North American non-financial companies. For example, as shown in Exhibit 2, investors on average recovered 46.9% of the value of senior secured debt that defaulted in 2016 but only 29.2% of the value of senior unsecured issues that defaulted that year.

Exhibit 2 Average Corporate Debt Recovery Rates Measured by Ultimate Recoveries

	Emergence Year*			Default Year		
Seniority Ranking	**2017**	**2016**	**1987–2017**	**2017**	**2016**	**1987–2017**
Bank loans	81.3%	72.6%	80.4%	80.2%	78.3%	80.4%
Senior secured bonds	52.3%	35.9%	62.3%	57.5%	46.9%	62.3%
Senior unsecured bonds	54.1%	11.7%	47.9%	47.4%	29.2%	47.9%
Subordinated bonds	4.5%	6.6%	28.0%	NA	8.0%	28.0%

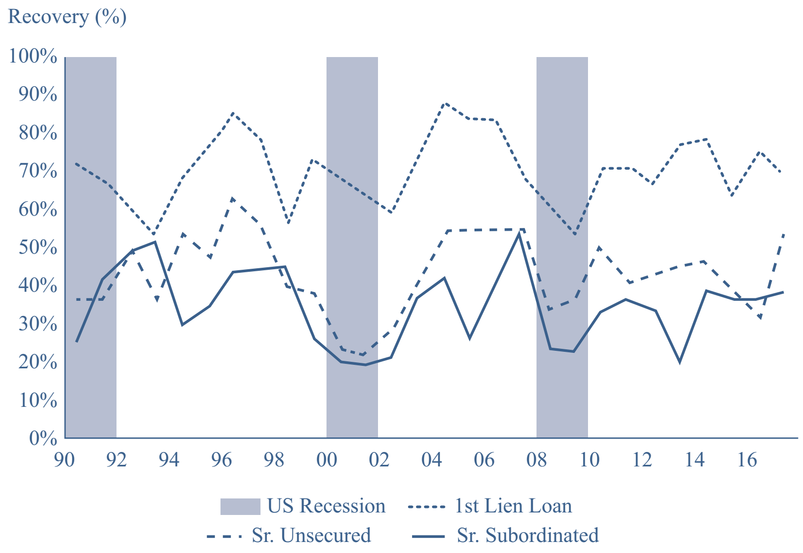
NA = not available.
*Emergence year is typically the year the defaulted company emerges from bankruptcy. Default year data refer to the recovery rate of debt that defaulted in that year (i.e., 2016 and 2017) or range of years (i.e., 1987–2017). Data are for North American nonfinancial companies.
Source: Moody's Investors Service's Ultimate Recovery Database.

A few things are worth noting:

1 **Recovery rates can vary widely by industry**. Companies that go bankrupt in industries that are in secular decline (e.g., newspaper publishing) will most likely have lower recovery rates than those that go bankrupt in industries merely suffering from a cyclical economic downturn.

2 **Recovery rates can also vary depending on when they occur in a credit cycle**. Credit cycles describe the changing availability—and pricing—of credit. When the economy is strong or improving, the willingness of lenders to extend credit on favorable terms is high. Conversely, when the economy is weak or weakening, lenders pull back, or "tighten" credit, by making it less available and more expensive. As shown in Exhibit 3, at or near the bottom of a credit cycle—which is almost always closely linked with an economic cycle—recoveries will tend to be lower than at other times in the credit cycle.

Exhibit 3 Global Recovery Rates by Seniority Ranking, 1990–2017



Source: Based on data from Moody's Investors Service's Ultimate Recovery Database.

3 **These recovery rates are averages**. In fact, there can be large variability, both across industries, as noted, as well as across companies within a given industry. Factors might include composition and proportion of debt across an issuer's capital structure. An abundance of secured debt will lead to smaller recovery rates on lower-ranked debt.

Understanding recovery rates is important because they are a key component of credit analysis and risk. Recall that the best measure of credit risk is expected loss—that is, probability of default times loss severity given default. And loss severity equals (1 – Recovery rate). Having an idea how much one can lose in the event of default is a critical factor in valuing credit, particularly lower-quality credit, as the default risk rises.

Priority of claims: Not always absolute. The priority of claims in bankruptcy—the idea that the highest-ranked creditors get paid out first, followed by the next level, and on down, like a waterfall—is well established and is often described as "absolute." In principle, in the event of bankruptcy or liquidation:

- Creditors with a secured claim have the right to the value of that specific property before any other claim. If the value of the pledged property is less than the amount of the claim, then the difference becomes a senior unsecured claim.
- Unsecured creditors have a right to be paid in full before holders of equity interests (common and preferred shareholders) receive value on their interests.
- Senior unsecured creditors take priority over all subordinated creditors. A creditor is senior unsecured unless expressly subordinated.

In practice, however, creditors with lower seniority and even shareholders may receive some consideration without more senior creditors being paid in full. Why might this be the case? In bankruptcy, there are different classes of claimants, and all classes that are impaired (that is, receive less than full claim) get to vote to confirm the plan of reorganization. This vote is subject to the absolute priority of claims. Either by consent of the various parties or by the judge's order, however, absolute priority may not be strictly enforced in the final plan. There may be disputes over the value of various assets in the bankruptcy estate (e.g., what is a plant, or a patent portfolio, worth?) or the present value or timing of payouts. For example, what is the value of the new debt I'm receiving for my old debt of a reorganized company before it emerges from bankruptcy?

Resolution of these disputes takes time, and cases can drag on for months and years. In the meantime, during bankruptcy, substantial expenses are being incurred for legal and accounting fees, and the value of the company may be declining as key employees leave, customers go elsewhere, and so on. Thus, to avoid the time, expense, and uncertainty over disputed issues, such as the value of property in the estate and the legality of certain claims, the various claimants have an incentive to negotiate and compromise. This frequently leads to creditors with lower seniority and other claimants (e.g., even shareholders) receiving more consideration than they are legally entitled to.

It's worth noting that in the United States, the bias is toward reorganization and recovery of companies in bankruptcy, whereas in other jurisdictions, such as the United Kingdom, the bias is toward liquidation of companies in bankruptcy and maximizing value to the banks and other senior creditors. It's also worth noting that bankruptcy and bankruptcy laws are very complex and can vary greatly by country, so it is difficult to generalize about how creditors will fare. As shown in the earlier chart, there is huge variability in recovery rates for defaulted debt. Every case is different.

EXAMPLE 3

Priority of Claims

1 Under which circumstance is a subordinated bondholder *most likely* to recover some value in a bankruptcy without a senior creditor getting paid in full? When:

- **A** absolute priority rules are enforced.
- **B** the various classes of claimants agree to it.
- **C** the company is liquidated rather than reorganized.

2 In the event of bankruptcy, claims at the same level of the capital structure are:

- **A** on an equal footing, regardless of size, maturity, or time outstanding.
- **B** paid in the order of maturity from shortest to longest, regardless of size or time outstanding.
- **C** paid on a first-in, first-out (FIFO) basis so that the longest-standing claims are satisfied first, regardless of size or maturity.

Solution to 1:

B is correct. All impaired classes get to vote on the reorganization plan. Negotiation and compromise are often preferable to incurring huge legal and accounting fees in a protracted bankruptcy process that would otherwise reduce the value of the estate for all claimants. This process may allow junior creditors (e.g., subordinated bondholders) to recover some value even though more senior creditors do not get paid in full.

Solution to 2:

A is correct. All claims at the same level of the capital structure are pari passu (on an equal footing).

4 RATING AGENCIES, CREDIT RATINGS, AND THEIR ROLE IN THE DEBT MARKETS

d compare and contrast corporate issuer credit ratings and issue credit ratings and describe the rating agency practice of "notching"

The major credit rating agencies—Moody's Investors Service ("Moody's"), Standard & Poor's ("S&P"), and Fitch Ratings ("Fitch")—play a central, if somewhat controversial, role in the credit markets. For the vast majority of outstanding bonds, at least two of the agencies provide ratings: a symbol-based measure of the potential risk of default of a particular bond or issuer of debt. In the public and quasi-public bond markets, underwritten by investment banks as opposed to privately placed, issuers won't offer, and investors won't buy, bonds that do not carry ratings from Moody's, S&P, or Fitch. This practice applies for all types of bonds: government or sovereign; entities with implicit or explicit guarantees from the government, such as Ginnie Mae in the United States and *Pfandbriefe* in Germany; supranational entities, such as the World Bank, which are owned by several governments; corporate; non-sovereign government; and mortgage- and asset-backed debt. How did the rating agencies attain such a dominant position in the credit markets? What are credit ratings, and what do they mean? How does the market use credit ratings? What are the risks of relying solely or excessively on credit ratings?

The history of the major rating agencies goes back more than 100 years. John Moody began publishing credit analysis and opinions on US railroads in 1909. S&P published its first ratings in 1916. They have grown in size and prominence since then. Many bond investors like the fact that there are independent analysts who meet with the issuer and often have access to material, non-public information, such as financial projections that investors cannot receive, to aid in the analysis. What has also proven very attractive to investors is that credit ratings provide direct and easy comparability of the relative credit riskiness of all bond issuers, within and across industries and bond types, although there is some debate about ratings comparability across the types of bonds. For instance, investigations conducted after the 2008–2009 global financial crisis suggested that, for a given rating category, municipal bonds have experienced a lower historical incidence of default than corporate debt.

Several factors have led to the near universal use of credit ratings in the bond markets and the dominant role of the major credit rating agencies. These factors include the following:

- Independent assessment of credit risk
- Ease of comparison across bond issuers, issues, and market segments
- Regulatory and statutory reliance and usage
- Issuer payment for ratings
- Huge growth of debt markets
- Development and expansion of bond portfolio management and the accompanying bond indexes

However, in the aftermath of the global financial crisis of 2008–2009, when the rating agencies were blamed for contributing to the crisis with their overly optimistic ratings on securities backed by subprime mortgages, attempts were made to reduce the role and dominant positions of the major credit rating agencies. New rules, regulations, and legislation were passed to require the agencies to be more transparent, reduce conflicts of interest, and stimulate more competition. The "issuer pay" model allows the distribution of ratings to a broad universe of investors and undoubtedly facilitated widespread reliance on ratings. Challenging the dominance of Moody's, S&P, and Fitch, additional credit rating agencies have emerged. Some credit rating agencies that are well-established in their home markets but are not so well known globally, such as Dominion Bond Rating Service (DBRS) in Canada and Japan Credit Rating Agency (JCR) in Japan, have tried to raise their profiles. The market dominance of the biggest credit rating agencies, however, remains largely intact.

4.1 Credit Ratings

The three major global credit rating agencies—Moody's, S&P, and Fitch—use similar, symbol-based ratings that are basically an assessment of a bond issue's risk of default. Exhibit 4 shows their long-term ratings ranked from highest to lowest. Ratings on short-term debt, although available, are not shown here.

Exhibit 4 Long-Term Rating Matrix: Investment Grade vs. Non-Investment Grade

		Moody's	S&P	Fitch
Investment Grade	High-Quality Grade	Aaa	AAA	AAA
		Aa1	AA+	AA+
		Aa2	AA	AA
		Aa3	AA–	AA–
	Upper-Medium Grade	A1	A+	A+
		A2	A	A
		A3	A–	A–
	Low-Medium Grade	Baa1	BBB+	BBB+
		Baa2	BBB	BBB
		Baa3	BBB–	BBB–
Non-Investment Grade ("Junk" or "High Yield")	Low Grade or Speculative Grade	Ba1	BB+	BB+
		Ba2	BB	BB
		Ba3	BB–	BB–
		B1	B+	B+
		B2	B	B
		B3	B–	B–
		Caa1	CCC+	CCC+
		Caa2	CCC	CCC
		Caa3	CCC–	CCC–
		Ca	CC	CC
		C	C	C
	Default	C	D	D

Bonds rated triple-A (Aaa or AAA) are said to be "of the highest quality, subject to the lowest level of credit risk" (Moody's Investors Service, see "Rating Symbols and Definitions", https://www.moodys.com/Pages/amr002002.aspx) and thus have extremely low probabilities of default. Double-A (Aa or AA) rated bonds are referred to as "high-quality grade" and are also regarded as having very low default risk. Bonds rated single-A are referred to as "upper-medium grade." Baa (Moody's) or BBB (S&P

and Fitch) are called "low-medium grade." Bonds rated Baa3/BBB– or higher are called "investment grade." Bonds rated Ba1 or lower by Moody's and BB+ or lower by S&P and Fitch, respectively, have speculative credit characteristics and increasingly higher default risk. As a group, these bonds are referred to in a variety of ways: "low grade," "speculative grade," "non-investment grade," "below investment grade," "high yield," and, in an attempt to reflect the extreme level of risk, some observers refer to these bonds as "junk bonds." The D rating is reserved for securities that are already in default in S&P's and Fitch's scales. For Moody's, bonds rated C are likely, but not necessarily, in default. Generally, issuers of bonds rated investment grade are more consistently able to access the debt markets and can borrow at lower interest rates than those rated below investment grade.

In addition, rating agencies will typically provide outlooks on their respective ratings—positive, stable, or negative—and may provide other indicators on the potential direction of their ratings under certain circumstances, such as "On Review for Downgrade" or "On CreditWatch for an Upgrade." It should also be noted that, in support of the ratings they publish, the rating agencies also provide extensive written commentary and financial analysis on the obligors they rate, as well as summary industry statistics.

4.2 Issuer vs. Issue Ratings

Rating agencies will typically provide both issuer and issue ratings, particularly as they relate to corporate debt. Terminology used to distinguish between issuer and issue ratings includes corporate family rating (CFR) and corporate credit rating (CCR) or issuer credit rating and issue credit rating. An issuer credit rating is meant to address an obligor's overall creditworthiness—its ability and willingness to make timely payments of interest and principal on its debt. The issuer credit rating usually applies to its senior unsecured debt.

Issue ratings refer to specific financial obligations of an issuer and take into consideration such factors as ranking in the capital structure (e.g., secured or subordinated). Although **cross-default provisions**, whereby events of default (such as non-payment of interest) on one bond trigger default on all outstanding debt, imply the same default probability for all issues, specific issues may be assigned different credit ratings—higher or lower—due to a rating adjustment methodology known as **notching**.

Notching

For the rating agencies, likelihood of default—default risk—is the primary factor in assigning their ratings. However, there are secondary factors as well. These factors include the priority of payment in the event of a default (e.g., secured versus senior unsecured versus subordinated) as well as potential loss severity in the event of default. Another factor considered by rating agencies is **structural subordination**, which can arise when a corporation with a holding company structure has debt at both its parent holding company and operating subsidiaries. Debt at the operating subsidiaries will get serviced by the cash flow and assets of the subsidiaries before funds can be passed ("upstreamed") to the holding company to service debt at that level.

Recognizing these different payment priorities, and thus the potential for higher (or lower) loss severity in the event of default, the rating agencies have adopted a notching process whereby their credit ratings on issues can be moved up or down from the issuer rating, which is usually the rating applied to its senior unsecured debt. As a general rule, the higher the senior unsecured rating, the smaller the notching adjustment. The reason behind this is that the higher the rating, the lower the perceived risk of default; so, the need to "notch" the rating to capture the potential difference in loss severity is greatly reduced. For lower-rated credits, however, the risk of default is greater and thus the potential difference in loss from a lower (or higher) priority ranking is a bigger consideration in assessing an issue's credit riskiness. Thus, the rating agencies will typically apply larger rating adjustments. For example, S&P applies the following notching guidelines:

> A key principle is that investment-grade ratings focus more on timeliness, while non-investment-grade ratings give additional weight to recovery. For example, subordinated debt can be rated up to two notches below a non-investment-grade corporate credit rating, but one notch at most if the corporate credit rating is investment grade. Conversely, … the 'AAA' rating category need not be notched at all, while at the 'CCC' level the gaps may widen.
>
> The rationale for this convention is straightforward: as default risk increases, the concern over what can be recovered takes on greater relevance and, therefore, greater rating significance. Accordingly, the ultimate recovery aspect of ratings is given more weight as one moves down the rating spectrum.(Standard & Poor's, "Rating the Issue," in *Corporate Ratings Criteria 2008* [New York: Standard & Poor's, 2008]: 65)

Exhibit 5 is an example of S&P's notching criteria, as applied to Infor Software Parent LLC, Inc. (Infor). Infor is a US-based global software and services company whose corporate credit rating from S&P is B–. Note how the company's senior secured bonds are rated B, whereas its senior unsecured bonds are rated two notches lower at CCC+ and its holding company debt is rated even one notch lower at CCC.

Exhibit 5 Infor S&P Ratings Detail (as of December 2018)

Corporate credit rating	B–/Stable
Senior secured (3 issues)	B
Senior unsecured (2 issues)	CCC+
Holding company debt (1 issue)	CCC

Source: Standard & Poor's Financial Services, LLC.

4.3 ESG Ratings

ESG investing has received a great deal of attention and acceptance as a viable investment strategy, even as a source of alpha. ESG investing involves asset selection and investment decision making based on disciplined evaluation of one or all of these considerations:

- Environmental themes, such as investing in companies that are responding to consumer demand for sustainable practices, the reduction of carbon emissions, and other environmental agendas.
- Social themes, such as investing in companies committed to a diverse and inclusive workplace, minimizing salary gap, and community involvement programs.
- Governance themes, such as investing in companies committed to diverse board composition and strong oversight.

MSCI Inc., a global provider of market indexes and analytic tools, has launched a set of ratings that aim to measure a company's attitudes, practices, and advances related to ESG. Their rules-based methodology aims to identify and track leaders and laggards in the space. Companies are evaluated according to their exposure to ESG risks and how well they manage those risks relative to peers. The MSCI ESG rating scale has seven tiers and spans AAA to CCC. From best to worst, leaders receive

AAA or AA; the next echelon constitutes an average rating of A, BBB, or BB; while laggards receive B or CCC. MSCI has also added countries, mutual funds, and ETFs to their ESG rating system. Traditional rating agencies, such as Moody's and S&P, are evaluating ESG for borrowers under these criteria too.

ESG investing allows investors to align their investment objectives with long-term trends considered important to the overall sustainability of a business. Some evidence suggests that funds that value 'ethics,' by construction, also outperform key market indexes.

4.4 Risks in Relying on Agency Ratings

e explain risks in relying on ratings from credit rating agencies

The dominant position of the rating agencies in the global debt markets, and the near-universal use of their credit ratings on debt securities, suggests that investors believe they do a good job assessing credit risk. In fact, with a few exceptions (e.g., too high ratings on US subprime mortgage-backed securities issued in the mid-2000s, which turned out to be much riskier than expected), their ratings have proved quite accurate as a relative measure of default risk. For example, Exhibit 6 shows historical S&P one-year global corporate default rates by rating category for the 20-year period from 1998 to 2017. It measures the percentage of issues that defaulted in a given calendar year based on how they were rated at the beginning of the year.

Exhibit 6 Global Corporate Annual Default Rates by Rating Category (%)

	AAA	AA	A	BBB	BB	B	CCC/C
1998	0.00	0.00	0.00	0.41	0.82	4.63	42.86
1999	0.00	0.17	0.18	0.20	0.95	7.29	33.33
2000	0.00	0.00	0.27	0.37	1.16	7.70	35.96
2001	0.00	0.00	0.27	0.34	2.96	11.53	45.45
2002	0.00	0.00	0.00	1.01	2.89	8.21	44.44
2003	0.00	0.00	0.00	0.23	0.58	4.07	32.73
2004	0.00	0.00	0.08	0.00	0.44	1.45	16.18
2005	0.00	0.00	0.00	0.07	0.31	1.74	9.09
2006	0.00	0.00	0.00	0.00	0.30	0.82	13.33
2007	0.00	0.00	0.00	0.00	0.20	0.25	15.24
2008	0.00	0.38	0.39	0.49	0.81	4.09	27.27
2009	0.00	0.00	0.22	0.55	0.75	10.94	49.46
2010	0.00	0.00	0.00	0.00	0.58	0.86	22.62
2011	0.00	0.00	0.00	0.07	0.00	1.67	16.30
2012	0.00	0.00	0.00	0.00	0.03	1.57	27.52
2013	0.00	0.00	0.00	0.00	0.10	1.64	24.50
2014	0.00	0.00	0.00	0.00	0.00	0.78	17.42
2015	0.00	0.00	0.00	0.00	0.16	2.40	26.51
2016	0.00	0.00	0.00	0.00	0.47	3.70	33.17
2017	0.00	0.00	0.00	0.00	0.08	0.98	26.23
Mean	0.00	0.03	0.07	0.19	0.69	3.82	27.98

Exhibit 6 (Continued)

	AAA	AA	A	BBB	BB	B	CCC/C
Max	0.00	0.38	0.39	1.01	2.96	11.53	45.45
Min	0.00	0.00	0.00	0.00	0.00	0.25	9.09

Source: Based on data from Standard & Poor's Financial Services, LLC.

As Exhibit 6 shows, the highest-rated bonds have extremely low default rates. With very few exceptions, the lower the rating, the higher the annual rate of default, with bonds rated CCC and lower experiencing the highest default rates by far.

Relying on credit rating agency ratings, however, also has limitations and risks, including the following:

- **Credit ratings can change over time**. Over a long time period (e.g., many years), credit ratings can migrate—move up or down—significantly from what they were at the time of bond issuance. Using Standard & Poor's data, Exhibit 7 shows the average three-year migration (or "transition") by rating from 1981 to 2017. Note that the higher the credit rating, the greater the ratings stability. Even for AAA rated credits, however, only about 65% of the time did ratings remain in that rating category over a three-year period. (Of course, AAA rated credits can have their ratings move in only one direction—down.) A very small fraction of AAA rated credits became non-investment grade or defaulted within three years. For single-B rated credits, only 41% of the time did ratings remain in that rating category over three-year periods. This observation about how credit ratings can change over time isn't meant to be a criticism of the rating agencies. It is meant to demonstrate that creditworthiness can and does change—up or down—and that bond investors should not assume an issuer's credit rating will remain the same from time of purchase through the entire holding period.

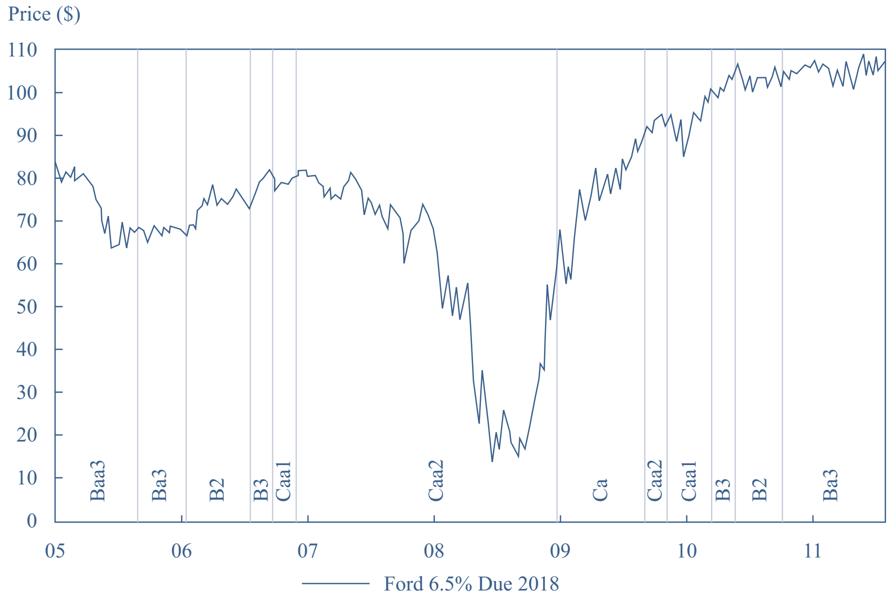
Exhibit 7 Average Three-Year Global Corporate Transition Rates, 1981–2017 (%)

From/To	AAA	AA	A	BBB	BB	B	CCC/C	D	NR
AAA	65.48	22.09	2.35	0.32	0.19	0.08	0.11	0.13	9.24
AA	1.21	66.14	18.53	2.06	0.35	0.22	0.03	0.12	11.33
A	0.06	4.07	68.85	11.72	1.30	0.44	0.09	0.25	13.21
BBB	0.02	0.28	8.42	64.66	7.11	1.64	0.30	0.87	16.70
BB	0.01	0.06	0.51	11.08	47.04	11.58	1.25	3.96	24.51
B	0.00	0.03	0.21	0.78	10.23	41.46	4.67	12.57	30.05
CCC/C	0.00	0.00	0.14	0.61	1.63	16.86	10.54	40.65	29.57

Notes: D = default. NR = not rated—that is, certain corporate issuers were no longer rated by S&P. This could occur for a variety of reasons, including issuers paying off their debt and no longer needing ratings.
Source: S&P Global Ratings, "2017 Annual Global Corporate Default Study and Rating Transitions" (5 April 2018): 53.

- **Credit ratings tend to lag the market's pricing of credit risk**. Bond prices and credit spreads frequently move more quickly because of changes in perceived creditworthiness than rating agencies change their ratings (or even outlooks) up or down. Bond prices and relative valuations can move every day, whereas bond ratings, appropriately, don't change that often. Even over long time periods, however, credit ratings can badly lag changes in bond prices. Exhibit 8 shows the price and Moody's rating of a bond from US automaker Ford Motor Company before, during, and after the global financial crisis in 2008. Note how the bond's price moved down sharply well before Moody's downgraded its credit rating—multiple times—and also how the bond's price began to recover—and kept recovering—well before Moody's upgraded its credit rating on Ford debt.

Exhibit 8 Historical Example: Ford Motor Company Senior Unsecured Debt: Price vs. Moody's Rating 2005–2011



Sources: Data based on Bloomberg Finance LP and Moody's Investors Service, Inc.

Moreover, particularly for certain speculative-grade credits, two bonds with similar ratings may trade at very different valuations. This is partly a result of the fact that credit ratings primarily try to assess the risk of default, whereas for low-quality credits, the market begins focusing more on expected loss (default probability times loss severity). So, bonds from two separate issuers with comparable (high) risk of default but different recovery rates may have similar ratings but trade at significantly different dollar prices.

Thus, bond investors who wait for rating agencies to change their ratings before making buy and sell decisions in their portfolios may be at risk of underperforming other investors who make portfolio decisions in advance of—or not solely based on—rating agency changes.

- **Rating agencies may make mistakes**. The mis-rating of billions of dollars of subprime-backed mortgage securities is one example. Other historical examples include the mis-ratings of US companies Enron and WorldCom and European issuer Parmalat. Like many investors, the rating agencies did not understand the accounting fraud being committed in those companies.
- **Some risks are difficult to capture in credit ratings**. Examples include litigation risk, such as that which can affect tobacco companies, or environmental and business risks faced by chemical companies and utility power plants. This

would also include the impact from natural disasters. Leveraged transactions, such as debt-financed acquisitions and large stock buybacks (share repurchases), are often difficult to anticipate and thus to capture in credit ratings.

As described, there are risks in relying on credit rating agency ratings when investing in bonds. Thus, while the credit rating agencies will almost certainly continue to play a significant role in the bond markets, it is important for investors to perform their own credit analyses and draw their own conclusions regarding the credit risk of a given debt issue or issuer.

EXAMPLE 4

Credit Ratings

1 Using the S&P rating scale, investment-grade bonds carry which of the following ratings?

A AAA to EEE

B BBB– to CCC

C AAA to BBB–

2 Using both Moody's and S&P ratings, which of the following pairs of ratings is considered high yield, also known as "below investment grade," "speculative grade," or "junk"?

A Baa1/BBB–

B B3/CCC+

C Baa3/BB+

3 What is the difference between an issuer rating and an issue rating?

A The issuer rating applies to all of an issuer's bonds, whereas the issue rating considers a bond's seniority ranking.

B The issuer rating is an assessment of an issuer's overall creditworthiness, whereas the issue rating is always higher than the issuer rating.

C The issuer rating is an assessment of an issuer's overall creditworthiness, typically reflected as the senior unsecured rating, whereas the issue rating considers a bond's seniority ranking (e.g., secured or subordinated).

4 Based on the practice of notching by the rating agencies, a subordinated bond from a company with an issuer rating of BB would likely carry what rating?

A B+

B BB

C BBB–

5 The fixed-income portfolio manager you work with asked you why a bond from an issuer you cover didn't rise in price when it was upgraded by Fitch from B+ to BB. Which of the following is the *most likely* explanation?

A Bond prices never react to rating changes.

B The bond doesn't trade often, so the price hasn't adjusted to the rating change yet.

C The market was expecting the rating change, and so it was already "priced in" to the bond.

6 Amalgamated Corp. and Widget Corp. each have bonds outstanding with similar coupons and maturity dates. Both bonds are rated B2, B–, and B by Moody's, S&P, and Fitch, respectively. The bonds, however, trade at very different prices—the Amalgamated bond trades at €89, whereas the Widget bond trades at €62. What is the *most likely* explanation of the price (and yield) difference?

- **A** Widget's credit ratings are lagging the market's assessment of the company's credit deterioration.
- **B** The bonds have similar risks of default (as reflected in the ratings), but the market believes the Amalgamated bond has a higher expected loss in the event of default.
- **C** The bonds have similar risks of default (as reflected in the ratings), but the market believes the Widget bond has a higher expected recovery rate in the event of default.

Solution to 1:

C is correct.

Solution to 2:

B is correct. Note that issuers with such ratings as Baa3/BB+ (answer C) are called "crossovers" because one rating is investment grade (the Moody's rating of Baa3) and the other is high yield (the S&P rating of BB+).

Solution to 3:

C is correct.

Solution to 4:

A is correct. The subordinated bond would have its rating notched lower than the company's BB rating, probably by two notches, reflecting the higher weight given to loss severity for below-investment-grade credits.

Solution to 5:

C is correct. The market was anticipating the rating upgrade and had already priced it in. Bond prices often do react to rating changes, particularly multi-notch ones. Even if bonds don't trade, their prices adjust based on dealer quotations given to bond pricing services.

Solution to 6:

A is correct. Widget's credit ratings are probably lagging behind the market's assessment of its deteriorating creditworthiness. Answers B and C both state the situation backwards. If the market believed that the Amalgamated bond had a higher expected loss given default, then that bond would be trading at a lower, not a higher, price. Similarly, if the market believed that the Widget bond had a higher expected recovery rate in the event of default, then that bond would be trading at a higher, not a lower, price.

TRADITIONAL CREDIT ANALYSIS: CORPORATE DEBT SECURITIES

5

f explain the four Cs (Capacity, Collateral, Covenants, and Character) of traditional credit analysis

The goal of credit analysis is to assess an issuer's ability to satisfy its debt obligations, including bonds and other indebtedness, such as bank loans. These debt obligations are contracts, the terms of which specify the interest rate to be paid, the frequency and timing of payments, the maturity date, and the covenants that describe the permissible and required actions of the borrower. Because corporate bonds are contracts, enforceable by law, credit analysts generally assume an issuer's willingness to pay and concentrate instead on assessing its ability to pay. Thus, the main focus in credit analysis is to understand a company's ability to generate cash flow over the term of its debt obligations. In so doing, analysts must assess both the credit quality of the company and the fundamentals of the industry in which the company operates. Traditional credit analysis considers the sources, predictability, and sustainability of cash generated by a company to service its debt obligations. This section will focus on corporate credit analysis; in particular, it will emphasize non-financial companies. Financial institutions have very different business models and funding profiles from industrial and utility companies.

5.1 Credit Analysis vs. Equity Analysis: Similarities and Differences

The description of credit analysis suggests credit and equity analyses should be very similar, and in many ways, they are. There are motivational differences, however, between equity and fixed-income investors that are an important aspect of credit analysis. Strictly speaking, management works for the shareholders of a company. Its primary objective is to maximize the value of the company for its owners. In contrast, management's legal duty to its creditors—including bondholders—is to meet the terms of the governing contracts. Growth in the value of a corporation from rising profits and cash flow accrues to the shareholders, while the best outcome for bondholders is to receive full, timely payment of interest and repayment of principal when due. Conversely, shareholders are more exposed to the decline in value if a company's earnings and cash flow decline because bondholders have a prior claim on cash flow and assets. But if a company's earnings and cash flow decline to the extent that it can no longer make its debt payments, then bondholders are at risk of loss as well.

In summary, in exchange for a prior claim on cash flow and assets, bondholders do not share in the growth in value of a company (except to the extent that its creditworthiness improves) but have downside risk in the event of default. In contrast, shareholders have theoretically unlimited upside opportunity, but in the event of default, their investment is typically wiped out before the bondholders suffer a loss. This is very similar to the type of payoff patterns seen in financial options. In fact, in recent years, credit risk models have been developed based on the insights of option pricing theory. Although it is beyond the scope of this present introduction to the subject, it is an expanding area of interest to both institutional investors and rating agencies.

Thus, although the analysis is similar in many respects for both equity and credit, equity analysts are interested in the strategies and investments that will increase a company's value and grow earnings per share. They then compare that earnings and growth potential with that of other companies in a given industry. Credit analysts will look more at the downside risk by measuring and assessing the sustainability of a company's cash flow relative to its debt levels and interest expense. Importantly for credit analysts, the balance sheet will show the composition of an issuer's debt—the overall amount, how much is coming due and when, and the distribution by seniority ranking. In general, equity analysts will focus more on income and cash flow statements, whereas credit analysts tend to focus more on the balance sheet and cash flow statements.

5.2 The Four Cs of Credit Analysis: A Useful Framework

Traditionally, many analysts evaluated creditworthiness based on what is often called the "four Cs of credit analysis":

- Capacity
- Collateral
- Covenants
- Character

Capacity refers to the ability of the borrower to make its debt payments on time; this is the focus of this section. **Collateral** refers to the quality and value of the assets supporting the issuer's indebtedness. **Covenants** are the terms and conditions of lending agreements that the issuer must comply with. **Character** refers to the quality of management. Each of these will now be covered in greater detail. Please note that the list of Cs is a convenient way to summarize the important aspects of the analysis; it is not a checklist to be applied mechanically, nor is it always exhaustive.

5.2.1 *Capacity*

Capacity is the ability of a borrower to service its debt. To determine that, credit analysis, in a process similar to equity analysis, starts with industry analysis and then turns to examination of the specific issuer (company analysis).

Industry structure. The Porter framework (Michael E. Porter, "The Five Competitive Forces That Shape Strategy," *Harvard Business Review* 86 [1; 2008]: 78–93) considers the effects of five competitive forces on an industry:

1 **Threat of entry.** Threat of entry depends on the extent of barriers to entry and the expected response from incumbents to new entrants. Industries with high entry barriers tend to be more profitable and have lower credit risk than industries with low entry barriers because incumbents do not need to hold down prices or take other steps to deter new entrants. High entry barriers can take many forms, including high capital investment, such as in aerospace; large, established distribution systems, such as in auto dealerships; patent protection, such as in technology or pharmaceutical industries; or a high degree of regulation, such as in utilities.

2 **Bargaining power of suppliers.** An industry that relies on just a few suppliers tends to be less profitable and to have greater credit risk than an industry that has multiple suppliers. Industries and companies with just a few suppliers have limited negotiating power to keep the suppliers from raising prices, whereas industries that have many suppliers can play them off against each other to keep prices in check.

3 **Bargaining power of customers.** Industries that rely heavily on just a few main customers have greater credit risk because the negotiating power lies with the buyers.

4 **Threat of substitutes.** Industries (and companies) that offer products and services that provide great value to their customers, and for which there are not good or cost-competitive substitutes, typically have strong pricing power, generate substantial cash flows, and represent less credit risk than other industries or companies. Certain (patent-protected) drugs are an example. Over time, however, disruptive technologies and inventions can increase substitution risk. For example, years ago, airplanes began displacing many trains and steamships. Newspapers were considered to have a nearly unassailable market position until television and then the internet became substitutes for how people received news and information. Over time, recorded music has shifted from records to tapes, to compact discs, to mp3s and other forms of digital media.

5 **Rivalry among existing competitors.** Industries with strong rivalry—because of numerous competitors, slow industry growth, or high barriers to exit—tend to have less cash flow predictability and, therefore, higher credit risk than industries with less competition. Regulation can affect the extent of rivalry and competition. For example, regulated utilities typically have a monopoly position in a given market, which results in relatively stable and predictable cash flows.

It is important to consider how companies in an industry generate revenues and earn profits. Is it an industry with high fixed costs and capital investment or one with modest fixed costs? These structures generate revenues and earn profits in very different ways. Two examples of industries with high fixed costs, also referred to as "having high operating leverage," are airlines and hotels. Many of their operating costs are fixed—running a hotel, flying a plane—so they cannot easily cut costs. If an insufficient number of people stay at a hotel or fly in a plane, fixed operating costs may not be covered and losses may result. With higher occupancy of a hotel or plane, revenues are higher, and it is more likely that fixed costs will be covered and profits earned.

Industry fundamentals. After understanding an industry's structure, the next step is to assess its fundamentals, including its sensitivity to macroeconomic factors, its growth prospects, its profitability, and its business need—or lack thereof—for high credit quality. Judgments about these can be made by looking at the following:

- *Cyclical or non-cyclical.* This is a crucial assessment because industries that are cyclical—that is, have greater sensitivity to broader economic performance—have more volatile revenues, margins, and cash flows and thus are inherently riskier than non-cyclical industries. Food producers, retailers, and health care companies are typically considered non-cyclical, whereas auto and steel companies can be very cyclical. Companies in cyclical industries should carry lower levels of debt relative to their ability to generate cash flow over an economic cycle than companies in less-cyclical or non-cyclical industries.
- *Growth prospects.* Although growth is typically a greater focus for equity analysts than for credit analysts, bond investors have an interest in growth as well. Industries that have little or no growth tend to consolidate via mergers and acquisitions. Depending upon how these are financed (e.g., using stock or debt) and the economic benefits (or lack thereof) of the merger, they may or may not be favorable to corporate bond investors. Weaker competitors in slow-growth industries may begin to struggle financially, adversely affecting their creditworthiness.
- *Published industry statistics.* Analysts can get an understanding of an industry's fundamentals and performance by researching statistics that are published by and available from a number of different sources, including the rating agencies, investment banks, industry publications, and frequently, government agencies.

Company fundamentals. Following analysis of an industry's structure and fundamentals, the next step is to assess the fundamentals of the company: the corporate borrower. Analysts should examine the following:

- Competitive position
- Track record/operating history
- Management's strategy and execution
- Ratios and ratio analysis
- When assessing the company fundamentals, analysts should also explore how the fundamentals are impacted by environmental, social, and governance (ESG) factors.

Competitive position. Based on their knowledge of the industry structure and fundamentals, analysts assess a company's competitive position within the industry. What is its market share? How has it changed over time: Is it increasing, decreasing, holding steady? Is it well above (or below) its peers? How does it compare with respect to cost structure? How might it change its competitive position? What sort of financing might that require?

Track record/Operating history. How has the company performed over time? It's useful to go back several years and analyze the company's financial performance, perhaps during times of both economic growth and contraction. What are the trends in revenues, profit margins, and cash flow? Capital expenditures represent what percentage of revenues? What are the trends on the balance sheet—use of debt versus equity? Was this track record developed under the current management team? If not, when did the current management team take over?

Management's strategy and execution. What is management's strategy for the company—to compete and to grow? Does it make sense, and is it plausible? How risky is it, and how differentiated is it from its industry peers? Is it venturing into unrelated businesses? Does the analyst have confidence in management's ability to execute? What is management's track record, both at this company and at previous ones?

Credit analysts also want to know and understand how management's strategy will affect its balance sheet. Does management plan to manage the balance sheet prudently, in a manner that doesn't adversely affect bondholders? Analysts can learn about management's strategy from reading comments, discussion, and analysis that are included with financial statements filed with appropriate regulators, listening to conference calls about earnings or other big announcements (e.g., acquisitions), going to company websites to find earnings releases and copies of slides of presentations at various industry conferences, visiting and speaking with the company, and so on.

EXAMPLE 5

Industry and Company Analysis

1 Given a hotel company, a chemical company, and a food retail company, which is *most likely* to be able to support a high debt load over an economic cycle?

   A The hotel company, because people need a place to stay when they travel.

   B The chemical company, because chemicals are a key input to many products.

  - **C** The food retail company, because such products as food are typically resistant to recessions.

2 Heavily regulated monopoly companies, such as utilities, often carry high debt loads. Which of the following statements about such companies is *most* accurate?

  - **A** Regulators require them to carry high debt loads.
  - **B** They generate strong and stable cash flows, enabling them to support high levels of debt.
  - **C** They are not very profitable and need to borrow heavily to maintain their plant and equipment.

3 XYZ Corp. manufactures a commodity product in a highly competitive industry in which no company has significant market share and where there are low barriers to entry. Which of the following *best* describes XYZ's ability to take on substantial debt?

  - **A** Its ability is very limited because companies in industries with those characteristics generally cannot support high debt loads.
  - **B** Its ability is high because companies in industries with those characteristics generally have high margins and cash flows that can support significant debt.
  - **C** We don't have enough information to answer the question.

Solution to 1:

C is correct. Food retail companies are considered non-cyclical, whereas hotel and chemical companies are more cyclical and thus more vulnerable to economic downturns.

Solution to 2:

B is correct. Because such monopolies' financial returns are generally dictated by the regulators, they generate consistent cash flows and are thus able to support high debt levels.

Solution to 3:

A is correct. Companies in industries with those characteristics typically have low margins and limited cash flow and thus cannot support high debt levels.

---

g calculate and interpret financial ratios used in credit analysis

h evaluate the credit quality of a corporate bond issuer and a bond of that issuer, given key financial ratios of the issuer and the industry

Ratios and ratio analysis. To provide context to the analysis and understanding of a company's fundamentals—based on the industry in which it operates, its competitive position, its strategy and execution—a number of financial measures derived from the company's principal financial statements are examined. Credit analysts calculate a number of ratios to assess the financial health of a company, identify trends over time, and compare companies across an industry to get a sense of relative creditworthiness. Note that typical values of these ratios vary widely from one industry to another because of different industry characteristics previously identified: competitive structure, economic cyclicality, regulation, and so on.

We will categorize the key credit analysis measures into three different groups:

- Profitability and cash flow

- Leverage
- Coverage

Profitability and cash flow measures. It is from profitability and cash flow generation that companies can service their debt. Credit analysts typically look at operating profit margins and operating income to get a sense of a company's underlying profitability and see how it varies over time. Operating income is defined as operating revenues minus operating expenses and is commonly referred to as "earnings before interest and taxes" (EBIT). Credit analysts focus on EBIT because it is useful to determine a company's performance prior to costs arising from its capital structure (i.e., how much debt it carries versus equity). And "before taxes" is used because interest expense is paid before income taxes are calculated.

Several measures of cash flow are used in credit analysis; some are more conservative than others because they make certain adjustments for cash that gets used in managing and maintaining the business or in making payments to shareholders. The cash flow measures and leverage and coverage ratios discussed next are non-IFRS in the sense that they do not have official IFRS definitions; the concepts, names, and definitions given should be viewed as one usage among several possible, in most cases.

- **Earnings before interest, taxes, depreciation, and amortization (EBITDA).** EBITDA is a commonly used measure of cash flow that takes operating income and adds back depreciation and amortization expense because those are non-cash items. This is a somewhat crude measure of cash flow because it excludes certain cash-related expenses of running a business, such as capital expenditures and changes in (non-cash) working capital. Thus, despite its popularity as a cash flow measure, analysts look at other measures in addition to EBITDA.
- **Funds from operations (FFO).** Standard & Poor's defines funds from operations as net income from continuing operations plus depreciation, amortization, deferred income taxes, and other non-cash items.
- **Free cash flow before dividends (FCF before dividends).** This measures excess cash flow generated by the company (excluding non-recurring items) before payments to shareholders or that could be used to pay down debt or pay dividends. It can be calculated as net income (excluding non-recurring items) plus depreciation and amortization minus increase (plus decrease) in non-cash working capital minus capital expenditures. This is, depending upon the treatment of dividends and interest in the cash flow statement, approximated by the cash flow from operating activities minus capital expenditures. Companies that have negative free cash flow before payments to shareholders will be consuming cash they have or will need to rely on additional financing—from banks, bond investors, or equity investors. This obviously represents higher credit risk.
- **Free cash flow after dividends (FCF after dividends).** This measure just takes free cash flow before dividends and subtracts dividend payments. If this number is positive, it represents cash that could be used to pay down debt or build up cash on the balance sheet. Either action may be viewed as deleveraging, which is favorable from a credit risk standpoint. Some credit analysts will calculate net debt by subtracting balance sheet cash from total debt, although they shouldn't assume the cash will be used to pay down debt. Actual debt paid down from free cash flow is a better indicator of deleveraging. Some analysts will also deduct stock buybacks to get the "truest" measure of free cash flow that can be used to de-lever on either a gross or net debt basis; however, others view stock buybacks (share repurchases) as more discretionary and as having less certain timing than dividends, and thus treat those two types of shareholder payments differently when calculating free cash flow.

Leverage ratios. A few measures of leverage are used by credit analysts. The most common are the debt/capital, debt/EBITDA, and measures of funds or cash flows/debt ratios. Note that many analysts adjust a company's reported debt levels for debt-like liabilities, such as underfunded pensions and other retiree benefits, as well as operating leases. When adjusting for leases, analysts will typically add back the imputed interest or rent expense to various cash flow measures.

- **Debt/capital**. Capital is calculated as total debt plus shareholders equity. This ratio shows the percentage of a company's capital base that is financed with debt. A lower percentage of debt indicates lower credit risk. This traditional ratio is generally used for investment-grade corporate issuers. Where goodwill or other intangible assets are significant (and subject to obsolescence, depletion, or impairment), it is often informative to also compute the debt to capital ratio after assuming a write-down of the after-tax value of such assets.
- **Debt/EBITDA.** This ratio is a common leverage measure. Analysts use it on a "snapshot" basis, as well as to look at trends over time and at projections and to compare companies in a given industry. Rating agencies often use it as a trigger for rating actions, and banks reference it in loan covenants. A higher ratio indicates more leverage and thus higher credit risk. Note that this ratio can be very volatile for companies with high cash flow variability, such as those in cyclical industries and with high operating leverage (fixed costs).
- **FFO/debt.** Credit rating agencies often use this leverage ratio. They publish key median and average ratios, such as this one, by rating category so analysts can get a sense of why an issuer is assigned a certain credit rating, as well as where that rating may migrate based on changes to such key ratios as this one. A higher ratio indicates greater ability to pay debt by funds from operations.
- **FCF after dividends/debt**. A higher ratio indicates that a greater amount of debt can be paid off from free cash flow after dividend payments.

Coverage ratios. Coverage ratios measure an issuer's ability to meet—to "cover"—its interest payments. The two most common are the EBITDA/interest expense and EBIT/interest expense ratios.

- **EBITDA/interest expense.** This measurement of interest coverage is a bit more liberal than the one that uses EBIT because it does not subtract out the impact of (non-cash) depreciation and amortization expense. A higher ratio indicates higher credit quality.
- **EBIT/interest expense.** Because EBIT does not include depreciation and amortization, it is considered a more conservative measure of interest coverage. This ratio is now used less frequently than EBITDA/interest expense.

Exhibit 9 is an example of key average credit ratios by rating category for industrial companies over the 12-month period 3Q2017–3Q2018, as calculated by Bloomberg Barclays Indices, using public company data. Note only a few AAA-rated corporations remain, so the small sample size can skew the average ratios of a few key credit metrics. That said, it should be clear that, overall, higher-rated issuers have stronger credit metrics.

Exhibit 9 Industrial Comparative Ratio Analysis

Credit Rating	EBITDA Margin (%)	Return on Capital (%)	EBIT Interest Coverage (x)	EBITDA Interest Coverage (x)	FFO/Debt (%)	Free Operations Cash Flow/Debt (%)	Debt/ EBITDA (x)	Debt/Debt plus Equity (%)
Aaa								
US	66.4	6.5	4.2	21.3	51.9	43.5	–0.2	43.3
Aa								
US	21.9	10.8	15.4	45.0	109.9	58.1	1.2	50.6
A								
US	26.0	13.5	13.3	18.9	49.1	31.8	1.8	51.2
Baa								
US	23.9	11.5	7.2	NA	40.7	20.3	3.9	49.4
Ba								
US	21.7	3.5	4.6	NA	27.7	11.0	4.1	64.0
B								
US	21.2	3.7	2.5	NA	20.3	1.8	5.2	69.3
Caa								
US	16.0	0.2	–0.6	1.3	10.0	–6.6	9.3	95.3

Note: As of 19 December 2018.
Source: Bloomberg Barclays Indices.

Comments on issuer liquidity. An issuer's access to liquidity is also an important consideration in credit analysis. Companies with high liquidity represent lower credit risk than those with weak liquidity, other factors being equal. The global financial crisis of 2008–2009 showed that access to liquidity via the debt and equity markets should not be taken for granted, particularly for companies that do not have strong balance sheets or steady operating cash flow.

When assessing an issuer's liquidity, credit analysts tend to look at the following:

- **Cash on the balance sheet.** Cash holdings provide the greatest assurance of having sufficient liquidity to make promised payments.
- **Net working capital.** The big US automakers used to have enormous negative working capital, despite having high levels of cash on the balance sheet. This proved disastrous when the global financial crisis hit in 2008 and the economy contracted sharply. Auto sales—and thus revenues—fell, the auto companies cut production, and working capital consumed billions of dollars in cash as accounts payable came due when the companies most needed liquidity.
- **Operating cash flow.** Analysts will project this figure out a few years and consider the risk that it may be lower than expected.
- **Committed bank lines.** Committed but untapped lines of credit provide contingent liquidity in the event that the company is unable to tap other, potentially cheaper, financing in the public debt markets.
- **Debt coming due and committed capital expenditures in the next one to two years.** Analysts will compare the sources of liquidity with the amount of debt coming due as well as with committed capital expenditures to ensure that companies can repay their debt and still invest in the business if the capital markets are somehow not available.

As will be discussed in more detail in the section on special considerations for high-yield credits, issuer liquidity is a bigger consideration for high-yield companies than for investment-grade companies.

EXAMPLE 6

Mallinckrodt PLC (Mallinckrodt) is an Ireland-incorporated specialty pharmaceutical company. As a credit analyst, you have been asked to assess its creditworthiness—on its own, compared to a competitor in its overall industry, and compared with a similarly rated company in a different industry. Using the financial statements provided in Exhibits 10 through 12 for the three years ending 31 December 2015, 2016, and 2017, address the following:

1. Calculate Mallinckrodt's operating profit margin, EBITDA, and free cash flow after dividends. Comment on what these measures indicate about Mallinckrodt's profitability and cash flow.
2. Determine Mallinckrodt's leverage ratios: debt/EBITDA, debt/capital, free cash flow after dividends/debt. Comment on what these leverage ratios indicate about Mallinckrodt's creditworthiness.
3. Calculate Mallinckrodt's interest coverage using both EBIT and EBITDA. Comment on what these coverage ratios indicate about Mallinckrodt's creditworthiness.
4. Using the credit ratios provided in Exhibit 11 on Johnson & Johnson, compare the creditworthiness of Mallinckrodt relative to Johnson & Johnson.
5. Compare the Exhibit 12 credit ratios of Luxembourg-based ArcelorMittal, one of the world's largest global steelmakers, with those of Mallinckrodt. Comment on the volatility of the credit ratios of the two companies. Which company looks to be more cyclical? What industry factors might explain some of the differences? In comparing the creditworthiness of these two companies, what other factors might be considered to offset greater volatility of credit ratios?

Exhibit 10a Mallinckrodt PLC Financial Statements

Consolidated Statements of Operations

(Dollars in millions, except per share amounts)	Year End 30 Sept.*	Years Ended 31 Dec.	
	2015	2016	2017
Net revenues	**2,923.1**	**3,399.5**	**3,221.6**
Operating expenses:			
Cost of sales	1,300.2	1,549.6	1,565.3
Research and development	203.3	267.0	277.3
Selling, general, and administrative expenses	1,023.8	1,070.3	920.9
Restructuring charges, net	45.0	33.0	31.2
Non-restructuring impairment charges	—	231.2	63.7
Gain on divestiture and license	(3.0)	—	(56.9)
Total operating expenses	**2,569.3**	**3,151.1**	**2,801.5**

(continued)

Exhibit 10a (Continued)

Consolidated Statements of Operations

(Dollars in millions, except per share amounts)	Year End 30 Sept.*	Years Ended 31 Dec.	
	2015	2016	2017
Operating income	**353.8**	**248.4**	**420.1**
Other (expense) income:			
Interest expense	(255.6)	(378.1)	(369.1)
Interest income	1.0	1.6	4.6
Other income (expense), net	8.1	(3.5)	6.0
Total other (expense) income, net	**(246.5)**	**(380.0)**	**(358.5)**
Income before income taxes and non-controlling interest	**107.3**	**(131.6)**	**61.6**
Provision (Benefit) for income taxes	(129.3)	(340.0)	(1,709.6)
Net income	**236.6**	**208.4**	**1,771.2**
Income from discontinued operations, net of income taxes	88.1	71.0	363.2
Net income attributable to common shareholders	**324.7**	**279.4**	**2,134.4**

*Mallinckrodt changed their fiscal year end from 30 September to 31 December.
Source: Company filings, Loomis, Sayles & Company.

Exhibit 10b Mallinckrodt PLC Financial Statements

Consolidated Balance Sheets

(Dollars in millions)	Year End 30 Sept.*	Years Ended 31 Dec.	
	2015	2016	2017
ASSETS			
Current assets:			
Cash and cash equivalents	365.9	342.0	1,260.9
Accounts receivable	489.6	431.0	445.8
Inventories	262.1	350.7	340.4
Deferred income taxes	139.2	—	—
Prepaid expenses and other current assets	194.4	131.9	84.1
Notes receivable	—	—	154.0
Current assets held for sale	394.9	310.9	—

Exhibit 10b (Continued)

Consolidated Balance Sheets

(Dollars in millions)	Year End 30 Sept.* 2015	Years Ended 31 Dec. 2016	2017
Total current assets	**1,846.1**	**1,566.5**	**2,285.2**
Property, plant, and equipment, net	793.0	881.5	966.8
Goodwill	3,649.4	3,498.1	3,482.7
Intangible assets, net	9,666.3	9,000.5	8,375.0
Other assets	225.7	259.7	171.2
Long-term assets held for sale	223.6	—	—
Total assets	**16,404.1**	**15,206.3**	**15,280.9**
LIABILITIES AND EQUITY			
Current liabilities:			
Current maturities of long-term debt	22.0	271.2	313.7
Accounts payable	116.8	112.1	113.3
Accrued payroll and payroll-related costs	95.0	76.1	98.5
Accrued interest	80.2	68.7	57.0
Income taxes payable	—	101.7	15.8
Accrued and other current liabilities	486.1	557.1	452.1
Current liabilities held for sale	129.3	120.3	—
Total current liabilities	**929.4**	**1,307.2**	**1,050.4**
Long-term debt	6,474.3	5,880.8	6,420.9
Pension and post-retirement benefits	114.2	136.4	67.1
Environmental liabilities	73.3	73.0	73.2
Deferred income taxes	3,117.5	2,398.1	689.0
Other income tax liabilities	121.3	70.4	94.1
Other liabilities	209.0	356.1	364.2
Long-term liabilities held for sale	53.9	—	—
Total liabilities	**11,092.9**	**10,222.0**	**8,758.9**
Shareholders' equity:			
Ordinary shares	23.5	23.6	18.4
Ordinary shares held in treasury at cost	(109.7)	(919.8)	(1,564.7)
Additional paid-in capital	5,357.6	5,424.0	5,492.6
Retained earnings	38.9	529.0	2,588.6
Accumulated other comprehensive income	0.9	(72.5)	(12.9)

(continued)

Exhibit 10b (Continued)

Consolidated Balance Sheets

(Dollars in millions)	Year End 30 Sept.* 2015	Years Ended 31 Dec. 2016	2017
Total shareholders' equity	5,311.2	4,984.3	6,522.0
Total liabilities and shareholders' equity	16,404.1	15,206.3	15,280.9

*Mallinckrodt changed their fiscal year end from 30 September to 31 December.
Source: Company filings, Loomis, Sayles & Company

Exhibit 10c Mallinckrodt PLC Financial Statements

Consolidated Statements of Cash Flow

(Dollars in millions)	Year End 30 Sept.* 2015	Years Ended 31 Dec. 2016	2017
Cash Flows from Operating Activities:			
Net income (loss)	324.7	279.4	2,134.4
Depreciation and amortization	672.5	831.7	808.3
Share-based compensation	117.0	45.4	59.2
Deferred income taxes	(191.6)	(528.3)	−1744.1
Non-cash impairment charges	—	231.2	63.7
Inventory provisions	—	8.5	34.1
Gain on disposal of discontinued operations	—	1.7	−418.1
Other non-cash items	(25.5)	45.5	−21.4
Change in working capital	33.4	153.7	−188.8
Net cash from operating activities	**930.5**	**1,068.8**	**727.3**
Cash Flows from Investing Activities:			
Capital expenditures	(148.0)	(199.1)	−186.1
Acquisitions and intangibles, net of cash acquired	(2,154.7)	(247.2)	−76.3
Proceeds from divestitures, net of cash	—	3.0	576.9
Other	3.0	(4.9)	3.9
Net cash from investing activities	**(2,299.7)**	**(448.2)**	**318.4**
Cash Flows from Financing Activities:			
Issuance of external debt	3,010.0	226.3	1,465
Repayment of external debt and capital leases	(1,848.4)	(525.7)	−917.2
Debt financing costs	(39.9)	—	−12.7
Proceeds from exercise of share options	34.4	10.8	4.1

Exhibit 10c (Continued)

Consolidated Statements of Cash Flow

(Dollars in millions)	Year End 30 Sept.* 2015	Years Ended 31 Dec. 2016	2017
Repurchase of shares	(92.2)	(536.3)	−651.7
Other	(28.1)	(21.8)	−17.7
Net cash from financing activities	**1,035.8**	**(846.7)**	**−130.2**
Effect of currency rate changes on cash	(11.6)	(1.2)	**2.5**
Net increase (decrease) in cash and cash equivalents	(345.0)	(227.3)	918
Cash and cash equivalents at beginning of period	777.6	588.4	361.1
Cash and cash equivalents at end of period	432.6	361.1	1,279.1

*Mallinckrodt changed their fiscal year end from 30 September to 31 December.
Source: Company filings, Loomis, Sayles & Company.

Exhibit 10d Mallinckrodt PLC Credit Ratios

	2015	2016	2017
Operating margin	12.1%	7.3%	13.0%
Debt/EBITDA	6.3x	5.7x	5.5x
EBITDA/Interest	4.0x	2.9x	3.3x
FCF/Debt	12.0%	14.1%	8.0%
Debt/Capital	55.0%	55.2%	50.8%

Source: Company filings, Loomis, Sayles & Company.

Exhibit 11 Johnson & Johnson's Credit Ratios

	2015	2016	2017
Operating profit margin	26.2%	29.5%	25.8%
Debt/EBITDA	0.9x	1.1x	1.4x
EBITDA/Interest	40.1x	34.4x	27.1x
FCF after dividends/Debt	81.1%	57.3%	51.4%
Debt/Capital	22.0%	28.1%	36.5%

Source: Company filings, Loomis, Sayles & Company.

Exhibit 12 ArcelorMittal Credit Ratios

	2015	2016	2017
Operating profit margin	0.3%	5.5%	7.7%
Debt/EBITDA	5.8x	2.3x	1.6x
EBITDA/Interest	2.5x	5.0x	9.2x
FCF after dividends/Debt	–2.8%	1.9%	13.5%
Debt/Capital	41.8%	29.7%	24.0%

Source: Company filings, Loomis, Sayles & Company.

Solutions:

1 Operating profit margin (%) = Operating income/Revenue

2015: 353.8/2,923.1 = 0.121 or 12.1%

2016: 248.4/3,399.5 = 0.073 or 7.3%

2017: 420.1/3,221.6 = 0.130 or 13.0%

EBITDA = Operating income + Depreciation and Amortization

2015: 353.8 + 672.5 = 1,026.3

2016: 248.4+ 831.7 = 1,080.1

2017: 420.1 + 808.3= 1,228.4

FCF after dividends = Cash flow from operations – Capital expenditures – Dividends

2015: 930.5 – 148 – 0 = 782.5

2016: 1,068.8 – 199.1 – 0 = 869.7

2017: 727.3 – 186.1 – 0 = 541.2

Operating profit margin decreased from 2015 to 2016 but increased from 2016 to 2017. Conversely, FCF after dividends increased from 2015 to 2016 but decreased from 2016 to 2017. EBITDA increased from 2015 to 2017. From 2015 to 2016, sales increased by 16.3% and operating expenses increased by 22.6%. As a result, operating profit margin decreased even though EBITDA and FCF increased. However, from 2016 to 2017, sales decreased by 5.2% and operating expenses decreased by 11.1%. As a result, operating profit margin and EBITDA increased, while FCF after dividends decreased.

2 Debt/EBITDA

Total debt = Short-term debt and Current portion of long-term debt + Long-term debt

2015: Debt: 22.0 + 6,474.3 = 6,496.3

Debt/EBITDA: 6,496.3/1,026.3 = 6.3x

2016: Debt: 271.2 + 5,880.8 = 6,152.0

Debt/EBITDA: 6,152.0/1,080.1 = 5.7x

2017: Debt: 313.7 + 6,420.9 = 6,734.6

Debt/EBITDA: 6,734.6/1,228. 4 = 5.5x

Debt/Capital (%)

Capital = Debt + Equity

2015: Capital: 6,496.3 + 5,311.2 = 11,807.5

Debt/Capital: 6,496.3/11,807.5 = 55.0%

2016: Capital: 6,152.0 + 4,984.3 = 11,136.3

Debt/Capital: 6,152.0/11,136.3 = 55.2%

2017: Capital: 6,734.6 + 6,522.0 = 13,256.6

Debt/Capital: 6,734.6/13,256.6 = 50.8%

FCF after dividends/Debt (%)

2015: 782.5/6,496.3 = 12.0%

2016: 869.7/6,152 = 14.1%

2017: 541.2/6,734.6 = 8.0%

Although debt/EBITDA and debt/capital improved between 2015 and 2017, the "FCF after dividends/Debt" deteriorated significantly as cash flow from operations declined as a result of the loss taken on disposal of discontinued operations. Given that the loss is most likely a non-recurring event, Mallinckrodt's creditworthiness likely improved over the 2015 to 2017 period.

3 EBIT/Interest expense

2015: 353.8/255.6 = 1.4x

2016: 248.4/378.1 = 0.7x

2017: 420.1/369.1 = 1.1x

EBITDA/Interest expense

2015: 1,026.3/255.6 = 4.0x

2016: 1,080.1/378.1= 2.9x

2017: 1,228.4/369.1 = 3.3x

Based on these coverage ratios, Mallinckrodt's creditworthiness declined from 2015 to 2016 and then showed modest improvement in 2017. The 2017 coverage ratios are still weaker than the 2015 coverage ratios, indicating that growth in EBIT and EBITDA are not keeping pace with the rising interest expense.

4 Johnson & Johnson (J&J) has a higher operating profit margin, better leverage ratios—lower debt/EBITDA, higher FCF after dividends/debt over the three years, lower debt/capital, and better interest coverage as measured by EBITDA/interest. Collectively, those ratios suggest J&J has higher credit quality than Mallinckrodt.

5 Mallinckrodt has both a higher and a less volatile operating profit margin than ArcelorMittal (Arcelor). However, while Mallinckrodt's leverage ratios have been deteriorating, Arcelor's have been improving. Based on the volatility of its cash flow and operating profit margin, Arcelor appears to be a much more cyclical credit. However, with its meaningfully lower debt levels, one could expect Arcelor to have a higher credit rating.

A steelmaker likely has a significant amount of long-term assets financed by debt. It is a highly competitive industry with little ability to distinguish products from other competitors. To mitigate the impact of its more volatile credit ratios, Arcelor might maintain high levels of liquidity. Its size and global diversity may also be a "plus." Given its size, it may be able to negotiate favorable supplier and customer contracts and keep costs down through economies of scale.

5.2.2 *Collateral*

Collateral, or asset value, analysis is typically emphasized more with lower credit quality companies. As discussed earlier, credit analysts focus primarily on probability of default, which is mostly about an issuer's ability to generate sufficient cash flow to support its debt payments, as well as its ability to refinance maturing debt. Only when the default probability rises to a sufficient level do analysts typically consider asset or collateral value in the context of loss severity in the event of default.

Analysts do think about the value and quality of a company's assets; however, these are difficult to observe directly. Factors to consider include the nature and amount of intangible assets on the balance sheet. Some assets, such as patents, are clearly valuable and can be sold if necessary to cover liabilities. Goodwill, on the other hand, is not considered a high-quality asset. In fact, sustained weak financial performance most likely implies that a company's goodwill will be written down, reinforcing its poor quality. Another factor to consider is the amount of depreciation an issuer takes relative to its capital expenditures: Low capital expenditures relative to depreciation expense could imply that management is insufficiently investing in its business, which will lead to lower-quality assets, potentially reduced future operating cash flow, and higher loss severity in the event of default.

A market-based signal that credit analysts use to impute the quality of a publicly traded company's assets, and its ability to support its debt, is equity market capitalization. For instance, a company whose stock trades below book value may have lower-quality assets than is suggested by the amount reported on the balance sheet.

As economies become more service- and knowledge-based and those types of companies issue debt, it's important to understand that these issuers rely more on human and intellectual capital than on "hard assets." In generating profits and cash flow, these companies are not as asset intensive. One example would be software companies. Another example would be investment management firms. Human- and intellectual-capital-based companies may generate a lot of cash flow, but their collateral value is questionable unless there are patents and other types of intellectual property and "intangible capital" that may not appear directly on the balance sheet but could be valuable in the event of financial distress or default.

Regardless of the nature of the business, the key point of collateral analysis is to assess the value of the assets relative to the issuer's level—and seniority ranking—of debt.

5.2.3 *Covenants*

Covenants are meant to protect creditors while also giving management sufficient flexibility to operate its business on behalf of and for the benefit of the shareholders. They are integral to credit agreements, whether they are bonds or bank loans, and they spell out what the issuer's management is (1) obligated to do and (2) limited in doing. The former are called "affirmative covenants," whereas the latter are called "negative" or "restrictive covenants." Obligations would include such duties as making interest and principal payments and filing audited financial statements on a timely basis. Covenants might also require a company to redeem debt in the event of the company being acquired, the change of control covenant, or to keep the ratio of debt to EBITDA below some prescribed amount. The limitations might include a cap on the amount of cash that can be paid out to shareholders relative to earnings, or perhaps on the amount of additional secured debt that can be issued. Covenant violations are a breach of contract and can be considered default events unless they are cured in a short time or a waiver is granted.

For corporate bonds, covenants are described in the bond **prospectus**, the document that is part of a new bond issue. The prospectus describes the terms of the bond issue, as well as supporting financial statements, to help investors perform their analyses and make investment decisions as to whether or not to submit orders to buy the new bonds. Actually, the **trust deed** or **bond indenture** is the governing legal credit agreement and is typically incorporated by reference in the prospectus.

Covenants are an important but underappreciated part of credit analysis. Strong covenants protect bond investors from the possibility of management taking actions that would hurt an issuer's creditworthiness. For example, without appropriate covenants, management might pay large dividends, undertake stock buybacks well in excess of free cash flow, sell the company in a leveraged buyout, or take on a lot of secured debt that structurally subordinates unsecured bondholders. All of these actions would enrich shareholders at the expense of bondholders. Recall that management works for the shareholders and that bonds are contracts, with management's only real obligation to creditors being to uphold the terms of the contract. The inclusion of covenants in the contract is intended to protect bondholders.

The bond-buying investor base is very large and diverse, particularly for investment-grade debt. It includes institutional investors, such as insurance companies, investment management firms, pension funds, mutual funds, hedge funds, sovereign wealth funds, and so on. Although there are some very large institutional investors, the buyer base is fragmented and does not—and legally cannot—act as a syndicate. Thus, bondholders are generally not able to negotiate strong covenants on most new bond issues. At the same time, issuers expect that borrowing from capital markets will carry lighter convenants than borrowing from banks. During weak economic or market conditions, however, covenants on new bond issues tend to be stronger because investors require additional incentive to lend money while seeking more protection. A few organized institutional investor groups are focused on strengthening covenants: the Credit Roundtable (thecreditroundtable.org) in the United States and the European Model Covenant Initiative in the United Kingdom.

Covenant language is often very technical and written in "legalese," so it can be helpful to have an in-house person with a legal background to review and interpret the specific covenant terms and wording. One might also use a third-party service specializing in covenant analysis, such as Covenant Review (www.covenantreview.com).

We will go into more detail on specific covenants in the section on special considerations for high-yield bonds.

5.2.4 *Character*

The character of a corporate borrower can be difficult to observe. The analysis of character as a factor in credit analysis dates to when loans were made to companies owned by individuals. Most corporate bond issuers are now publicly owned by shareholders or privately owned by pools of capital, such as private equity firms. Management often has little ownership in a corporation, so analysis and assessment of character is different than it would be for owner-managed firms. Credit analysts can make judgments about management's character in the following ways:

- An assessment of the soundness of management's strategy.
- Management's track record in executing past strategies, particularly if they led to bankruptcy or restructuring. A company run by executives whose prior positions/ventures resulted in significant distress might still be able to borrow in the debt markets, but it would likely have to borrow on a secured basis and/or pay a higher interest rate.
- Use of aggressive accounting policies and/or tax strategies. Examples might include using a significant amount of off-balance-sheet financing, capitalizing versus immediately expensing items, recognizing revenue prematurely, and/or frequently changing auditors. These are potential warning flags to other behaviors or actions that may adversely impact an issuer's creditworthiness.
- Any history of fraud or malfeasance—a major warning flag to credit analysts.

- Previous poor treatment of bondholders—for example, management actions that resulted in major credit rating downgrades. These actions might include a debt-financed acquisition, a large special dividend to shareholders, or a major debt-financed stock buyback program.

EXAMPLE 7

The Four Cs

1 Which of the following would not be a bond covenant?
   - A The issuer must file financial statements with the bond trustee on a timely basis.
   - B The company can buy back as much stock as it likes.
   - C If the company offers security to any creditors, it must offer security to this bond issue.

2 Why should credit analysts be concerned if a company's stock trades below book value?
   - A It means the company is probably going bankrupt.
   - B It means the company will probably incur lots of debt to buy back its undervalued stock.
   - C It's a signal that the company's asset value on its balance sheet may be impaired and have to be written down, suggesting less collateral protection for creditors.

3 If management is of questionable character, how can investors incorporate this assessment into their credit analysis and investment decisions?
   - A They can choose not to invest based on the increased credit risk.
   - B They can insist on getting collateral (security) and/or demand a higher return.
   - C They can choose not to invest or insist on additional security and/or higher return.

Solution to 1:

B is correct. Covenants describe what the borrower is (1) obligated to do or (2) limited in doing. It's the absence of covenants that would permit a company to buy back as much stock as it likes. A requirement that the company offer security to this bond issue if it offers security to other creditors (answer C) is referred to as a "negative pledge."

Solution to 2:

C is correct.

Solution to 3:

C is correct. Investors can always say no if they are not comfortable with the credit risk presented by a bond or issuer. They may also decide to lend to a borrower with questionable character only on a secured basis and/or demand a higher return for the perceived higher risk.

CREDIT RISK VS. RETURN: YIELDS AND SPREADS

6

i. describe macroeconomic, market, and issuer-specific factors that influence the level and volatility of yield spreads

> *The material in this section applies to all bonds subject to credit risk. For simplicity, in what follows all such bonds are sometimes referred to as "corporate" bonds.*

As in other types of investing, taking more risk in credit offers higher potential return but with more volatility and less certainty of earning that return. Using credit ratings as a proxy for risk, Exhibit 13 shows the composite yield-to-maturity for bonds of all maturities within each rating category in the US and European bond markets according to Bloomberg Barclays, one of the largest providers of fixed-income market indexes. For non-investment-grade bonds, the yield-to-call (YTC) or yield-to-worst (YTW) is reported as many contain optionality.

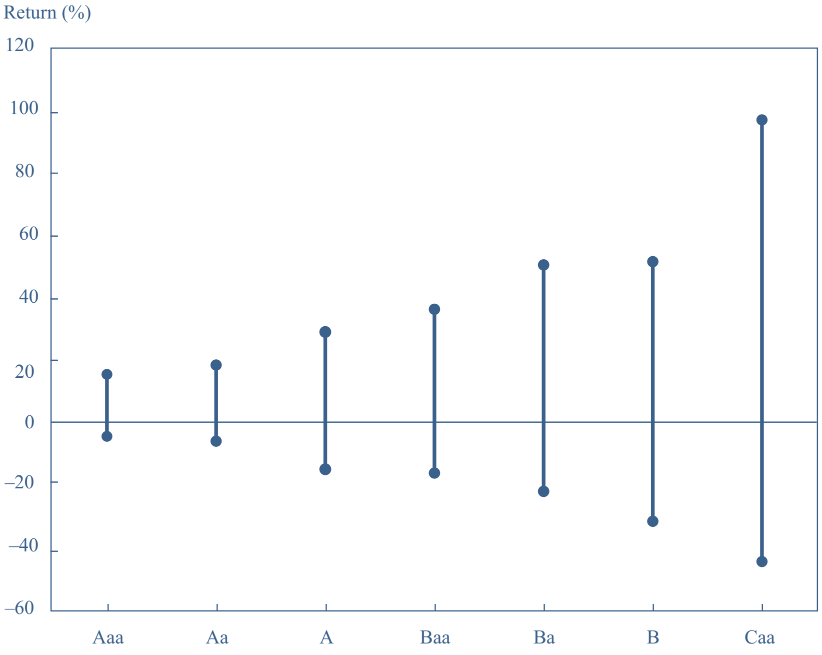
Exhibit 13 Corporate Yields by Rating Category (%)

Bloomberg Barclays Indices	Investment Grade				Non-Investment Grade			
	AAA	AA	A	BBB	BB	B	CCC	CC–D
US	3.63	3.52	3.86	4.35	5.14	6.23	8.87	19.51
Pan European*	1.25	0.76	1.18	1.67	2.92	5.63	8.78	54.95

*Pan European yields may be "artificially low" due to the ECB's extraordinary corporate bond Quantitative Easing (QE) Program.
Note: Data as of 30 September 2018.
Source: Bloomberg Barclays Indices.

Note that the lower the credit quality, the higher the quoted yield. The realized yield, or return, will almost always be different because of changes in interest rates, re-investment of coupons, holding period changes, and the credit-related risks discussed earlier. In addition to the absolute returns, the volatility of returns will also vary by rating. Trailing 12-month returns by credit rating category and the volatility (standard deviation) of those returns are shown in Exhibit 14.

Exhibit 14 US Credit Trailing 12-Month Returns by Rating Category, 31 December 1996–30 September 2018



	Aaa	Aa	A	Baa	Ba	B	Caa
Avg. Return:	4.82%	5.60%	5.78%	6.53%	7.98%	6.79%	8.26%
Max. Return:	16.08%	19.30%	29.39%	36.43%	50.80%	52.10%	97.47%
Min. Return:	–4.17%	–5.47%	–14.98%	–15.70%	–22.12%	–31.47%	–44.35%
Avg. Volatility:	3.43%	4.14%	4.76%	5.15%	5.75%	7.13%	11.00%

Sources: Bloomberg Barclays Indices and Loomis, Sayles & Company.

As shown in the exhibit, the higher the credit risk, the greater the return potential and the higher the volatility of that return. This pattern is consistent with other types of investing that involve risk and return (although average returns on single-B rated bonds appear anomalous in this example).

For extremely liquid bonds that are deemed to have virtually no default risk (e.g., German government bonds, or *Bunds*), the yield for a given maturity is a function of real interest rates plus an expected inflation rate premium. The yield on corporate bonds will include an additional risk premium that provides the investor with compensation for the credit and liquidity risks and possibly the tax impact of holding a specific bond. Changes in any of these components will alter the yield, price, and return on the bond. In general, however, it is not possible to directly observe the market's assessment of the components separately—analysts can only observe the total yield spread.

Spreads on all corporate bonds can be affected by a number of factors, with lower-quality issuers typically experiencing greater spread volatility. These factors, which are frequently linked, include the following:

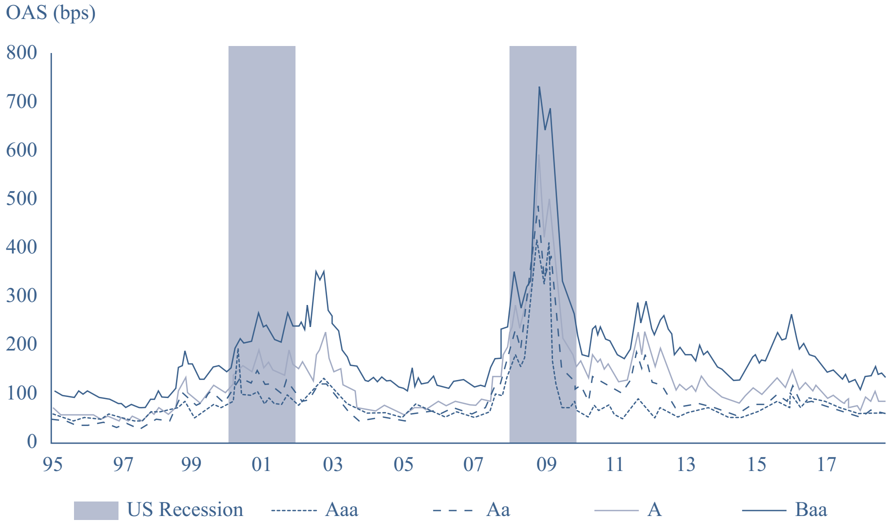
- **Credit cycle.** As the credit cycle improves, credit spreads will narrow. Conversely, a deteriorating credit cycle will cause credit spreads to widen. Spreads are tightest at or near the top of the credit cycle, when financial markets believe risk is low; they are widest at or near the bottom of the credit cycle, when financial markets believe risk is high.

- **Broader economic conditions.** Not surprisingly, weakening economic conditions will push investors to desire a greater risk premium and drive overall credit spreads wider. Conversely, a strengthening economy will cause credit spreads to narrow because investors anticipate credit measures will improve due to rising corporate cash flow, thus reducing the risk of default. In a steady, low-volatility environment, credit spreads will typically also narrow as investors tend to "reach for yield."
- **Funding availability in financial sector.** Bonds trade primarily over the counter, so investors need broker–dealers to commit capital for market-making purposes. Episodes of financial and regulatory stress have the potential to greatly reduce the total capital available for making markets (to facilitate trading) and the willingness to buy/sell credit-risky bonds. Future regulatory reform may well lead to persistent or even permanent reductions in broker-provided capital. Funding stresses would naturally translate into wider spreads.
- **General market supply and demand.** In periods of heavy new issue supply, credit spreads will widen if there is insufficient demand. In periods of high demand for bonds, spreads will move tighter.
- **Financial performance of the issuer.** Corporate bond spreads will be impacted by earnings releases, news, and other developments associated with the issuer. Earlier we explained the "four Cs of credit analysis": capacity, collateral, covenants, and character. Announcement and disclosures by the issuer will impact investors' view of the issuer's financial performance and ability to service and repay its debt. Good news increases the attractiveness of buying and holding bonds issued by the corporation, which raises bond prices and narrows spreads; bad news will have the opposite effect.

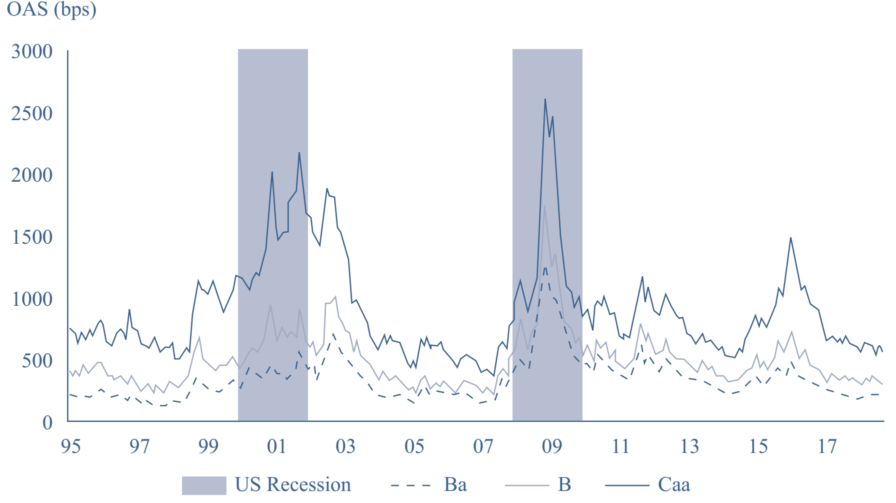
A number of these factors played a role during the global financial crisis of 2008–2009, causing spreads to widen dramatically, as shown in Exhibit 15, before narrowing sharply as governments intervened and markets stabilized. This is shown in two panels—one for investment grade, another for high yield—because of the much greater spread volatility in high-yield bonds, particularly CCC rated credits. This spread volatility is reflected in the different spread ranges on the y-axes. OAS is option-adjusted spread, which incorporates the value of the embedded call option in certain corporate bonds that issuers have the right to exercise before maturity.

Exhibit 15 US Investment-Grade and High-Yield Corporate Spreads

A. Investment-Grade Corporate Spreads



B. High-Yield Corporate Spreads



Sources: Bloomberg Barclays Indices and Loomis, Sayles & Company.

EXAMPLE 8

Yield Spreads

1 Which bonds are likely to exhibit the greatest spread volatility?

- **A** Bonds from issuers rated AA
- **B** Bonds from issuers rated BB
- **C** Bonds from issuers rated A

2 If investors become increasingly worried about the economy—say, as shown by declining stock prices—what is the *most likely* impact on credit spreads?

A There will be no change to credit spreads. They aren't affected by equity markets.

B Narrower spreads will occur. Investors will move out of equities into debt securities.

C Wider spreads will occur. Investors are concerned about weaker creditworthiness.

Solution to 1:

B is correct. Lower-quality bonds exhibit greater spread volatility than higher-quality bonds. All of the factors that affect spreads—the credit cycle, economic conditions, financial performance, market-making capacity, and supply/demand conditions—will tend to have a greater impact on the pricing of lower-quality credits.

Solution to 2:

C is correct. Investors will require higher yields as compensation for the credit deterioration—including losses—that is likely to occur in a weakening economy.

6.1 Credit Risk vs. Return: The Price Impact of Spread Changes

We have discussed how yield spreads on credit-risky debt obligations, such as corporate bonds, can fluctuate based on a number of factors, including changes in the market's view of issuer-specific or idiosyncratic risk. The next question to consider is how these spread changes affect the price of and return on these bonds.

Although bond investors do concern themselves with default risks, recall that the probability of default for higher-quality bonds is typically very low: For investment-grade bonds, annual defaults are nearly always well below 1% (recall Exhibit 6). On the other hand, default rates can be very high for lower-quality issuers, although they can vary widely depending upon the credit cycle, among other things. What most investors in investment-grade debt focus on more than default risk is spread risk—that is, the effect on prices and returns from changes in spreads.

The price impact from spread changes is driven by two main factors: (1) the modified duration (price sensitivity with respect to changes in interest rates) of the bond and (2) the magnitude of the spread change. The effect on return to the bondholder depends on the holding period used for calculating the return.

The simplest example is that of a small, instantaneous change in the yield spread. In this case, the price impact (i.e., the percentage change in price, including accrued interest) can be approximated by

$$\%\Delta PV^{\text{Full}} \approx -\text{AnnModDur} \times \Delta\text{Spread},$$

where AnnModDur is the annual modified duration. The negative sign in this equation reflects the fact that because bond prices and yields move in opposite directions, lower spreads have a positive impact on bond prices and thus returns, whereas higher spreads have a negative impact on bond returns. Note that if the spread change is expressed in basis points, then the price impact will also be in basis points, whereas if the spread change is expressed as a decimal, the price impact will also be expressed as a decimal.

For larger spread changes (and thus larger yield changes), the impact of convexity needs to be incorporated into the approximation:

$$\%\Delta PV^{\text{Full}} \approx -(\text{AnnModDur} \times \Delta\text{Spread}) + \tfrac{1}{2}\text{AnnConvexity} \times (\Delta\text{Spread})^2.$$

In this case, one must be careful to ensure that convexity (denoted by AnnConvexity) is appropriately scaled to be consistent with the way the spread change is expressed. In general, for bonds without embedded options, one can scale convexity so that it has the same order of magnitude as the duration squared and then express the spread change as a decimal. For example, for a bond with duration of 5.0 and reported convexity of 0.235, one would re-scale convexity to 23.5 before applying the formula. For a 1% (i.e., 100 bps) increase in spread, the result would be

$$\%\Delta PV^{Full} = (-5.0 \times 0.01) + ½ \times 23.5 \times (0.01)^2 = -0.048825 \text{ or } -4.8825\%.$$

The price impact of instantaneous spread changes is illustrated in Exhibit 16 using two bonds from British Telecom, the UK telecommunications company. The bonds, denominated in British pounds, are priced to provide a certain spread over British government bonds (gilts) of a similar maturity. From the starting spread, in increments of 25 bps and for both wider and narrower spreads, the new price and actual return for each spread change are calculated. In addition, the exhibit shows the approximate returns with and without the convexity term. As can be seen, the approximation using only duration is reasonably accurate for small spread changes; for larger changes, however, the convexity term generally provides a meaningful improvement.

Exhibit 16 Impact of Duration on Price for a Given Change in Spread

Issuer: British Telecom, 5.75%, 12/07/2028

Price: £122.978	Modified Duration: 7.838	Spread to Gilt Curve: 150.7 b.p.
Accrued interest: 0.958	Convexity: 77.2	YTM (conv): 3.16

	Scenarios								
Spread Δ (b.p.)	−100	−75	−50	−25	0	25	50	75	100
Spread (b.p.)	50.7	75.7	100.7	125.7	150.7	175.7	200.7	225.7	250.7
Price (£)	131.62	129.12	126.68	124.29	122.98	119.69	117.47	115.30	113.18
Price + Accrued (£)	132.58	130.08	127.64	125.25	123.94	120.65	118.43	116.26	114.14
Price Δ (£)	8.64	6.14	3.70	1.31	0.00	−3.29	−5.51	−7.68	−9.80
Return (%)									
Actual	6.97%	4.96%	2.99%	1.06%	0.00%	−2.65%	−4.44%	−6.20%	−7.91%
Approx: Dur only	7.84%	5.88%	3.92%	1.96%	0.00%	−1.96%	−3.92%	−5.88%	−7.84%
Approx: Dur & Cvx	8.22%	6.10%	4.02%	1.98%	0.00%	−1.94%	−3.82%	−5.66%	−7.45%

Issuer: British Telecom, 3.625%, 21/11/2047

Price: £94.244	Modified Duration: 17.144	Spread to Gilt Curve: 210.8 b.p.
Accrued interest: 2.185	Convexity: 408.4	YTM (conv): 4.11

	Scenarios								
Spread Δ (b.p.)	−100	−75	−50	−25	0	25	50	75	100
Spread (b.p.)	110.8	135.8	160.8	185.8	210.8	235.8	260.8	285.8	310.8
Price (£)	111.28	106.38	101.77	97.41	93.24	89.41	85.75	82.30	79.04
Price + Accrued (£)	113.47	108.57	103.96	99.60	95.43	91.60	87.94	84.48	81.22
Price Δ (£)	18.04	13.14	8.53	4.17	0.00	−3.83	−7.49	−10.95	−14.21
Return (%)									
Actual	18.90%	13.77%	8.93%	4.37%	0.00%	−4.02%	−7.85%	−11.47%	−14.89%

Exhibit 16 (Continued)

	Scenarios								
Approx: Dur only	17.14%	12.86%	8.57%	4.29%	0.00%	−4.29%	−8.57%	−12.86%	−17.14%
Approx: Dur & Cvx	19.19%	14.01%	9.08%	4.41%	0.00%	−4.16%	−8.06%	−11.71%	−15.10%

Note: Settle date is 13 December 2018.
Source: Bloomberg Finance, L.P.

Note that the price change for a given spread change is higher for the longer-duration bond—in this case, the 2047 maturity British Telecom bond—than for the shorter-duration, 2028 maturity British Telecom bond. Longer-duration corporate bonds are referred to as having "higher spread sensitivity"; that is, their prices, and thus returns, are more volatile with respect to changes in spread. It is essentially the same concept as duration for any bond: The longer the duration of a bond, the greater the price volatility for a given change in interest rates/yields.

In addition, investors want to be compensated for the fact that the further time one is from a bond's maturity (i.e., the longer the bond), the greater the uncertainty about an issuer's future creditworthiness. Based on credit analysis, an investor might be confident that an issuer's risk of default is relatively low in the near term; however, looking many years into the future, the investor's uncertainty grows because of factors that are increasingly difficult, if not impossible, to forecast (e.g., poor management strategy or execution, technological obsolescence, natural or man-made disasters, corporate leveraging events). This increase in credit risk over time can be seen in Exhibit 17. Note that in this Standard & Poor's study, one-year default rates for the 2017 issuance pool are 0% for all rating categories of BB or higher. The three-year default rates for bonds issued in 2015 are materially higher, and the observed defaults include bonds originally rated up to BBB (i.e., low investment grade). The 5-year default rates for bonds issued in 2013 are higher than the 3-year default rates, and the defaults also include bonds initially rated as high as BBB. In addition to the risk of default rising over time, the data also show quite conclusively that the lower the credit rating, the higher the risk of default. Finally, note the very high risk of default for bonds rated CCC or lower over all time horizons. This is consistent with Exhibit 7 presented earlier, which showed significant three-year ratings variability ("migration"), with much of the migration to lower credit ratings (i.e., higher risk of default).

Exhibit 17 Default Rate by Rating Category (%) (Non-financials)

Credit Rating	1 Year (2017 pool)	3 Year (2015 pool)	10 Year (2013 pool)
AAA	0.00	0.00	0.00
AA	0.00	0.00	0.00
A	0.00	0.00	0.00
BBB	0.00	0.08	0.27
BB	0.10	2.46	3.33

(continued)

Exhibit 17 (Continued)

Credit Rating	1 Year (2017 pool)	3 Year (2015 pool)	10 Year (2013 pool)
B	0.95	10.11	12.90
CCC/C	27.15	41.43	44.70

Source: Based on data from S&P Global Ratings, "2017 Annual Global Corporate Default Study and Rating Transitions" (5 April 2018).

EXAMPLE 9

Price Impact

Calculate the price impact on a 10-year corporate bond with a 4.75% coupon priced at 100, with an instantaneous 50 bps widening in spread due to the issuer's announcement that it was adding substantial debt to finance an acquisition (which resulted in a two-notch downgrade by the rating agencies). The bond has a modified duration of 7.9, and its convexity is 74.9.

Solution:

The impact from the 50 bps spread widening is:

$$\begin{aligned}\text{Price impact} &\approx -(\text{AnnModDur} \times \Delta\text{Spread}) + \tfrac{1}{2}\,\text{AnnConvexity} \times (\Delta\text{Spread})^2\\ &= -(0.0050 \times 7.9) + (0.5 \times 74.9) \times (0.0050)^2\\ &= -0.0386, \text{ or } -3.86\%.\end{aligned}$$

Because yields and bond prices move in opposite directions, the wider spread caused the bond price to fall. Using a bond-pricing calculator, the exact return is −3.85%, so this approximation was very accurate.

In summary, spread changes can have a significant impact on the price and performance of credit-risky bonds over a given holding period, and the higher the modified duration of the bond(s), the greater the price impact from changes in spread. Wider spreads hurt bond performance, whereas narrower spreads help bond performance. For bond investors who actively manage their portfolios (i.e., don't just buy bonds and hold them to maturity), forecasting spread changes and expected credit losses on both individual bonds and their broader portfolios is an important strategy for enhancing investment performance.

7 HIGH-YIELD, SOVEREIGN, AND NON-SOVEREIGN CREDIT ANALYSIS

j explain special considerations when evaluating the credit of high-yield, sovereign, and non-sovereign government debt issuers and issues

Thus far, we have focused primarily on basic principles of credit analysis and investing with emphasis on higher-quality, investment-grade corporate bonds. Although many of these principles are applicable to other credit-risky segments of the bond market, some differences in credit analysis need to be considered. This section focuses on special considerations in evaluating the credit of debt issuers from the following three market segments: high-yield corporate bonds, sovereign bonds, and non-sovereign government bonds.

7.1 High Yield

Recall that high-yield, or non-investment-grade, corporate bonds are those rated below Baa3/BBB– by the major rating agencies. These bonds are sometimes referred to as "junk bonds" because of the higher risk inherent in their weak balance sheets and/or poor or less-proven business prospects.

Companies are rated below investment grade for many reasons, including the following:

- Highly leveraged capital structure
- Weak or limited operating history
- Limited or negative free cash flow
- Highly cyclical business
- Poor management
- Risky financial policies
- Lack of scale and/or competitive advantages
- Large off-balance-sheet liabilities
- Declining industry (e.g., newspaper publishing)

Companies with weak balance sheets and/or business profiles have lower margin for error and greater risk of default relative to higher-quality investment-grade names. And the higher risk of default means more attention must be paid to recovery analysis (or loss severity, in the event of default). Consequently, high-yield analysis typically is more in-depth than investment-grade analysis and thus has special considerations. This includes the following:

- Greater focus on issuer liquidity and cash flow
- Detailed financial projections
- Detailed understanding and analysis of the debt structure
- Understanding of an issuer's corporate structure
- Covenants
- Equity-like approach to high-yield analysis

Liquidity

Liquidity—that is, having cash and/or the ability to generate or raise cash—is important to all issuers. It is absolutely critical for high-yield companies. Investment-grade companies typically have substantial cash on their balance sheets, generate a lot of cash from operations relative to their debt (or else they wouldn't be investment grade!), and/or are presumed to have alternate sources of liquidity, such as bank lines and commercial paper. For these reasons, investment-grade companies can more easily roll over (refinance) maturing debt. On the other hand, high-yield companies may not have those options available. For example, there is no high-yield commercial paper market, and bank credit facilities often carry tighter restrictions for high-yield companies. Both bad company-specific news and difficult financial market conditions can lead to high-yield companies being unable to access the debt markets. And although the vast majority of investment-grade corporate debt issuers have publicly traded equity and can thus use that equity as a financing option, many high-yield companies are privately held and thus don't have access to public equity markets.

Thus, issuer liquidity is a key focus in high-yield analysis. Sources of liquidity, from strongest to weakest, are the following:

1 Cash on the balance sheet
2 Working capital
3 Operating cash flow
4 Bank credit facilities
5 Equity issuance
6 Asset sales

Cash on the balance sheet is easy to see and self-evident as a source for repaying debt, although a portion of it may be 'trapped' overseas for certain tax, business, accounting, or regulatory reasons and thus not easily accessible. As mentioned earlier, working capital can be a large source or use of liquidity, depending on its amount, its use in a company's cash-conversion cycle, and its role in a company's operations. Operating cash flow is a ready source of liquidity as sales turn to receivables, which turn to cash over a fairly short time period. Bank lines, or credit facilities, can be an important source of liquidity, though there may be some covenants relating to the use of the bank lines that are crucial to know and will be covered a little later. Equity issuance may not be a reliable source of liquidity because an issuer is private or because of poor market conditions if a company does have publicly traded equity. Asset sales are the least reliable source of liquidity because both the potential value and the actual time of closing can be highly uncertain.

The amount of these liquidity sources should be compared with the amount and timing of upcoming debt maturities. A large amount of debt coming due in the next 6–12 months alongside low sources of liquidity will be a warning flag for bond investors and could push an issuer into default because investors may choose not to buy new bonds intended to pay off the existing debt. Insufficient liquidity—that is, running out of cash or no longer having access to external financing to refinance or pay off existing debt—is the principal reason issuers default. Although liquidity is important for industrial companies, it is an absolute necessity for financial firms, as seen in the case of Lehman Brothers and other troubled firms during the global financial crisis of 2008. Financial institutions are highly levered and often highly dependent on funding longer-term assets with short-term liabilities.

Financial Projections

Because high-yield companies have less room for error, it's important to forecast, or project, future earnings and cash flow out several years, perhaps including several scenarios, to assess whether the issuer's credit profile is stable, improving, or declining and thus whether it needs other sources of liquidity or is at risk of default. Ongoing capital expenditures and working capital changes should be incorporated as well. Special emphasis should be given to realistic "stress" scenarios that could expose a borrower's vulnerabilities.

Debt Structure

High-yield companies tend to have many layers of debt in their capital structures, with varying levels of seniority and, therefore, different potential recovery rates in the event of default. (Recall the historical table of default recovery rates based on seniority in Exhibit 2.) A high-yield issuer will often have at least some of the following types of obligations in its debt structure:

- (Secured) Bank debt
- Second lien debt
- Senior unsecured debt
- Subordinated debt, which may include convertible bonds
- Preferred stock

The lower the ranking in the debt structure, the lower the credit rating and the lower the expected recovery in the event of default. In exchange for these associated higher risks, investors will normally demand higher yields.

As discussed earlier, a standard leverage calculation used by credit analysts is debt/EBITDA and is quoted as a multiple (e.g., "5.2x levered"). For an issuer with several layers of debt with different expected recovery rates, high-yield analysts should calculate leverage at each level of the debt structure. Example 10 shows calculations of gross leverage, as measured by debt/EBITDA, at each level of the debt structure and net leverage for the entire debt structure. Gross leverage calculations do not adjust debt for cash on hand. Net leverage adjusts debt by subtracting cash from total debt.

EXAMPLE 10

Debt Structure and Leverage

Hexion Inc. is a specialty chemical company. It has a complicated, high-yield debt structure, consisting of first lien debt (loans and bonds), secured bonds, second lien bonds, and senior unsecured debt, due to a series of mergers as well as a leveraged buyout in 2005. Exhibit 18 is a simplified depiction of the company's debt structure, as well as some key credit-related statistics.

Exhibit 18 Hexion Inc. Debt and Leverage Structure as of Year-End 2017

Financial Information ($ millions)	
Cash	$115
Total debt	$3,668
Net debt	$3,553
Interest expense	$329
EBITDA	$365

Debt Structure ($ millions)	
First lien debt (loans and bonds)	$2,607
Secured bonds	$225
Second lien bonds	$574

(continued)

Exhibit 18 (Continued)

Debt Structure ($ millions)	
Senior unsecured bonds	$263
TOTAL DEBT	$3,669

Source: Company filings, Loomis, Sayles & Company.

Using the information provided, address the following:

1 Calculate gross leverage, as measured by debt/EBITDA, through each level of debt, including total debt.

2 Calculate the net leverage, as measured by (Debt – Cash)/EBITDA, for the total debt structure.

3 Why might Hexion have so much secured debt relative to unsecured debt (both senior and subordinated)? (*Note*: This question draws on concepts from earlier sections.)

Solutions to 1 and 2:

	Gross Leverage (Debt/EBITDA)	Net Leverage (Debt – Cash)/ EBITDA
Secured debt leverage		
(First lien + Secured debt)/EBITDA		
(2,607 + 225)/365	7.8x	
Second lien leverage		
(First lien + Secured debt + Second lien debt)/EBITDA		
(2,607 + 225 + 574)/365	9.3x	
Total leverage (includes unsecured)		
(Total debt/EBITDA)		
3,669/365	10.1x	
Net leverage (leverage net of cash through entire debt structure)		
(Total debt – Cash)/EBITDA		9.7x

Solution to 3:

Hexion might have that much secured debt because (1) it was less expensive than issuing additional unsecured debt on which investors would have demanded a higher yield and/or (2) given the riskiness of the business (chemicals are a cyclical business), the high leverage of the business model, and the riskiness of the balance sheet (lots of debt from a leveraged buyout), investors would only be willing to lend the company money on a secured basis.

High-yield companies that have a lot of secured debt (typically bank debt) relative to unsecured debt are said to have a "top-heavy" capital structure. With this structure, there is less capacity to take on more bank debt in the event of financial stress. Along with the often more stringent covenants associated with bank debt and its generally shorter maturity compared with other types of debt, this means that these issuers are more susceptible to default, as well as to lower recovery for the various less secured creditors.

Corporate Structure

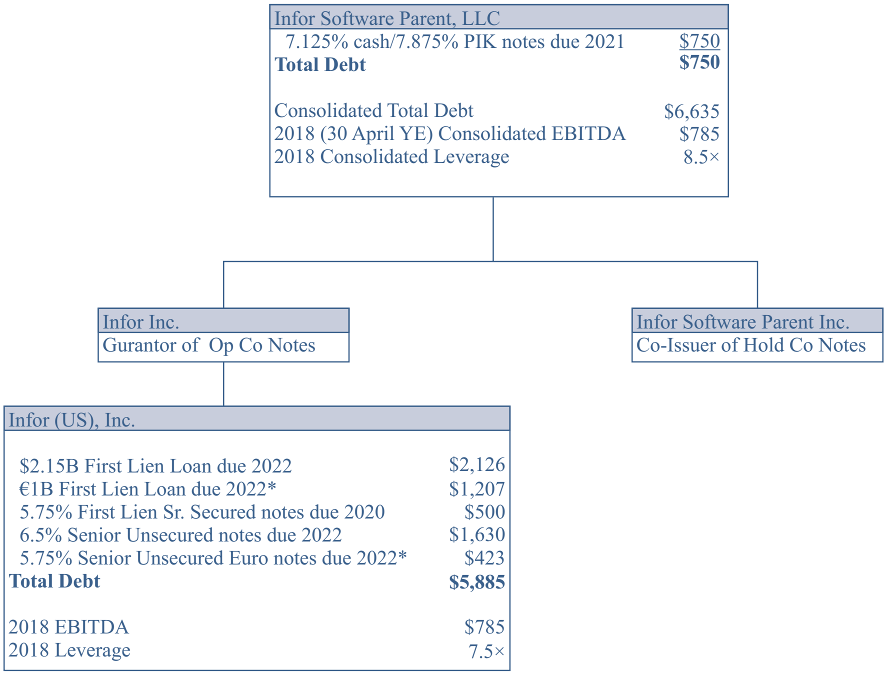
Many debt-issuing corporations, including high-yield companies, utilize a holding company structure with a parent and several operating subsidiaries. Knowing where an issuer's debt resides (parent versus subsidiaries) and how cash can move from subsidiary to parent ("upstream") and vice versa ("downstream") are critical to the analysis of high-yield issuers.

In a holding company structure, the parent owns stock in its subsidiaries. Typically, the parent doesn't generate much of its own earnings or cash flow but instead receives dividends from its subsidiaries. The subsidiaries' dividends are generally paid out of earnings after they satisfy all their other obligations, such as debt payments. To the extent that their earnings and cash flow are weak, subsidiaries may be limited in their ability to pay dividends to the parent. Moreover, subsidiaries that carry a lot of their own debt may have restrictions or limitations on how much cash they can provide to the parent via dividends or in another way, such as through an intercompany loan. These restrictions and limitations on cash moving between parent and subsidiaries can have a major impact on their respective abilities to meet their debt obligations. The parent's reliance on cash flow from its subsidiaries means the parent's debt is structurally subordinated to the subsidiaries' debt and thus will usually have a lower recovery rating in default.

For companies with very complex holding companies, there may also be one or more intermediate holding companies, each carrying their own debt, and in some cases, they may not own 100% of the subsidiaries' stock. Often times a default in one subsidiary may not trigger a cross-default. This structure is sometimes seen in high-yield companies that have been put together through many mergers and acquisitions or that were part of a leveraged buyout.

Exhibit 19 shows the capital structure of Infor, Inc. (Infor), a high-yield software and services company highlighted earlier as an example of the credit rating agency notching process. Infor's capital structure consists of a parent company that has debt—in this case, convertible senior notes—as well as a subsidiary with multiple layers of outstanding debt by seniority.

Exhibit 19 Infor's Capital Structure


Infor Software Parent, LLC
7.125% cash/7.875% PIK notes due 2021 $750
Total Debt $750
Consolidated Total Debt $6,635
2018 (30 April YE) Consolidated EBITDA $785
2018 Consolidated Leverage 8.5×
Infor Inc.
Gurantor of Op Co Notes
Infor Software Parent Inc.
Co-Issuer of Hold Co Notes
Infor (US), Inc.
$2.15B First Lien Loan due 2022 $2,126
€1B First Lien Loan due 2022* $1,207
5.75% First Lien Sr. Secured notes due 2020 $500
6.5% Senior Unsecured notes due 2022 $1,630
5.75% Senior Unsecured Euro notes due 2022* $423
Total Debt $5,885
2018 EBITDA $785
2018 Leverage 7.5×

---

*Conversion from euro to US dollar.

Source: Company filing, Loomis, Sayles & Company.

Thus, high-yield investors should analyze and understand an issuer's corporate structure, including the distribution of debt between the parent and its subsidiaries. Leverage ratios should be calculated at each of the debt-issuing entities, as well as on a consolidated basis.

Also important is that although the debt of an operating subsidiary may be "closer to" and better secured by particular assets of the subsidiary, the credit quality of a parent company might still be higher. The parent company could, while being less directly secured by any particular assets, still benefit from the diversity and availability of all the cash flows in the consolidated system. In short, credit quality is not simply an automatic analysis of debt provisions and liens.

Covenant Analysis

As discussed earlier, analysis of covenants is very important for all bonds. It is especially important for high-yield credits because of their reduced margin of safety. Key covenants for high-yield issuers may include the following:

- Change of control put
- Restricted payments
- Limitations on liens and additional indebtedness
- Restricted versus unrestricted subsidiaries

Under the **change of control put**, in the event of an acquisition (a "change of control"), bondholders have the right to require the issuer to buy back their debt (a "put option"), often at par or at some small premium to par value. This covenant is intended to protect creditors from being exposed to a weaker, more indebted borrower as a result of acquisition. For investment-grade issuers, this covenant typically has a two-pronged test: acquisition of the borrower and a consequent downgrade to a high-yield rating.

The **restricted payments** covenant is meant to protect creditors by limiting how much cash can be paid out to shareholders over time. The restricted payments "basket" is typically sized relative to an issuer's cash flow and debt outstanding—or is being raised—and is an amount that can grow with retained earnings or cash flow, giving management more flexibility to make payouts.

The **limitations on liens** covenant is meant to put limits on how much secured debt an issuer can have. This covenant is important to unsecured creditors who are structurally subordinated to secured creditors; the higher the amount of debt that is layered ahead of them, the less they stand to recover in the event of default.

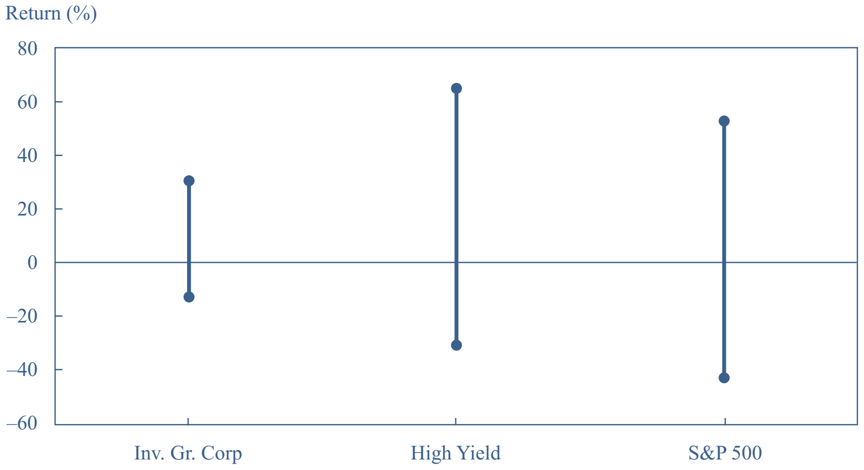
With regard to restricted versus unrestricted subsidiaries, issuers may classify certain of their subsidiaries as restricted and others as unrestricted as it pertains to offering guarantees for their holding company debt. These subsidiary guarantees can be very useful to holding company creditors because they put their debt on equal standing (pari passu) with debt at the subsidiaries instead of with structurally subordinated debt. Restricted subsidiaries should be thought of as those that are designated to help service parent-level debt, typically through guarantees. They tend to be an issuer's larger subsidiaries and have significant assets, such as plants and other facilities, and/or cash flow. There may be tax or legal (e.g., country of domicile) reasons why certain subsidiaries are restricted while others are not. Analysts should carefully read the definitions of restricted versus unrestricted subsidiaries in the indenture because sometimes the language is so loosely written that the company can reclassify subsidiaries from one type to another with a simple vote by a board of directors or trustees.

For high-yield investors, it is also important to know what covenants are in an issuer's bank credit agreements. These agreements are typically filed with the securities commission in the country where the loan document was drafted. Bank covenants can be more restrictive than bond covenants and may include so-called **maintenance covenants**, such as leverage tests, whereby the ratio of, say, debt/EBITDA may not exceed "x" times. In the event a covenant is breached, the bank is likely to block further loans under the agreement until the covenant is cured. If not cured, the bank may accelerate full payment of the facility, triggering a default.

Equity-like approach to high-yield analysis

High-yield bonds are sometimes thought of as a "hybrid" between higher-quality bonds, such as investment-grade corporate debt, and equity securities. Their more volatile price and spread movements are less influenced by interest rate changes than are higher-quality bonds, and they show greater correlation with movements in equity markets. Indeed, as shown in Exhibit 20, historical returns on high-yield bonds and the standard deviation of those returns fall somewhere between investment-grade bonds and equities.

Exhibit 20 US Trailing 12-Month Returns by Asset Class, 31 December 1988–30 September 2018



	Inv. Gr. Corp	High Yield	S&P 500
Avg. Return:	7.00%	8.86%	11.39%
Max. Return:	31.07%	64.95%	53.62%
Min. Return:	–13.82%	–31.23%	–43.32%
Avg. Volatility:	4.71%	6.79%	12.91%

Sources: Bloomberg Barclays Indices; Haver Analytics; and Loomis, Sayles & Company.

Consequently, an equity market-like approach to analyzing a high-yield issuer can be useful. One approach is to calculate an issuer's enterprise value. Enterprise value (EV) is usually calculated by adding equity market capitalization and total debt and then subtracting excess cash. Enterprise value is a measure of what a business is worth (before any takeover premium) because an acquirer of the company would have to either pay off or assume the debt and it would receive the acquired company's cash.

Bond investors like using EV because it shows the amount of equity "cushion" beneath the debt. It can also give a sense of (1) how much more leverage management might attempt to put on a company in an effort to increase equity returns or (2) how likely—and how expensive—a credit-damaging leveraged buyout might be. Similar to how stock investors look at equity multiples, bond investors may calculate and compare EV/EBITDA and debt/EBITDA across several issuers as part of their analysis. Narrow differences between EV/EBITDA and debt/EBITDA for a given issuer indicate a small equity cushion and, therefore, potentially higher risk for bond investors.

7.2 Sovereign Debt

Governments around the world issue debt to help finance their general operations, including such current expenses as wages for government employees, and investments in such long-term assets as infrastructure and education. Government bonds in developed countries have traditionally been viewed as the default risk-free rate off of which all other credits are priced. Fiscal challenges in developed countries exacerbated by the 2008 global financial crisis and the 2011–2012 eurozone crisis, however, have called into question the notion of a "risk-free rate," even for some of the highest-quality government borrowers. As their capital markets have developed, an increasing number of sovereign governments have been able to issue debt in foreign markets (generally denominated in a currency other than that of the sovereign government, often the US dollar or euro) as well as debt in the domestic market (issued in the sovereign government's own currency). Generally, sovereign governments with weak currencies can access foreign debt markets only by issuing bonds in foreign currencies that are viewed to be safer stores of value. Debt issued in the domestic market is somewhat easier to service because the debt is typically denominated in the country's own currency, subject to its own laws, and money can be printed to service the sovereign government's domestic debt. Twenty years ago, many emerging market countries could only issue debt in foreign markets because a domestic market did not exist. Today, many are able to issue debt domestically and have successfully built yield curves of domestic bonds across the maturity spectrum. All sovereign governments are best able to service foreign and domestic debt if they run "twin surpluses"—that is, a government budget surplus as well as a current account surplus.

Despite ongoing financial globalization and the development of domestic bond markets, sovereign government defaults occur. Defaults are often precipitated by such events as war, political upheaval, major currency devaluation, a sharp deterioration in trade, or dramatic price declines in a country's key commodity exports. Default risks for some developed countries escalated after 2009 as government revenues dropped precipitously following the global financial crisis of 2008, expenditures surged, and financial markets focused on the long-term sustainability of public finances, given aging populations and rising social security needs. Some of the weaker and more highly indebted members of the eurozone became unable to access the debt markets at economical rates and had to seek loans from the International Monetary Fund (IMF) and the European Union. These weaker governments had previously been able to borrow at much lower rates because of their membership in the European Union and adoption of the euro. Intra-eurozone yield spreads widened and countries were shut out of markets, however, as the global financial crisis exacted a high toll on their public finances and, in some cases, their banking systems, which became contingent liabilities for the sovereigns. In Ireland, the government guaranteed most bank liabilities, undermining the country's own fiscal metrics.

Like corporate analysis, sovereign credit analysis is based on a combination of qualitative and quantitative factors. Ultimately, the two key issues for sovereign analysis are: (1) a government's ability to pay and (2) its willingness to pay. Willingness to pay is important because of the principle of sovereign immunity, where investors are generally unable to force a sovereign to pay its debts. Sovereign immunity prevents governments from being sued. To date, global initiatives aimed at creating a mechanism for orderly sovereign restructurings and defaults have not found traction.

To illustrate the most important considerations in sovereign credit analysis, we present a basic framework for evaluating sovereign credit and assigning sovereign debt ratings (this outline was developed from the detailed exposition of Standard & Poor's methodology given in S&P Global Ratings, "Sovereign Rating Methodology" [18 December 2017]). The framework highlights the specific characteristics analysts should expect in a high-quality sovereign credit. Some of these are self-explanatory (e.g., absence of corruption). For others, a brief rationale and/or range of values is included to clarify interpretation. Most, but not all, of these items are included in rating agency Standard & Poor's methodology.

Institutional and economic profile

- *Institutional assessment*
  - Institutions' ability to deliver sound public finances and balanced economic growth.
  - Effectiveness and predictability of policymaking institutions.
  - Track record of managing previous political, economic, and/or financial crises.
  - Ability and willingness to implement reforms to address fiscal challenges.

  - Transparent and accountable institutions with low perceived level of corruption.
  - Independence of statistical offices and media.
  - Checks and balances between institutions.
  - Unbiased enforcement of contracts and respect for rule of law and property rights.
  - Debt repayment culture.
  - Potential external and domestic security risks.
- *Economic assessment*
  - Income per capita: More prosperous countries generally have a broader and deeper tax base with which to support debt.
  - Trend growth prospects: Creditworthiness is supported by sustainable and durable trend growth across business cycles.
  - Diversity and stability of growth: Sovereigns exposed to economic concentration are more vulnerable. A narrow economy tends to show higher volatility in growth and can impair a government's balance sheet.

Flexibility and performance profile

- *External assessment*
  - Status of currency: Sovereigns that control a reserve currency or a very actively traded currency are able to use their own currency in many international transactions and are less vulnerable to adverse shifts in global investor portfolios.
  - External liquidity: Countries with a substantial supply of foreign currency (foreign exchange reserves plus current account receipts) relative to projected funding needs in foreign currency (current account payments plus debt maturities) are less vulnerable to interruption of external liquidity.
  - External debt: Countries with low foreign debt relative to current account receipts are better able to service their foreign debt. This is similar to a coverage ratio for a corporation.
- *Fiscal assessment*
  - Fiscal performance and flexibility: Trend change in net general government debt as a percentage of GDP. Stable or declining debt as a percentage of GDP indicates a strong credit; a rising ratio can prove unsustainable and is, therefore, a sign of diminishing creditworthiness.
  - Long-term fiscal trends: Perceived willingness and ability to increase revenue or cut expenditure to ensure debt service.
  - Debt burden and structure: Net general government debt of less than 30% is good; more than 100% is poor. General government interest expense as a percentage of revenue: Less than 5% is good; greater than 15%is poor.
  - Ability to access funding, manage both the amortization profile and contingent liabilities arising from the financial sector, public enterprises, and guarantees.
- *Monetary assessment*
  - Ability to use monetary policy tailored to domestic economic objectives (e.g., growth) to address imbalances or shocks.

- Exchange rate regime: Sovereigns with a reserve currency have the most flexibility. A freely floating currency allows maximum effectiveness for monetary policy. A fixed-rate regime limits effectiveness and flexibility. A hard peg, such as a currency board or monetary union, affords no independent monetary policy.
- Credibility of monetary policy: Measured by track record of low and stable inflation. Credible monetary policy is supported by an operationally and legally independent central bank with a clear mandate. The central bank's ability to be a lender of last resort to the financial system also enhances stability.
- Confidence in the central bank provides a foundation for confidence in the currency as a store of value and for the central bank to effectively manage policy through a crisis.
- Most effective policy transmission occurs in systems with sound banking systems and well-developed domestic capital markets, including active money market and corporate bond markets, such that policymakers credibly enact policy relying on market-based policy tools (e.g., open market operations) over administrative policy tools (e.g., reserve requirements).

In light of a sovereign government's various powers—taxation, regulation, monetary policy, and ultimately, the sovereign's ability to "print money" to repay debt—within its own economy, it is virtually always at least as good a credit in its domestic currency as it is in foreign currency. Thus, credit rating agencies often distinguish between domestic and foreign bonds, with domestic bond ratings sometimes one notch higher. Of course, if a sovereign government were to rely heavily on printing money to repay debt, it would fuel high inflation or hyperinflation and increase default risk on domestic debt as well.

EXAMPLE 11

Sovereign Debt

Exhibit 21 shows several key sovereign statistics for Portugal.

Exhibit 21 Key Sovereign Statistics for Portugal

€ (billions), except where noted	2006	2008	2010	2012	2014	2015	2016	2017
Nominal GDP	160.3	171.2	172.6	168.4	173.1	179.8	186.5	194.6
Population (millions)	10.6	10.6	10.6	10.5	10.4	10.3	10.3	10.3
Unemployment (%)	8.6	8.5	12	15.6	13.9	12.4	11.1	8.9
Exports as share GDP (%)	22.2	22.6	21.3	26.8	27.8	27.6	26.8	28.3
Current account as share GDP (%)	−10.7	−12.6	−10	−2.1	0.2	0.3	0.7	0.6
Government revenues	64.8	70.7	71.8	72.2	77.2	78.8	79.9	83.1
Government expenditures	71.4	77.1	88.7	81.7	89.6	86.7	83.5	88.9
Budget balance (surplus/deficit)	−6.5	−6.4	−16.9	−9.5	−12.4	−7.9	−3.6	−5.8
Government interest payments	4.2	5.3	5.2	8.2	8.5	8.2	7.8	7.4
Primary balance (surplus/deficit)	−2.2	−1.1	−11.7	−1.3	−3.9	0.3	4.2	1.6

(continued)

Exhibit 21 (Continued)

€ (billions), except where noted	2006	2008	2010	2012	2014	2015	2016	2017
Government debt	102.4	123.1	161.3	212.6	226	231.5	241	242.8
Interest rate on new debt (%)	3.9	4.5	5.4	3.4	3.8	2.4	3.2	3.1

Sources: Haver Analytics, Eurostat, and Instituto Nacional de Estatística (Portugal).

1 Calculate the government debt/GDP for Portugal for the years 2014–2017 as well as for the years 2006, 2008, 2010, and 2012.

2 Calculate GDP/capita for the same periods.

3 Based on those calculations, as well as other data from Exhibit 21, what can you say about Portugal's credit trend?

Solutions to 1 and 2:

	2006	2008	2010	2012	2014	2015	2016	2017
Gross government debt/GDP	64%	72%	93%	126%	131%	129%	129%	125%
GDP/capita	15,123	16,151	16,283	16,038	16,644	17,456	18,107	18,893

Solution to 3:

The credit trend is stabilizing. Government debt/GDP is declining ever so slightly after peaking in 2014. The government is running a modest budget deficit with a primary balance that is in surplus for the past three years. Portugal is running a very small current account surplus, reducing its reliance on external funding, and has increased its exports as a share of GDP. Unemployment, while still fairly high, has fallen in the past several years. Interest payments on government debt have started to decline, both as a percentage of GDP and in absolute terms. The interest rate on new government debt has stabilized, perhaps benefitting from the European Central Bank's quantitative easing policies. Taken together, there are strong indications that the Portuguese government's credit situation has stabilized and may be expected to improve further if current trends are sustained.

7.3 Non-Sovereign Government Debt

Sovereigns are the largest issuers of government debt, but non-sovereign—sometimes called sub-sovereign or local—governments and the quasi-government entities that are created by governments issue bonds as well. The non-sovereign or local governments include governments of states, provinces, regions, and cities. For example, the City of Tokyo (Tokyo Metropolitan Government) has debt outstanding, as does the Lombardy region in Italy, the City of Buenos Aires in Argentina, and the State of California in the United States. Local government bonds may be referred to as municipal bonds.

However, when people talk about municipal bonds, they are usually referring to US municipal bonds, which represent one of the largest bond markets. As of year-end 2017, the US municipal bond market was approximately $3.9 trillion in size, roughly 9% of the total US bond market (Securities Industry and Financial Markets Association [SIFMA], "Outstanding U.S. Bond Market Data," as of 2Q 2018). The US municipal bond market is composed of both tax-exempt and, to a lesser extent, taxable bonds issued by state and city governments and their agencies. Municipal borrowers may also issue bonds on behalf of private entities, such as non-profit colleges or hospitals.

Historically, for any given rating category, these bonds have much lower default rates than corporate bonds with the same ratings. For example, according to Moody's Investors Service ("US Municipal Bond Defaults and Recoveries, 1970–2017"), the 10-year average cumulative default rate from 1970 through 2017 was 0.17% for municipal bonds, compared with a 10.24% 10-year average cumulative default rate for all corporate debt.

The majority of local government bonds, including municipal bonds, are either general obligation bonds or revenue bonds. General obligation (GO) bonds are unsecured bonds issued with the full faith and credit of the issuing non-sovereign government. These bonds are supported by the taxing authority of the issuer. Revenue bonds are issued for specific project financing (e.g., financing for a new sewer system, a toll road, bridge, hospital, or sports arena).

The credit analysis of GO bonds has some similarities to sovereign debt analysis (e.g., the ability to levy and collect taxes and fees to help service debt) but also some differences. For example, almost without exception, US municipalities must balance their operating budgets (i.e., exclusive of long-term capital projects) annually. Non-sovereign governments are unable to use monetary policy the way many sovereigns can.

The economic analysis of non-sovereign government GO bonds, including US municipal bonds, focuses on employment, per capita income (and changes in it over time), per capita debt (and changes in it over time), the tax base (depth, breadth, diversification, stability, etc.), demographics, and net population growth, as well as an analysis of whether the area represented by the non-sovereign government has the infrastructure and location to attract and support new jobs. Analysis should look at the volatility and variability of revenues during times of both economic strength and weakness. An overreliance on one or two types of tax revenue—particularly a volatile one, such as capital gains taxes or sales taxes—can signal increased credit risk. Pensions and other post-retirement obligations may not show up directly on the non-sovereign government's balance sheet, and many of these entities have underfunded pensions that need to be addressed. Adding the unfunded pension and post-retirement obligations to the debt reveals a more realistic picture of the issuer's debt and longer-term obligations. The relative ease or difficulty in managing the annual budgeting process and the government's ability to operate consistently within its budget are also important credit analysis considerations.

Disclosure by non-sovereign governments varies widely, with some of the smaller issuers providing limited financial information. Reporting requirements are inconsistent, so the financial reports may not be available for six months or more after the closing of a reporting period.

Exhibit 22 compares several key debt statistics from two of the larger states in the United States: Illinois and Texas. Illinois has the lowest credit ratings of any of the states, whereas Texas has one of the highest. Note the higher debt burden (and lower ranking) across several measures: Total debt, Debt/Capita, Debt/Personal income, and debt as a percentage of state GDP. When including net pension liabilities of government employees and retirees, the debt burdens are even greater, especially in the case of Illinois. What is not shown here is that Illinois also has a higher tax burden and greater difficulty balancing its budget on an annual basis than Texas does.

Exhibit 22 Municipal Debt Comparison: Illinois vs. Texas

	Illinois	Texas
Ratings:		
Moody's	Baa3	Aaa
S&P	BBB–	AAA

(continued)

Exhibit 22 (Continued)

	Illinois	Texas
Fitch	BBB	AAA
Unemployment rate (%)*	4.20	3.70
Median household income ($)**	$61,229	$57,051
Debt burden, net ($/rank)***		
Total (millions)	37,374 (5)	11,603 (13)
Per capita	2,919 (6)	410 (42)
As a percentage of 2016 personal income	5.60 (5)	0.90 (42)
As a percentage of 2016 GDP	4.70 (6)	0.73 (42)
ANPL, net ($/rank)****		
Total (millions)	250,136 (1)	140,253 (3)
Per capita	19,539 (1)	4,955 (19)
As a percentage of 2017 personal income	37.00 (1)	10.60 (19)
As a percentage of 2017 GDP	30.50 (1)	8.30 (20)

* *Source:* Bureau of Labor Statistics, data as of October 2018.
** *Source:* US Census Bureau, data as of 2017.
*** *Source:* Moody's Investors Service, Inc., debt data as of 2017.
**** *Source:* Moody's Investors Service, Inc., adjusted net pension liability data as of 2017.

Revenue bonds, which are issued to finance a specific project, have a higher degree of risk than GO bonds because they are dependent on a single source of revenue. The analysis of these bonds is a combination of an analysis of the project and the finances around the particular project. The project analysis focuses on the need and projected utilization of the project, as well as on the economic base supporting the project. The financial analysis has some similarities to the analysis of a corporate bond in that it is focused on operating results, cash flow, liquidity, capital structure, and the ability to service and repay the debt. A key credit measure for revenue-backed non-sovereign government bonds is the debt-service-coverage (DSC) ratio, which measures how much revenue is available to cover debt payments (principal and interest) after operating expenses. Many revenue bonds have a minimum DSC ratio covenant; the higher the DSC ratio, the stronger the creditworthiness.

SUMMARY

We introduced basic principles of credit analysis. We described the importance of the credit markets and credit and credit-related risks. We discussed the role and importance of credit ratings and the methodology associated with assigning ratings, as well as the risks of relying on credit ratings. We covered the key components of credit analysis and the financial measure used to help assess creditworthiness.

We also discussed risk versus return when investing in credit and how spread changes affect holding period returns. In addition, we addressed the special considerations to take into account when doing credit analysis of high-yield companies, sovereign borrowers, and non-sovereign government bonds.

- Credit risk is the risk of loss resulting from the borrower failing to make full and timely payments of interest and/or principal.
- The key components of credit risk are risk of default and loss severity in the event of default. The product of the two is expected loss. Investors in higher-quality bonds tend not to focus on loss severity because default risk for those securities is low.
- Loss severity equals (1 – Recovery rate).
- Credit-related risks include downgrade risk (also called credit migration risk) and market liquidity risk. Either of these can cause yield spreads—yield premiums—to rise and bond prices to fall.
- Downgrade risk refers to a decline in an issuer's creditworthiness. Downgrades will cause its bonds to trade with wider yield spreads and thus lower prices.
- Market liquidity risk refers to a widening of the bid–ask spread on an issuer's bonds. Lower-quality bonds tend to have greater market liquidity risk than higher-quality bonds, and during times of market or financial stress, market liquidity risk rises.
- The composition of an issuer's debt and equity is referred to as its "capital structure." Debt ranks ahead of all types of equity with respect to priority of payment, and within the debt component of the capital structure, there can be varying levels of seniority.
- With respect to priority of claims, secured debt ranks ahead of unsecured debt, and within unsecured debt, senior debt ranks ahead of subordinated debt. In the typical case, all of an issuer's bonds have the same probability of default due to cross-default provisions in most indentures. Higher priority of claim implies higher recovery rate—lower loss severity—in the event of default.
- For issuers with more complex corporate structures—for example, a parent holding company that has operating subsidiaries—debt at the holding company is structurally subordinated to the subsidiary debt, although the possibility of more diverse assets and earnings streams from other sources could still result in the parent having higher effective credit quality than a particular subsidiary.
- Recovery rates can vary greatly by issuer and industry. They are influenced by the composition of an issuer's capital structure, where in the economic and credit cycle the default occurred, and what the market's view of the future prospects are for the issuer and its industry.
- The priority of claims in bankruptcy is not always absolute. It can be influenced by several factors, including some leeway accorded to bankruptcy judges, government involvement, or a desire on the part of the more senior creditors to settle with the more junior creditors and allow the issuer to emerge from bankruptcy as a going concern, rather than risking smaller and delayed recovery in the event of a liquidation of the borrower.
- Credit rating agencies, such as Moody's, Standard & Poor's, and Fitch, play a central role in the credit markets. Nearly every bond issued in the broad debt markets carries credit ratings, which are opinions about a bond issue's creditworthiness. Credit ratings enable investors to compare the credit risk of debt issues and issuers within a given industry, across industries, and across geographic markets.

- Bonds rated Aaa to Baa3 by Moody's and AAA to BBB– by Standard & Poor's (S&P) and/or Fitch (higher to lower) are referred to as "investment grade." Bonds rated lower than that—Ba1 or lower by Moody's and BB+ or lower by S&P and/or Fitch—are referred to as "below investment grade" or "speculative grade." Below-investment-grade bonds are also called "high-yield" or "junk" bonds.
- The rating agencies rate both issuers and issues. Issuer ratings are meant to address an issuer's overall creditworthiness—its risk of default. Ratings for issues incorporate such factors as their rankings in the capital structure.
- The rating agencies will notch issue ratings up or down to account for such factors as capital structure ranking for secured or subordinated bonds, reflecting different recovery rates in the event of default. Ratings may also be notched due to structural subordination.
- Rating agencies incorporate ESG factors into their ratings of firms. Some have launched a set of ratings that aim to measure a company's attitudes, practices, and advances related to ESG. They identify and track leaders and laggards in the space. Companies are evaluated according to their exposure to ESG risks and how well they manage those risks relative to peers.
- There are risks in relying too much on credit agency ratings. Creditworthiness may change over time, and initial/current ratings do not necessarily reflect the creditworthiness of an issuer or bond over an investor's holding period. Valuations often adjust before ratings change, and the notching process may not adequately reflect the price decline of a bond that is lower ranked in the capital structure. Because ratings primarily reflect the probability of default but not necessarily the severity of loss given default, bonds with the same rating may have significantly different expected losses (default probability times loss severity). And like analysts, credit rating agencies may have difficulty forecasting certain credit-negative outcomes, such as adverse litigation and leveraging corporate transactions, and such low probability/high severity events as earthquakes and hurricanes.
- The role of corporate credit analysis is to assess the company's ability to make timely payments of interest and to repay principal at maturity.
- Credit analysis is similar to equity analysis. It is important to understand, however, that bonds are contracts and that management's duty to bondholders and other creditors is limited to the terms of the contract. In contrast, management's duty to shareholders is to act in their best interest by trying to maximize the value of the company—perhaps even at the expense of bondholders at times.
- Credit analysts tend to focus more on the downside risk given the asymmetry of risk/return, whereas equity analysts focus more on upside opportunity from earnings growth, and so on.
- The "4 Cs" of credit—capacity, collateral, covenants, and character—provide a useful framework for evaluating credit risk.
- Credit analysis focuses on an issuer's ability to generate cash flow. The analysis starts with an industry assessment—structure and fundamentals—and continues with an analysis of an issuer's competitive position, management strategy, and track record.
- Credit measures are used to calculate an issuer's creditworthiness, as well as to compare its credit quality with peer companies. Key credit ratios focus on leverage and interest coverage and use such measures as EBITDA, free cash flow, funds from operations, interest expense, and balance sheet debt.

- An issuer's ability to access liquidity is also an important consideration in credit analysis.
- The higher the credit risk, the greater the offered/required yield and potential return demanded by investors. Over time, bonds with more credit risk offer higher returns but with greater volatility of return than bonds with lower credit risk.
- The yield on a credit-risky bond comprises the yield on a default risk–free bond with a comparable maturity plus a yield premium, or "spread," that comprises a credit spread and a liquidity premium. That spread is intended to compensate investors for credit risk—risk of default and loss severity in the event of default—and the credit-related risks that can cause spreads to widen and prices to decline—downgrade or credit migration risk and market liquidity risk.

$$\text{Yield spread} = \text{Liquidity premium} + \text{Credit spread}.$$

- In times of financial market stress, the liquidity premium can increase sharply, causing spreads to widen on all credit-risky bonds, with lower-quality issuers most affected. In times of credit improvement or stability, however, credit spreads can narrow sharply as well, providing attractive investment returns.
- The impact of spread changes on holding period returns for credit-risky bonds is a product of two primary factors: the basis point spread change and the sensitivity of price to yield as reflected by (end-of-period) modified duration and convexity. Spread narrowing enhances holding period returns, whereas spread widening has a negative impact on holding period returns. Longer-duration bonds have greater price and return sensitivity to changes in spread than shorter-duration bonds.

$$\text{Price impact} \approx -(\text{AnnModDur} \times \Delta\text{Spread}) + ½\text{AnnConvexity} \times (\Delta\text{Spread})^2.$$

- For high-yield bonds, with their greater risk of default, more emphasis should be placed on an issuer's sources of liquidity and its debt structure and corporate structure. Credit risk can vary greatly across an issuer's debt structure depending on the seniority ranking. Many high-yield companies have complex capital structures, resulting in different levels of credit risk depending on where the debt resides.
- Covenant analysis is especially important for high-yield bonds. Key covenants include payment restrictions, limitation on liens, change of control, coverage maintenance tests (often limited to bank loans), and any guarantees from restricted subsidiaries. Covenant language can be very technical and legalistic, so it may help to seek legal or expert assistance.
- An equity-like approach to high-yield analysis can be helpful. Calculating and comparing enterprise value with EBITDA and debt/EBITDA can show a level of equity "cushion" or support beneath an issuer's debt.
- Sovereign credit analysis includes assessing both an issuer's ability and willingness to pay its debt obligations. Willingness to pay is important because, due to sovereign immunity, a sovereign government cannot be forced to pay its debts.
- In assessing sovereign credit risk, a helpful framework is to focus on five broad areas: (1) institutional effectiveness and political risks, (2) economic structure and growth prospects, (3) external liquidity and international investment position, (4) fiscal performance, flexibility, and debt burden, and (5) monetary flexibility.
- Among the characteristics of a high-quality sovereign credit are the absence of corruption and/or challenges to political framework; governmental checks and balances; respect for rule of law and property rights; commitment to honor

debts; high per capita income with stable, broad-based growth prospects; control of a reserve or actively traded currency; currency flexibility; low foreign debt and foreign financing needs relative to receipts in foreign currencies; stable or declining ratio of debt to GDP; low debt service as a percentage of revenue; low ratio of net debt to GDP; operationally independent central bank; track record of low and stable inflation; and a well-developed banking system and active money market.

- Non-sovereign or local government bonds, including municipal bonds, are typically either general obligation bonds or revenue bonds.
- General obligation (GO) bonds are backed by the taxing authority of the issuing non-sovereign government. The credit analysis of GO bonds has some similarities to sovereign analysis—debt burden per capita versus income per capita, tax burden, demographics, and economic diversity. Underfunded and "off-balance-sheet" liabilities, such as pensions for public employees and retirees, are debt-like in nature.
- Revenue-backed bonds support specific projects, such as toll roads, bridges, airports, and other infrastructure. The creditworthiness comes from the revenues generated by usage fees and tolls levied.

PRACTICE PROBLEMS

1 The risk that a bond's creditworthiness declines is *best* described by:

A credit migration risk.

B market liquidity risk.

C spread widening risk.

2 Stedsmart Ltd and Fignermo Ltd are alike with respect to financial and operating characteristics, except that Stedsmart Ltd has less publicly traded debt outstanding than Fignermo Ltd. Stedsmart Ltd is *most likely* to have:

A no market liquidity risk.

B lower market liquidity risk.

C higher market liquidity risk.

3 In the event of default, the recovery rate of which of the following bonds would *most likely* be the highest?

A First mortgage debt

B Senior unsecured debt

C Junior subordinate debt

4 During bankruptcy proceedings of a firm, the priority of claims was not strictly adhered to. Which of the following is the *least likely* explanation for this outcome?

A Senior creditors compromised.

B The value of secured assets was less than the amount of the claims.

C A judge's order resulted in actual claims not adhering to strict priority of claims.

5 A fixed-income analyst is *least likely* to conduct an independent analysis of credit risk because credit rating agencies:

A may at times mis-rate issues.

B often lag the market in pricing credit risk.

C cannot foresee future debt-financed acquisitions.

6 If goodwill makes up a large percentage of a company's total assets, this *most likely* indicates that:

A the company has low free cash flow before dividends.

B there is a low likelihood that the market price of the company's common stock is below book value.

C a large percentage of the company's assets are not of high quality.

7 In order to analyze the **collateral** of a company, a credit analyst should assess the:

A cash flows of the company.

B soundness of management's strategy.

C value of the company's assets in relation to the level of debt.

8 In order to determine the **capacity** of a company, it would be *most* appropriate to analyze the:

A company's strategy.

© 2021 CFA Institute. All rights reserved.

B growth prospects of the industry.

C aggressiveness of the company's accounting policies.

9 A credit analyst is evaluating the credit worthiness of three companies: a construction company, a travel and tourism company, and a beverage company. Both the construction and travel and tourism companies are cyclical, whereas the beverage company is non-cyclical. The construction company has the highest debt level of the three companies. The highest credit risk is *most likely* exhibited by the:

A construction company.

B beverage company.

C travel and tourism company.

10 Based on the information provided in Exhibit 1, the EBITDA interest coverage ratio of Adidas AG is *closest* to:

A 16.02x.

B 23.34x.

C 37.98x.

Exhibit 1 Adidas AG Excerpt from Consolidated Income Statement in a given year (€ in millions)

Gross profit	12,293
Royalty and commission income	154
Other operating income	56
Other operating expenses	9,843
Operating profit	2,660
Interest income	64
Interest expense	166
Income before taxes	2,558
Income taxes	640
Net income	1,918
Additional information:	
Depreciation and amortization: €1,214 million	

Source: Adidas AG Annual Financial Statements, December 2019.

11 The following information is from the annual report of Adidas AG for December 2019:

- Depreciation and amortization: €1,214 million
- Total assets: €20,640 million
- Total debt: €4,364 million
- Shareholders' equity: €7,058 million

The debt/capital of Adidas AG is *closest* to:

A 21.14%.

B 38.21%.

C 61.83%.

12 Funds from operations (FFO) of Pay Handle Ltd (a fictitious company) increased in 20X1. In 20X1, the total debt of the company remained unchanged while additional common shares were issued. Pay Handle Ltd's ability to service its debt in 20X1, as compared to 20X0, *most likely*:

A improved.

B worsened.

C remained the same.

13 Based on the information in Exhibit 2, GZ Group's (a hypothetical company) credit risk is *most likely*:

A lower than the industry.

B higher than the industry.

C the same as the industry.

Exhibit 2 European Food, Beverage, and Tobacco Industry and GZ Group Selected Financial Ratios for 20X0

	Total Debt/Total Capital (%)	FFO/Total Debt (%)	Return on Capital (%)	Total Debt/ EBITDA (x)	EBITDA Interest Coverage (x)
GZ Group	47.1	77.5	19.6	1.2	17.7
Industry median	**42.4**	**23.6**	**6.55**	**2.85**	**6.45**

14 Based on the information in Exhibit 3, the credit rating of DCM Group (a hypothetical company in the European food & beverage sector) is *most likely*:

A lower than AB plc.

B higher than AB plc.

C the same as AB plc.

Exhibit 3 DCM Group and AB plc Selected Financial Ratios for 20X0

Company	Total Debt/ Total Capital (%)	FFO/Total Debt (%)	Return on Capital (%)	Total Debt/EBITDA (x)	EBITDA Interest Coverage (x)
AB plc	0.2	84.3	0.1	1.0	13.9
DCM Group	42.9	22.9	8.2	3.2	3.2
European Food, Beverage, and Tobacco median	**42.4**	**23.6**	**6.55**	**2.85**	**6.45**

15 Holding all other factors constant, the *most likely* effect of low demand and heavy new issue supply on bond yield spreads is that yield spreads will:

A widen.

B tighten.

C not be affected.

16 Credit risk of a corporate bond is *best* described as the:

A risk that an issuer's creditworthiness deteriorates.

B probability that the issuer fails to make full and timely payments.

C risk of loss resulting from the issuer failing to make full and timely payments.

17 The risk that the price at which investors can actually transact differs from the quoted price in the market is called:

A spread risk.

B credit migration risk.

C market liquidity risk.

18 Loss severity is *best* described as the:

A default probability multiplied by the loss given default.

B portion of a bond's value recovered by bondholders in the event of default.

C portion of a bond's value, including unpaid interest, an investor loses in the event of default.

19 The two components of credit risk are default probability and:

A spread risk.

B loss severity.

C market liquidity risk.

20 For a high-quality debt issuer with a large amount of publicly traded debt, bond investors tend to devote *most* effort to assessing the issuer's:

A default risk.

B loss severity.

C market liquidity risk.

21 The expected loss for a given debt instrument is estimated as the product of default probability and:

A (1 + Recovery rate).

B (1 – Recovery rate).

C 1/(1 + Recovery rate).

22 The priority of claims for senior subordinated debt is:

A lower than for senior unsecured debt.

B the same as for senior unsecured debt.

C higher than for senior unsecured debt.

23 A senior unsecured credit instrument holds a higher priority of claims than one ranked as:

A mortgage debt.

B second lien loan.

C senior subordinated.

24 In a bankruptcy proceeding, when the absolute priority of claims is enforced:

A senior subordinated creditors rank above second lien holders.

B preferred equity shareholders rank above unsecured creditors.

C creditors with a secured claim have the first right to the value of that specific property.

25 In the event of default, which of the following is *most likely* to have the highest recovery rate?

A Second lien

B Senior unsecured

C Senior subordinated

26 The process of moving credit ratings of different issues up or down from the issuer rating in response to different payment priorities is *best* described as:

A notching.

B structural subordination.

C cross-default provisions.

27 The factor considered by rating agencies when a corporation has debt at both its parent holding company and operating subsidiaries is *best* referred to as:

A credit migration risk.

B corporate family rating.

C structural subordination.

28 Which type of security is *most likely* to have the same rating as the issuer?

A Preferred stock

B Senior secured bond

C Senior unsecured bond

29 Which of the following corporate debt instruments has the highest seniority ranking?

A Second lien

B Senior unsecured

C Senior subordinated

30 An issuer credit rating usually applies to a company's:

A secured debt.

B subordinated debt.

C senior unsecured debt.

31 The rating agency process whereby the credit ratings on issues are moved up or down from the issuer rating *best* describes:

A notching.

B pari passu ranking.

C cross-default provisions.

32 The notching adjustment for corporate bonds rated Aa2/AA is *most likely*:

A larger than the notching adjustment for corporate bonds rated B2/B.

B the same as the notching adjustment for corporate bonds rated B2/B.

C smaller than the notching adjustment for corporate bonds rated B2/B.

33 Which of the following statements about credit ratings is *most accurate*?

A Credit ratings can migrate over time.

B Changes in bond credit ratings precede changes in bond prices.

C Credit ratings are focused on expected loss rather than risk of default.

34 Which industry characteristic *most likely* has a positive effect on a company's ability to service debt?

A Low barriers to entry in the industry

B High number of suppliers to the industry

C Broadly dispersed market share among large number of companies in the industry

35 When determining the capacity of a borrower to service debt, a credit analyst should begin with an examination of:

A industry structure.

B industry fundamentals.

C company fundamentals.

36 Which of the following accounting issues should *mostly likely* be considered a character warning flag in credit analysis?

A Expensing items immediately

B Changing auditors infrequently

C Significant off-balance-sheet financing

37 In credit analysis, capacity is *best* described as the:

A quality of management.

B ability of the borrower to make its debt payments on time.

C quality and value of the assets supporting an issuer's indebtedness.

38 Among the four Cs of credit analysis, the recognition of revenue prematurely *most likely* reflects a company's:

A character.

B covenants.

C collateral.

Use the following Exhibit for Questions 39 and 40

Exhibit 4 Industrial Comparative Ratio Analysis, Year 20XX

	EBITDA Margin (%)	Return on Capital (%)	EBIT/ Interest Expense (×)	EBITDA/ Interest Expense (×)	Debt/ EBITDA (×)	Debt/ Capital (%)
Company A	25.1	25.0	15.9	19.6	1.6	35.2
Company B	29.6	36.3	58.2	62.4	0.5	15.9
Company C	21.8	16.6	8.9	12.4	2.5	46.3

39 Based on only the leverage ratios in Exhibit 4, the company with the *highest* credit risk is:

A Company A.

B Company B.

C Company C.

40 Based on only the coverage ratios in Exhibit 4, the company with the *highest* credit quality is:

A Company A.

B Company B.

C Company C.

Use the following Exhibits for Questions 41 and 42

Exhibit 5 Consolidated Income Statement (£ millions)

	Company X	Company Y
Net revenues	50.7	83.7
Operating expenses	49.6	70.4
Operating income	1.1	13.3
Interest income	0.0	0.0
Interest expense	0.6	0.8
Income before income taxes	0.5	12.5
Provision for income taxes	–0.2	–3.5
Net income	0.3	9.0

Exhibit 6 Consolidated Balance Sheets (£ millions)

	Company X	Company Y
ASSETS		
Current assets	10.3	21.9
Property, plant, and equipment, net	3.5	20.1
Goodwill	8.3	85.0
Other assets	0.9	5.1
Total assets	23.0	132.1
LIABILITIES AND SHAREHOLDERS' EQUITY		
Current liabilities		
Accounts payable and accrued expenses	8.4	16.2
Short-term debt	0.5	8.7
Total current liabilities	8.9	24.9
Long-term debt	11.7	21.1
Other non-current liabilities	1.1	22.1

(continued)

Exhibit 6 (Continued)

	Company X	Company Y
Total liabilities	21.7	68.1
Total shareholders' equity	1.3	64.0
Total liabilities and shareholders' equity	23.0	132.1

Exhibit 7 Consolidated Statements of Cash Flow (£ millions)

	Company X	Company Y
CASH FLOWS FROM OPERATING ACTIVITIES		
Net income	0.3	9.0
Depreciation	1.0	3.8
Goodwill impairment	2.0	1.6
Changes in working capital	0.0	–0.4
Net cash provided by operating activities	3.3	14.0
CASH FLOWS FROM INVESTING ACTIVITIES		
Additions to property and equipment	–1.0	–4.0
Additions to marketable securities	–0.1	0.0
Proceeds from sale of property and equipment	0.2	2.9
Proceeds from sale of marketable securities	0.3	0.0
Net cash used in investing activities	–0.6	–1.1
CASH FLOWS FROM FINANCING ACTIVITIES		
Repurchase of common stock	–1.5	–4.0
Dividends to shareholders	–0.3	–6.1
Change in short-term debt	0.0	–3.4
Additions to long-term debt	3.9	3.9
Reductions in long-term debt	–3.4	–2.5
Net cash–financing activities	–1.3	–12.1
NET INCREASE IN CASH AND CASH EQUIVALENTS	1.4	0.8

41 Based on Exhibits 5–7, in comparison to Company X, Company Y has a higher:

A debt/capital.

B debt/EBITDA.

C free cash flow after dividends/debt.

42 Based on Exhibits 5–7, in comparison to Company Y, Company X has greater:

A leverage.

B interest coverage.

C operating profit margin.

43 Credit yield spreads *most likely* widen in response to:

A high demand for bonds.

B weak performance of equities.

C strengthening economic conditions.

44 The factor that *most likely* results in corporate credit spreads widening is:

A an improving credit cycle.

B weakening economic conditions.

C a period of high demand for bonds.

45 Credit spreads are *most likely* to widen:

A in a strengthening economy.

B as the credit cycle improves.

C in periods of heavy new issue supply and low borrower demand.

46 Which of the following factors in credit analysis is more important for general obligation non-sovereign government debt than for sovereign debt?

A Per capita income

B Power to levy and collect taxes

C Requirement to balance an operating budget

47 In contrast to high-yield credit analysis, investment-grade analysis is *more likely* to rely on:

A spread risk.

B an assessment of bank credit facilities.

C matching of liquidity sources to upcoming debt maturities.

48 Which of the following factors would *best* justify a decision to avoid investing in a country's sovereign debt?

A Freely floating currency

B A population that is not growing

C Suitable checks and balances in policymaking

SOLUTIONS

1 A is correct. Credit migration risk or downgrade risk refers to the risk that a bond issuer's creditworthiness may deteriorate or migrate lower. The result is that investors view the risk of default to be higher, causing the spread on the issuer's bonds to widen.

2 C is correct. Market liquidity risk refers to the risk that the price at which investors transact may be different from the price indicated in the market. Market liquidity risk is increased by (1) less debt outstanding and/or (2) a lower issue credit rating. Because Stedsmart Ltd is comparable to Fignermo Ltd except for less publicly traded debt outstanding, it should have higher market liquidity risk.

3 A is correct. First mortgage debt is senior secured debt and has the highest priority of claims. First mortgage debt also has the highest expected recovery rate. First mortgage debt refers to the pledge of specific property. Neither senior unsecured nor junior subordinate debt has any claims on specific assets.

4 B is correct. Whether or not secured assets are sufficient for the claims against them does not influence priority of claims. Any deficiency between pledged assets and the claims against them becomes senior unsecured debt and still adheres to the guidelines of priority of claims.

5 C is correct. Both analysts and rating agencies have difficulty foreseeing future debt-financed acquisitions.

6 C is correct. Goodwill is viewed as a lower quality asset compared with tangible assets that can be sold and more easily converted into cash.

7 C is correct. The value of assets in relation to the level of debt is important to assess the collateral of the company—that is, the quality and value of the assets that support the debt levels of the company.

8 B is correct. The growth prospects of the industry provide the analyst insight regarding the capacity of the company.

9 A is correct. The construction company is both highly leveraged, which increases credit risk, and in a highly cyclical industry, which results in more volatile earnings.

10 B is correct. The interest expense is €166 million and EBITDA = Operating profit + Depreciation and amortization = €2,660 + 1,214 million = €3,874 million. EBITDA interest coverage = EBITDA/Interest expense = 3,874/166 = 23.34 times.

11 B is correct. Total debt is €4,364 million with Total capital = Total debt + Shareholders' equity = €4,364 + 7,058 = €11,422 million. The Debt/Capital = 4,364/11,422 = 38.21%.

12 A is correct. If the debt of the company remained unchanged but FFO increased, more cash was available to service debt compared to the previous year. Additionally, debt/capital improved, which implies that the ability of Pay Handle Ltd to service their debt also improved.

13 A is correct. Based on four of the five credit ratios, GZ Group's credit quality is superior to that of the industry.

14 A is correct. DCM Group has more financial leverage and less interest coverage than AB plc, which implies greater credit risk.

15 A is correct. Low demand implies wider yield spreads, and heavy supply widens spreads even further.

16 C is correct. Credit risk is the risk of loss resulting from the borrower failing to make full and timely payments of interest and/or principal.

17 C is correct. Market liquidity risk is the risk that the price at which investors can actually transact—buying or selling—may differ from the price indicated in the market.

18 C is correct. Loss severity is the portion of a bond's value (including unpaid interest) an investor loses in the event of default.

19 B is correct. The two components of credit risk are default probability and loss severity. In the event of default, loss severity is the portion of a bond's value (including unpaid interest) an investor loses. A and C are incorrect because spread and market liquidity risk are credit-related risks, not components of credit risk.

20 A is correct. Credit risk has two components: default risk and loss severity. Because default risk is quite low for most high-quality debt issuers, bond investors tend to focus more on this likelihood and less on the potential loss severity.

21 B is correct. The expected loss for a given debt instrument is the default probability multiplied by the loss severity given default. The loss severity is often expressed as (1 – Recovery rate).

22 A is correct. Senior subordinated debt is ranked lower than senior unsecured debt and thus has a lower priority of payment.

23 C is correct. The highest-ranked unsecured debt is senior unsecured debt. Lower-ranked debt includes senior subordinated debt. A and B are incorrect because mortgage debt and second lien loans are secured and higher ranked.

24 C is correct. According to the absolute priority of claims, in the event of bankruptcy, creditors with a secured claim have the right to the value of that specific property before any other claim.

25 A is correct. A second lien has a secured interest in the pledged assets. Second lien debt ranks higher in priority of payment than senior unsecured and senior subordinated debt and thus would most likely have a higher recovery rate.

26 A is correct. Notching is the process for moving ratings up or down relative to the issuer rating when rating agencies consider secondary factors, such as priority of claims in the event of a default and the potential loss severity.

27 C is correct. Structural subordination can arise when a corporation with a holding company structure has debt at both its parent holding company and operating subsidiaries. Debt at the operating subsidiaries is serviced by the cash flow and assets of the subsidiaries before funds are passed to the parent holding company.

28 C is correct. The issuer credit rating usually applies to its senior unsecured debt.

29 A is correct. Second lien debt is secured debt, which is senior to unsecured debt and to subordinated debt.

30 C is correct. An issuer credit rating usually applies to its senior unsecured debt.

31 A is correct. Recognizing different payment priorities, and thus the potential for higher (or lower) loss severity in the event of default, the rating agencies have adopted a notching process whereby their credit ratings on issues can be moved up or down from the issuer rating (senior unsecured).

32 C is correct. As a general rule, the higher the senior unsecured rating, the smaller the notching adjustment. Thus, for corporate bonds rated Aa2/AA, the rating agencies will typically apply smaller rating adjustments, or notches, to the related issue.

33 A is correct. Credit migration is the risk that a bond issuer's creditworthiness deteriorates, or migrates lower. Over time, credit ratings can migrate significantly from what they were at the time a bond was issued. An investor should not assume that an issuer's credit rating will remain the same from the time of purchase through the entire holding period.

34 B is correct. An industry with a high number of suppliers reduces the suppliers' negotiating power, thus helping companies control expenses and aiding in the servicing of debt.

35 A is correct. Credit analysis starts with industry structure—for example, by looking at the major forces of competition, followed by an analysis of industry fundamentals—and then turns to examination of the specific issuer.

36 C is correct. Credit analysts can make judgments about management's character by evaluating the use of aggressive accounting policies, such as timing revenue recognition. This activity is a potential warning flag for other behaviors or actions that may adversely affect an issuer's creditworthiness.

37 B is correct. Capacity refers to the ability of a borrower to service its debt. Capacity is determined through credit analysis of an issuer's industry and of the specific issuer.

38 A is correct. Credit analysts can make judgments about management's character in a number of ways, including by observing its use of aggressive accounting policies and/or tax strategies. An example of this aggressiveness is recognizing revenue prematurely.

39 C is correct. Debt/capital and debt/EBITDA are used to assess a company's leverage. Higher leverage ratios indicate more leverage and thus higher credit risk. Company C's debt/capital (46.3%) and debt/EBITDA (2.5×) are higher than those for Companies A and B.

40 B is correct. The EBITDA/interest expense and EBIT/interest expense are coverage ratios. Coverage ratios measure an issuer's ability to meet its interest payments. A higher ratio indicates better credit quality. Company B's EBITDA/interest expense (62.4×) and EBIT/interest expense (58.2×) are higher than those for Companies A and C.

41 C is correct because Company Y has a higher ratio of free cash flow after dividends to debt than Company X, not lower, as shown in the following table.

$$\text{Free cash flow after dividends as a \% of debt} = \frac{\text{FCF after dividends}}{\text{Debt}}$$

	Company X	**Company Y**
Cash flow from operations	£3.3	£14.0
Less		
Net capital expenditures	−0.8	−1.1
Dividends	−0.3	−6.1
Free cash flow after dividends	£2.2	£6.8
Debt	£12.2	£29.8
Free cash flow after dividends as a % of debt	(2.2/12.2) × 100	(6.8/29.8) × 100
Free cash flow after dividends as a % of debt	18.0%	22.8%

A is incorrect. Company Y has a lower debt/capital than Company X, as shown in the following table.

$$\text{Debt divided by Capital (\%)} = \frac{\text{Debt}}{(\text{Debt} + \text{Equity})}$$

	Company X	**Company Y**
Debt	£12.2	£29.8
Capital		
Debt	12.2	29.8
+ Equity	1.3	64.0
Capital	£13.5	£93.8
Debt/Capital (%)	(12.2/13.5) × 100	(29.8/93.8) × 100
Debt/Capital (%)	90.4%	31.8%

B is incorrect because Company Y has a lower debt/EBITDA than Company Y, not higher, as shown in the following table.

	Company X	**Company Y**
Operating income	£1.1	£13.3
EBIT	£1.1	£13.3
plus		
Depreciation	1.0	3.8
Amortization	0.0	0.0
EBITDA	£2.1	£17.1
Debt	£12.2	£29.8
Debt/EBITDA	12.2/2.1	29.8/17.1
Debt/EBITDA	5.81	1.74

42 A is correct. Compared with Company Y, based on both their debt/capital and their ratios of free cash flow after dividends to debt, which are measures of leverage commonly used in credit analysis, Company X is more highly leveraged, as shown in the following table.

$$\text{Debt divided by Capital (\%)} = \frac{\text{Debt}}{(\text{Debt} + \text{Equity})}$$

	Company X	**Company Y**
Debt	£2.2	£29.8
Capital		
Debt	2.2	29.8
+ Equity	4.3	64.0
Capital	£6.5	£93.8

(continued)

Debt/Capital (%)	(12.2/13.5) × 100	(29.8/93.8) × 100
Debt/Capital (%)	90.4%	31.8%

$$\text{Free cash flow after dividends as a \% of debt} = \frac{\text{FCF after dividends}}{\text{Debt}}$$

	Company X	Company Y
Cash flow from operations	£3.3	£14.0
Less		
Net capital expenditures	–0.8	–1.1
Dividends	–0.3	–6.1
Free cash flow after dividends	£2.2	£6.8
Debt	£12.2	£29.8
Free cash flow after dividends as a % of debt	(2.2/12.2) × 100	(6.8/29.8) × 100
Free cash flow after dividends as a % of debt	18.0%	22.8%

43 B is correct. In weak financial markets, including weak markets for equities, credit spreads will widen.

44 B is correct. Weakening economic conditions will push investors to desire a greater risk premium and drive overall credit spreads wider.

45 C is correct. In periods of heavy new issue supply, credit spreads will widen if demand is insufficient.

46 C is correct. Non-sovereign governments typically must balance their operating budgets and lack the discretion to use monetary policy as many sovereigns can.

47 A is correct. Most investors in investment-grade debt focus on spread risk—that is, the effect of changes in spreads on prices and returns—while in high-yield analysis, the focus on default risk is relatively greater.

48 B is correct. Among the most important considerations in sovereign credit analysis is growth and age distribution of population. A relatively young and growing population contributes to growth in GDP and an expanding tax base and relies less on social services, pensions, and health care relative to an older population.

Derivatives

STUDY SESSION

Study Session 15	Derivatives

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to demonstrate a working knowledge of the analysis of derivatives, including forwards, futures, options, and swaps.

Derivatives—financial instruments whose prices are derived from the value of some underlying asset—have become increasingly important for managing financial risk, exploiting investment opportunities, and creating synthetic asset class exposure. As in other security markets, arbitrage and market efficiency play a critical role in establishing prices for these securities.

© 2021 CFA Institute. All rights reserved.

DERIVATIVES

STUDY SESSION

15

Derivatives

This study session builds the conceptual framework for understanding the basic derivatives and derivative markets. Essential features and valuation concepts for forward commitments such as forwards, futures, and swaps and contingent claims such as options are introduced.

READING ASSIGNMENTS

Reading 45	Derivative Markets and Instruments by Don M. Chance, PhD, CFA
Reading 46	Basics of Derivative Pricing and Valuation by Don M. Chance, PhD, CFA

© 2021 CFA Institute. All rights reserved.

READING

45

Derivative Markets and Instruments

by Don M. Chance, PhD, CFA

Don M. Chance, PhD, CFA, is at Louisiana State University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** define a derivative and distinguish between exchange-traded and over-the-counter derivatives;
☐	**b.** contrast forward commitments with contingent claims;
☐	**c.** define forward contracts, futures contracts, options (calls and puts), swaps, and credit derivatives and compare their basic characteristics;
☐	**d.** determine the value at expiration and profit from a long or a short position in a call or put option;
☐	**e.** describe purposes of, and controversies related to, derivative markets;
☐	**f.** explain arbitrage and the role it plays in determining prices and promoting market efficiency.

DERIVATIVES: INTRODUCTION, DEFINITIONS, AND USES

1

a define a derivative and distinguish between exchange-traded and over-the-counter derivatives;

b contrast forward commitments with contingent claims;

Equity, fixed-income, currency, and commodity markets are facilities for trading the basic assets of an economy. Equity and fixed-income securities are claims on the assets of a company. Currencies are the monetary units issued by a government or central bank. Commodities are natural resources, such as oil or gold. These underlying assets are said to trade in **cash markets** or **spot markets** and their prices are sometimes referred to as **cash prices** or **spot prices**, though we usually just refer to them as stock prices, bond prices, exchange rates, and commodity prices. These markets

© 2019 CFA Institute. All rights reserved.

exist around the world and receive much attention in the financial and mainstream media. Hence, they are relatively familiar not only to financial experts but also to the general population.

Somewhat less familiar are the markets for **derivatives**, which are financial instruments that derive their values from the performance of these basic assets. This reading is an overview of derivatives. Subsequent readings will explore many aspects of derivatives and their uses in depth. Among the questions that this first reading will address are the following:

- What are the defining characteristics of derivatives?
- What purposes do derivatives serve for financial market participants?
- What is the distinction between a forward commitment and a contingent claim?
- What are forward and futures contracts? In what ways are they alike and in what ways are they different?
- What are swaps?
- What are call and put options and how do they differ from forwards, futures, and swaps?
- What are credit derivatives and what are the various types of credit derivatives?
- What are the benefits of derivatives?
- What are some criticisms of derivatives and to what extent are they well founded?
- What is arbitrage and what role does it play in a well-functioning financial market?

This reading is organized as follows. Section 1 explores the definition and uses of derivatives and establishes some basic terminology. Section 2 describes derivatives markets. Sections 3–9 categorize and explain types of derivatives. Sections 10 and 11 discuss the benefits and criticisms of derivatives, respectively. Section 12 introduces the basic principles of derivative pricing and the concept of arbitrage.

Derivatives: Definitions and Uses

The most common definition of a derivative reads approximately as follows:

> *A derivative is a financial instrument that derives its performance from the performance of an underlying asset.*

This definition, despite being so widely quoted, can nonetheless be a bit troublesome. For example, it can also describe mutual funds and exchange-traded funds, which would never be viewed as derivatives even though they derive their values from the values of the underlying securities they hold. Perhaps the distinction that best characterizes derivatives is that they usually *transform* the performance of the underlying asset before paying it out in the derivatives transaction. In contrast, with the exception of expense deductions, mutual funds and exchange-traded funds simply pass through the returns of their underlying securities. This transformation of performance is typically understood or implicit in references to derivatives but rarely makes its way into the formal definition. In keeping with customary industry practice, this characteristic will be retained as an implied, albeit critical, factor distinguishing derivatives from mutual funds and exchange-traded funds and some other straight pass-through instruments. Also, note that the idea that derivatives take their *performance* from an underlying asset encompasses the fact that derivatives take their value and certain other characteristics from the underlying asset. Derivatives strategies perform in ways that are derived from the underlying and the specific features of derivatives.

Derivatives are similar to insurance in that both allow for the transfer of risk from one party to another. As everyone knows, insurance is a financial contract that provides protection against loss. The party bearing the risk purchases an insurance policy, which transfers the risk to the other party, the insurer, for a specified period of time. The risk itself does not change, but the party bearing it does. Derivatives allow for this same type of transfer of risk. One type of derivative in particular, the put option, when combined with a position exposed to the risk, functions almost exactly like insurance, but all derivatives can be used to protect against loss. Of course, an insurance contract must specify the underlying risk, such as property, health, or life. Likewise, so do derivatives. As noted earlier, derivatives are associated with an underlying asset. As such, the so-called "underlying asset" is often simply referred to as the **underlying**, whose value is the source of risk. In fact, the underlying need not even be an asset itself. Although common derivatives underlyings are equities, fixed-income securities, currencies, and commodities, other derivatives underlyings include interest rates, credit, energy, weather, and even other derivatives, all of which are not generally thought of as assets. Thus, like insurance, derivatives pay off on the basis of a source of risk, which is often, but not always, the value of an underlying asset. And like insurance, derivatives have a definite life span and expire on a specified date.

Derivatives are created in the form of legal contracts. They involve two parties—the buyer and the seller (sometimes known as the writer)—each of whom agrees to do something for the other, either now or later. The buyer, who purchases the derivative, is referred to as the **long** or the holder because he owns (holds) the derivative and holds a long position. The seller is referred to as the **short** because he holds a short position.[1]

A derivative contract always defines the rights and obligations of each party. These contracts are intended to be, and almost always are, recognized by the legal system as commercial contracts that each party expects to be upheld and supported in the legal system. Nonetheless, disputes sometimes arise, and lawyers, judges, and juries may be required to step in and resolve the matter.

There are two general classes of derivatives. Some provide the ability to lock in a price at which one might buy or sell the underlying. Because they force the two parties to transact in the future at a previously agreed-on price, these instruments are called **forward commitments**. The various types of forward commitments are called forward contracts, futures contracts, and swaps. Another class of derivatives provides *the right but not the obligation* to buy or sell the underlying at a pre-determined price. Because the choice of buying or selling versus doing nothing depends on a particular random outcome, these derivatives are called **contingent claims**. The primary contingent claim is called an **option**. The types of derivatives will be covered in more detail later in this reading and in considerably more depth later in the curriculum.

The existence of derivatives begs the obvious question of what purpose they serve. If one can participate in the success of a company by holding its equity, what reason can possibly explain why another instrument is required that takes its value from the performance of the equity? Although equity and other fundamental markets exist and usually perform reasonably well without derivative markets, it is possible that derivative markets can *improve* the performance of the markets for the underlyings. As you will see later in this reading, that is indeed true in practice.

Derivative markets create beneficial opportunities that do not exist in their absence. Derivatives can be used to create strategies that cannot be implemented with the underlyings alone. For example, derivatives make it easier to go short, thereby benefiting from a decline in the value of the underlying. In addition, derivatives, in and of themselves, are characterized by a relatively high degree of leverage, meaning that participants in derivatives transactions usually have to invest only a small amount of their own capital relative to the value of the underlying. As such, small movements in the underlying can lead to fairly large movements in the amount of money made or lost on the derivative. Derivatives generally trade at lower transaction costs than comparable spot market transactions, are often more liquid than their underlyings, and offer a simple, effective, and low-cost way to transfer risk. For example, a shareholder of a company can reduce or even completely eliminate the market exposure by trading a derivative on the equity. Holders of fixed-income securities can use derivatives to reduce or completely eliminate interest rate risk, allowing them to focus on the credit risk. Alternatively, holders of fixed-income securities can reduce or eliminate the credit risk, focusing more on the interest rate risk. Derivatives permit such adjustments easily and quickly. These features of derivatives are covered in more detail later in this reading.

---

[1] In the financial world, the *long* always benefits from an increase in the value of the instrument he owns, and the *short* always benefits from a decrease in the value of the instrument he has sold. Think of the long as having possession of something and the short as having incurred an obligation to deliver that something.

The types of performance transformations facilitated by derivatives allow market participants to practice more effective risk management. Indeed, the entire field of derivatives, which at one time was focused mostly on the instruments themselves, is now more concerned with the *uses* of the instruments. Just as a carpenter uses a hammer, nails, screws, a screwdriver, and a saw to build something useful or beautiful, a financial expert uses derivatives to manage risk. And just as it is critically important that a carpenter understand how to use these tools, an investment practitioner must understand how to properly use derivatives. In the case of the carpenter, the result is building something useful; in the case of the financial expert, the result is managing financial risk. Thus, like tools, derivatives serve a valuable purpose but like tools, they must be used carefully.

The practice of risk management has taken a prominent role in financial markets. Indeed, whenever companies announce large losses from trading, lending, or operations, stories abound about how poorly these companies managed risk. Such stories are great attention grabbers and a real boon for the media, but they often miss the point that risk management does not guarantee that large losses will not occur. Rather, **risk management** *is the process by which an organization or individual defines the level of risk it wishes to take, measures the level of risk it is taking, and adjusts the latter to equal the former*. Risk management never offers a guarantee that large losses will not occur, and it does not eliminate the possibility of total failure. To do so would typically require that the amount of risk taken be so small that the organization would be effectively constrained from pursuing its primary objectives. Risk taking is inherent in all forms of economic activity and life in general. The possibility of failure is never eliminated.

EXAMPLE 1

Characteristics of Derivatives

1 Which of the following is the best example of a derivative?

- **A** A global equity mutual fund
- **B** A non-callable government bond
- **C** A contract to purchase Apple Computer at a fixed price

2 Which of the following is **not** a characteristic of a derivative?

- **A** An underlying
- **B** A low degree of leverage
- **C** Two parties—a buyer and a seller

3 Which of the following statements about derivatives is **not** true?

A They are created in the spot market.

B They are used in the practice of risk management.

C They take their values from the value of something else.

Solution to 1:

C is correct. Mutual funds and government bonds are not derivatives. A government bond is a fundamental asset on which derivatives might be created, but it is not a derivative itself. A mutual fund can technically meet the definition of a derivative, but as noted in the reading, derivatives transform the value of a payoff of an underlying asset. Mutual funds merely pass those payoffs through to their holders.

Solution to 2:

B is correct. All derivatives have an underlying and must have a buyer and a seller. More importantly, derivatives have high degrees of leverage, not low degrees of leverage.

Solution to 3:

A is correct. Derivatives are used to practice risk management and they take (derive) their values from the value of something else, the underlying. They are not created in the spot market, which is where the underlying trades.

Note also that risk management is a dynamic and ongoing process, reflecting the fact that the risk assumed can be difficult to measure and is constantly changing. As noted, derivatives are tools, indeed *the* tools that make it easier to manage risk. Although one can trade stocks and bonds (the underlyings) to adjust the level of risk, it is almost always more effective to trade derivatives.

Risk management is addressed more directly elsewhere in the CFA curriculum, but the study of derivatives necessarily entails the concept of risk management. In an explanation of derivatives, the focus is usually on the instruments and it is easy to forget the overriding objective of managing risk. Unfortunately, that would be like a carpenter obsessed with his hammer and nails, forgetting that he is building a piece of furniture. It is important to always try to keep an eye on the objective of managing risk.

THE STRUCTURE OF DERIVATIVE MARKETS

2

a define a derivative and distinguish between exchange-traded and over-the-counter derivatives;

b contrast forward commitments with contingent claims;

Having an understanding of equity, fixed-income, and currency markets is extremely beneficial—indeed, quite necessary—in understanding derivatives. One could hardly consider the wisdom of using derivatives on a share of stock if one did not understand the equity markets reasonably well. As you likely know, equities trade on organized exchanges as well as in over-the-counter (OTC) markets. These exchange-traded equity markets—such as the Deutsche Börse, the Tokyo Stock Exchange, and the New York Stock Exchange and its Eurex affiliate—are formal organizational structures that bring buyers and sellers together through market makers, or dealers, to facilitate transactions. Exchanges have formal rule structures and are required to comply with all securities laws.

OTC securities markets operate in much the same manner, with similar rules, regulations, and organizational structures. At one time, the major difference between OTC and exchange markets for securities was that the latter brought buyers and sellers together in a physical location, whereas the former facilitated trading strictly in an electronic manner. Today, these distinctions are blurred because many organized securities exchanges have gone completely to electronic systems. Moreover, OTC securities markets can be formally organized structures, such as NASDAQ, or can merely refer to informal networks of parties who buy and sell with each other, such as the corporate and government bond markets in the United States.

The derivatives world also comprises organized exchanges and OTC markets. Although the derivatives world is also moving toward less distinction between these markets, there are clear differences that are important to understand.

2.1 Exchange-Traded Derivatives Markets

Derivative instruments are created and traded either on an exchange or on the OTC market. Exchange-traded derivatives are standardized, whereas OTC derivatives are customized. To standardize a derivative contract means that its terms and conditions are precisely specified by the exchange and there is very limited ability to alter those terms. For example, an exchange might offer trading in certain types of derivatives that expire only on the third Friday of March, June, September, and December. If a party wanted the derivative to expire on any other day, it would not be able to trade such a derivative on that exchange, nor would it be able to persuade the exchange to create it, at least not in the short run. If a party wanted a derivative on a particular entity, such as a specific stock, that party could trade it on that exchange only if the exchange had specified that such a derivative could trade. Even the magnitudes of the contracts are specified. If a party wanted a derivative to cover €150,000 and the exchange specified that contracts could trade only in increments of €100,000, the party could do nothing about it if it wanted to trade that derivative on that exchange.

This standardization of contract terms facilitates the creation of a more liquid market for derivatives. If all market participants know that derivatives on the euro trade in 100,000-unit lots and that they all expire only on certain days, the market functions more effectively than it would if there were derivatives with many different unit sizes and expiration days competing in the same market at the same time. This standardization makes it easier to provide liquidity. Through designated market makers, derivatives exchanges guarantee that derivatives can be bought and sold.[2]

The cornerstones of the exchange-traded derivatives market are the market makers (or dealers) and the speculators, both of whom typically own memberships on the exchange.[3] The market makers stand ready to buy at one price and sell at a higher price. With standardization of terms and an active market, market makers are often able to buy and sell almost simultaneously at different prices, locking in small, short-term profits—a process commonly known as scalping. In some cases, however, they are unable to do so, thereby forcing them to either hold exposed positions or find other parties with whom they can trade and thus lay off (get rid of) the risk. This is when speculators come in. Although speculators are market participants who are willing to take risks, it is important to understand that being a speculator does not mean the reckless assumption of risk. Although speculators will take large losses at times, good speculators manage those risks by watching their exposures, absorbing market information, and observing the flow of orders in such a manner that they are able to survive and profit. Often, speculators will hedge their risks when they become uncomfortable.

---

[2] It is important to understand that merely being able to buy and sell a derivative, or even a security, does not mean that liquidity is high and that the cost of liquidity is low. Derivatives exchanges guarantee that a derivative can be bought and sold, but they do not guarantee the price. The ask price (the price at which the market maker will sell) and the bid price (the price at which the market maker will buy) can be far apart, which they will be in a market with low liquidity. Hence, such a market can have liquidity, loosely defined, but the cost of liquidity can be quite high. The factors that can lead to low liquidity for derivatives are similar to those for securities: little trading interest and a high level of uncertainty.

[3] Exchanges are owned by their *members*, whose memberships convey the right to trade. In addition, some exchanges are themselves publicly traded corporations whose members are shareholders, and there are also non-member shareholders.

Standardization also facilitates the creation of a clearing and settlement operation. **Clearing** refers to the process by which the exchange verifies the execution of a transaction and records the participants' identities. **Settlement** refers to the related process in which the exchange transfers money from one participant to the other or from a participant to the exchange or vice versa. This flow of money is a critical element of derivatives trading. Clearly, there would be no confidence in markets in which money is not efficiently collected and disbursed. Derivatives exchanges have done an excellent job of clearing and settlement, especially in comparison to securities exchanges. Derivatives exchanges clear and settle all contracts overnight, whereas most securities exchanges require two business days.

The clearing and settlement process of derivative transactions also provides a credit guarantee. If two parties engage in a derivative contract on an exchange, one party will ultimately make money and the other will lose money. Derivatives exchanges use their clearinghouses to provide a guarantee to the winning party that if the loser does not pay, the clearinghouse will pay the winning party. The clearinghouse is able to provide this credit guarantee by requiring a cash deposit, usually called the **margin bond** or **performance bond**, from the participants to the contract. Derivatives clearinghouses manage these deposits, occasionally requiring additional deposits, so effectively that they have never failed to pay in the nearly 100 years they have existed. We will say more about this process later and illustrate how it works.

Exchange markets are said to have **transparency**, which means that full information on all transactions is disclosed to exchanges and regulatory bodies. All transactions are centrally reported within the exchanges and their clearinghouses, and specific laws require that these markets be overseen by national regulators. Although this would seem a strong feature of exchange markets, there is a definite cost. Transparency means a loss of privacy: National regulators can see what transactions have been done. Standardization means a loss of flexibility: A participant can do only the transactions that are permitted on the exchange. Regulation means a loss of both privacy and flexibility. It is not that transparency or regulation is good and the other is bad. It is simply a trade-off.

Derivatives exchanges exist in virtually every developed (and some emerging market) countries around the world. Some exchanges specialize in derivatives and others are integrated with securities exchanges.

Although there have been attempts to create somewhat non-standardized derivatives for trading on an exchange, such attempts have not been particularly successful. Standardization is a critical element by which derivatives exchanges are able to provide their services. We will look at this point again when discussing the alternative to standardization: customized OTC derivatives.

2.2 Over-the-Counter Derivatives Markets

The OTC derivatives markets comprise an informal network of market participants that are willing to create and trade virtually any type of derivative that can legally exist. The backbone of these markets is the set of dealers, which are typically banks. Most of these banks are members of a group called the International Swaps and Derivatives Association (ISDA), a worldwide organization of financial institutions that engage in derivative transactions, primarily as dealers. As such, these markets are sometimes called *dealer markets*. Acting as principals, these dealers informally agree to buy and sell various derivatives. It is *informal* because the dealers are not obligated to do so. Their participation is based on a desire to profit, which they do by purchasing at one price and selling at a higher price. Although it might seem that a dealer who can "buy low, sell high" could make money easily, the process in practice is not that simple. Because OTC instruments are not standardized, a dealer cannot expect to buy a derivative at one price and simultaneously sell it to a different party who happens to want to buy the same derivative at the same time and at a higher price.

To manage the risk they assume by buying and selling customized derivatives, OTC derivatives dealers typically hedge their risks by engaging in alternative but similar transactions that pass the risk on to other parties. For example, if a company comes to a dealer to buy a derivative on the euro, the company would effectively be transferring the risk of the euro to the dealer. The dealer would then attempt to lay off (get rid of) that risk by engaging in an alternative but similar transaction that would transfer the risk to another party. This hedge might involve another derivative on the euro or it might simply be a transaction in the euro itself. Of course, that begs the question of why the company could not have laid off the risk itself and avoided the dealer. Indeed, some can and do, but laying off risk is not simple. Unable to find identical offsetting transactions, dealers usually have to find *similar* transactions with which they can lay off the risk. Hedging one derivative with a different kind of derivative on the same underlying is a similar but not identical transaction. It takes specialized knowledge and complex models to be able to do such transactions effectively, and dealers are more capable of doing so than are ordinary companies. Thus, one might think of a dealer as a middleman, a sort of financial wholesaler using its specialized knowledge and resources to facilitate the transfer of risk. In the same manner that one could theoretically purchase a consumer product from a manufacturer, a network of specialized middlemen and retailers is often a more effective method.

Because of the customization of OTC derivatives, there is a tendency to think that the OTC market is less liquid than the exchange market. That is not necessarily true. Many OTC instruments can easily be created and then essentially offset by doing the exact opposite transaction, often with the same party. For example, suppose Corporation A buys an OTC derivative from Dealer B. Before the expiration date, Corporation A wants to terminate the position. It can return to Dealer B and ask to sell a derivative with identical terms. Market conditions will have changed, of course, and the value of the derivative will not be the same, but the transaction can be conducted quite easily with either Corporation A or Dealer B netting a gain at the expense of the other. Alternatively, Corporation A could do this transaction with a different dealer, the result of which would remove exposure to the underlying risk but would leave two transactions open and some risk that one party would default to the other. In contrast to this type of OTC liquidity, some exchange-traded derivatives have very little trading interest and thus relatively low liquidity. Liquidity is always driven by trading interest, which can be strong or weak in both types of markets.

OTC derivative markets operate at a lower degree of regulation and oversight than do exchange-traded derivative markets. In fact, until around 2010, it could largely be said that the OTC market was essentially unregulated. OTC transactions could be executed with only the minimal oversight provided through laws that regulated the parties themselves, not the specific instruments. Following the financial crisis of 2007–2009, new regulations began to blur the distinction between OTC and exchange-listed markets. In both the United States (the Wall Street Reform and Consumer Protection Act of 2010, commonly known as the Dodd–Frank Act) and Europe (the Regulation of the European Parliament and of the Council on OTC Derivatives, Central

Counterparties, and Trade Repositories), regulations are changing the characteristics of OTC markets. In general, world policy-makers have advanced an agenda to make global derivatives markets more resilient and robust, pursuing increased transparency and lowered systemic risk.

When the full implementation of these new laws takes place, a number of OTC transactions will have to be cleared through central clearing agencies, information on most OTC transactions will need to be reported to regulators, and entities that operate in the OTC market will be more closely monitored. There are, however, quite a few exemptions that cover a significant percentage of derivative transactions. Clearly, the degree of OTC regulation, although increasing in recent years, is still lighter than that of exchange-listed market regulation. Many transactions in OTC markets will retain a degree of privacy with lower transparency, and most importantly, the OTC markets will remain considerably more flexible than the exchange-listed markets.

EXAMPLE 2

Exchange-Traded versus Over-the-Counter Derivatives

1 Which of the following characteristics is **not** associated with exchange-traded derivatives?

   A Margin or performance bonds are required.

   B The exchange guarantees all payments in the event of default.

   C All terms except the price are customized to the parties' individual needs.

2 Which of the following characteristics is associated with over-the-counter derivatives?

   A Trading occurs in a central location.

   B They are more regulated than exchange-listed derivatives.

   C They are less transparent than exchange-listed derivatives.

3 Market makers earn a profit in both exchange and over-the-counter derivatives markets by:

   A charging a commission on each trade.

   B a combination of commissions and markups.

   C buying at one price, selling at a higher price, and hedging any risk.

4 Which of the following statements *most* accurately describes exchange-traded derivatives relative to over-the-counter derivatives? Exchange-traded derivatives are more likely to have:

   A greater credit risk.

   B standardized contract terms.

   C greater risk management uses.

Solution to 1:

C is correct. Exchange-traded contracts are standardized, meaning that the exchange determines the terms of the contract except the price. The exchange guarantees against default and requires margins or performance bonds.

Solution to 2:

C is correct. OTC derivatives have a lower degree of transparency than exchange-listed derivatives. Trading does not occur in a central location but, rather, is quite dispersed. Although new national securities laws are tightening the regulation of OTC derivatives, the degree of regulation is less than that of exchange-listed derivatives.

Solution to 3:

C is correct. Market makers buy at one price (the bid), sell at a higher price (the ask), and hedge whatever risk they otherwise assume. Market makers do not charge a commission. Hence, A and B are both incorrect.

Solution to 4:

B is correct. Standardization of contract terms is a characteristic of exchange-traded derivatives. A is incorrect because credit risk is well-controlled in exchange markets. C is incorrect because the risk management uses are not limited by being traded over the counter.

3 TYPES OF DERIVATIVES: INTRODUCTION, FORWARD CONTRACTS

b contrast forward commitments with contingent claims;

c define forward contracts, futures contracts, options (calls and puts), swaps, and credit derivatives and compare their basic characteristics;

As previously stated, derivatives fall into two general classifications: forward commitments and contingent claims. The factor that distinguishes forward commitments from contingent claims is that forward commitments *obligate* the parties to engage in a transaction at a future date on terms agreed upon in advance, whereas contingent claims provide one party the *right but not the obligation* to engage in a future transaction on terms agreed upon in advance.

3.1 Forward Commitments

Forward commitments are contracts entered into at one point in time that require both parties to engage in a transaction at a later point in time (the expiration) on terms agreed upon at the start. The parties establish the identity and quantity of the underlying, the manner in which the contract will be executed or settled when it expires, and the fixed price at which the underlying will be exchanged. This fixed price is called the **forward price**.

As a hypothetical example of a forward contract, suppose that today Markus and Johannes enter into an agreement that Markus will sell his BMW to Johannes for a price of €30,000. The transaction will take place on a specified date, say, 180 days from today. At that time, Markus will deliver the vehicle to Johannes's home and Johannes will give Markus a bank-certified check for €30,000. There will be no recourse, so if the vehicle has problems later, Johannes cannot go back to Markus for compensation. It should be clear that both Markus and Johannes must do their due diligence and carefully consider the reliability of each other. The car could have serious quality issues and Johannes could have financial problems and be unable to pay the €30,000. Obviously, the transaction is essentially unregulated. Either party could renege on his obligation, in response to which the other party could go to court, provided a formal contract exists and is carefully written. Note finally that one of the two parties is likely to end up gaining and the other losing, depending on the secondary market price of this type of vehicle at expiration of the contract.

This example is quite simple but illustrates the essential elements of a forward contract. In the financial world, such contracts are very carefully written, with legal provisions that guard against fraud and require extensive credit checks. Now let us take a deeper look at the characteristics of forward contracts.

3.1.1 *Forward Contracts*

The following is the formal definition of a forward contract:

> *A forward contract is an over-the-counter derivative contract in which two parties agree that one party, the buyer, will purchase an underlying asset from the other party, the seller, at a later date at a fixed price they agree on when the contract is signed.*

In addition to agreeing on the price at which the underlying asset will be sold at a later date, the two parties also agree on several other matters, such as the specific identity of the underlying, the number of units of the underlying that will be delivered, and where the future delivery will occur. These are important points but relatively minor in this discussion, so they can be left out of the definition to keep it uncluttered.

As noted earlier, a forward contract is a commitment. Each party agrees that it will fulfill its responsibility at the designated future date. Failure to do so constitutes a default and the non-defaulting party can institute legal proceedings to enforce performance. It is important to recognize that although either party could default to the other, only one party at a time can default. The party owing the greater amount could default to the other, but the party owing the lesser amount cannot default because its claim on the other party is greater. The amount owed is always based on the net owed by one party to the other.

To gain a better understanding of forward contracts, it is necessary to examine their payoffs. As noted, forward contracts—and indeed all derivatives—take (derive) their payoffs from the performance of the underlying asset. To illustrate the payoff of a forward contract, start with the assumption that we are at time $t = 0$ and that the forward contract expires at a later date, time $t = T$.[4] The spot price of the underlying asset at time 0 is S_0 and at time T is S_T. Of course, when we initiate the contract at time 0, we do not know what S_T will ultimately be. Remember that the two parties, the buyer and the seller, are going long and short, respectively.

At time $t = 0$, the long and the short agree that the short will deliver the asset to the long at time T for a price of $F_0(T)$. The notation $F_0(T)$ denotes that this value is established at time 0 and applies to a contract expiring at time T. $F_0(T)$ is the forward price.

So, let us assume that the buyer enters into the forward contract with the seller for a price of $F_0(T)$, with delivery of one unit of the underlying asset to occur at time T. Now, let us roll forward to time T, when the price of the underlying is S_T. The long is obligated to pay $F_0(T)$, for which he receives an asset worth S_T. If $S_T > F_0(T)$, it is clear that the transaction has worked out well for the long. He paid $F_0(T)$ and receives something of greater value. Thus, the contract effectively pays off $S_T - F_0(T)$ to the long, which is the value of the contract at expiration. The short has the mirror image of the long. He is required to deliver the asset worth S_T and accept a smaller amount, $F_0(T)$. The contract has a payoff for him of $F_0(T) - S_T$, which is negative. Even if the asset's value, S_T, is less than the forward price, $F_0(T)$, the payoffs are still $S_T - F_0(T)$ for the long and $F_0(T) - S_T$ for the short. We can consolidate these results by writing the short's payoff as the negative of the long's, $-[S_T - F_0(T)]$, which serves as a useful reminder that the long and the short are engaged in a zero-sum game, which is a type of competition in which one participant's gains are the other's losses. Although both lose a modest amount in the sense of both having some costs to engage in the transaction, these costs are relatively small and worth ignoring for our purposes at this time. In addition, it is worthwhile to note how derivatives transform the performance of the underlying. The gain from owning the underlying would be $S_T - S_0$, whereas the gain from owning the forward contract would be $S_T - F_0(T)$. Both figures are driven by S_T, the price of the underlying at expiration, but they are not the same.

---

[4] Such notations as $t = 0$ and $t = T$ are commonly used in explaining derivatives. To indicate that $t = 0$ simply means that we initiate a contract at an imaginary time designated like a counter starting at zero. To indicate that the contract expires at $t = T$ simply means that at some future time, designated as T, the contract expires. Time T could be a certain number of days from now or a fraction of a year later or T years later. We will be more specific in later readings that involve calculations. For now, just assume that $t = 0$ and $t = T$ are two dates—the initiation and the expiration—of the contract.

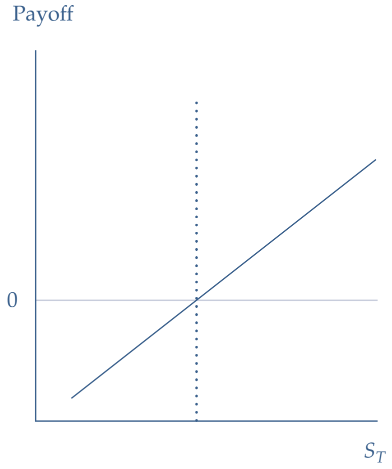
For an example, a buyer enters a forward contract to buy gold at a price of $F_0(T)$ = \$1,312.90 per ounce four months from now. The spot price of gold is S_0 = \$1,207.40 per ounce. Four months in the future, the price of the underlying gold is S_T = \$1,275.90 per ounce. The buyer's gain from the forward contract, the payoff from the contract, is the value of gold (at maturity) less the forward price: $S_T - F_0(T)$ = 1,275.90 – 1,312.90 = –\$37.00 per ounce. Because the value of gold when the contract matures is less than the forward price, $S_T < F_0(T)$, the buyer has incurred a loss. Notably, the forward contract seller has a contract payoff, +\$37.00, that is the negative of that of the contract buyer. The gain on owning the underlying, which is $S_T - S_0$ = 1,275.90 – 1,207.40 = \$68.40, differs from the gain (–\$37.00) on the forward contract.

The buyer also enters a forward contract to buy oil at a price of $F_0(T)$ = \$71.86 per barrel four months from now. The spot price of oil is S_0 = \$71.11 per barrel. Four months in the future, the price of the underlying oil is S_T = \$80.96 per barrel. The buyer's gain from the forward contract, the payoff from the contract, is the value of oil less the forward price: $S_T - F_0(T)$ = 80.96 – 71.86 = \$9.10 per barrel. Unlike the forward contract on gold above, because the value of oil when the contract matures is greater than the forward price, $S_T > F_0(T)$, the buyer of the forward contract realizes a gain.

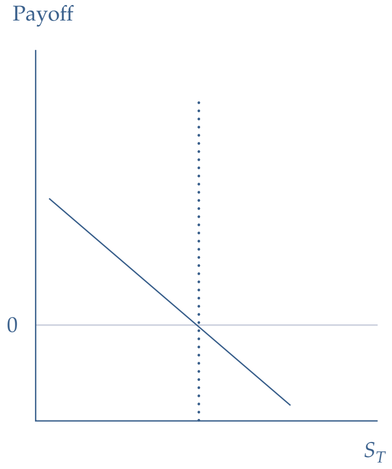
Exhibit 1 illustrates the payoffs from both buying and selling a forward contract.

Exhibit 1 Payoffs from a Forward Contract

A. Payoff from Buying = $S_T - F_0(T)$



B. Payoff from Selling = $-[S_T - F_0(T)]$



The long hopes the price of the underlying will rise above the forward price, $F_0(T)$, whereas the short hopes the price of the underlying will fall below the forward price. Except in the extremely rare event that the underlying price at T equals the forward price, there will ultimately be a winner and a loser.

An important element of forward contracts is that no money changes hands between parties when the contract is initiated. Unlike in the purchase and sale of an asset, there is no value exchanged at the start. The buyer does not pay the seller some money and obtain something. In fact, forward contracts have zero value at the start. They are neither assets nor liabilities. As you will learn in later readings, their values will deviate from zero later as prices move. Forward contracts will almost always have non-zero values at expiration.

As noted previously, the primary purpose of derivatives is for risk management. Although the uses of forward contracts are covered in depth later in the curriculum, there are a few things to note here about the purposes of forward contracts. It should be apparent that locking in the future buying or selling price of an underlying asset can be extremely attractive for some parties. For example, an airline anticipating the purchase of jet fuel at a later date can enter into a forward contract to buy the fuel at a price agreed upon when the contract is initiated. In so doing, the airline has hedged its cost of fuel. Thus, forward contracts can be structured to create a perfect hedge, providing an assurance that the underlying asset can be bought or sold at a price known when the contract is initiated. Likewise, speculators, who ultimately assume the risk laid off by hedgers, can make bets on the direction of the underlying asset without having to invest the money to purchase the asset itself.

Finally, forward contracts need not specifically settle by delivery of the underlying asset. They can settle by an exchange of cash. These contracts—called **non-deliverable forwards** (NDFs), **cash-settled forwards**, or **contracts for differences**—have the same economic effect as do their delivery-based counterparts. For example, for a physical delivery contract, if the long pays $F_0(T)$ and receives an asset worth S_T, the contract is worth $S_T - F_0(T)$ to the long at expiration. A non-deliverable forward contract would have the short simply pay cash to the long in the amount of $S_T - F_0(T)$. The long would not take possession of the underlying asset, but if he wanted the asset, he could purchase it in the market for its current price of S_T. Because he received a cash settlement in the amount of $S_T - F_0(T)$, in buying the asset the long would have to pay out only $S_T - [S_T - F_0(T)]$, which equals $F_0(T)$. Thus, the long could acquire the asset, effectively paying $F_0(T)$, exactly as the contract promised. Transaction costs do make cash settlement different from physical delivery, but this point is relatively minor and can be disregarded for our purposes here.

As previously mentioned, forward contracts are OTC contracts. There is no formal forward contract exchange. Nonetheless, there are exchange-traded variants of forward contracts, which are called futures contracts or just futures.

4 TYPES OF DERIVATIVES: FUTURES

b contrast forward commitments with contingent claims;

c define forward contracts, futures contracts, options (calls and puts), swaps, and credit derivatives and compare their basic characteristics;

Futures contracts are specialized versions of forward contracts that have been standardized and that trade on a futures exchange. By standardizing these contracts and creating an organized market with rules, regulations, and a central clearing facility, the futures markets offer an element of liquidity and protection against loss by default.

Formally, a futures contract is defined as follows:

> *A futures contract is a standardized derivative contract created and traded on a futures exchange in which two parties agree that one party, the buyer, will purchase an underlying asset from the other party, the seller, at a later date and at a price agreed on by the two parties when the contract is initiated and in which there is a daily settling of gains and losses and a credit guarantee by the futures exchange through its clearinghouse.*

First, let us review what standardization means. Recall that in forward contracts, the parties customize the contract by specifying the underlying asset, the time to expiration, the delivery and settlement conditions, and the quantity of the underlying, all according to whatever terms they agree on. These contracts are not traded on an exchange. As noted, the regulation of OTC derivatives markets is increasing, but these contracts are not subject to the traditionally high degree of regulation that applies to securities and futures markets. Futures contracts first require the existence of a futures exchange, a legally recognized entity that provides a market for trading these contracts. Futures exchanges are highly regulated at the national level in all countries. These exchanges specify that only certain contracts are authorized for trading. These contracts have specific underlying assets, times to expiration, delivery and settlement conditions, and quantities. The exchange offers a facility in the form of a physical location and/or an electronic system as well as liquidity provided by authorized market makers.

Probably the most important distinctive characteristic of futures contracts is the daily settlement of gains and losses and the associated credit guarantee provided by the exchange through its clearinghouse. When a party buys a futures contract, it commits to purchase the underlying asset at a later date and at a price agreed upon when the contract is initiated. The counterparty (the seller) makes the opposite commitment, an agreement to sell the underlying asset at a later date and at a price agreed upon when the contract is initiated. The agreed-upon price is called the **futures price**. Identical contracts trade on an ongoing basis at different prices, reflecting the passage of time and the arrival of new information to the market. Thus, as the futures price changes, the parties make and lose money. Rising (falling) prices, of course, benefit (hurt) the long and hurt (benefit) the short. At the end of each day, the clearinghouse engages in a practice called **mark to market**, also known as the **daily settlement**. The clearinghouse determines an average of the final futures trades of the day and designates that price as the **settlement price**. All contracts are then said to be *marked to the settlement price*. For example, if the long purchases the contract during the day at a futures price of £120 and the settlement price at the end of the day is £122, the long's account would be marked for a gain of £2. In other words, the long has made a profit of £2 and that amount is credited to his account, with the money coming from the account of the short, who has lost £2. Naturally, if the futures price decreases, the long loses money and is charged with that loss, and the money is transferred to the account of the short.[5]

The account is specifically referred to as a **margin** account. Of course, in equity markets, margin accounts are commonly used, but there are significant differences between futures margin accounts and equity margin accounts. Equity margin accounts involve the extension of credit. An investor deposits part of the cost of the stock and borrows the remainder at a rate of interest. With futures margin accounts, both parties deposit a required minimum sum of money, but the remainder of the price is not borrowed. This required margin is typically less than 10% of the futures price, which is considerably less than in equity margin trading. In the example above, let us assume that the required margin is £10, which is referred to as the **initial margin**. Both the long and the short put that amount into their respective margin accounts. This money is deposited there to support the trade, not as a form of equity, with the remaining amount borrowed. There is no formal loan created as in equity markets. A futures margin is more of a performance bond or good faith deposit, terms that were previously mentioned. It is simply an amount of money put into an account that covers possible future losses.

Associated with each initial margin is another figure called the **maintenance margin**. The maintenance margin is the amount of money that each participant must maintain in the account after the trade is initiated, and it is always significantly lower than the initial margin. Let us assume that the maintenance margin in this example is £6. If the buyer's account is marked to market with a credit of £2, his margin balance moves to £12, while the seller's account is charged £2 and his balance moves to £8. The clearinghouse then compares each participant's balance with the maintenance margin. At this point, both participants more than meet the maintenance margin.

---

[5] The actual amount of money charged and credited depends on the contract size and the number of contracts. A price of £120 might actually refer to a contract that has a standard size of £100,000. Thus, £120 might actually mean 120% of the standard size, or £120,000. In addition, the parties are likely to hold more than one contract. Hence, the gain of £2 referred to in the text might really mean £2,000 (122% minus 120% times the £100,000 standard size) times the number of contracts held by the party.

Let us say, however, that the price continues to move in the long's favor and, therefore, against the short. A few days later, assume that the short's balance falls to £4, which is below the maintenance margin requirement of £6. The short will then get a **margin call**, which is a request to deposit additional funds. The amount that the short has to deposit, however, is *not* the £2 that would bring his balance up to the maintenance margin. Instead, the short must deposit enough funds to bring the balance up to the initial margin. So, the short must come up with £6. The purpose of this rule is to get the party's position significantly above the minimum level and provide some breathing room. If the balance were brought up only to the maintenance level, there would likely be another margin call soon. A party can choose not to deposit additional funds, in which case the party would be required to close out the contract as soon as possible and would be responsible for any additional losses until the position is closed.

As with forward contracts, neither party pays any money to the other when the contract is initiated. Value accrues as the futures price changes, but at the end of each day, the mark-to-market process settles the gains and losses, effectively resetting the value for each party to zero.

The clearinghouse moves money between the participants, crediting gains to the winners and charging losses to the losers. By doing this on a daily basis, the gains and losses are typically quite small, and the margin balances help ensure that the clearinghouse will collect from the party losing money. As an extra precaution, in fast-moving markets, the clearinghouse can make margin calls during the day, not just at the end of the day. Yet there still remains the possibility that a party could default. A large loss could occur quickly and consume the entire margin balance, with additional money owed.[6] If the losing party cannot pay, the clearinghouse provides a guarantee that it will make up the loss, which it does by maintaining an insurance fund. If that fund were depleted, the clearinghouse could levy a tax on the other market participants, though that has never happened.

Some futures contracts contain a provision limiting price changes. These rules, called **price limits**, establish a band relative to the previous day's settlement price, within which all trades must occur. If market participants wish to trade at a price above the upper band, trading stops, which is called **limit up**, until two parties agree on a trade at a price lower than the upper limit. Likewise, if market participants wish to trade at a price below the lower band, which is called **limit down**, no trade can take place until two parties agree to trade at a price above the lower limit. When the market hits these limits and trading stops, it is called **locked limit**. Typically, the exchange rules provide for an expansion of the limits the next day. These price limits, which may be somewhat objectionable to proponents of free markets, are important in helping the clearinghouse manage its credit exposure. Just because two parties wish to trade a futures contract at a price beyond the limits does not mean they should be allowed to do so. The clearinghouse is a third participant in the contract, guaranteeing to each party that it ensures against the other party defaulting. Therefore, the clearinghouse has a vested interest in the price and considerable exposure. Sharply moving prices make it more difficult for the clearinghouse to collect from the parties losing money.

Most participants in futures markets buy and sell contracts, collecting their profits and incurring their losses, with no ultimate intent to make or take delivery of the underlying asset. For example, the long may ultimately sell her position before expiration. When a party re-enters the market at a later date but before expiration and engages in the opposite transaction—a long selling her previously opened contract or a short buying her previously opened contract—the transaction is referred to as an offset. The clearinghouse marks the contract to the current price relative to the previous settlement price and closes out the participant's position.

---

[6] For example, let us go back to when the short had a balance of £4, which is £2 below the maintenance margin and £6 below the initial margin. The short will get a margin call, but suppose he elects not to deposit additional funds and requests that his position be terminated. In a fast-moving market, the price might increase more than £4 before his broker can close his position. The remaining balance of £4 would then be depleted, and the short would be responsible for any additional losses.

At any given time, the number of outstanding contracts is called the **open interest**. Each contract counted in the open interest has a long and a corresponding short. The open interest figure changes daily as some parties open up new positions, while other parties offset their old positions. It is theoretically possible that all longs and shorts offset their positions before expiration, leaving no open interest when the contract expires, but in practice there is nearly always some open interest at expiration, at which time there is a final delivery or settlement.

When discussing forward contracts, we noted that a contract could be written such that the parties engage in physical delivery or cash settlement at expiration. In the futures markets, the exchange specifies whether physical delivery or cash settlement applies. In physical delivery contracts, the short is required to deliver the underlying asset at a designated location and the long is required to pay for it. Delivery replaces the mark-to-market process on the final day. It also ensures an important principle that you will use later: *The futures price converges to the spot price at expiration.* Because the short delivers the actual asset and the long pays the current spot price for it, the futures price at expiration has to be the spot price at that time. Alternatively, a futures contract initiated right at the instant of expiration is effectively a spot transaction and, therefore, the futures price at expiration must equal the spot price. Following this logic, in cash settlement contracts, there is a final mark to market, with the futures price formally set to the spot price, thereby ensuring automatic convergence.

In discussing forward contracts, we described the process by which they pay off as the spot price at expiration minus the forward price, $S_T - F_0(T)$, the former determined at expiration and the latter agreed upon when the contract is initiated. Futures contracts basically pay off the same way, but there is a slight difference. Let us say the contract is initiated on Day 0 and expires on Day T. The intervening days are designated Days 1, 2, …, T. The initial futures price is designated $f_0(T)$ and the daily settlement prices on Days 1, 2, …, T are designated $f_1(T), f_2(T), \ldots, f_T(T)$. There are, of course, futures prices within each trading day, but let us focus only on the settlement prices for now. For simplicity, let us assume that the long buys at the settlement price on Day 0 and holds the position all the way to expiration. Through the mark-to-market process, the cash flows to the account of the long will be

$$\begin{aligned}
&f_1(T) - f_0(T) \text{ on Day 1}\\
&f_2(T) - f_1(T) \text{ on Day 2}\\
&f_3(T) - f_2(T) \text{ on Day 3}\\
&\ldots\\
&f_T(T) - f_{T-1}(T) \text{ on Day } T
\end{aligned}$$

These add up to

$$f_T(T) - f_0(T) \text{ on Day } T.$$

And because of the convergence of the final futures price to the spot price,

$$f_T(T) - f_0(T) = S_T - f_0(T),$$

which is the same as with forward contracts.[7] Note, however, that the timing of these profits is different from that of forwards. Forward contracts realize the full amount, $S_T - f_0(T)$, at expiration, whereas futures contracts realize this amount in parts on a day-to-day basis. Naturally, the time value of money principle says that these are not equivalent amounts of money. But the differences tend to be small, particularly in low-interest-rate environments, some of these amounts are gains and some are losses, and most futures contracts have maturities of less than a year.

---

[7] Because of this equivalence, we will not specifically illustrate the profit graphs of futures contracts. You can generally treat them the same as those of forwards, which were shown in Exhibit 1.

But the near equivalence of the profits from a futures and a forward contract disguises an important distinction between these types of contracts. In a forward contact, with the entire payoff made at expiration, a loss by one party can be large enough to trigger a default. Hence, forward contracts are subject to default and require careful consideration of the credit quality of the counterparties. Because futures contracts settle gains and collect losses daily, the amounts that could be lost upon default are much smaller and naturally give the clearinghouse much greater flexibility to manage the credit risk it assumes.

Unlike forward markets, futures markets are highly regulated at the national level. National regulators are required to approve new futures exchanges and even new contracts proposed by existing exchanges as well as changes in margin requirements, price limits, and any significant changes in trading procedures. Violations of futures regulations can be subject to governmental prosecution. In addition, futures markets are far more transparent than forward markets. Futures prices, volume, and open interest are widely reported and easily obtained. Futures prices of nearby expiring contracts are often used as proxies for spot prices, particularly in decentralized spot markets, such as gold, which trades in spot markets all over the world.

In spite of the advantages of futures markets over forward markets, forward markets also have advantages over futures markets. Transparency is not always a good thing. Forward markets offer more privacy and fewer regulatory encumbrances. In addition, forward markets offer more flexibility. With the ability to tailor contracts to the specific needs of participants, forward contracts can be written exactly the way the parties want. In contrast, the standardization of futures contracts makes it more difficult for participants to get exactly what they want, even though they may get close substitutes. Yet, futures markets offer a valuable credit guarantee.

Like forward markets, futures markets can be used for hedging or speculation. For example, a jewelry manufacturer can buy gold futures, thereby hedging the price it will have to pay for one of its key inputs. Although it is more difficult to construct a futures strategy that hedges perfectly than to construct a forward strategy that does so, futures offer the benefit of the credit guarantee. It is not possible to argue that futures are better than forwards or vice versa. Market participants always trade off advantages against disadvantages. Some participants prefer futures, and some prefer forwards. Some prefer one over the other for certain risks and the other for other risks. Some might use one for a particular risk at a point in time and a different instrument for the same risk at another point in time. The choice is a matter of taste and constraints.

The third and final type of forward commitment we will cover is swaps. They go a step further in committing the parties to buy and sell something at a later date: They obligate the parties to a sequence of multiple purchases and sales.

5

TYPES OF DERIVATIVES: SWAPS

b contrast forward commitments with contingent claims;

c define forward contracts, futures contracts, options (calls and puts), swaps, and credit derivatives and compare their basic characteristics;

The concept of a swap is that two parties exchange (swap) a series of cash flows. One set of cash flows is variable or floating and will be determined by the movement of an underlying asset or rate. The other set of cash flows can be variable and determined by a different underlying asset or rate, or it can be fixed. Formally, a swap is defined as follows:

> *A swap is an over-the-counter derivative contract in which two parties agree to exchange a series of cash flows whereby one party pays a variable series that will be determined by an underlying asset or rate and the other party pays either (1) a variable series determined by a different underlying asset or rate or (2) a fixed series.*

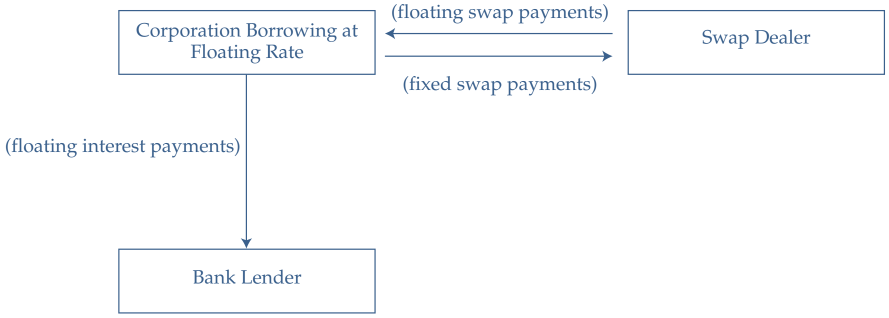
As with forward contracts, swap contracts also contain other terms—such as the identity of the underlying, the relevant payment dates, and the payment procedure—that are negotiated between the parties and written into the contract. A swap is a bit more like a forward contract than a futures contract in that it is an OTC contract, so it is privately negotiated and subject to default. Nonetheless, the similarities between futures and forwards apply to futures and swaps and, indeed, combinations of futures contracts expiring at different dates are often compared to swaps.

As with forward contracts, either party can default but only one party can default at a particular time. The money owed is always based on the net owed by one party to the other. Hence, the party owing the lesser amount cannot default to the party owing the greater amount. Only the latter can default, and the amount it owes is the net of what it owes and what is owed to it, which is also true with forwards.

Swaps are relatively young financial instruments, having been created only in the early 1980s. Thus, it may be somewhat surprising to learn that the swap is the most widely used derivative, a likely result of its simplicity and embracement by the corporate world. The most common swap is the **fixed-for-floating interest rate swap**. In fact, this type of swap is so common that it is often called a "plain vanilla swap" or just a "vanilla swap," owing to the notion that vanilla ice cream is considered plain (albeit tasty).

Let us examine a scenario in which the vanilla interest rate swap is frequently used. Suppose a corporation borrows from a bank at a floating rate. It would prefer a fixed rate, which would enable it to better anticipate its cash flow needs in making its interest payments.[8] The corporation can effectively convert its floating-rate loan to a fixed-rate loan by adding a swap, as shown in Exhibit 2.

Exhibit 2 Using an Interest Rate Swap to Convert a Floating-Rate Loan to a Fixed-Rate Loan



---

[8] Banks prefer to make floating-rate loans because their own funding is typically short term and at floating rates. Thus, their borrowing rates reset frequently, giving them a strong incentive to pass that risk on to their customers through floating-rate loans.

The interest payments on the loan are tied to a specific floating rate. For a dollar-based loan, that rate has historically been US dollar Libor.[9] The payments would be based on the rate from the Libor market on a specified reset date times the loan balance times a factor reflecting the number of days in the current interest calculation period. The actual payment is made at a later date. Thus, for a loan balance of, say, $10 million with monthly payments, the rate might be based on Libor on the first business day of the month, with interest payable on the first business day of the next month, which is the next reset date, and calculated as $10 million times the rate times 30/360. The 30/360 convention, an implicit assumption of 30 days in a month, is common but only one of many interest calculation conventions used in the financial world. Often, "30" is replaced by the exact number of days since the last interest payment. The use of a 360-day year is a common assumption in the financial world, which originated in the pre-calculator days when an interest rate could be multiplied by a number like 30/360, 60/360, 90/360, etc., more easily than if 365 were used.

Whatever the terms of the loan are, the terms of the swap are typically set to match those of the loan. Thus, a Libor-based loan with monthly payments based on the 30/360 convention would be matched with a swap with monthly payments based on Libor and the 30/360 convention and the same reset and payment dates. Although the loan has an actual balance (the amount owed by borrower to creditor), the swap does not have such a balance owed by one party to the other. Thus, it has no principal, but it does have a balance of sorts, called the **notional principal**, which ordinarily matches the loan balance. A loan with only one principal payment, the final one, will be matched with a swap with a fixed notional principal. An amortizing loan, which has a declining principal balance, will be matched with a swap with a pre-specified declining notional principal that matches the loan balance.

As with futures and forwards, no money changes hands at the start; thus, the value of a swap when initiated must be zero. The fixed rate on the swap is determined by a process that forces the value to zero, a procedure that will be covered later in the curriculum. As market conditions change, the value of a swap will deviate from zero, being positive to one party and negative to the other.

As with forward contracts, swaps are subject to default, but because the notional amount of a swap is not typically exchanged, the credit risk of a swap is much less than that of a loan.[10] The only money passing from one party to the other is the net difference between the fixed and floating interest payments. In fact, the parties do not even pay each other. Only one party pays the other, as determined by the net of the greater amount owed minus the lesser amount. This does not mean that swaps are not subject to a potentially large amount of credit risk. At a given point in time, one party could default, effectively owing the value of all remaining payments, which could substantially exceed the value that the non-defaulting party owes to the defaulting party. Thus, there is indeed credit risk in a swap. This risk must be managed by careful analysis before the transaction and by the potential use of such risk-mitigating measures as collateral.

There are also interest rate swaps in which one party pays on the basis of one interest rate and the other party pays on the basis of a different interest rate. For example, one party might make payments at Libor, whereas the other might make payments on the basis of the U. S. Treasury bill rate. The difference between Libor and the T-bill rate, often called the TED spread (T-bills versus Eurodollar), is a measure of the credit risk premium of London banks, which have historically borrowed short term at Libor, versus that of the U. S. government, which borrows short term at the T-bill rate. This transaction is called a basis swap. There are also swaps in which the floating rate is set as an average rate over the period, in accordance with the convention for many loans. Some swaps, called overnight indexed swaps, are tied to a Fed funds–type rate, reflecting the rate at which banks borrow overnight. As we will cover later, there are many other different types of swaps that are used for a variety of purposes. The plain vanilla swap is merely the simplest and most widely used.

---

[9] Libor is being phased out, as the panel of banks will no longer be required to submit quotations after 2021. In anticipation of this, market participants and regulators have been working to develop alternative reference rates.

[10] It is possible that the notional principal will be exchanged in a currency swap, whereby each party makes a series of payments to the other in different currencies. Whether the notional principal is exchanged depends on the purpose of the swap. This point will be covered later in the curriculum. At this time, you should see that it would be fruitless to exchange notional principals in an interest rate swap because that would mean each party would give the other the same amount of money when the transaction is initiated and re-exchange the same amount of money when the contract terminates.

Because swaps, forwards, and futures are forward commitments, they can all accomplish the same thing. One could create a series of forwards or futures expiring at a set of dates that would serve the same purpose as a swap. Although swaps are better suited for risks that involve multiple payments, at its most fundamental level, a swap is more or less just a series of forwards and, acknowledging the slight differences discussed above, more or less just a series of futures.

EXAMPLE 3

Forward Contracts, Futures Contracts, and Swaps

1 Which of the following characterizes forward contracts and swaps but **not** futures?

  A They are customized.

  B They are subject to daily price limits.

  C Their payoffs are received on a daily basis.

2 Which of the following distinguishes forwards from swaps?

  A Forwards are OTC instruments, whereas swaps are exchange traded.

  B Forwards are regulated as futures, whereas swaps are regulated as securities.

  C Swaps have multiple payments, whereas forwards have only a single payment.

3 Which of the following occurs in the daily settlement of futures contracts?

  A Initial margin deposits are refunded to the two parties.

  B Gains and losses are reported to other market participants.

  C Losses are charged to one party and gains credited to the other.

Solution to 1:

A is correct. Forwards and swaps are OTC contracts and, therefore, are customized. Futures are exchange traded and, therefore, are standardized. Some futures contracts are subject to daily price limits and their payoffs are received daily, but these characteristics are not true for forwards and swaps.

Solution to 2:

C is correct. Forwards and swaps are OTC instruments and both are regulated as such. Neither is regulated as a futures contract or a security. A swap is a series of multiple payments at scheduled dates, whereas a forward has only one payment, made at its expiration date.

Solution to 3:

C is correct. Losses and gains are collected and distributed to the respective parties. There is no specific reporting of these gains and losses to anyone else. Initial margin deposits are not refunded and, in fact, additional deposits may be required.

This material completes our introduction to forward commitments. All forward commitments are firm contracts. The parties are required to fulfill the obligations they agreed to. The benefit of this rigidity is that neither party pays anything to the other when the contract is initiated. If one party needs some flexibility, however, it can get it by agreeing to pay the other party some money when the contract is initiated. When the contract expires, the party who paid at the start has some flexibility in deciding whether to buy the underlying asset at the fixed price. Thus, that party did not actually agree to do anything. It had a choice. This is the nature of contingent claims.

6

CONTINGENT CLAIMS: OPTIONS

b contrast forward commitments with contingent claims;

c define forward contracts, futures contracts, options (calls and puts), swaps, and credit derivatives and compare their basic characteristics;

d determine the value at expiration and profit from a long or a short position in a call or put option;

A **contingent claim** is a derivative in which the outcome or payoff is dependent on the outcome or payoff of an underlying asset. Although this characteristic is also associated with forward commitments, a contingent claim has come to be associated with a *right*, but not an *obligation*, to make a final payment contingent on the performance of the underlying. Given that the holder of the contingent claim has a choice, the term *contingent claim* has become synonymous with the term *option*. The holder has a choice of whether or not to exercise the option. This choice creates a payoff that transforms the underlying payoff in a more pronounced manner than does a forward, futures, or swap. Those instruments provide linear payoffs: As the underlying goes up (down), the derivative gains (loses). The further up (down) the underlying goes, the more the derivative gains (loses). Options are different in that they limit losses in one direction. In addition, options can pay off as the underlying goes down. Hence, they transform the payoffs of the underlying into something quite different.

6.1 Options

We might say that an option, as a contingent claim, grants the right but not the obligation to buy an asset at a later date and at a price agreed on when the option is initiated. But there are so many variations of options that we cannot settle on this statement as a good formal definition. For one thing, options can also grant the right to sell instead of the right to buy. Moreover, they can grant the right to buy or sell earlier than at expiration. So, let us see whether we can combine these points into an all-encompassing definition of an option.

> *An option is a derivative contract in which one party, the buyer, pays a sum of money to the other party, the seller or writer, and receives the right to either buy or sell an underlying asset at a fixed price either on a specific expiration date or at any time prior to the expiration date.*

Unfortunately, even that definition does not cover every unique aspect of options. For example, options can be created in the OTC market and customized to the terms of each party, or they can be created and traded on options exchanges and standardized. As with forward contracts and swaps, customized options are subject to default, are less regulated, and are less transparent than exchange-traded derivatives. Exchange-traded options are protected against default by the clearinghouse of the options exchange and are relatively transparent and regulated at the national level. As noted in the definition above, options can be terminated early or at their expirations. When an option is terminated, either early or at expiration, the holder of the option chooses whether to exercise it. If he exercises it, he either buys or sells the underlying asset, but he does not have both rights. The right to buy is one type of option, referred to as a **call** or **call option**, whereas the right to sell is another type of option, referred to as a **put** or **put option**. With one very unusual and advanced exception that we do not cover, an option is either a call or a put, and that point is made clear in the contract.

An option is also designated as exercisable early (before expiration) or only at expiration. Options that can be exercised early are referred to as **American-style**. Options that can be exercised only at expiration are referred to as **European-style**. *It is extremely important that you do not associate these terms with where these options are traded*. Both types of options trade on all continents.

As with forwards and futures, an option can be exercised by physical delivery or cash settlement, as written in the contract. For a call option with physical delivery, upon exercise the underlying asset is delivered to the call buyer, who pays the call seller the exercise price. For a put option with physical delivery, upon exercise the put buyer delivers the underlying asset to the put seller and receives the strike price. For a cash settlement option, exercise results in the seller paying the buyer the cash equivalent value as if the asset were delivered and paid for.

The fixed price at which the underlying asset can be purchased is called the **exercise price** (also called the "strike price," the "strike," or the "striking price"). This price is somewhat analogous to the forward price because it represents the price at which the underlying will be purchased or sold if the option is exercised. The forward price, however, is set in the pricing of the contract such that the contract value at the start is zero. The strike price of the option is chosen by the participants. The actual price or value of the option is an altogether different concept.

As noted, the buyer pays the writer a sum of money called the **option premium**, or just the "premium." It represents a fair price of the option, and in a well-functioning market, it would be the value of the option. Consistent with everything we know about finance, it is the present value of the cash flows that are expected to be received by the holder of the option during the life of the option. At this point, we will not get into how this price is determined, but you will learn that later. For now, there are some fundamental concepts you need to understand, which form a basis for understanding how options are priced and why anyone would use an option.

Because the option buyer (the long) does not have to exercise the option, beyond the initial payment of the premium, there is no obligation of the long to the short. Thus, only the short can default, which would occur if the long exercises the option and the short fails to do what it is supposed to do. Thus, in contrast to forwards and swaps, in which either party could default to the other, default in options is possible only from the short to the long.

Ruling out the possibility of default for now, let us examine what happens when an option expires. Using the same notation used previously, let S_T be the price of the underlying at the expiration date, T, and X be the exercise price of the option. Remember that a call option allows the holder, or long, to pay X and receive the underlying. It should be obvious that the long would exercise the option at expiration if S_T is greater than X, meaning that the underlying value is greater than what he would pay to obtain the underlying. Otherwise, he would simply let the option expire. Thus, on the expiration date, the option is described as having a payoff of $Max(0,S_T - X)$.

Because the holder of the option would be entitled to exercise it and claim this amount, it also represents the value of the option at expiration. Let us denote that value as c_T. Thus,

$$c_T = Max(0,S_T - X) \qquad \text{(payoff to the call buyer),}$$

which is read as "take the maximum of either zero or $S_T - X$." For example, suppose you buy a call option with an exercise price of 50 and an expiration of three months for a premium of 1.50 when the stock is trading at 45. At expiration, consider the outcomes when the stock's price is 45, 50, or 55. The buyer's payoffs would be:

For $S_T = 45$, payoff $= c_T = Max(0,S_T - X) = Max(0,45 - 50) = \text{Max}(0, -5) = 0$.

For $S_T = 50$, payoff $= c_T = Max(0,S_T - X) = Max(0,50 - 50) = \text{Max}(0, 0) = 0$.

For $S_T = 55$, payoff $= c_T = Max(0,S_T - X) = Max(0,55 - 50) = \text{Max}(0, 5) = 5$.

Thus, if the underlying value exceeds the exercise price ($S_T > X$), then the option value is positive and equal to $S_T - X$. The call option is then said to be **in the money**. If the underlying value is less than the exercise price ($S_T < X$), then $S_T - X$ is negative; zero is greater than a negative number, so the option value would be zero. When the underlying value is less than the exercise price, the call option is said to be **out of the money**. When $S_T = X$, the call option is said to be **at the money**, although at the money is, for all practical purposes, out of the money because the value is still zero.

This payoff amount is also the value of the option at expiration. It represents value because it is what the option is worth at that point. If the holder of the option sells it to someone else an instant before expiration, it should sell for that amount because the new owner would exercise it and capture that amount. To the seller, the value of the option at that point is $-Max(0,S_T - X)$, which is negative to the seller if the option is in the money and zero otherwise.

Using the payoff value and the price paid for the option, we can determine the profit from the strategy, which is denoted with the Greek symbol Π. Let us say the buyer paid c_0 for the option at time 0. Then the profit is

$$\Pi = Max(0,S_T - X) - c_0 \qquad \text{(profit to the call buyer),}$$

Continuing with the example with underlying prices at expiration of 45, 50, or 55, the call buyer's profit would be:

For $S_T = 45$, profit $= Max(0,S_T - X) - c_0 = Max(0,45 - 50) - 1.50 = -1.50$.

For $S_T = 50$, profit $= Max(0,S_T - X) - c_0 = Max(0,50 - 50) - 1.50 = -1.50$.

For $S_T = 55$, profit $= Max(0,S_T - X) - c_0 = Max(0,55 - 50) - 1.50 = 3.50$.

To the seller, who received the premium at the start, the payoff is

$$-c_T = -Max(0,S_T - X) \qquad \text{(payoff to the call seller),}$$

At expiration, the call seller's payoffs are:

For $S_T = 45$, payoff $= -c_T = -Max(0,S_T - X) = -Max(0,45 - 50) = 0$.

For $S_T = 50$, payoff $= -c_T = -Max(0,S_T - X) = -Max(0,50 - 50) = 0$.

For $S_T = 55$, payoff $= -c_T = -Max(0,S_T - X) = -Max(0,55 - 50) = -5$.

The call seller's profit is

$\Pi = -Max(0,S_T - X) + c_0$ (profit to the call seller),

Finally, at expiration, the call seller's profit for each underlying price at expiration are:

For $S_T = 45$, profit $= -Max(0,S_T - X) + c_0 = -Max(0,45 - 50) + 1.50 = 1.50$.

For $S_T = 50$, profit $= -Max(0,S_T - X) + c_0 = -Max(0,50 - 50) + 1.50 = 1.50$.

For $S_T = 55$, profit $= -Max(0,S_T - X) + c_0 = -Max(0,55 - 50) + 1.50 = -3.50$.

For any given price at expiration, the call seller's payoff or profit is equal to the negative of the call buyer's payoff or profit.

EXAMPLE 4

Call Option Payoffs and Profit at Expiration

Consider a call option selling for \$7 in which the exercise price is \$100 and the price of the underlying is \$98.

1 Determine the value at expiration and the profit for a call buyer under the following outcomes:

  A The price of the underlying at expiration is \$102.

  B The price of the underlying at expiration is \$94.

2 Determine the value at expiration and the profit for a call seller under the following outcomes:

  A The price of the underlying at expiration is \$91.

  B The price of the underlying at expiration is \$101.

Solution to 1:

A If the price of the underlying at expiration is \$102,

$$\begin{aligned}\text{The call buyer's value at expiration} &= c_T = Max(0,S_T - X)\\ &= Max(0,102 - 100) = \$2.\\ \text{The call buyer's profit} &= \Pi = c_T - c_0 = 2 - 7 = -\$5.\end{aligned}$$

B If the price of the underlying at expiration is \$94,

$$\begin{aligned}\text{The call buyer's value at expiration} &= c_T = Max(0,S_T - X)\\ &= Max(0,94 - 100) = \$0.\\ \text{The call buyer's profit} &= \Pi = c_T - c_0 = 0 - 7 = -\$7.\end{aligned}$$

Solution to 2:

A If the price of the underlying at expiration is \$91,

$$\begin{aligned}\text{The call seller's value at expiration} &= -c_T = -Max(0,S_T - X)\\ &= -Max(0,91 - 100) = \$0.\\ \text{The call seller's profit} &= \Pi = -c_T + c_0 = 0 + 7 = \$7\end{aligned}$$

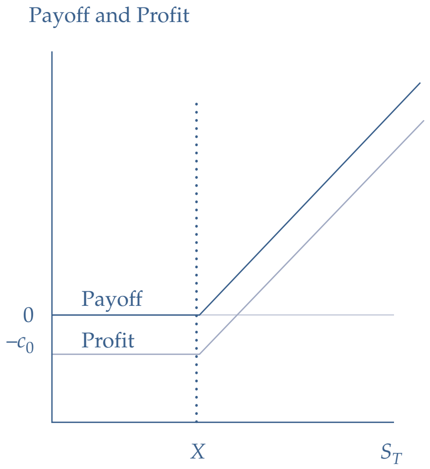
B If the price of the underlying at expiration is \$101,

$$\begin{aligned}\text{The call seller's value at expiration} &= -c_T = -Max(0,S_T - X)\\ &= -Max(0,101 - 100) = -\$1.\\ \text{The call seller's profit} &= \Pi = -c_T + c_0 = -1 + 7 = \$6.\end{aligned}$$

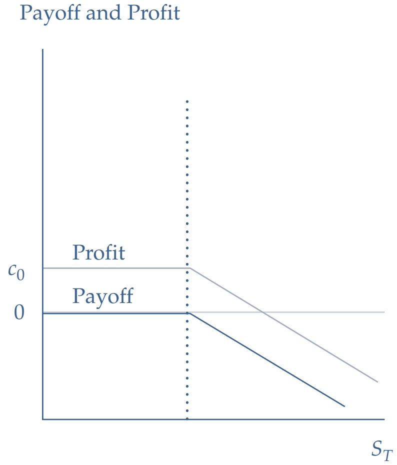
Exhibit 3 illustrates the payoffs and profits to the call buyer and seller as graphical representations of these equations, with the payoff or value at expiration indicated by the dark line and the profit indicated by the light line. Note in Panel A that the buyer has no upper limit on the profit and has a fixed downside loss limit equal to the premium paid for the option. Such a condition, with limited loss and unlimited gain, is a temptation to many unsuspecting investors, but keep in mind that the graph does not indicate the frequency with which gains and losses will occur. Panel B is the mirror image of Panel A and shows that the seller has unlimited losses and limited gains. One might suspect that selling a call is, therefore, the worst investment strategy possible. Indeed, it is a risky strategy, but at this point these are only simple strategies. Other strategies can be added to mitigate the seller's risk to a substantial degree.

Exhibit 3 Payoff and Profit from a Call Option

A. Payoff and Profit from Buying



B. Payoff and Profit from Selling



Now let us consider put options. Recall that a put option allows its holder to sell the underlying asset at the exercise price. Thus, the holder should exercise the put at expiration if the underlying asset is worth less than the exercise price ($S_T < X$). In that case, the put is said to be in the money. If the underlying asset is worth the same as the exercise price ($S_T = X$), meaning the put is at the money, or more than the exercise price ($S_T > X$), meaning the put is out of the money, the option holder would not exercise it and it would expire with zero value. Thus, the payoff to the put holder is

$$p_T = Max(0,X - S_T) \qquad \text{(payoff to the put buyer),}$$

If the put buyer paid p_0 for the put at time 0, the profit is

$$\Pi = Max(0,X - S_T) - p_0 \qquad \text{(profit to the put buyer),}$$

And for the seller, the payoff is

$$-p_T = -Max(0,X - S_T) \qquad \text{(payoff to the put seller),}$$

And the profit is

$$\Pi = -Max(0,X - S_T) + p_0 \qquad \text{(profit to the put seller),}$$

To illustrate the payoffs and profit to a put buyer and put seller, assume the put price (p_0) is 1.50, the exercise price (X) is 20.00, and the stock price at expiration (S_T) is either 18.00 or 22.00. The put buyer's payoff is:

For $S_T = 18$, payoff $= p_T = Max(0,X - S_T) = Max(0,20 - 18) = 2$.

For $S_T = 22$, payoff $= p_T = Max(0,X - S_T) = Max(0,20 - 22) = 0$.

The put buyer's profit is:

For $S_T = 18$, profit $= Max(0,X - S_T) - p_0 = Max(0,20 - 18) - 1.50 = 0.50$.

For $S_T = 22$, profit $= Max(0,X - S_T) - p_0 = Max(0,20 - 22) - 1.50 = -1.50$.

The put seller's payoff is:

For $S_T = 18$, payoff $= -p_T = -Max(0,X - S_T) = -Max(0,20 - 18) = -2$.

For $S_T = 22$, payoff $= -p_T = -Max(0,X - S_T) = -Max(0,20 - 22) = 0$.

Finally, the put seller's profit is:

For $S_T = 18$, profit $= -Max(0,X - S_T) + p_0 = -Max(0,20 - 18) + 1.50 = -0.50$.

For $S_T = 22$, profit $= -Max(0,X - S_T) + p_0 = -Max(0,20 - 22) + 1.50 = 1.50$.

For a given stock price at expiration, the put seller's payoff or profit are the negative of the put buyer's payoff or profit.

EXAMPLE 5

Put Option Payoffs and Profit at Expiration

Consider a put option selling for $4 in which the exercise price is $60 and the price of the underlying is $62.

1 Determine the value at expiration and the profit for a put buyer under the following outcomes:

- **A** The price of the underlying at expiration is $62.
- **B** The price of the underlying at expiration is $55.

2 Determine the value at expiration and the profit for a put seller under the following outcomes:

- **A** The price of the underlying at expiration is $51.
- **B** The price of the underlying at expiration is $68.

Solution to 1:

A If the price of the underlying at expiration is $62,

$$\begin{aligned}\text{The put buyer's value at expiration} &= p_T = Max(0,\mathrm{X} - S_T)\\ &= Max(0,60 - 62) = \$0.\\ \text{The put buyer's profit} &= \Pi = p_T - p_0 = 0 - 4 = -\$4.\end{aligned}$$

B If the price of the underlying at expiration is $55,

$$\begin{aligned}\text{The put buyer's value at expiration} &= p_T = Max(0,X - S_T)\\ &= Max(0,60 - 55) = \$5.\\ \text{The put buyer's profit} &= \Pi = p_T - p_0 = 5 - 4 = \$1.\end{aligned}$$

Solution to 2:

A If the price of the underlying at expiration is \$51,

$$\begin{aligned}\text{The put seller's value at expiration} &= -p_T = -Max(0,X-S_T)\\ &= -Max(0,60-51) = -\$9.\\ \text{The put seller's profit} &= \Pi = -p_T + p_0 = -9 + 4 = -\$5.\end{aligned}$$

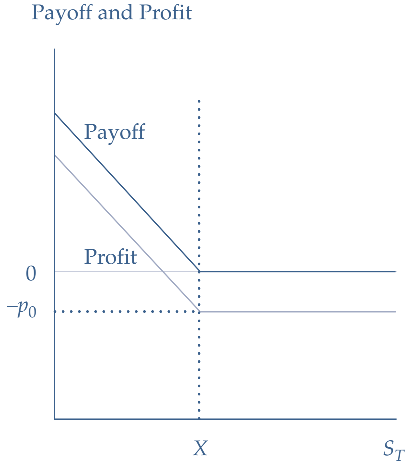
B If the price of the underlying at expiration is \$68,

$$\begin{aligned}\text{The put seller's value at expiration} &= -p_T = -Max(0,X-S_T)\\ &= -Max(0,60-68) = \$0.\\ \text{The put seller's profit} &= \Pi = -p_T + p_0 = 0 + 4 = \$4.\end{aligned}$$

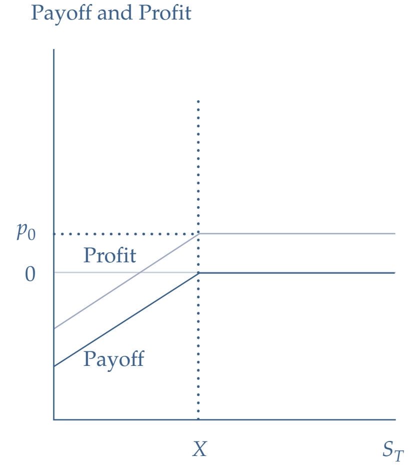
Exhibit 4 illustrates the payoffs and profits to the buyer and seller of a put.

Exhibit 4 Payoff and Profit from a Put Option

A. Payoff and Profit from Buying



B. Payoff and Profit from Selling



The put buyer has a limited loss, and although the gain is limited by the fact that the underlying value cannot go below zero, the put buyer does gain more the lower the value of the underlying. In this manner, we see how a put option is like insurance. Bad outcomes for the underlying trigger a payoff for both the insurance policy and the put, whereas good outcomes result only in loss of the premium. The put seller, like the insurer, has a limited gain and a loss that is larger the lower the value of the underlying. As with call options, these graphs must be considered carefully because they do not indicate the frequency with which gains and losses will occur. At this point, it should be apparent that buying a call option is consistent with a bullish point of view and buying a put option is consistent with a bearish point of view. Moreover, in contrast to forward commitments, which have payoffs that are linearly related to the payoffs of the underlying (note the straight lines in Exhibit 1), contingent claims have payoffs that are non-linear in relation to the underlying. There is linearity over a range—say, from 0 to X or from X upward or downward—but over the entire range of values for the underlying, the payoffs of contingent claims cannot be depicted with a single straight line.

We have seen only a snapshot of the payoff and profit graphs that can be created with options. Calls can be combined with puts, the underlying asset, and other calls or puts with different expirations and exercise prices to create a diverse set of payoff and profit graphs, some of which are covered later in the curriculum.

Before leaving options, let us again contrast the differences between options and forward commitments. With forward commitments, the parties agree to trade an underlying asset at a later date and at a price agreed upon when the contract is initiated. Neither party pays any cash to the other at the start. With options, the buyer pays cash to the seller at the start and receives the right, but not the obligation, to buy (if a call) or sell (if a put) the underlying asset at expiration at a price agreed upon (the exercise price) when the contract is initiated. In contrast to forwards, futures, and swaps, options do have value at the start: the premium paid by buyer to seller. That premium pays for the *right*, eliminating the *obligation*, to trade the underlying at a later date, as would be the case with a forward commitment.

Although there are numerous variations of options, most have the same essential features described here. There is, however, a distinctive family of contingent claims that emerged in the early 1990s and became widely used and, in some cases, heavily criticized. These instruments are known as credit derivatives.

7

CONTINGENT CLAIMS: CREDIT DERIVATIVES

b contrast forward commitments with contingent claims;

c define forward contracts, futures contracts, options (calls and puts), swaps, and credit derivatives and compare their basic characteristics;

Credit risk is surely one of the oldest risks known to mankind. Human beings have been lending things to each other for thousands of years, and even the most primitive human beings must have recognized the risk of lending some of their possessions to their comrades. Until the last 20 years or so, however, the management of credit risk was restricted to simply doing the best analysis possible before making a loan, monitoring the financial condition of the borrower during the loan, limiting the exposure to a given party, and requiring collateral. Some modest forms of insurance against credit risk have existed for a number of years, but insurance can be a slow and cumbersome way of protecting against credit loss. Insurance is typically highly regulated, and insurance laws are usually very consumer oriented. Thus, credit insurance as a financial product has met with only modest success.

In the early 1990s, however, the development of the swaps market led to the creation of derivatives that would hedge credit risk. These instruments came to be known as **credit derivatives**, and they avoided many of the regulatory constraints of the traditional insurance industry. Here is a formal definition:

> *A credit derivative is a class of derivative contracts between two parties, a credit protection buyer and a credit protection seller, in which the latter provides protection to the former against a specific credit loss.*

One of the first credit derivatives was a **total return swap**, in which the underlying is typically a bond or loan, in contrast to, say, a stock or stock index. The credit protection buyer offers to pay the credit protection seller the total return on the underlying bond. This total return consists of all interest and principal paid by the borrower plus any changes in the bond's market value. In return, the credit protection seller typically pays the credit protection buyer either a fixed or a floating rate of interest. Thus, if the bond defaults, the credit protection seller must continue to make its promised payments, while receiving a very small return or virtually no return from the credit protection buyer. If the bond incurs a loss, as it surely will if it defaults, the credit protection seller effectively pays the credit protection buyer.

Another type of credit derivative is the **credit spread option**, in which the underlying is the credit (yield) spread on a bond, which is the difference between the bond's yield and the yield on a benchmark default-free bond. As you will learn in the fixed-income material, the credit spread is a reflection of investors' perception of credit risk. Because a credit spread option requires a credit spread as the underlying, this type of derivative works only with a traded bond that has a quoted price. The credit protection buyer selects the strike spread it desires and pays the option premium to the credit protection seller. At expiration, the parties determine whether the option is in the money by comparing the bond's yield spread with the strike chosen, and if it is, the credit protection seller pays the credit protection buyer the established payoff. Thus, this instrument is essentially a call option in which the underlying is the credit spread.

A third type of credit derivative is the **credit-linked note (CLN)**. With this derivative, the credit protection buyer holds a bond or loan that is subject to default risk (the underlying reference security) and issues its own security (the credit-linked note) with the condition that if the bond or loan it holds defaults, the principal payoff on the credit-linked note is reduced accordingly. Thus, the buyer of the credit-linked note effectively insures the credit risk of the underlying reference security.

These three types of credit derivatives have had limited success compared with the fourth type of credit derivative, the **credit default swap (CDS)**. The credit default swap, in particular, has achieved much success by capturing many of the essential features of insurance while avoiding the high degree of consumer regulations that are typically associated with traditional insurance products.

In a CDS, one party—the credit protection buyer, who is seeking credit protection against a third party—makes a series of regularly scheduled payments to the other party, the credit protection seller. The seller makes no payments until a credit event occurs. A declaration of bankruptcy is clearly a credit event, but there are other types of credit events, such as a failure to make a scheduled payment or an involuntary restructuring. The CDS contract specifies what constitutes a credit event, and the industry has a procedure for declaring credit events, though that does not guarantee the parties will not end up in court arguing over whether something was or was not a credit event.

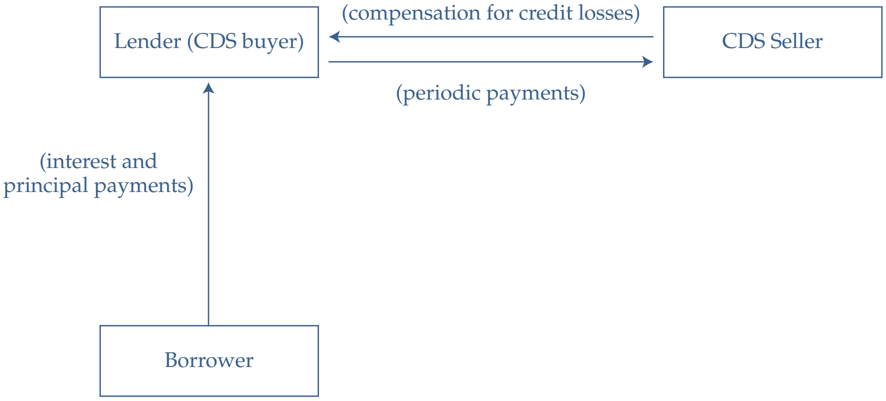
Formally, a credit default swap is defined as follows:

> *A credit default swap is a derivative contract between two parties, a credit protection buyer and a credit protection seller, in which the buyer makes a series of cash payments to the seller and receives a promise of compensation for credit losses resulting from the default of a third party.*

A CDS is conceptually a form of insurance. Sellers of CDSs, oftentimes banks or insurance companies, collect periodic payments and are required to pay out if a loss occurs from the default of a third party. These payouts could take the form of restitution of the defaulted amount or the party holding the defaulting asset could turn it over to the CDS seller and receive a fixed amount. The most common approach is for the payout to be determined by an auction to estimate the market value of the defaulting debt. Thus, CDSs effectively provide coverage against a loss in return for the protection buyer paying a premium to the protection seller, thereby taking the form of insurance against credit loss. Although insurance contracts have certain legal characteristics that are not found in credit default swaps, the two instruments serve similar purposes and operate in virtually the same way: payments made by one party in return for a promise to cover losses incurred by the other.

Exhibit 5 illustrates the typical use of a CDS by a lender. The lender is exposed to the risk of non-payment of principal and interest. The lender lays off this risk by purchasing a CDS from a CDS seller. The lender—now the CDS buyer—promises to make a series of periodic payments to the CDS seller, who then stands ready to compensate the CDS buyer for credit losses.

Exhibit 5 Using a Credit Default Swap to Hedge the Credit Risk of a Loan


Lender (CDS buyer)
(compensation for credit losses)
CDS Seller
(periodic payments)
(interest and principal payments)
Borrower

Clearly, the CDS seller is betting on the borrower's not defaulting or—more generally, as insurance companies operate—that the total payouts it is responsible for are less than the total payments collected. Of course, most insurance companies are able to do this by having reliable actuarial statistics, diversifying their risk, and selling some of the risk to other insurance companies. Actuarial statistics are typically quite solid. Average claims for life, health, and casualty insurance are well documented, and insurers can normally set premiums to cover losses and operate at a reasonable profit. Although insurance companies try to manage some of their risks at the micro level (e.g., charging smokers more for life and health insurance), most of their risk management is at the macro level, wherein they attempt to make sure their risks are not concentrated. Thus, they avoid selling too much homeowners insurance to individuals in tornado-prone areas. If they have such an exposure, they can use the reinsurance market to sell some of the risk to other companies that are not overexposed to that risk. Insurance companies attempt to diversify their risks and rely on the principle of uncorrelated risks, which plays such an important role in portfolio management. A well-diversified insurance company, like a well-diversified portfolio, should be able to earn a return commensurate with its assumed risk in the long run.

Credit default swaps should operate the same way. Sellers of CDSs should recognize when their credit risk is too concentrated. When that happens, they become buyers of CDSs from other parties or find other ways to lay off the risk. Unfortunately, during the financial crisis that began in 2007, many sellers of CDSs failed to recognize the high correlations among borrowers whose debt they had guaranteed. One well-known CDS seller, AIG, is a large and highly successful traditional insurance company that got into the business of selling CDSs. Many of these CDSs insured against mortgages. With the growth of the subprime mortgage market, many of these CDS-insured mortgages had a substantial amount of credit risk and were often poorly documented. AIG and many other CDS sellers were thus highly exposed to systemic credit contagion, a situation in which defaults in one area of an economy ripple into another, accompanied by bank weaknesses and failures, rapidly falling equity markets, rising credit risk premiums, and a general loss of confidence in the financial system and the economy. These presumably well-diversified risks guaranteed by CDS sellers, operating as though they were insurance companies, ultimately proved to be poorly diversified. Systemic financial risks can spread more rapidly than fire, health, and casualty risks. Virtually no other risks, except those originating from wars or epidemics, spread in the manner of systemic financial risks.

Thus, to understand and appreciate the importance of the CDS market, it is necessary to recognize how that market can fail. The ability to separate and trade risks is a valuable one. Banks can continue to make loans to their customers, thereby satisfying the customers' needs, while laying off the risk elsewhere. In short, parties not wanting to bear certain risks can sell them to parties wanting to assume certain risks. If all parties do their jobs correctly, the markets and the economy work more efficiently. If, as in the case of certain CDS sellers, not everyone does a good job of managing risk, there can be serious repercussions. In the case of AIG and some other companies, taxpayer bailouts were the ultimate price paid to keep these large institutions afloat so that they could continue to provide their other critical services to consumers. The rules proposed in the new OTC derivatives market regulations—which call for greater regulation and transparency of OTC derivatives and, in particular, CDSs—have important implications for the future of this market and these instruments.

EXAMPLE 6

Options and Credit Derivatives

1 An option provides which of the following?

- **A** Either the right to buy or the right to sell an underlying
- **B** The right to buy and sell, with the choice made at expiration
- **C** The obligation to buy or sell, which can be converted into the right to buy or sell

2 Which of the following is **not** a characteristic of a call option on a stock?

- **A** A guarantee that the stock will increase
- **B** A specified date on which the right to buy expires
- **C** A fixed price at which the call holder can buy the stock

3 A credit derivative is which of the following?

- **A** A derivative in which the premium is obtained on credit

B A derivative in which the payoff is borrowed by the seller

C A derivative in which the seller provides protection to the buyer against credit loss from a third party

Solution to 1:

A is correct. An option is strictly the right to buy (a call) or the right to sell (a put). It does not provide both choices or the right to convert an obligation into a right.

Solution to 2:

A is correct. A call option on a stock provides no guarantee of any change in the stock price. It has an expiration date, and it provides for a fixed price at which the holder can exercise the option, thereby purchasing the stock.

Solution to 3:

C is correct. Credit derivatives provide a guarantee against loss caused by a third party's default. They do not involve borrowing the premium or the payoff.

8 TYPES OF DERIVATIVES: ASSET-BACKED SECURITIES AND HYBRIDS

b contrast forward commitments with contingent claims;

c define forward contracts, futures contracts, options (calls and puts), swaps, and credit derivatives and compare their basic characteristics;

Although these instruments are covered in more detail in the fixed-income material, we would be remiss if we failed to include them with derivatives. But we will give them only light coverage here.

As discussed earlier, derivatives take (derive) their value from the value of the underlying, as do mutual funds and exchange-traded funds (ETFs). A mutual fund or an ETF holding bonds is virtually identical to the investor holding the bonds directly. Asset-backed securities (ABSs) take this concept a step further by altering the payment streams. ABSs typically divide the payments into slices, called tranches, in which the priority of claims has been changed from equivalent to preferential. For example, in a bond mutual fund or an ETF, all investors in the fund have equal claims, and so the rate of return earned by each investor is exactly the same. If a portfolio of the same bonds were assembled into an ABS, some investors in the ABS would have claims that would supersede those of other investors. The differential nature of these claims becomes relevant when either prepayments or defaults occur.

Prepayments mostly affect only mortgages. When a portfolio of mortgages is assembled into an ABS, the resulting instrument is called a **collateralized mortgage obligation** (CMO). Commonly but not always, the credit risk has been reduced or eliminated, perhaps by a CDS, as discussed earlier. When homeowners pay off their mortgages early due to refinancing at lower rates, the holders of the mortgages suffer losses. They expected to receive a stream of returns that is now terminated. The funds that were previously earning a particular rate will now have to be invested to earn a lower rate. These losses are the mirror images of the gains homeowners make when they proudly proclaim that they refinanced their mortgages and substantially lowered their payments.

CMOs partition the claims against these mortgages into different tranches, which are typically called A, B, and C. Class C tranches bear the first wave of prepayments until that tranche has been completely repaid its full principal investment. At that point, the Class B tranche holders bear the next prepayments until they have been fully repaid. The Class A tranche holders then bear the next wave of prepayments.[11] Thus, the risk faced by the various tranche holders is different from that of a mutual fund or ETF, which would pass the returns directly through such that investors would all receive the same rates of return. Therefore, the expected returns of CMO tranches vary and are commensurate with the prepayment risk they assume. Some CMOs are also characterized by credit risk, perhaps a substantial amount, from subprime mortgages.

When bonds or loans are assembled into ABSs, they are typically called **collateralized bond obligations** (CBOs) or **collateralized loan obligations** (CLOs). These instruments (known collectively as **collateralized debt obligations**, or CDOs) do not traditionally have much prepayment risk but they do have credit risk and oftentimes a great deal of it. The CDO structure allocates this risk to tranches that are called senior, mezzanine, or junior tranches (the last sometimes called equity tranches). When defaults occur, the junior tranches bear the risk first, followed by the mezzanine tranches, and then the senior tranches. The expected returns of the tranches vary according to the perceived credit risk, with the senior tranches having the highest credit quality and the junior the lowest. Thus, the senior tranches have the lowest expected returns and the junior tranches have the highest.

An asset-backed security is formally defined as follows:

> *An asset-backed security is a derivative contract in which a portfolio of debt instruments is assembled and claims are issued on the portfolio in the form of tranches, which have different priorities of claims on the payments made by the debt securities such that prepayments or credit losses are allocated to the most-junior tranches first and the most-senior tranches last.*

ABSs seem to have only an indirect and subtle resemblance to options, but they are indeed options. They promise to make a series of returns that are typically steady. These returns can be lowered if prepayments or defaults occur. Thus, they are contingent on prepayments and defaults. Take a look again at Exhibit 4, Panel B (the profit and payoff of a short put option). If all goes well, there is a fixed return. If something goes badly, the return can be lowered, and the worse the outcome, the lower the return. Thus, holders of ABSs have effectively written put options.

This completes the discussion of contingent claims. Having now covered forward commitments and contingent claims, the final category of derivative instruments is more or less just a catch-all category in case something was missed.

8.1 Hybrids

The instruments just covered encompass all the fundamental instruments that exist in the derivatives world. Yet, the derivatives world is truly much larger than implied by what has been covered here. We have not covered and will touch only lightly on the many hybrid instruments that combine derivatives, fixed-income securities, currencies, equities, and commodities. For example, options can be combined with bonds to form either callable bonds or convertible bonds. Swaps can be combined with options to form swap payments that have upper and lower limits. Options can be combined with futures to obtain options on futures. Options can be created with swaps as the underlying to form swaptions. Some of these instruments will be covered later. For now, you should just recognize that the possibilities are almost endless.

---

[11] The reference to only three tranches is just a general statement. There are many more types of tranches. Our discussion of the three classes is for illustrative purposes only and serves to emphasize that there are high-priority claims, low-priority claims, and other claims somewhere in the middle.

We will not address these hybrids directly, but some are covered elsewhere in the curriculum. The purpose of discussing them here is for you to realize that derivatives create possibilities not otherwise available in their absence. This point will lead to a better understanding of why derivatives exist, a topic we will get to very shortly.

EXAMPLE 7

Forward Commitments versus Contingent Claims

1 Which of the following is **not** a forward commitment?
   - **A** An agreement to take out a loan at a future date at a specific rate
   - **B** An offer of employment that must be accepted or rejected in two weeks
   - **C** An agreement to lease a piece of machinery for one year with a series of fixed monthly payments

2 Which of the following statements is true about contingent claims?
   - **A** Either party can default to the other.
   - **B** The payoffs are linearly related to the performance of the underlying.
   - **C** The most the long can lose is the amount paid for the contingent claim.

Solution to 1:

B is correct. Both A and C are commitments to engage in transactions at future dates. In fact, C is like a swap because the party agrees to make a series of future payments and in return receives temporary use of an asset whose value could vary. B is a contingent claim. The party receiving the employment offer can accept it or reject it if there is a better alternative.

Solution to 2:

C is correct. The maximum loss to the long is the premium. The payoffs of contingent claims are not linearly related to the underlying, and only one party, the short, can default.

9

DERIVATIVES UNDERLYINGS

b contrast forward commitments with contingent claims;

c define forward contracts, futures contracts, options (calls and puts), swaps, and credit derivatives and compare their basic characteristics;

Before discussing the purposes and benefits of derivatives, we need to clarify some points that have been implied so far. We have alluded to certain underlying assets, this section will briefly discuss the underlyings more directly.

9.1 Equities

Equities are one of the most popular categories of underlyings on which derivatives are created. There are two types of equities on which derivatives exist: individual stocks and stock indexes. Derivatives on individual stocks are primarily options. Forwards, futures, and swaps on individual stocks are not widely used. Index derivatives in the form of options, forwards, futures, and swaps are very popular. Index swaps, more often called equity swaps, are quite popular and permit investors to pay the return on one stock index and receive the return on another index or a fixed rate. They can be very useful in asset allocation strategies by allowing an equity manager to increase or reduce exposure to an equity market or sector without trading the individual securities.

In addition, options on stocks are frequently used by companies as compensation and incentives for their executives and employees. These options are granted to provide incentives to work toward driving the stock price up and can result in companies paying lower cash compensation. Some companies also issue warrants, which are options sold to the public that allow the holders to exercise them and buy shares directly from the companies.

9.2 Fixed-Income Instruments and Interest Rates

Options, forwards, futures, and swaps on bonds are widely used. The problem with creating derivatives on bonds, however, is that there are almost always many issues of bonds. A single issuer, whether it is a government or a private borrower, often has more than one bond issue outstanding. For futures contracts, with their standardization requirements, this problem is particularly challenging. What does it mean to say that a futures contract is on a German bund, a US Treasury note, or a UK gilt? The most common solution to this problem is to allow multiple issues to be delivered on a single futures contract. This feature adds some interesting twists to the pricing and trading strategies of these instruments.

Until now, we have referred to the underlying as an *asset*. Yet, one of the largest derivative underlyings is not an asset. It is simply an interest rate. An interest rate is not an asset. One cannot hold an interest rate or place it on a balance sheet as an asset. Although one can hold an instrument that pays an interest rate, the rate itself is not an asset. But there are derivatives in which the rate, not the instrument that pays the rate, is the underlying. In fact, we have already covered one of these derivatives: The plain vanilla interest rate swap in which Libor is the underlying.[12] Instead of a swap, an interest rate derivative could be an option. For example, a call option on 90-day Libor with a strike of 5% would pay off if at expiration Libor exceeds 5%. If Libor is below 5%, the option simply expires unexercised.

Interest rate derivatives are the most widely used derivatives. With that in mind, we will be careful in using the expression *underlying asset* and will use the more generic *underlying*.

9.3 Currencies

Currency risk is a major factor in global financial markets, and the currency derivatives market is extremely large. Options, forwards, futures, and swaps are widely used. Currency derivatives can be complex, sometimes combining elements of other underlyings. For example, a currency swap involves two parties making a series of interest rate payments to each other in different currencies. Because interest rates and currencies are both subject to change, a currency swap has two sources of risk. Although this instrument may sound extremely complicated, it merely reflects the fact that companies operating across borders are subject to both interest rate risk and currency risk and currency swaps are commonly used to manage those risks.

---

[12] As you will see later, there are also futures in which the underlying is an interest rate (Eurodollar futures) and forwards in which the underlying is an interest rate (forward rate agreements, or FRAs).

9.4 Commodities

Commodities are resources, such as food, oil, and metals, that humans use to sustain life and support economic activity. Because of the economic principle of comparative advantage, countries often specialize in the production of certain resources. Thus, the commodities market is extremely large and subject to an almost unimaginable array of risks. One need only observe how the price of oil moves up as tension builds in the Middle East or how the price of orange juice rises on a forecast of cold weather in Florida.

Commodity derivatives are widely used to speculate in and manage the risk associated with commodity price movements. The primary commodity derivatives are futures, but forwards, swaps, and options are also used. The reason that futures are in the lead in the world of commodities is simply history. The first futures markets were futures on commodities. The first futures exchange, the Chicago Board of Trade, was created in 1848, and until the creation of currency futures in 1972, there were no futures on any underlying except commodities.

There has been a tendency to think of the commodities world as somewhat separate from the financial world. Commodity traders and financial traders were quite different groups. Since the creation of financial futures, however, commodity and financial traders have become relatively homogeneous. Moreover, commodities are increasingly viewed as an important asset class that should be included in investment strategies because of their ability to help diversify portfolios.

9.5 Credit

As we previously discussed, credit is another underlying and quite obviously not an asset. Credit default swaps (CDSs) and collateralized debt obligations (CDOs) were discussed extensively in an earlier section. These instruments have clearly established that credit is a distinct underlying that has widespread interest from a trading and risk management perspective. In addition, to the credit of a single entity, credit derivatives are created on multiple entities. CDOs themselves are credit derivatives on portfolios of credit risks. In recent years, indexes of CDOs have been created, and instruments based on the payoffs of these CDO indexes are widely traded.

9.6 Other

This category is included here to capture some of the really unusual underlyings. One in particular is weather. Although weather is hardly an asset, it is certainly a major force in how some entities perform. For example, a ski resort needs snow, farmers need an adequate but not excessive amount of rain, and public utilities experience strains on their capacity during temperature extremes. Derivatives exist in which the payoffs are measured as snowfall, rainfall, and temperature. Although these derivatives have not been widely used—because of some complexities in pricing, among other things—they continue to exist and may still have a future. In addition, there are derivatives on electricity, which is also not an asset. It cannot be held in the traditional sense because it is created and consumed almost instantaneously. Another unusual type of derivative is based on disasters in the form of insurance claims.

Financial institutions will continue to create derivatives on all types of risks and exposures. Most of these derivatives will fail because of little trading interest, but a few will succeed. If that speaks badly of derivatives, it must be remembered that most small businesses fail, most creative ideas fail, and most people who try to become professional entertainers or athletes fail. It is the sign of a healthy and competitive system that only the very best survive.

THE SIZE OF THE DERIVATIVES MARKET

In case anyone thinks that the derivatives market is not large enough to justify studying, we should consider how big the market is. Unfortunately, gauging the size of the derivatives market is not a simple task. OTC derivatives contracts are private transactions. No reporting agency gathers data, and market size is not measured in traditional volume-based metrics, such as shares traded in the stock market. Complicating things further is the fact that derivatives underlyings include equities, fixed-income securities, interest rates, currencies, commodities, and a variety of other underlyings. All these underlyings have their own units of measurement. Hence, measuring how "big" the underlying derivatives markets are is like trying to measure how much fruit consumers purchase; the proverbial mixing of apples, oranges, bananas, and all other fruits.

The exchange-listed derivatives market reports its size in terms of volume, meaning the number of contracts traded. Exchange-listed volume, however, is an inconsistent number. For example, US Treasury bond futures contracts trade in units covering $100,000 face value. Eurodollar futures contracts trade in units covering $1,000,000 face value. Crude oil trades in 1,000-barrel (42 gallons each) units. Yet, one traded contract of each gets equal weighting in volume totals.

FIA (a global trade organization for futures, options, and centrally cleared derivatives markets) publishes detailed information about the industry. For 2017, global trading volume was 14.8 billion futures contracts, 10.4 billion options contracts, and a combined total of 25.2 billion contracts. Of nine futures instrument types, the largest three categories were interest rate futures (21.4% of global trading), equity index futures (16.9%), and currency futures (14.6%). Of nine options instrument types, the largest three categories were equity index options (48.4% of global trading), individual equity options (33.5%), and currency options (7.9%). Because options and futures contracts are typically short lived, the number of open-interest contracts is substantially less than the amount of trading volume. For the end of December 2017, open-interest futures contracts were 0.24 billion, open-interest option contracts were 0.60 billion, and the combined total was 0.84 billion contracts.

OTC volume is even more difficult to measure. There is no count of the number of contracts that trade. In fact, *volume* is an almost meaningless concept in OTC markets because any notion of volume requires a standardized size. If a customer goes to a swaps dealer and enters into a swap to hedge a $50 million loan, there is no measure of how much volume that transaction generated. The $50 million swap's notional principal, however, does provide a measure to some extent. Forwards, swaps, and OTC options all have notional principals, so they can be measured in that manner. Another measure of the size of the derivatives market is the market value of these contracts. As noted, forwards and swaps start with zero market value, but their market value changes as market conditions change. Options do not start with zero market value and almost always have a positive market value until expiration, when some options expire out of the money.

The OTC industry has taken both of these concepts—notional principal and market value—as measures of the size of the market. Notional principal is probably a more accurate measure. The amount of a contract's notional principal is unambiguous: It is written into the contract and the two parties cannot disagree over it. Yet, notional principal terribly overstates the amount of money actually at risk. For example, a $50 million notional principal swap will have nowhere near $50 million at risk. The payments on such a swap are merely the net of two opposite series of interest payments on $50 million. The market value of such a swap is the present value of one stream of payments minus the present value of the other. This market value figure will always be well below the notional principal. Thus, market value seems like a better measure except that, unlike notional principal, it is not unambiguous. Market value requires measurement, and two parties can disagree on the market value of the same transaction.

Notional principal and market value estimates for the global OTC derivatives market are collected semi-annually by the Bank for International Settlements of Basel, Switzerland, and published on its website (http://www.bis.org/statistics/derstats.htm). At the end of 2017, notional principal was more than $532 trillion and market value was about $27 trillion. A figure of $600 trillion is an almost unfathomable number and, as noted, is a misleading measure of the amount of money at risk.[13] The market value figure of $11 trillion is a much more realistic measure, but as noted, it is less accurate, relying on estimates provided by banks. Interest rate contracts constituted 81.9% of the total notional amount outstanding. The relative sizes of the other categories were 16.4% for foreign exchange contracts and less than 2% for credit derivatives, equity-linked contracts, and commodity contracts.

The exchange-listed and OTC markets use different measures and each of those measures is subject to severe limitations. About all we can truly say for sure about the derivatives market is, "It is big."

10 THE PURPOSES AND BENEFITS OF DERIVATIVES

e describe purposes of, and controversies related to, derivative markets;

Economic historians know that derivatives markets have existed since at least the Middle Ages. It is unclear whether derivatives originated in the Asian rice markets or possibly in medieval trade fairs in Europe. We do know that the origin of modern futures markets is the creation of the Chicago Board of Trade in 1848. To understand why derivatives markets exist, it is useful to take a brief look at why the Chicago Board of Trade was formed.

In the middle of the 19th century, midwestern America was rapidly becoming the center of agricultural production in the United States. At the same time, Chicago was evolving into a major American city, a hub of transportation and commerce. Grain markets in Chicago were the central location to which midwestern farmers brought their wheat, corn, and soybeans to sell. Unfortunately, most of these products arrived at approximately the same time of the year, September through November. The storage facilities in Chicago were strained beyond capacity. As a result, prices would fall tremendously and some farmers reportedly found it more economical to dump their grains in the Chicago River rather than transport them back to the farm. At other times of the year, prices would rise steeply. A group of businessmen saw this situation as unnecessary volatility and a waste of valuable produce. To deal with this problem, they created the Chicago Board of Trade and a financial instrument called the "to-arrive" contract. A farmer could sell a to-arrive contract at any time during the year. This contract fixed the price of the farmer's grain on the basis of delivery in Chicago at a specified later date. Grain is highly storable, so farmers can hold on to the grain and deliver it at almost any later time. This plan substantially reduced seasonal market volatility and made the markets work much better for all parties.

The traders in Chicago began to trade these contracts, speculating on movements in grain prices. Soon, it became apparent that an important and fascinating market had developed. Widespread hedging and speculative interest resulted in substantial market growth, and about 80 years later, a clearinghouse and a performance guarantee were added, thus completing the evolution of the to-arrive contract into today's modern futures contract.

---

[13] To put it in perspective, it would take almost 17 million years for a clock to tick off 532 trillion seconds!

Many commodities and all financial assets that underlie derivatives contracts are not seasonally produced. Hence, this initial motivation for futures markets is only a minor advantage of derivatives markets today. But there are many reasons why derivative markets serve an important and useful purpose in contemporary finance.

10.1 Risk Allocation, Transfer, and Management

Until the advent of derivatives markets, risk management was quite cumbersome. Setting the actual level of risk to the desired level of risk required engaging in transactions in the underlyings. Such transactions typically had high transaction costs and were disruptive of portfolios. In many cases, it is quite difficult to fine-tune the level of risk to the desired level. From the perspective of a risk taker, it was quite costly to buy risk because a large amount of capital would be required.

Derivatives solve these problems in a very effective way: They allow trading the risk without trading the instrument itself. For example, consider a stockholder who wants to reduce exposure to a stock. In the pre-derivatives era, the only way to do so was to sell the stock. Now, the stockholder can sell futures, forwards, calls, or swaps, or buy put options, all while retaining the stock. For a company founder, these types of strategies can be particularly useful because the founder can retain ownership and probably board membership. Many other excellent examples of the use of derivatives to transfer risk are covered elsewhere in the curriculum. The objective at this point is to establish that derivatives provide an effective method of transferring risk from parties who do not want the risk to parties who do. In this sense, risk allocation is improved within markets and, indeed, the entire global economy.

The overall purpose of derivatives is to obtain more effective risk management within companies and the entire economy. Although some argue that derivatives do not serve this purpose very well (we will discuss this point in Section 11), for now you should understand that derivatives can improve the allocation of risk and facilitate more effective risk management for both companies and economies.

10.2 Information Discovery

One of the advantages of futures markets has been described as *price discovery*. A futures price has been characterized by some experts as a revelation of some information about the future. Thus, a futures price is sometimes thought of as predictive. This statement is not strictly correct because futures prices are not really forecasts of future spot prices. They provide only a little more information than do spot prices, but they do so in a very efficient manner. The markets for some underlyings are highly decentralized and not very efficient. For example, what is gold worth? It trades in markets around the world, but probably the best place to look is at the gold futures contract expiring soonest. What is the value of the S&P 500 Index when the US markets are not open? As it turns out, US futures markets open before the US stock market opens. The S&P 500 futures price is frequently viewed as an indication of where the stock market will open.

Derivative markets can, however, convey information not impounded in spot markets. By virtue of the fact that derivative markets require less capital, information can flow into the derivative markets before it gets into the spot market. The difference may well be only a matter of minutes or possibly seconds, but it can provide the edge to astute traders.

Finally, we should note that futures markets convey another simple piece of information: What price would one accept to avoid uncertainty? If you hold a stock worth $40 and could hedge the next 12 months' uncertainty, what locked-in price should you expect to earn? As it turns out, it should be the price that guarantees the risk-free rate minus whatever dividends would be paid on the stock. Derivatives—specifically, futures, forwards, and swaps—reveal the price that the holder of an asset could take and avoid the risk.

What we have said until now applies to futures, forwards, and swaps. What about options? As you will learn later, given the underlying and the type of option (call or put), an option price reflects two characteristics of the option (exercise price and time to expiration), three characteristics of the underlying (price, volatility, and cash flows it might pay), and one general macroeconomic factor (risk-free rate). Only one of these factors, volatility, is not relatively easy to identify. But with the available models to price the option, we can infer what volatility people are using from the actual market prices at which they execute trades. That volatility, called **implied volatility**, measures the expected risk of the underlying. It reflects the volatility that investors use to determine the market price of the option. Knowing the expected risk of the underlying asset is an extremely useful piece of information. In fact, for options on broad-based market indexes, such as the S&P 500, the implied volatility is a good measure of the general level of uncertainty in the market. Some experts have even called it a measure of fear. Thus, options provide information about what investors think of the uncertainty in the market, if not their fear of it.[14]

In addition, options allow the creation of trading strategies that cannot be done by using the underlying. As the exhibits on options explained, these strategies provide asymmetrical performance: limited movement in one direction and movement in the other direction that changes with movements in the underlying.

10.3 Operational Advantages

We noted earlier that derivatives have lower transaction costs than the underlying. The transaction costs of derivatives can be high relative to the value of the derivatives, but these costs are typically low relative to the value of the underlying. Thus, an investor who wants to take a position in, say, an equity market index would likely find it less costly to use the futures to get a given degree of exposure than to invest directly in the index to get that same exposure.

Derivative markets also typically have greater liquidity than the underlying spot markets, a result of the smaller amount of capital required to trade derivatives than to get the equivalent exposure directly in the underlying. Futures margin requirements and option premiums are quite low relative to the cost of the underlying.

One other extremely valuable operational advantage of derivative markets is the ease with which one can go short. With derivatives, it is nearly as easy to take a short position as to take a long position, whereas for the underlying asset, it is almost always much more difficult to go short than to go long. In fact, for many commodities, short selling is nearly impossible.

---

[14] The Chicago Board Options Exchange publishes a measure of the implied volatility of the S&P 500 Index option, which is called the VIX (volatility index). The VIX is widely followed and is cited as a measure of investor uncertainty and sometimes fear.

10.4 Market Efficiency

In the study of portfolio management, you learn that an efficient market is one in which no single investor can consistently earn returns in the long run in excess of those commensurate with the risk assumed. Of course, endless debates occur over whether equity markets are efficient. No need to resurrect that issue here, but let us proceed with the assumption that equity markets—and, in fact, most free and competitive financial markets—are reasonably efficient. This assumption does not mean that abnormal returns can never be earned, and indeed prices do get out of line with fundamental values. But competition, the relatively free flow of information, and ease of trading tend to bring prices back in line with fundamental values. Derivatives can make this process work even more rapidly.

When prices deviate from fundamental values, derivative markets offer less costly ways to exploit the mispricing. As noted earlier, less capital is required, transaction costs are lower, and short selling is easier. We also noted that as a result of these features, it is possible, indeed likely, that fundamental value will be reflected in the derivatives markets before it is restored in the underlying market. Although this time difference could be only a matter of minutes, for a trader seeking abnormal returns, a few minutes can be a valuable opportunity.

All these advantages of derivatives markets make the financial markets in general function more effectively. Investors are far more willing to trade if they can more easily manage their risk, trade at lower cost and with less capital, and go short more easily. This increased willingness to trade increases the number of market participants, which makes the market more liquid. A very liquid market may not automatically be an efficient market, but it certainly has a better chance of being one.

Even if one does not accept the concept that financial markets are efficient, it is difficult to say that markets are not more effective and competitive with derivatives. Yet, many blame derivatives for problems in the market. Let us take a look at these arguments.

CRITICISMS AND MISUSES OF DERIVATIVES

11

e describe purposes of, and controversies related to, derivative markets;

The history of financial markets is filled with extreme ups and downs, which are often called bubbles and crashes. Bubbles occur when prices rise for a long time and appear to exceed fundamental values. Crashes occur when prices fall rapidly. Although bubbles, if they truly exist, are troublesome, crashes are even more so because nearly everyone loses substantial wealth in a crash. A crash is then typically followed by a government study commissioned to find the causes of the crash. In the last 30 years, almost all such studies have implicated derivatives as having some role in causing the crash. Of course, because derivatives are widely used and involve a high degree of leverage, it is a given that they would be seen in a crash. It is unclear whether derivatives are the real culprit or just the proverbial smoking gun used by someone to do something wrong.

The two principal arguments against derivatives are that they are such speculative devices that they effectively permit legalized gambling and that they destabilize the financial system. Let us look at these points more closely.

11.1 Speculation and Gambling

As noted earlier, derivatives are frequently used to manage risk. In many contexts, this use involves hedging or laying off risk. Naturally, for hedging to work, there must be speculators. Someone must accept the risk. Derivatives markets are unquestionably attractive to speculators. All the benefits of derivatives draw speculators in large numbers, and indeed they should. The more speculators that participate in the market, the cheaper it is for hedgers to lay off risk. These speculators take the form of hedge funds and other professional traders who willingly accept risk that others need to shed. In recent years, the rapid growth of these types of investors has been alarming to some but almost surely has been beneficial for all investors.

Unfortunately, the general image of speculators is not a good one. Speculators are often thought to be short-term traders who attempt to exploit temporary inefficiencies, caring little about long-term fundamental values. The profits from short-term trading are almost always taxed more heavily than the profits from long-term trading, clearly targeting and in some sense punishing speculators. Speculators are thought to engage in price manipulation and to trade at extreme prices.[15] All of this type of trading is viewed more or less as just a form of gambling.

Yet, there are notable differences between gambling and speculation. Gambling typically benefits only a limited number of participants and does not generally help society as a whole. But derivatives trading brings extensive benefits to financial markets, as explained earlier, and thus does benefit society as a whole. In short, the benefits of derivatives are broad, whereas the benefits of gambling are narrow.

Nonetheless, the argument that derivatives are a form of legalized gambling will continue to be made. Speculation and gambling are certainly both forms of financial risk taking, so these arguments are not completely off base. But insurance companies speculate on loss claims, mutual funds that invest in stocks speculate on the performance of companies, and entrepreneurs go up against tremendous odds to speculate on their own ability to create successful businesses. These so-called speculators are rarely criticized for engaging in a form of legalized gambling, and indeed entrepreneurs are praised as the backbone of the economy. Really, all investment is speculative. So, why is speculation viewed as such a bad thing by so many? The answer is unclear.

11.2 Destabilization and Systemic Risk

The arguments against speculation through derivatives often go a step further, claiming that it is not merely speculation or gambling per se but rather that it has destabilizing consequences. Opponents of derivatives claim that the very benefits of derivatives (low cost, low capital requirements, ease of going short) result in an excessive amount of speculative trading that brings instability to the market. They argue that speculators use large amounts of leverage, thereby subjecting themselves and their creditors to substantial risk if markets do not move in their hoped-for direction. Defaults by speculators can then lead to defaults by their creditors, their creditors' creditors, and so on. These effects can, therefore, be systemic and reflect an epidemic contagion whereby instability can spread throughout markets and an economy, if not the entire world. Given that governments often end up bailing out some banks and insurance companies, society has expressed concern that the risk managed with derivatives must be controlled.

---

[15] Politicians and regulators have been especially critical of energy market speculators. Politicians, in particular, almost always blame rising oil prices on speculators, although credit is conspicuously absent for falling oil prices.

This argument is not without merit. Such effects occurred in the Long-Term Capital Management fiasco of 1998 and again in the financial crisis of 2008, in which derivatives, particularly credit default swaps, were widely used by many of the problem entities. Responses to such events typically take the course of calling for more rules and regulations restricting the use of derivatives, requiring more collateral and credit mitigation measures, backing up banks with more capital, and encouraging, if not requiring, OTC derivatives to be centrally cleared like exchange-traded derivatives.

In response, however, we should note that financial crises—including the South Sea and Mississippi bubbles and the stock market crash of 1929, as well as a handful of economic calamities of the 19th and 20th centuries—have existed since the dawn of capitalism. Some of these events preceded the era of modern derivatives markets, and others were completely unrelated to the use of derivatives. Some organizations, such as Orange County, California, in 1994–1995, have proved that derivatives are not required to take on excessive leverage and nearly bring the entity to ruin. Proponents of derivatives argue that derivatives are but one of many mechanisms through which excessive risk can be taken. Derivatives may seem dangerous, and they can be if misused, but there are many ways to take on leverage that look far less harmful but can be just as risky.

Another criticism of derivatives is simply their complexity. Many derivatives are extremely complex and require a high-level understanding of mathematics. The financial industry employs many mathematicians, physicists, and computer scientists. This single fact has made many distrust derivatives and the people who work on them. It is unclear why this reason has tarnished the reputation of the derivatives industry. Scientists work on complex problems in medicine and engineering without public distrust. One explanation probably lies in the fact that scientists create models of markets by using scientific principles that often fail. To a physicist modeling the movements of celestial bodies, the science is reliable and the physicist is unlikely to misapply the science. The same science applied to financial markets is far less reliable. Financial markets are driven by the actions of people who are not as consistent as the movements of celestial bodies. When financial models fail to work as they should, the scientists are often blamed for either building models that are too complex and unable to accurately capture financial reality or misusing those models, such as using poor estimates of inputs. And derivatives, being so widely used and heavily leveraged, are frequently in the center of it all.

EXAMPLE 8

Purposes and Controversies of Derivative Markets

1. Which of the following is **not** an advantage of derivative markets?
   - **A** They are less volatile than spot markets.
   - **B** They facilitate the allocation of risk in the market.
   - **C** They incur lower transaction costs than spot markets.
2. Which of the following pieces of information is **not** conveyed by at least one type of derivative?
   - **A** The volatility of the underlying
   - **B** The most widely used strategy of the underlying
   - **C** The price at which uncertainty in the underlying can be eliminated
3. Which of the following responds to the criticism that derivatives can be destabilizing to the underlying market?

A Market crashes and panics have occurred since long before derivatives existed.

B Derivatives are sufficiently regulated that they cannot destabilize the spot market.

C The transaction costs of derivatives are high enough to keep their use at a minimum level.

Solution to 1:

A is correct. Derivative markets are not by nature more or less volatile than spot markets. They facilitate risk allocation by making it easier and less costly to transfer risk, and their transaction costs are lower than those of spot markets.

Solution to 2:

B is correct. Options do convey the volatility of the underlying, and futures, forwards, and swaps convey the price at which uncertainty in the underlying can be eliminated. Derivatives do not convey any information about the use of the underlying in strategies.

Solution to 3:

A is correct. Derivatives regulation is not more and is arguably less than spot market regulation, and the transaction costs of derivatives are not a deterrent to their use; in fact, derivatives are widely used. Market crashes and panics have a very long history, much longer than that of derivatives.

An important element of understanding and using derivatives is having a healthy respect for their power. Every day, we use chemicals, electricity, and fire without thinking about their dangers. We consume water and drive automobiles, both of which are statistically quite dangerous. Perhaps these risks are underappreciated, but it is more likely the case that most adults learn how to safely use chemicals, electricity, fire, water, and automobiles. Of course, there are exceptions, many of which are foolish, and foolishness is no stranger to the derivatives industry. The lesson here is that derivatives can make our financial lives better, but like chemicals, electricity, and all the rest, we need to know how to use them safely, which is why they are an important part of the CFA curriculum.

Later in the curriculum, you will learn a great deal about how derivatives are priced. At this point, we introduce the pricing of derivatives. This material not only paves the way for a deeper understanding of derivatives but also complements earlier material by helping you understand how derivatives work.

12

ELEMENTARY PRINCIPLES OF DERIVATIVE PRICING

f explain arbitrage and the role it plays in determining prices and promoting market efficiency.

Pricing and valuation are fundamental elements of the CFA Program. The study of fixed-income and equity securities, as well as their application in portfolio management, is solidly grounded on the principle of valuation. In valuation, the question is simple: What is something worth? Without an answer to that question, one can hardly proceed to use that *something* wisely.

Determining what a derivative is worth is similar to determining what an asset is worth. As you learn in the fixed-income and equity readings, value is the present value of future cash flows, with discounting done at a rate that reflects both the opportunity cost of money and the risk. Derivatives valuation applies that same principle but in a somewhat different way.

Think of a derivative as *attached* to an underlying. We know that the derivative *derives* its value from the value of the underlying. If the underlying's value changes, so should the value of the derivative. The underlying takes its value from the discounted present value of the expected future cash flows it offers, with discounting done at a rate reflecting the investor's risk tolerance. But if the value of the underlying is embedded in the value of the derivative, it would be double counting to discount the derivative's expected future cash flows at a risky discount rate. That effect has already been incorporated into the value of the underlying, which goes into the value of the derivative.

Derivatives usually take their values from the underlying by constructing a hypothetical combination of the derivatives and the underlyings that eliminates risk. This combination is typically called a **hedge portfolio**. With the risk eliminated, it follows that the hedge portfolio should earn the risk-free rate. A derivative's value is the price of the derivative that forces the hedge portfolio to earn the risk-free rate.

This principle of derivative valuation relies completely on the ability of an investor to hold or store the underlying asset. Let us take a look at what that means.

12.1 Storage

As noted previously, the first derivatives were agricultural commodities. Most of these commodities can be stored (i.e., held) for a period of time. Some extreme cases, such as oil and gold, which are storable for millions of years, are excellent examples of fully storable commodities. Grains, such as wheat and corn, can be stored for long but not infinite periods of time. Some commodities, such as bananas, are storable for relatively short periods of time. In the CFA Program, we are more interested in financial assets. Equities and currencies have perpetual storability, whereas bonds are storable until they mature.

Storage incurs costs. Commodity storage costs can be quite expensive. Imagine storing 1,000 kilograms of gold or a million barrels of oil. Financial assets, however, have relatively low storage costs. Some assets pay returns during storage. Stocks pay dividends and bonds pay interest. The net of payments offered minus storage costs plays a role in the valuation of derivatives.

An example earlier in this reading illustrates this point. Suppose an investor holds a dividend-paying stock and wants to eliminate the uncertainty of its selling price over a future period of time. Suppose further that the investor enters into a forward contract that commits him to deliver the stock at a later date, for which he will receive a fixed price. With uncertainty eliminated, the investor should earn the risk-free rate, but in fact, he does not. He earns more because while holding the stock, he collects dividends. Therefore, he should earn the risk-free rate *minus* the dividend yield, a concept known as the cost of carry, which will be covered in great detail in later readings. The cost of carry *plus* the dividends he earns effectively means that he makes the risk-free rate. Now, no one is claiming that this is a good way to earn the risk-free rate. There are many better ways to do that, but this strategy could be executed. There is one and only one forward price that guarantees that this strategy earns a return of the risk-free rate minus the dividend yield, or the risk-free rate after accounting for the dividends collected. If the forward price at which contracts are created does not equal this price, investors can take advantage of this discrepancy by engaging in arbitrage, which is discussed in the next section.

Forwards, futures, swaps, and options are all priced in this manner. Hence, they rely critically on the ability to store or hold the asset. Some underlyings are not storable. We previously mentioned electricity. It is produced and consumed almost instantaneously. Weather is also not storable. Fresh fish have very limited storability. Although this absence of storability may not be the reason, derivative markets in these types of underlyings have not been particularly successful, whereas those in underlyings that are more easily storable have often been successful.

The opposite of storability is the ability to go short—that is, to borrow the underlying, sell it, and buy it back later. We discussed earlier that short selling of some assets can be difficult. It is not easy to borrow oil or soybeans. There are ways around this constraint, but derivatives valuation is generally much easier when the underlying can be shorted. This point is discussed in more depth later in the curriculum.

12.2 Arbitrage

What we have been describing is the foundation of the principle of **arbitrage**. In well-functioning markets with low transaction costs and a free flow of information, the same asset cannot sell for more than one price. If it did, someone would buy it in the cheaper market and sell it in the more expensive market, earning a riskless profit. The combined actions of all parties doing this would push up the lower price and push down the higher price until they converged. For this reason, arbitrage is often referred to as the **law of one price**. Of course, for arbitrage to be feasible, the ability to purchase and sell short the asset is important.

Obviously, this rule does not apply to all markets. The same consumer good can easily sell for different prices, which is one reason why people spend so much time shopping on the internet. The costs associated with purchasing the good in the cheaper market and selling it in the more expensive market can make the arbitrage not worthwhile. The absence of information on the very fact that different prices exist would also prevent the arbitrage from occurring. Although the internet and various price-comparing websites reduce these frictions and encourage all sellers to offer competitive prices, consumer goods are never likely to be arbitragable.[16]

Financial markets, of course, are a different matter. Information on securities prices around the world is quite accessible and relatively inexpensive. Most financial markets are fairly competitive because dealers, speculators, and brokers attempt to execute trades at the best prices. Arbitrage is considered a dependable rule in the financial markets. Nonetheless, there are people who purport to make a living as arbitrageurs. How could they exist? To figure that out, first consider some examples of arbitrage.

The simplest case of an arbitrage might be for the same stock to sell at different prices in two markets. If the stock were selling at $52 in one market and $50 in another, an arbitrageur would buy the stock at $50 in the one market and sell it at $52 in the other. This trade would net an immediate $2 profit at no risk and would not require the commitment of any of the investor's capital. This outcome would be a strong motivation for all arbitrageurs, and their combined actions would force the lower price up and the higher price down until the prices converged.

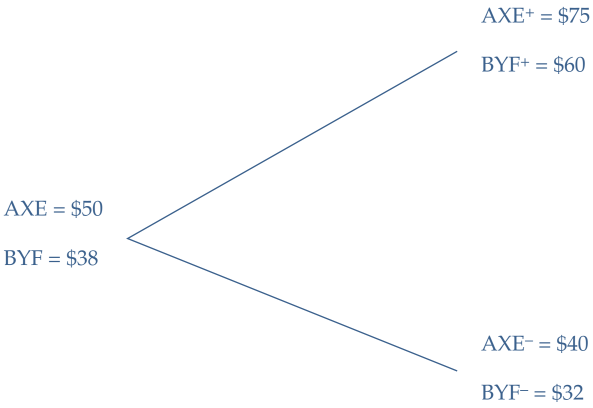
But what would be the final price? It is entirely possible that $50 is the true fundamental value and $52 is too high. Or $52 could be the true fundamental value and $50 is too low. Or the true fundamental value could lie somewhere between the two. Arbitrage does not tell us the true fundamental value. It is not an *absolute* valuation methodology, such as the discounted cash flow equity valuation model. It is a *relative* valuation methodology. It tells us the correct price of one asset or derivative *relative to* another asset or derivative.

---

[16] If the same consumer good sells for different prices in markets with a relatively free flow of information (e.g., via price-comparing websites), it still may not be possible to truly arbitrage. Buying the good at a lower price and selling it at a higher price but less than the price of the most expensive seller may not be practical, but the most expensive seller may be driven out of business. When everyone knows what everyone else is charging, the same effect of arbitrage can still occur.

Now, consider another situation, illustrated in Exhibit 6. Observe that we have one stock, AXE Electronics, that today is worth \$50 and one period later will be worth either \$75 or \$40. We will denote these prices as AXE = \$50, AXE^+ = \$75, and AXE^- = \$40. Another stock, BYF Technology, is today worth \$38 and one period later will be worth \$60 or \$32. Thus, BYF = \$38, BYF^+ = \$60, and BYF^- = \$32. Assume that the risk-free borrowing and lending rate is 4%. Also assume no dividends are paid on either stock during the period covered by this example.

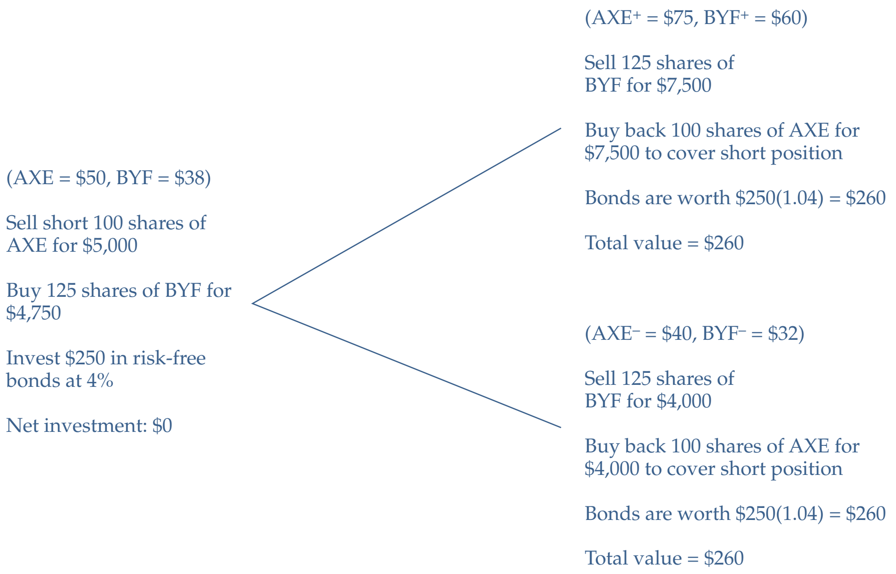
Exhibit 6 Arbitrage Opportunity with Stock AXE, Stock BYF, and a Risk-Free Bond



Note: The risk-free rate is 4%.

The opportunity exists to make a profit at no risk without committing any of our funds, as demonstrated in Exhibit 7. Suppose we borrow 100 shares of stock AXE, which is selling for \$50, and sell them short, thereby receiving \$5,000. We take \$4,750 and purchase 125 shares of stock BYF (125 × \$38 = \$4,750). We invest the remaining \$250 in risk-free bonds at 4%. This transaction will not require us to use any funds of our own: The short sale will be sufficient to fund the investment in BYF and leave money to invest in risk-free bonds.

Exhibit 7 Execution of Arbitrage Transaction with Stock AXE, Stock BYF, and a Risk-Free Bond



(AXE = \$50, BYF = \$38)

Sell short 100 shares of AXE for \$5,000

Buy 125 shares of BYF for \$4,750

Invest \$250 in risk-free bonds at 4%

Net investment: \$0

(AXE^+ = \$75, BYF^+ = \$60)

Sell 125 shares of BYF for \$7,500

Buy back 100 shares of AXE for \$7,500 to cover short position

Bonds are worth \$250(1.04) = \$260

Total value = \$260

(AXE^- = \$40, BYF^- = \$32)

Sell 125 shares of BYF for \$4,000

Buy back 100 shares of AXE for \$4,000 to cover short position

Bonds are worth \$250(1.04) = \$260

Total value = \$260

If the top outcome in Exhibit 7 occurs, we sell the 125 shares of BYF for 125 × \$60 = \$7,500. This amount is sufficient to buy back the 100 shares of AXE, which is selling for \$75. But we will also have the bonds, which are worth \$250 × 1.04 = \$260. If the bottom outcome occurs, we sell the 125 shares of BYF for 125 × \$32 = \$4,000—enough money to buy back the 100 shares of AXE, which is selling for \$40. Again, we will have the risk-free bonds, worth \$260. Regardless of the outcome, we end up with \$260.

Recall that we invested no money of our own and end up with a sure \$260. It should be apparent that this transaction is extremely attractive, so everyone would do it. The combined actions of multiple investors would drive down the price of AXE and/or drive up the price of BYF until an equilibrium is reached, at which point this transaction would no longer be profitable. As noted earlier, we cannot be sure of the correct fundamental price, but let us assume that BYF's price remains constant. Then AXE would fall to \$47.50. Alternatively, if we assume that AXE's price remains constant, then the price of BYF would rise to \$40. These values are obtained by noting that the prices for both outcomes occur according to the ratio 1.25 (\$75/\$60 = 1.25; \$40/\$32 = 1.25). Thus, their initial prices should be consistent with that ratio. If BYF is \$38, AXE should be \$38 × 1.25 = \$47.50. If AXE is \$50, BYF should be \$40.00 because \$40.00 × 1.25 = \$50. Of course, the two prices could settle in between. Arbitrage is only a relative pricing method. It prices the two stocks in relation to each other but does not price either on the basis of its own fundamentals.

Of course, this example is extremely simplified. Clearly, a stock price can change to more than two other prices. Also, if a given stock is at one price, another stock may be at any other price. We have created a simple case here to illustrate a point. But as you will learn later in the curriculum, when derivatives are involved, the simplification here is relatively safe. As we know, the price of a derivative is determined by the price of the underlying. Hence, when the underlying is at one particular price, the derivative's price will be determined by that price. The two assets need not be two stocks; one can be a stock and the other can be a derivative on the stock.

To see that point, consider another type of arbitrage opportunity that involves a forward contract. Recall from the previous example that at the start, AXE sells for \$50. Suppose we borrow \$50 at 4% interest by issuing a risk-free bond, use the money to buy one share of stock AXE, and simultaneously enter into a forward contract to sell this share at a price of \$54 one period later. The stock will then move to either \$75 or \$40 in the next period. The forward contract requires that we deliver the stock and accept \$54 for it. And of course, we will owe \$50 × 1.04 = \$52 on the loan.

Now consider the two outcomes. Regardless of the outcome, the end result is the same. The forward contract fixes the delivery price of the stock at \$54:

AXE goes to \$75	
Deliver stock to settle forward contract	+ \$54
Pay back loan	– \$52
Net	+ \$2
AXE goes to \$40	
Deliver stock to settle forward contract	+ \$54
Pay back loan	– \$52
Net	+ \$2

In either case, we made \$2, free and clear. In fact, we can even accommodate the possibility of more than two future prices for AXE and we will always make \$2.[17] The key point is that we faced no risk and did not have to invest any of our own money, but ended up with \$2, which is clearly a good trade. The \$2 is an arbitrage profit. But where did it originate?

It turns out that the forward price, \$54, was an inappropriate price given current market conditions. In fact, it was just an arbitrary price made up to illustrate the point. To eliminate the opportunity to earn the \$2 profit, the forward price should be \$52, which is equal, not coincidentally, to the amount owed on the loan. It is also no coincidence that \$52 is the price of the asset increased by the rate of interest. We will cover this point later in the curriculum, but for now consider that you have just seen your first derivative pricing model.[18]

Of course, many market participants would do this transaction as long as it generated an arbitrage profit. These forces of arbitrage would either push the forward price down or the stock price up, or both, until an equilibrium is reached that eliminates the opportunity to profit at no risk with no commitment of one's own funds.

To summarize, the forces of arbitrage in financial markets assure us that the same asset cannot sell for different prices, nor can two equivalent combinations of assets that produce the same results sell for different prices. Realistically, some arbitrage opportunities can exist on a temporary basis, but they will be quickly exploited, bringing relative prices back in line with each other. Other apparent arbitrage opportunities will be too small to warrant exploiting.

Not to be naive, however, we must acknowledge that there is a large industry of people who call themselves arbitrageurs. So, how can such an industry exist if there are no opportunities for riskless profit? One explanation is that most of the arbitrage transactions are more complex than the simple examples used here. Many involve estimating information, which can result in differing opinions. Arbitrage involving options, for example, usually requires an estimate of a stock's volatility. Different participants have different opinions about the volatility. It is quite possible that the two counterparties trading with each other believe that each is arbitraging against the other.[19]

---

[17] A good study suggestion is to try this example with any future stock price. You should get the same result, a \$2 risk-free profit.

[18] This illustration is the quick look at forward pricing alluded to in Section 2.

But more importantly, the absence of arbitrage opportunities is upheld, ironically, only if participants believe that arbitrage opportunities do exist. If traders believe that no opportunities exist to earn arbitrage profits, then traders will not follow market prices and compare those prices with what they ought to be. Thus, eliminating arbitrage opportunities requires that participants be alert in watching for arbitrage opportunities. In other words, strange as it may sound, disbelief and skepticism concerning the absence of arbitrage opportunities are required for the no-arbitrage rule to be upheld.

Markets in which arbitrage opportunities are either nonexistent or quickly eliminated are relatively efficient markets. Recall that efficient markets are those in which it is not possible to consistently earn returns in excess of those that would be fair compensation for the risk assumed. Although abnormal returns can be earned in a variety of ways, arbitrage profits are definitely examples of abnormal returns. Thus, they are the most egregious violations of the principle of market efficiency.

Throughout the derivatives component of the CFA curriculum, we will use the principle of arbitrage as a dominant theme and assume that arbitrage opportunities cannot exist for any significant length of time nor can any one investor consistently capture them. Thus, prices must conform to models that assume no arbitrage. But we do not want to take the absence of arbitrage opportunities so seriously that we give up and believe that arbitrage opportunities never exist. Otherwise, they will arise and someone else will take them. Consider the rule of arbitrage a law that will be broken from time to time but one that holds far more often than not and one that should be understood and respected.

EXAMPLE 9

Arbitrage

1 Which of the following is a result of arbitrage?

- **A** The law of one price
- **B** The law of similar prices
- **C** The law of limited profitability

2 When an arbitrage opportunity exists, what happens in the market?

- **A** The combined actions of all arbitrageurs force the prices to converge.
- **B** The combined actions of arbitrageurs result in a locked-limit situation.
- **C** The combined actions of all arbitrageurs result in sustained profits to all.

3 Which of the following accurately defines arbitrage?

- **A** An opportunity to make a profit at no risk
- **B** An opportunity to make a profit at no risk and with the investment of no capital

---

[19] In reality, many of the transactions that arbitrageurs do are not really arbitrage. They are quite speculative. For example, many people call themselves arbitrageurs because they buy companies that are potential takeover targets and sell the companies they think will be the buyers. This transaction is not arbitrage by any stretch of the definition. Some transactions are called "risk arbitrage," but this term is an oxymoron. As an investment professional, you should simply be prepared for such misuses of words, which simply reflect the flexibility of language.

  C An opportunity to earn a return in excess of the return appropriate for the risk assumed

4 Which of the following ways best describes how arbitrage contributes to market efficiency?

  A Arbitrage penalizes those who trade too rapidly.

  B Arbitrage equalizes the risks taken by all market participants.

  C Arbitrage improves the rate at which prices converge to their relative fair values.

Solution to 1:

A is correct. Arbitrage forces equivalent assets to have a single price. There is nothing called the law of similar prices or the law of limited profitability.

Solution to 2:

A is correct. Prices converge because of the heavy demand for the cheaper asset and the heavy supply of the more expensive asset. Profits are not sustained, and, in fact, they are eradicated as prices converge. Locked-limit is a condition in the futures market and has nothing to do with arbitrage.

Solution to 3:

B is correct. An opportunity to profit at no risk could merely describe the purchase of a risk-free asset. An opportunity to earn a return in excess of the return appropriate for the risk assumed is a concept studied in portfolio management and is often referred to as an abnormal return. It is certainly desirable but is hardly an arbitrage because it requires the assumption of risk and the investment of capital. Arbitrage is risk free and requires no capital because selling the overpriced asset produces the funds to buy the underpriced asset.

Solution to 4:

C is correct. Arbitrage imposes no penalties on rapid trading; in fact, it tends to reward those who trade rapidly to take advantage of arbitrage opportunities. Arbitrage has no effect of equalizing risk among market participants. Arbitrage does result in an acceleration of price convergence to fair values relative to instruments with equivalent payoffs.

SUMMARY

This first reading on derivatives introduces you to the basic characteristics of derivatives, including the following points:

- A derivative is a financial instrument that derives its performance from the performance of an underlying asset.
- The underlying asset, called the underlying, trades in the cash or spot markets and its price is called the cash or spot price.
- Derivatives consist of two general classes: forward commitments and contingent claims.
- Derivatives can be created as standardized instruments on derivatives exchanges or as customized instruments in the over-the-counter market.

- Exchange-traded derivatives are standardized, highly regulated, and transparent transactions that are guaranteed against default through the clearinghouse of the derivatives exchange.
- Over-the-counter derivatives are customized, flexible, and more private and less regulated than exchange-traded derivatives, but are subject to a greater risk of default.
- A forward contract is an over-the-counter derivative contract in which two parties agree that one party, the buyer, will purchase an underlying asset from the other party, the seller, at a later date and at a fixed price they agree upon when the contract is signed.
- A futures contract is similar to a forward contract but is a standardized derivative contract created and traded on a futures exchange. In the contract, two parties agree that one party, the buyer, will purchase an underlying asset from the other party, the seller, at a later date and at a price agreed on by the two parties when the contract is initiated. In addition, there is a daily settling of gains and losses and a credit guarantee by the futures exchange through its clearinghouse.
- A swap is an over-the-counter derivative contract in which two parties agree to exchange a series of cash flows whereby one party pays a variable series that will be determined by an underlying asset or rate and the other party pays either a variable series determined by a different underlying asset or rate or a fixed series.
- An option is a derivative contract in which one party, the buyer, pays a sum of money to the other party, the seller or writer, and receives the right to either buy or sell an underlying asset at a fixed price either on a specific expiration date or at any time prior to the expiration date.
- A call is an option that provides the right to buy the underlying.
- A put is an option that provides the right to sell the underlying.
- Credit derivatives are a class of derivative contracts between two parties, the credit protection buyer and the credit protection seller, in which the latter provides protection to the former against a specific credit loss.
- A credit default swap is the most widely used credit derivative. It is a derivative contract between two parties, a credit protection buyer and a credit protection seller, in which the buyer makes a series of payments to the seller and receives a promise of compensation for credit losses resulting from the default of a third party.
- An asset-backed security is a derivative contract in which a portfolio of debt instruments is assembled and claims are issued on the portfolio in the form of tranches, which have different priorities of claims on the payments made by the debt securities such that prepayments or credit losses are allocated to the most-junior tranches first and the most-senior tranches last.
- Derivatives can be combined with other derivatives or underlying assets to form hybrids.
- Derivatives are issued on equities, fixed-income securities, interest rates, currencies, commodities, credit, and a variety of such diverse underlyings as weather, electricity, and disaster claims.
- Derivatives facilitate the transfer of risk, enable the creation of strategies and payoffs not otherwise possible with spot assets, provide information about the spot market, offer lower transaction costs, reduce the amount of capital required, are easier than the underlyings to go short, and improve the efficiency of spot markets.

- Derivatives are sometimes criticized for being a form of legalized gambling and for leading to destabilizing speculation, although these points can generally be refuted.
- Derivatives are typically priced by forming a hedge involving the underlying asset and a derivative such that the combination must pay the risk-free rate and do so for only one derivative price.
- Derivatives pricing relies heavily on the principle of storage, meaning the ability to hold or store the underlying asset. Storage can incur costs but can also generate cash, such as dividends and interest.
- Arbitrage is the condition that two equivalent assets or derivatives or combinations of assets and derivatives sell for different prices, leading to an opportunity to buy at the low price and sell at the high price, thereby earning a risk-free profit without committing any capital.
- The combined actions of arbitrageurs bring about a convergence of prices. Hence, arbitrage leads to the law of one price: Transactions that produce equivalent results must sell for equivalent prices.

PRACTICE PROBLEMS

1 A derivative is *best* described as a financial instrument that derives its performance by:

- **A** passing through the returns of the underlying.
- **B** replicating the performance of the underlying.
- **C** transforming the performance of the underlying.

2 Derivatives are similar to insurance in that both:

- **A** have an indefinite life span.
- **B** allow for the transfer of risk from one party to another.
- **C** allow for the transformation of the underlying risk itself.

3 A beneficial opportunity created by the derivatives market is the ability to:

- **A** adjust risk exposures to desired levels.
- **B** generate returns proportional to movements in the underlying.
- **C** simultaneously take long positions in multiple highly liquid fixed-income securities.

4 Compared with exchange-traded derivatives, over-the-counter derivatives would *most likely* be described as:

- **A** standardized.
- **B** less transparent.
- **C** more transparent.

5 Exchange-traded derivatives are:

- **A** largely unregulated.
- **B** traded through an informal network.
- **C** guaranteed by a clearinghouse against default.

6 The clearing and settlement process of an exchange-traded derivatives market:

- **A** provides a credit guarantee.
- **B** provides transparency and flexibility.
- **C** takes longer than that of most securities exchanges.

7 Which of the following statements *best* portrays the full implementation of post-financial-crisis regulations in the OTC derivatives market?

- **A** Transactions are no longer private.
- **B** Most transactions need to be reported to regulators.
- **C** All transactions must be cleared through central clearing agencies.

8 A characteristic of forward commitments is that they:

- **A** provide linear payoffs.
- **B** do not depend on the outcome or payoff of an underlying asset.
- **C** provide one party the right to engage in future transactions on terms agreed on in advance.

9 In contrast to contingent claims, forward contracts:

- **A** have their prices chosen by the participants.
- **B** could end in default by either party.

© 2019 CFA Institute. All rights reserved.

C can be exercised by physical or cash delivery.

10 Which of the following statements *best* describes the payoff from a forward contract?

A The buyer has more to gain going long than the seller has to lose going short.

B The buyer profits if the price of the underlying at expiration exceeds the forward price.

C The gains from owning the underlying versus owning the forward contract are equivalent.

11 Which of the following statements regarding the settlement of forward contracts is correct?

A Contract settlement by cash has different economic effects from those of a settlement by delivery.

B Non-deliverable forwards and contracts for differences have distinct settlement procedures.

C At cash settlement, when the long party acquires the asset in the market, it effectively pays the forward price.

12 A futures contract is *best* described as a contract that is:

A standardized.

B subject to credit risk.

C marked to market throughout the trading day.

13 Which of the following statements explains a characteristic of futures price limits? Price limits:

A help the clearinghouse manage its credit exposure.

B can typically be expanded intra-day by willing traders.

C establish a band around the final trade of the previous day.

14 Which of the following statements describes an aspect of margin accounts for futures?

A The maintenance margin is always less than the initial margin.

B The initial margin required is typically at least 10% of the futures price.

C A margin call requires a deposit sufficient to raise the account balance to the maintenance margin.

15 Which of the following factors is shared by forwards and futures contracts?

A Timing of profits

B Flexible settlement arrangements

C Nearly equivalent profits by expiration

16 Which of the following derivatives is classified as a contingent claim?

A Futures contracts

B Interest rate swaps

C Credit default swaps

17 In contrast to contingent claims, forward commitments provide the:

A right to buy or sell the underlying asset in the future.

B obligation to buy or sell the underlying asset in the future.

C promise to provide credit protection in the event of default.

18 Which of the following derivatives provide payoffs that are non-linearly related to the payoffs of the underlying?

- **A** Options
- **B** Forwards
- **C** Interest-rate swaps

19 An interest rate swap is a derivative contract in which:

- **A** two parties agree to exchange a series of cash flows.
- **B** the credit seller provides protection to the credit buyer.
- **C** the buyer has the right to purchase the underlying from the seller.

20 Forward commitments subject to default are:

- **A** forwards and futures.
- **B** futures and interest rate swaps.
- **C** interest rate swaps and forwards.

21 A swap is:

- **A** more like a forward than a futures contract.
- **B** subject to simultaneous default by both parties.
- **C** based on an exchange of two series of fixed cash flows.

22 A plain vanilla interest rate swap is also known as:

- **A** a basis swap.
- **B** a fixed-for-floating swap.
- **C** an overnight indexed swap.

23 The notional principal of a swap is:

- **A** not exchanged in the case of an interest rate swap.
- **B** a fixed amount whenever it is matched with a loan.
- **C** equal to the amount owed by one swap party to the other.

24 Which of the following derivatives is *least likely* to have a value of zero at initiation of the contract?

- **A** Futures
- **B** Options
- **C** Forwards

25 The buyer of an option has a contingent claim in the sense that the option creates:

- **A** a right.
- **B** an obligation.
- **C** a linear payoff with respect to gains and losses of the underlying.

26 Which of the following options grants the holder the right to purchase the underlying prior to expiration?

- **A** American-style put option
- **B** European-style call option
- **C** American-style call option

27 A credit derivative is a derivative contract in which the:

- **A** clearinghouse provides a credit guarantee to both the buyer and the seller.
- **B** seller provides protection to the buyer against the credit risk of a third party.
- **C** the buyer and seller provide a performance bond at initiation of the contract.

28 The junior and senior tranches of an asset-backed security:

A have equivalent expected returns.

B have claims on separate underlying portfolios.

C may be differentially impacted by prepayments or credit losses.

29 In a declining interest rate environment, compared with a CMO's Class A tranche, its Class C tranche will be repaid:

A earlier.

B at the same pace.

C later.

30 For a given CDO, which of the following tranches is *most likely* to have the highest expected return?

A Equity

B Senior

C Mezzanine

31 Which of the following derivatives allows an investor to pay the return on a stock index and receive a fixed rate?

A Equity swap

B Stock warrant

C Index futures contract

32 Which of the following is *most likely* the underlying of a plain vanilla interest rate swap?

A 180-day Libor

B 10-year US Treasury bond

C Bloomberg Barclay's US Aggregate Bond Index

33 Currency swaps are:

A rarely used.

B commonly used to manage interest rate risk.

C executed by two parties making a series of interest rate payments in the same currency.

34 Which of the following statements regarding commodity derivatives is correct?

A The primary commodity derivatives are futures.

B Commodities are subject to a set of well-defined risk factors.

C Commodity traders and financial traders today are distinct groups within the financial world.

35 Compared with the underlying spot market, derivative markets are *more likely* to have:

A greater liquidity.

B higher transaction costs.

C higher capital requirements.

36 Which of the following characteristics is *least likely* to be a benefit associated with using derivatives?

A More effective management of risk

B Payoffs similar to those associated with the underlying

C Greater opportunities to go short compared with the spot market

37 Which of the following statements *best* represents information discovery in the futures market?

A The futures price is predictive.

B Information flows more slowly into the futures market than into the spot market.

C The futures market reveals the price that the holder of the asset can take to avoid uncertainty.

38 The derivative markets tend to:

A transfer liquidity from the broader financial markets.

B not reflect fundamental value after it is restored in the underlying market.

C offer a less costly way to exploit mispricing in comparison to other free and competitive financial markets.

39 Which of the following statements *most likely* contributes to the view that derivatives have some role in causing financial crashes?

A Derivatives are the primary means by which leverage and related excessive risk is brought into financial markets.

B Growth in the number of investors willing to speculate in derivatives markets leads to excessive speculative trading.

C Restrictions on derivatives, such as enhanced collateral requirements and credit mitigation measures, in the years leading up to crashes introduce market rigidity.

40 In contrast to gambling, derivatives speculation:

A has a positive public image.

B is a form of financial risk taking.

C benefits the financial markets and thus society.

41 Derivatives may contribute to financial contagion because of the:

A centrally cleared nature of OTC derivatives.

B associated significant costs and high capital requirements.

C reliance by derivatives speculators on large amounts of leverage.

42 The complex nature of derivatives has led to:

A reliable financial models of derivatives markets.

B widespread trust in applying scientific principles to derivatives.

C financial industry employment of mathematicians and physicists.

43 Which of the following is *most likely* to be a destabilizing consequence of speculation using derivatives?

A Increased defaults by speculators and creditors

B Market price swings resulting from arbitrage activities

C The creation of trading strategies that result in asymmetric performance

44 The law of one price is *best* described as:

A the true fundamental value of an asset.

B earning a risk-free profit without committing any capital.

C two assets that will produce the same cash flows in the future must sell for equivalent prices.

45 Arbitrage opportunities exist when:

A two identical assets or derivatives sell for different prices.

B combinations of the underlying asset and a derivative earn the risk-free rate.

C arbitrageurs simultaneously buy takeover targets and sell takeover acquirers.

For questions 46–49, consider a call option selling for $4 in which the exercise price is $50

46 Determine the value at expiration and the profit for a *buyer* if the price of the underlying at expiration is $55.

- **A** $5
- **B** $1
- **C** –$1

47 Determine the value at expiration and the profit for a *buyer* if the price of the underlying at expiration is $48.

- **A** –$4
- **B** $0
- **C** $2

48 Determine the value at expiration and the profit for a *seller* if the price of the underling at expiration is $49.

- **A** $4
- **B** $0
- **C** –$1

49 Determine the value at expiration and the profit for a *seller* if the price of the underling at expiration is $52.

- **A** –$2
- **B** $5
- **C** $2

For questions 50–52, consider the following scenario

Suppose you believe that the price of a particular underlying, currently selling at $99, is going to increase substantially in the next six months. You decide to purchase a call option expiring in six months on this underlying. The call option has an exercise price of $105 and sells for $7.

50 Determine the profit if the price of the underlying six months from now is $99.

- **A** $6
- **B** $0
- **C** –$7

51 Determine the profit if the price of the underlying six months from now is $112.

- **A** $7
- **B** $0
- **C** –$3

52 Determine the profit if the price of the underlying six months from now is $115.

- **A** $0

B $3

C –$3

For questions 53–55, consider the following scenario

Suppose you believe that the price of a particular underlying, currently selling at $99, is going to decrease substantially in the next six months. You decide to purchase a put option expiring in six months on this underlying. The put option has an exercise price of $95 and sells for $5.

53 Determine the profit for you if the price of the underlying six months from now is $100.

A $0

B $5

C –$5

54 Determine the profit for you if the price of the underlying six months from now is $95.

A $0

B $5

C –$5

55 Determine the profit for you if the price of the underlying six months from now is $85.

A $10

B $5

C $0

SOLUTIONS

1 C is correct. A derivative is a financial instrument that transforms the performance of the underlying. The transformation of performance function of derivatives is what distinguishes it from mutual funds and exchange traded funds that pass through the returns of the underlying.

A is incorrect because derivatives, in contrast to mutual funds and exchange traded funds, do not simply pass through the returns of the underlying at payout. B is incorrect because a derivative transforms rather than replicates the performance of the underlying.

2 B is correct. Insurance is a financial contract that provides protection against loss. The party bearing the risk purchases an insurance policy, which transfers the risk to the other party, the insurer, for a specified period of time. The risk itself does not change, but the party bearing it does. Derivatives allow for this same type of risk transfer.

A is incorrect because derivatives, like insurance, have a definite, as opposed to indefinite, life span and expire on a specified date.

C is incorrect because both derivatives and insurance allow for the transfer of risk from one party (the purchaser of the insurance policy or of a derivative) to another party (the insurer or a derivative seller), for a specified period of time. The risk itself does not change, but the party bearing it does.

3 A is correct. Derivatives allow market participants to practice more effective risk management, a process by which an organization, or individual, defines the level of risk it wishes to take, measures the level of risk it is taking, and adjusts the latter to equal the former.

B is incorrect because derivatives are characterized by a relatively high degree of leverage, meaning that participants in derivatives transactions usually have to invest only a small amount, as opposed to a large amount, of their own capital relative to the value of the underlying. This allows participants to generate returns that are disproportional, as opposed to proportional, to movements in the underlying.

C is incorrect because derivatives are not needed to copy strategies that can be implemented with the underlying on a standalone basis. Rather, derivatives can be used to create strategies that cannot be implemented with the underlying alone. Simultaneously taking long positions in multiple highly liquid fixed-income securities is a strategy that can be implemented with the underlying securities on a standalone basis.

4 B is correct. Over-the counter-derivatives markets are customized and mostly unregulated. As a result, over-the-counter markets are less transparent in comparison with the high degree of transparency and standardization associated with exchange-traded derivative markets.

A is incorrect because exchange-traded derivatives are standardized, whereas over-the counter derivatives are customized. C is incorrect because exchange-traded derivatives are characterized by a high degree of transparency because all transactions are disclosed to exchanges and regulatory agencies, whereas over-the-counter derivatives are relatively opaque.

5 C is correct. Exchanged-traded derivatives are guaranteed by a clearinghouse against default.

A is incorrect because traded derivatives are characterized by a relatively high degree of regulation. B is incorrect because the terms of exchange-traded derivatives terms are specified by the exchange.

6 A is correct. The clearing and settlement process of derivative transactions provides a credit guarantee.

B is incorrect because although the exchange markets are said to have transparency, they also involve standardization. That entails a loss of flexibility, with participants limited to only those transactions permitted on the exchange.

C is incorrect because derivatives exchanges clear and settle all contracts overnight, which is faster than most securities exchanges, which require two business days.

7 B is correct. With full implementation of these regulations in the OTC derivatives market, most OTC transactions need to be reported to regulators.

A is incorrect because although under full implementation of the regulations information on most OTC transactions needs to be reported to regulators, many transactions retain a degree of privacy with lower transparency.

C is incorrect because although under full implementation of new regulations a number of OTC transactions have to be cleared through central clearing agencies, there are exemptions that cover a significant percentage of derivative transactions.

8 A is correct because forward commitments provide linear payoffs.

B is incorrect because forward commitments depend on the outcome or payoff of an underlying asset.

C is incorrect because forward commitments obligate parties to make (not provide the right to engage) a final payment contingent on the performance of the underlying.

9 B is correct. In a forward contract, either party could default, whereas in a contingent claim, default is possible only from the short to the long.

A is incorrect because the forward price is set in the pricing of the contract such that the starting contract value is zero, unlike contingent claims, under which parties can select any starting value.

C is incorrect because both forward contracts and contingent claims can be settled by either physical or cash delivery.

10 B is correct. The buyer is obligated to pay the forward price $F_0(T)$ at expiration and receives an asset worth S_T, the price of the underlying. The contract effectively pays off $S_T - F_0(T)$, the value of the contract at expiration. The buyer therefore profits if $S_T > F_0(T)$.

A is incorrect because the long and the short are engaged in a zero-sum game. This is a type of competition in which one participant's gains are the other's losses, with their payoffs effectively being mirror images.

C is incorrect because although the gain from owning the underlying and the gain from owning the forward are both driven by S_T, the price of the underlying at expiration, they are not the same value. The gain from owning the underlying would be $S_T - S_0$, the change in its price, whereas the gain from owning the forward would be $S_T - F_0(T)$, the value of the forward at expiration.

11 C is correct. In the case of cash settlement, the long can acquire the asset, effectively paying the forward price, $F_0(T)$.

A is incorrect because forward contracts settled by cash or by delivery have the same economic effect.

B is incorrect because both non-deliverable forwards and contracts for differences can settle by an exchange of cash.

12 A is correct. A futures contract is a standardized derivative contract.

B is incorrect because through its clearinghouse the futures exchange provides a credit guarantee that it will make up a loss in the event a losing party cannot pay.

C is incorrect because a futures contract is marked to market at the end of each day, a process in which the futures clearinghouse determines an average of the final futures trade of the day and designates that price as the settlement price.

13 A is correct. Price limits are important in helping the clearinghouse manage its credit exposure. Sharply moving prices make it more difficult for the clearinghouse to collect from parties losing money.

B is incorrect because typically the exchange rules allow for an expansion of price limits the next day (not intra-day) if traders are willing.

C is incorrect because price limits establish a band relative to the previous day's settlement price (not final trade).

14 A is correct. The maintenance margin is always significantly lower than the initial margin.

B is incorrect because the initial margin required is typically at most (not at least) 10% of the futures price.

C is incorrect because a margin call requires a deposit large enough to bring the balance up to the initial (not maintenance) margin.

15 C is correct. Comparing the derivatives, forward and futures contracts have nearly equivalent profits by the time of expiration of the forward.

A is incorrect because the timing of profits for a futures contract is different from that of forwards. Forwards realize the full amount at expiration, whereas futures contracts realize their profit in parts on a day-to-day basis.

B is incorrect because the settlement arrangements for the forwards can be agreed on at initiation and written in the contract based on the desires of the engaged parties. However, in the case of a futures contract, the exchange (not the engaged parties) specifies whether physical delivery or cash settlement applies.

16 C is correct. A credit default swap (CDS) is a derivative in which the credit protection seller provides protection to the credit protection buyer against the credit risk of a separate party. CDS are classified as a contingent claim.

A is incorrect because futures contracts are classified as forward commitments. B is incorrect because interest rate swaps are classified as forward commitments.

17 B is correct. Forward commitments represent an obligation to buy or sell the underlying asset at an agreed upon price at a future date.

A is incorrect because the right to buy or sell the underlying asset is a characteristic of contingent claims, not forward commitments. C is incorrect because a credit default swap provides a promise to provide credit protection to the credit protection buyer in the event of a credit event such as a default or credit downgrade and is classified as a contingent claim.

18 A is correct. Options are classified as a contingent claim which provides payoffs that are non-linearly related to the performance of the underlying.

B is incorrect because forwards are classified as a forward commitment, which provides payoffs that are linearly related to the performance of the underlying. C is incorrect because interest-rate swaps are classified as a forward commitment, which provides payoffs that are linearly related to the performance of the underlying.

19 A is correct. An interest rate swap is defined as a derivative in which two parties agree to exchange a series of cash flows: One set of cash flows is variable, and the other set can be variable or fixed.

B is incorrect because a credit derivative is a derivative contract in which the credit protection seller provides protection to the credit protection buyer. C is incorrect because a call option gives the buyer the right to purchase the underlying from the seller.

20 C is correct. Interest rate swaps and forwards are over-the-counter contracts that are privately negotiated and are both subject to default. Futures contracts are traded on an exchange, which provides a credit guarantee and protection against default.

A is incorrect because futures are exchange-traded contracts which provide daily settlement of gains and losses and a credit guarantee by the exchange through its clearinghouse. B is incorrect because futures are exchange-traded contracts which provide daily settlement of gains and losses and a credit guarantee by the exchange through its clearinghouse.

21 A is correct. A swap is a bit more like a forward contract than a futures contract in that it is an OTC contract, so it is privately negotiated and subject to default.

B is incorrect because in a swap, although either party can default, only one party can do so at a particular time. Money owed is based on the net owed by one party to the other, and only the party owing the greater amount can default to the counterparty owing the lesser amount.

C is incorrect because a swap involves an exchange between parties in which at least one party pays a variable series of cash flows determined by an underlying asset or rate.

22 B is correct. A plain vanilla swap is a fixed-for-floating interest rate swap, which is the most common type of swap.

A is incorrect because a basis swap is a transaction based on the TED spread (T-bills versus Eurodollars) and is not the same as a plain vanilla swap.

C is incorrect because an overnight indexed swap is a swap that is tied to a federal funds type of rate, reflecting the rate at which banks borrow overnight, and is not the same as a plain vanilla swap.

23 A is correct. The notional principal of a swap is not exchanged in the case of an interest rate swap.

B is incorrect because an amortizing loan will be matched with a swap with a pre-specified declining (not fixed) notional principal that matches the loan balance.

C is incorrect because the notional principal is equal to the loan balance. Although the loan has an actual balance (the amount owed by the borrower to the creditor), the swap does not have such a balance owed by one swap party to the other.

24 B is correct. The buyer of the option pays the option premium to the seller of the option at the initiation of the contract. The option premium represents the value of the option, whereas futures and forwards have a value of zero at the initiation of the contract.

A is incorrect because no money changes hands between parties at the initiation of the futures contract, thus the value of the futures contract is zero at initiation. C is incorrect because no money changes hands between parties at the initiation of the forward contract, thus the value of the forward contract is zero at initiation.

25 A is correct. A contingent claim, a derivative in which the outcome or payoff depends on the outcome or payoff of an underlying asset, has come to be associated with a right, but not an obligation, to make a final payment contingent on the performance of the underlying.

B is incorrect because an option, as a contingent claim, grants the right but not the obligation to buy or sell the underlying at a later date.

C is incorrect because the holder of an option has a choice of whether to exercise the option. This choice creates a payoff that transforms the underlying payoff in a more pronounced manner than does a forward, futures, or swap, which provide linear payoffs. Options are different in that they limit losses in one direction.

26 C is correct. The right to buy the underlying is referred to as a call option. Furthermore, options that can be exercised prior to the expiration date are referred to as American-style options.

A is incorrect because a put option grants the holder the right to sell, as opposed to buy, the underlying.

B is incorrect because European-style options can only be exercised at expiration.

27 B is correct. A credit derivative is a derivative contract in which the credit protection seller provides protection to the credit protection buyer against the credit risk of a third party.

A is incorrect because the clearinghouse provides a credit guarantee to both the buyer and the seller of a futures contract, whereas a credit derivative is between two parties, in which the credit protection seller provides a credit guarantee to the credit protection buyer. C is incorrect because futures contracts require that both the buyer and the seller of the futures contract provide a cash deposit for a portion of the futures transaction into a margin account, often referred to as a performance bond or good faith deposit.

28 C is correct. An asset-backed security is a derivative contract in which a portfolio of debt instruments is assembled and claims are issued on the portfolio in the form of tranches, which have different priorities of claims on the payments made by the debt securities such that prepayments or credit losses are allocated to the most junior tranches first and the most senior tranches last.

A is incorrect because the expected returns of the tranches vary according to the perceived credit risk, with the senior tranches having the highest credit quality and the junior tranches the lowest. Thus, the senior tranches have the lowest expected returns and the junior tranches have the highest. Notably, in a bond mutual fund or an ETF, all investors in the fund have equal claims, and so the rate of return earned by each investor is the same.

B is incorrect because an asset-backed security is a derivative contract in which a single portfolio of securities is assembled and claims are issued on the portfolio in the form of tranches.

29 A is correct. Lower interest rates entice homeowners to pay off their mortgages early because they can refinance at lower rates. The most junior tranche in a CMO will bear the first wave of prepayments until that tranche has been completely repaid its full principal investment. At that point, the next tranche will bear prepayments until that tranche has been fully repaid. Therefore, the Class C tranche of a CMO will be repaid before the more senior Class A tranche.

B is incorrect because the tranches, which have different priorities of claims on the principal payments made by the underlying mortgages, will see prepayments allocated to the most junior tranches first and the most senior tranches last.

C is incorrect because the most junior tranche in a CMO will bear the first wave of prepayments until that tranche has been completely repaid its full principal investment. At that point, the next tranche will bear prepayments until that tranche has been fully repaid. Therefore, the Class C tranche will be repaid prior to, not after, the Class A tranche.

30 A is correct. The expected returns of the tranches vary according to the perceived credit risk, with the senior tranches having the highest credit quality and the junior tranches the lowest. Thus, the senior tranches have the lowest expected returns and the junior tranches have the highest. The most junior tranche is sometimes called the "equity tranche."

B is incorrect because the senior tranches in a CDO have the lowest expected returns and the junior (or equity) tranches have the highest.

C is incorrect because the senior tranches in a CDO have the lowest expected returns and the junior (or equity) tranches have the highest. A mezzanine tranche is intermediate between the senior and junior tranches.

31 A is correct. Equity swaps, also known as index swaps, are quite popular and permit investors to pay the return on one stock index and receive the return on another index or a fixed rate.

B is incorrect because warrants are options that are sold directly to the public, allowing holders to exercise and buy shares directly from the company as opposed to using stock indexes to determine returns.

C is incorrect because although index derivatives in the form of options, forwards, futures, and swaps are very popular, paying the return on a stock index and receiving a fixed rate describes an equity swap (or index swap), not a futures contract.

32 A is correct. In a plain vanilla interest rate swap, an interest rate, such as Libor, serves as the underlying. A plain vanilla interest rate swap is one of many derivatives in which a rate, not the instrument that pays the rate, is the underlying.

B is incorrect because a plain vanilla interest rate swap is one of many derivatives in which a rate, not an instrument that pays a rate, is the underlying.

C is incorrect because a plain vanilla interest rate swap is one of many derivatives in which a rate, not an instrument (or index) that pays a rate, is the underlying.

33 B is correct. Because interest rates and currencies are both subject to change, a currency swap has two sources of risk. Furthermore, companies operating across borders are subject to both interest rate risk and currency risk, and currency swaps are commonly used to manage these risks.

A is incorrect because currency risk is a major factor in global financial markets, and the currency derivatives market is extremely large, as opposed to small.

C is incorrect because a currency swap is executed by two parties making a series of interest rate payments to each other in different currencies, as opposed to the same currency.

34 A is correct. The primary commodity derivatives are futures, but forwards, swaps, and options are also used.

B is incorrect because the commodity market is extremely large and subject to an almost unimaginable array of risks.

C is incorrect because commodity and financial traders have become relatively homogeneous since the creation of financial futures. Historically, commodity traders and financial traders were quite different groups, and there used to be a tendency to think of the commodity world as somewhat separate from the financial world.

35 A is correct. Derivative markets typically have greater liquidity than the underlying spot market as a result of the lower capital required to trade derivatives compared with the underlying. Derivatives also have lower transaction costs and lower capital requirements than the underlying.

B is incorrect because transaction costs for derivatives are lower than the underlying spot market. C is incorrect because derivatives markets have lower capital requirements than the underlying spot market.

36 B is correct. One of the benefits of derivative markets is that derivatives create trading strategies not otherwise possible in the underlying spot market, thus providing opportunities for more effective risk management than simply replicating the payoff of the underlying.

A is incorrect because effective risk management is one of the primary purposes associated with derivative markets. C is incorrect because one of the operational advantages associated with derivatives is that it is easier to go short compared to the underlying spot market.

37 C is correct. The futures market reveals the price that the holder of an asset could take and avoid the risk of uncertainty.

A is incorrect because although the futures price is sometimes thought of as predictive, it provides only a little more information than does a spot price and is not really a forecast of the futures spot price.

B is incorrect because by virtue of the fact that the futures market requires less capital, information can flow into the futures market before it gets into the spot market.

38 C is correct. When prices deviate from fundamental values, derivative markets offer a less costly way to exploit mispricing in comparison to other free and competitive financial markets.

A is incorrect because derivative markets tend to transfer liquidity to (not from) the broader financial markets, because investors are far more willing to trade if they can more easily manage their risk, trade at lower cost and with less capital, and go short more easily. An increased willingness to trade leads to a more liquid market.

B is incorrect because it is likely (not unlikely) that fundamental value will be reflected in the derivative markets both before and after it is restored in the underlying market owing to lower capital requirements and transaction costs in the derivative markets.

39 B is correct. Opponents of derivatives claim that excessive speculative trading brings instability to the markets. Defaults by speculators can lead to defaults by their creditors, their creditors' creditors, and so on.

A is incorrect because derivatives are one of many mechanisms through which excessive risk can be taken. There are many ways to take on leverage that look far less harmful but can be just as risky.

C is incorrect because responses to crashes and crises typically call for more rules and regulations restricting the use of derivatives, such as requiring more collateral and credit mitigation measures. Such rules and regulations are generally implemented after a crash and are directed at limiting government bailouts of the costs from derivatives risks.

40 C is correct. Derivatives trading brings extensive benefits to financial markets (low costs, low capital requirements, ease of going short, etc.) and thus benefits society as a whole. Gambling, on the other hand, typically benefits only a limited number of participants.

A is incorrect because the general image of speculators is not a good one. Speculators are often thought to be short-term traders who attempt to exploit temporary inefficiencies, caring little about long-term fundamental values.

B is incorrect because speculation and gambling are both forms of financial risk taking.

41 C is correct. Opponents argue that speculators use large amounts of leverage, thereby subjecting themselves and their creditors to substantial risk if markets do not move in their hoped-for direction. Defaults by speculators can then lead to defaults by their creditors, their creditors' creditors, and so on. These effects can, therefore, be systemic and reflect an epidemic contagion whereby instability can spread throughout markets and an economy, if not the entire world.

A is incorrect because central clearing of OTC derivatives, similar to how exchange-traded derivatives are cleared, is intended to lessen the risk of contagion.

B is incorrect because it is derivatives' low cost and low capital requirements, not high cost and high capital requirements, that opponents point to as contributing to an excessive amount of speculative trading that brings instability to the markets.

42 C is correct. Many derivatives are extremely complex and require a high-level understanding of mathematics. As a result, the financial industry employs many mathematicians, physicists, and computer scientists.

A is incorrect because scientists create models of markets by using scientific principles that often fail. For example, to a physicist modeling the movements of celestial bodies, the science is reliable and the physicist is unlikely to misapply the science. The same science applied to financial markets is far less reliable. Financial markets are driven by the actions of people who are not as consistent as the movements of celestial bodies.

B is incorrect because the complex nature of derivatives has made many distrust, as opposed to trust, derivatives, the people who work with them, and the scientific methods they use.

43 A is correct. The benefits of derivatives, such as low transaction costs, low capital requirements, use of leverage, and the ease in which participants can go short, also can result in excessive speculative trading. These activities can lead to defaults on the part of speculators and creditors.

B is incorrect because arbitrage activities tend to bring about a convergence of prices to intrinsic value. C is incorrect because asymmetric performance is not itself destabilizing.

44 C is correct. The law of one price occurs when market participants engage in arbitrage activities so that identical assets sell for the same price in different markets.

A is incorrect because the law of one price refers to identical assets. B is incorrect because it refers to arbitrage not the law of one price.

45 A is correct. Arbitrage opportunities exist when the same asset or two equivalent combinations of assets that produce the same results sell for different prices. When this situation occurs, market participants would buy the asset in the cheaper market and simultaneously sell it in the more expensive market, thus earning a riskless arbitrage profit without committing any capital.

B is incorrect because it is not the definition of an arbitrage opportunity. C is incorrect because it is not the definition of an arbitrage opportunity.

46 B is correct. $C_T = Max(0,S_T - X) = Max(0,55 - 50) = 5$

$\Pi = C_T - C_0 = 5 - 4 = 1$

47 A is correct. $C_T = Max(0,S_T - X) = Max(0,48 - 50) = 0$

$\Pi = C_T - C_0 = 0 - 4 = -4$

48 A is correct. $-C_T = -Max(0,S_T - X) = -Max(0,49 - 50) = 0$

$\Pi = -C_T + C_0 = -0 + 4 = 4$

49 C is correct. $-C_T = -Max(0,S_T - X) = -Max(0,52 - 50) = -2$

$\Pi = -C_T + C_0 = -2 + 4 = 2$

50 C is correct. $C_T = Max(0,S_T - X) = Max(0,99 - 105) = 0$

$\Pi = C_T - C_0 = 0 - 7 = -7$

51 B is correct. $C_T = Max(0,S_T - X) = Max(0,112 - 105) = 7$

$\Pi = C_T - C_0 = 7 - 7 = 0$

Note: $112 is the breakeven price

52 B is correct. $C_T = Max(0,S_T - X) = Max(0,115 - 105) = 10$

$\Pi = C_T - C_0 = 10 - 7 = 3$

53 C is correct. $P_T = Max(0, X - S_T) = Max(0,95 - 100) = 0$

$\Pi = P_{\mathrm{T}} - P_0 = 0 - 5 = -5$

54 C is correct. $P_T = Max(0, X - S_T) = Max(0,95 - 95) = 0$

$\Pi = P_T - P_0 = 0 - 5 = -5$

55 B is correct. $P_T = Max(0, X - S_T) = Max(0,95 - 85) = 10$

$\Pi = P_T - P_0 = 10 - 5 = 5$

READING

46

Basics of Derivative Pricing and Valuation

by Don M. Chance, PhD, CFA

Don M. Chance, PhD, CFA, is at Louisiana State University (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** explain how the concepts of arbitrage, replication, and risk neutrality are used in pricing derivatives;
☐	**b.** explain the difference between value and price of forward and futures contracts;
☐	**c.** calculate a forward price of an asset with zero, positive, or negative net cost of carry;
☐	**d.** explain how the value and price of a forward contract are determined at expiration, during the life of the contract, and at initiation;
☐	**e.** describe monetary and nonmonetary benefits and costs associated with holding the underlying asset and explain how they affect the value and price of a forward contract;
☐	**f.** define a forward rate agreement and describe its uses;
☐	**g.** explain why forward and futures prices differ;
☐	**h.** explain how swap contracts are similar to but different from a series of forward contracts;
☐	**i.** explain the difference between value and price of swaps;
☐	**j.** explain the exercise value, time value, and moneyness of an option;
☐	**k.** identify the factors that determine the value of an option and explain how each factor affects the value of an option;
☐	**l.** explain put–call parity for European options;
☐	**m.** explain put–call–forward parity for European options;
☐	**n.** explain how the value of an option is determined using a one-period binomial model;
☐	**o.** explain under which circumstances the values of European and American options differ.

© 2019 CFA Institute. All rights reserved.

1 INTRODUCTION

It is important to understand how prices of derivatives are determined. Whether one is on the buy side or the sell side, a solid understanding of pricing financial products is critical to effective investment decision making. After all, one can hardly determine what to offer or bid for a financial product, or any product for that matter, if one has no idea how its characteristics combine to create value.

Understanding the pricing of financial assets is important. Discounted cash flow methods and models, such as the capital asset pricing model and its variations, are useful for determining the prices of financial assets. The unique characteristics of derivatives, however, pose some complexities not associated with assets, such as equities and fixed-income instruments. Somewhat surprisingly, however, derivatives also have some simplifying characteristics. For example, as we will see in this reading, in well-functioning derivatives markets the need to determine risk premiums is obviated by the ability to construct a risk-free hedge. Correspondingly, the need to determine an investor's risk aversion is irrelevant for derivative pricing, although it is certainly relevant for pricing the underlying.

The purpose of this reading is to establish the foundations of derivative pricing on a basic conceptual level. The following topics are covered:

- How does the pricing of the underlying asset affect the pricing of derivatives?
- How are derivatives priced using the principle of arbitrage?
- How are the prices and values of forward contracts determined?
- How are futures contracts priced differently from forward contracts?
- How are the prices and values of swaps determined?
- How are the prices and values of European options determined?
- How does American option pricing differ from European option pricing?

This reading is organized as follows. Sections 2–4 explore two related topics, the pricing of the underlying assets on which derivatives are created and the principle of arbitrage. Sections 5–7 describe the pricing and valuation of forwards, futures, and swaps. Sections 8–12 introduce the pricing and valuation of options. Section 13 provides a summary.

2 BASIC DERIVATIVE CONCEPTS, PRICING THE UNDERLYING

d explain how the value and price of a forward contract are determined at expiration, during the life of the contract, and at initiation;

In this section, we will briefly review the concepts associated with derivatives, the types of derivatives, and the pricing principles of the underlying assets. We will also look at arbitrage, a critical concept that links derivative pricing to the price of the underlying.

2.1 Basic Derivative Concepts

The definition of a derivative is as follows:

> *A derivative is a financial instrument that derives its performance from the performance of an underlying asset.*

A derivative is created as a contract between two parties, the buyer and the seller. Derivatives trade in markets around the world, which include organized exchanges, where highly standardized and regulated versions exist, and over-the-counter markets, where customized and more lightly regulated versions trade. The basic characteristics of derivatives that influence pricing are not particularly related to where the derivatives trade, but are critically dependent on the types of derivatives.

The two principal types of derivatives are forward commitments and contingent claims. A forward commitment is an obligation to engage in a transaction in the spot market at a future date at terms agreed upon today.[1] By entering into a forward commitment, a party locks in the terms of a transaction that he or she will conduct later. The word "commitment" is critical here. A forward contract is a firm obligation.

There are three types of forward commitments: forward contracts, futures contracts, and swap contracts. These contracts can be referred to more simply as forwards, futures, and swaps.

> *A* **forward contract** *is an over-the-counter derivative contract in which two parties agree that one party, the buyer, will purchase an underlying asset from the other party, the seller, at a later date at a fixed price they agree upon when the contract is signed.*
>
> *A* **futures contract** *is a standardized derivative contract created and traded on a futures exchange in which two parties agree that one party, the buyer, will purchase an underlying asset from the other party, the seller, at a later date at a price agreed upon by the two parties when the contract is initiated and in which there is a daily settling of gains and losses and a credit guarantee by the futures exchange through its clearinghouse.*
>
> *A* **swap contract** *is an over-the-counter derivative contract in which two parties agree to exchange a series of cash flows whereby one party pays a variable series that will be determined by an underlying asset or rate and the other party pays either 1) a variable series determined by a different underlying asset or rate or 2) a fixed series.*

As these definitions illustrate, forwards and futures are similar. They both establish the terms of a spot transaction that will occur at a later date. Forwards are customized, less transparent, less regulated, and subject to higher counterparty default risk. Futures are standardized, more transparent, more regulated, and generally immune to counterparty default. A swap is equivalent to a series of forward contracts, a point that will be illustrated later.

A contingent claim is a derivative in which the outcome or payoff is determined by the outcome or payoff of an underlying asset, conditional on some event occurring. Contingent claims include options, credit derivatives, and asset-backed securities. Because credit derivatives and asset-backed securities are highly specialized, this reading will focus only on options.

Recall the definition of an option:

> *An* **option** *is a derivative contract in which one party, the buyer, pays a sum of money to the other party, the seller or writer, and receives the right to either buy or sell an underlying asset at a fixed price either on a specific expiration date or at any time prior to the expiration date.*

---

[1] Remember that the term "spot market" refers to the market in which the underlying trades. A transaction in the spot market involves a buyer paying for an asset and receiving it right away or at least within a few days, given the normal time required to settle a financial transaction.

Options can be either customized over-the-counter contracts or standardized and traded on exchanges.

Because derivatives take their prices from the price of the underlying, it is important to first understand how the underlying is priced. We will approach the underlying from a slightly different angle, one that emphasizes the often-subtle costs of holding the underlying, which turn out to play a major role in derivative pricing.

2.2 Pricing the Underlying

The four main types of underlying on which derivatives are based are equities, fixed-income securities/interest rates, currencies, and commodities. Equities, fixed-income securities (but not interest rates), currencies, and commodities are all assets. An interest rate is not an asset, but it can be structured as the underlying of a derivative.[2]

Consider a generic underlying asset. This asset is something of value that you can own. Some assets are financial assets, such as equities, bonds, and currencies, and some are real assets, such as commodities (e.g., gold, oil, and agricultural products) and certain physical objects (e.g., houses, automobiles, and computers).

The price of a financial asset is often determined using a present value of future cash flows approach. The value of the financial asset is the expected future price plus any interim payments such as dividends or coupon interest discounted at a rate appropriate for the risk assumed. Such a definition presumes a period of time over which an investor anticipates holding an asset, known as the holding period. The investor forecasts the price expected to prevail at the end of the holding period as well as any cash flows that are expected to be earned over the holding period. He then takes that predicted future price and expected cash flows and finds their current value by discounting them to the present. Thereby, the investor arrives at a fundamental value for the asset and will compare that value with its current market price. Based on any differential relative to the cost of trading and his confidence in his valuation model, he will make a decision about whether to trade.

2.2.1 *The Formation of Expectations*

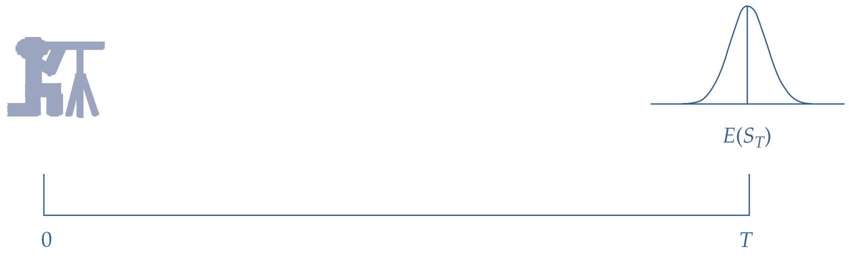
Let us first assume that the underlying does not pay interest or dividends, nor does it have any other cash flows attributable to holding the asset. Exhibit 1 illustrates the basic idea behind the valuation process. Using a probability distribution, the investor forecasts the future over a holding period spanning time 0 to time T. The center of the distribution is the expected price of the asset at time T, which we denote as $E(S_T)$, and represents the investor's prediction of the spot price at T. The investor knows there is risk, so this prediction is imperfect—hence the reason for the probability distribution. Nonetheless, at time 0 the investor makes her best prediction of the spot price at time T, which becomes the foundation for determining what she perceives to be the value of the asset.[3]

---

[2] This is a good example of why it is best not to use the term "underlying *asset*" when speaking of derivatives. Not all derivatives have underlying assets, but all have underlyings, some of which are not assets. Some other examples of non-asset underlyings used in derivatives are weather, insurance claims, and shipping rates. There are also some derivatives in which the underlying is another derivative.

[3] The distribution shown here is symmetrical and relatively similar to a normal distribution, but this characterization is for illustrative purposes only. We are making no assumptions about symmetry or normality at this point.

Exhibit 1 The Formation of Expectations for an Asset



2.2.2 *The Required Rate of Return on the Underlying Asset*

To determine the value of the asset, this prediction must be converted into its price or present value. The specific procedure is to discount this expected future price, but that is the easy part. Determining the rate at which to discount the expected future price is the hard part. We use the symbol k to denote this currently unknown discount rate, which is often referred to as the required rate of return and sometimes the expected rate of return or just the expected return. At a minimum, that rate will include the risk-free rate of interest, which we denote as r. This rate represents the opportunity cost, or so-called time value of money, and reflects the price of giving up your money today in return for receiving more money later.

2.2.3 *The Risk Aversion of the Investor*

At this point, we must briefly discuss an important characteristic of investors: their degree of risk aversion. We can generally characterize three potential types of investors by how they feel about risk: risk averse, risk neutral, or risk seeking.

Risk-neutral investors are willing to engage in risky investments for which they expect to earn only the risk-free rate. Thus, they do not expect to earn a premium for bearing risk. For risk-averse investors, however, risk is undesirable, so they do not consider the risk-free rate an adequate return to compensate them for the risk. Thus, risk-averse investors require a risk premium, which is an increase in the expected return that is sufficient to justify the acceptance of risk. All things being equal, an investment with a higher risk premium will have a lower price. It is very important to understand, however, that risk premiums are not automatically earned. They are merely expectations. Actual outcomes can differ. Clearly stocks that decline in value did not earn risk premiums, even though someone obviously bought them with the expectation that they would. Nonetheless, risk premiums must exist in the long run or risk-averse investors would not accept the risk.

The third type of investor is one we must mention but do not treat as realistic. Risk seekers are those who prefer risk over certainty and will pay more to invest when there is risk, implying a negative risk premium. We almost always assume that investors prefer certainty over uncertainty, so we generally treat a risk-seeking investor as just a theoretical possibility and not a practical reality.[4]

---

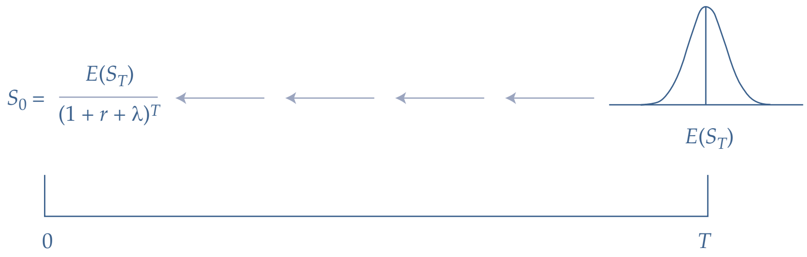
[4] People who gamble in casinos or play lotteries appear to be risk-seekers, given the advantage of the casino or the lottery organizer, but they are merely earning utility from the game itself, not necessarily from the expected financial outcome.

We will assume that investors are risk averse. To justify taking risk, risk-averse investors require a risk premium. We will use the Greek symbol λ (lambda) to denote the risk premium.[5]

2.2.4 *The Pricing of Risky Assets*

Exhibit 2 illustrates the process by which an investor obtains the current price, S_0, by discounting the expected future price of an asset with no interim cash flows, $E(S_T)$, by r (the risk-free rate) plus λ (the risk premium) over the period from 0 to T.

Exhibit 2 Discounting the Expected Future Price to Obtain the Current Price



2.2.5 *Other Benefits and Costs of Holding an Asset*

Many assets generate benefits and some incur costs to their owners. Some of these costs are monetary and others are nonmonetary. The dividends paid by companies and coupon interest paid by borrowers on their bonds represent obvious benefits to the holders of these securities. With currencies representing investments that earn the risk-free rate in a foreign country, they too generate benefits in the form of interest. Barring default, interest payments on bonds and currencies are relatively certain, so we will treat them as such. Dividend payments are not certain, but dividends do tend to be fairly predictable. As such, we will make an assumption common to most derivative models that dividends are certain.[6]

There is substantial evidence that some commodities generate a benefit that is somewhat opaque and difficult to measure. This benefit is called the **convenience yield**. It represents a nonmonetary advantage of holding the asset. For most financial assets, convenience yields are either nonexistent or extremely limited. Financial assets do not possess beauty that might make a person enjoy owning them just to look at them. Convenience yields are primarily associated with commodities and generally exist as a result of difficulty in either shorting the commodity or unusually tight supplies. For example, if a commodity cannot be sold short without great difficulty or cost, the holder of the commodity has an advantage if market conditions suggest that the commodity should be sold. Also, if a commodity is in short supply, the holders of the commodity can sometimes extract a price premium that is believed by some to be higher than what would be justified in well-functioning markets. The spot price of the commodity could even be above the market's expectation of its future price, a condition that would seem to imply a negative expected return. This scenario raises the question of why anyone would want to hold the commodity if its expected return is negative. The convenience yield provides a possible explanation that attributes an implied but non-financial expected return to the advantage of holding a commodity in short supply. The holder of the commodity has the ability to sell it when market conditions suggest that selling is advisable and short selling is difficult.

---

[5] Although the risk-free rate is invariant with a country's economy, the risk premium varies with the amount of risk taken. Thus, while the risk-free rate is the same when applied to every investment, the risk premium is not the same for every investment.

[6] Some derivative models incorporate uncertain dividends and interest, but those are beyond the scope of this introductory reading.

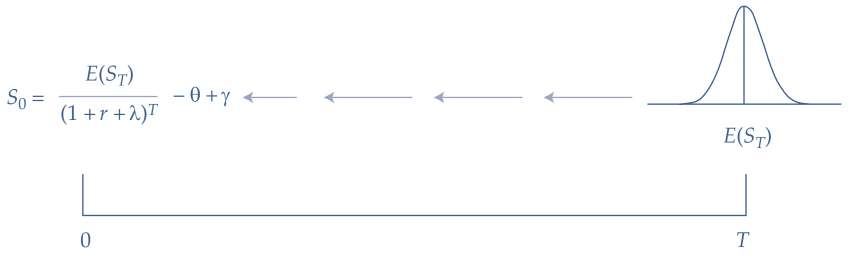
One cost incurred in owning commodities is the cost of storage. One could hardly own gold, oil, or wheat without incurring some costs in storing these assets. There are also costs incurred in protecting and insuring some commodities against theft or destruction. Depending on the commodity, these costs can be quite significant. For financial assets, however, the storage costs are so low that we can safely ignore them.

Finally, there is the opportunity cost of the money invested. If a person buys an asset, he forgoes interest on his money. The effect on this interest is reflected by compounding the price paid for the asset to a future value at the risk-free rate of interest. Thus, an investor who buys a stock that costs £50 in a market in which the risk-free rate is 4% will effectively have paid £50 × 1.04 = £52 a year later. Of course, the stock could be worth any value at that time, and any gain or loss should be determined in comparison to the effective price paid of £52.

As we described earlier, we determine the current price of an asset by discounting the expected future price by the sum of the risk-free rate (r) plus the risk premium (λ). When we introduce costs and benefits of holding the asset, we have to make an adjustment. With the exception of this opportunity cost of money, we will incorporate the effect of these costs and benefits by determining their value at the end of the holding period. Under the assumption that these costs and benefits are certain, we can then discount them at the risk-free rate to obtain their present value. There is a logic to doing it this way (i.e., finding their future value and discounting back to the present, as opposed to finding their present value directly). By finding their future value, we are effectively saying that the costs and benefits adjust the expected payoff at the end of the holding period. But because they are certain, we can discount their effects at the risk-free rate. So we have effectively just found their present value. The net effect is that the costs reduce the current price and the benefits increase the current price. We use the symbol θ (theta) to denote the present value of the costs and γ (gamma) as the present value of any benefits.

The net of the costs and benefits is often referred to by the term **carry**, or sometimes **cost of carry**. The holding, storing, or "carrying" of an asset is said to incur a net cost that is essentially what it takes to "carry" an asset. Exhibit 3 illustrates the effect in which the carry adjusts the price of an asset in the valuation process.

Exhibit 3 Pricing an Asset That Incurs Costs and Generates Benefits



$$S_0 = \frac{E(S_T)}{(1 + r + \lambda)^T} - \theta + \gamma$$

EXAMPLE 1

Pricing the Spot Asset

1 Which of the following factors does **not** affect the spot price of an asset that has no interim costs or benefits?
  - A The time value of money
  - B The risk aversion of investors
  - C The price recently paid by other investors

2 Which of the following does **not** represent a benefit of holding an asset?
  - A The convenience yield
  - B An optimistic expected outlook for the asset
  - C Dividends if the asset is a stock or interest if the asset is a bond

Solution to 1:

C is correct. The price recently paid by other investors is past information and does not affect the spot price. The time value of money and the risk aversion of investors determine the discount rate. Only current information is relevant as investors look ahead, not back.

Solution to 2:

B is correct. An optimistic forecast for the asset is not a benefit of holding the asset, but it does appear in the valuation of the asset as a high expected price at the horizon date. Convenience yields and dividends and interest are benefits of holding the asset.

To recap, although the various underlyings differ with respect to the specifics of pricing, all of them are based on expectations, risk, and the costs and benefits of holding a specific underlying. Understanding how assets are priced in the spot market is critical to understanding how derivatives are priced. To understand derivative pricing, it is necessary to establish a linkage between the derivative market and the spot market. That linkage occurs through arbitrage.

3

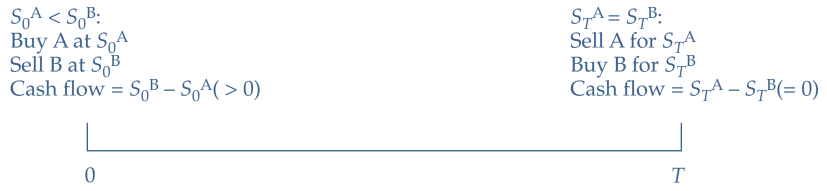
THE PRINCIPLE OF ARBITRAGE

a explain how the concepts of arbitrage, replication, and risk neutrality are used in pricing derivatives;

Arbitrage is a type of transaction undertaken when two assets or portfolios produce identical results but sell for different prices. If a trader buys the asset or portfolio at the cheaper price and sells it at the more expensive price, she will generate a net inflow of funds at the start. Because the two assets or portfolios produce identical results, a long position in one and a short position in the other means that at the end of the holding period, the payoffs offset. Hence, no money is gained or lost at the end of the holding period, so there is no risk. The net effect is that the arbitrageur receives money at the start and never has to pay out any money later. Such a situation amounts to free money, like walking down the street, finding money on the ground, and never having to give it up. Exhibit 4 illustrates this process for assets A and B, which have no dividends or other benefits or costs and pay off identically but sell for different prices, with $S_0^A < S_0^B$.

Exhibit 4 Executing an Arbitrage

Given: Assets A and B produce the same values at time T but at time 0, A is selling for less than B.



$S_0^A < S_0^B$:
Buy A at S_0^A
Sell B at S_0^B
Cash flow = $S_0^B - S_0^A (> 0)$

$S_T^A = S_T^B$:
Sell A for S_T^A
Buy B for S_T^B
Cash flow = $S_T^A - S_T^B (= 0)$

3.1 The (In)Frequency of Arbitrage Opportunities

When arbitrage opportunities exist, traders exploit them very quickly. The combined actions of many traders engaging in the same transaction of buying the low-priced asset or portfolio and selling the high-priced asset or portfolio results in increased demand and an increasing price for the former and decreased demand and a decreasing price for the latter. This market activity will continue until the prices converge. Assets that produce identical results can thus have only one true market price. This rule is called the "law of one price." With virtually all market participants alert for the possibility of earning such profits at no risk, it should not be surprising that arbitrage opportunities are rare.

In practice, prices need not converge precisely, or even all that quickly, because the transaction cost of exploiting an opportunity could exceed the benefit. For example, say you are walking down the sidewalk of the Champs-Élysées in Paris and notice a €1 coin on the sidewalk. You have a bad back, and it would take some effort to bend over. The transaction cost of exploiting this opportunity without any risk could exceed the benefit of the money. Some arbitrage opportunities represent such small discrepancies that they are not worth exploiting because of transaction costs.

Significant arbitrage opportunities, however, will be exploited. A significant opportunity arises from a price differential large enough to overcome the transaction costs. Any such price differential will continue to be exploited until the opportunity disappears. Thus, if you find a €10 note on the Champs-Élysées sidewalk, there is a good chance you will find it worth picking up (even with your bad back), and even if you do not pick it up, it will probably not be there for long. With enough people alert for such opportunities, only a few will arise, and the ones that do will be quickly exploited and disappear. In this manner, arbitrage makes markets work much more efficiently.

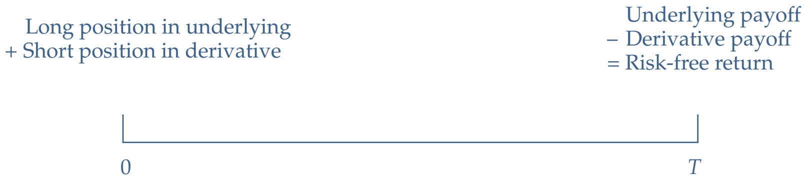
3.2 Arbitrage and Derivatives

It may be difficult to conceive of many investments that would produce identical payoffs. Even similar companies such as McDonalds and Burger King, which are in the same line of business, do not perform identically. Their performance may be correlated, but each has its own unique characteristics. For equity securities and with no derivatives involved, about the only such situation that could exist in reality is a stock that trades simultaneously in two different markets, such as Royal Dutch Shell, which trades in Amsterdam and London but is a single company. Clearly there can be only one price. If those two markets operate in different currencies, the currency-adjusted prices should be the same. Bonds issued by the same borrower are also potentially arbitrageable. All bonds of an issuer will be priced off of the term structure of interest rates. Because of this common factor, bonds of different maturities can be arbitraged against each other. But in general, two securities are unlikely to perform identically.

The picture changes, however, if we introduce derivatives. For most derivatives, the payoffs come (derive) directly from the value of the underlying at the expiration of the derivative. Although no one can predict with certainty the value of the underlying at expiration, as soon as that value is determined, the value of the derivative at expiration becomes certain. So, while the performance of McDonalds' stock may have a strong correlation to the performance of Burger King's stock, neither completely determines the other. But derivatives on McDonalds' stock and derivatives on Burger King's stock are completely determined by their respective stocks. All of the uncertainty in a derivative comes from the uncertainty in the underlying. As a result, the price of the derivative is tied to the price of the underlying. That being the case, the derivative can be used to hedge the underlying, or vice versa.

Exhibit 5 illustrates this point. When a long position in the underlying is combined with a short position in the derivative to produce a perfect hedge, all of the risk is eliminated and the position should earn the risk-free rate. If not, arbitrageurs begin to trade. If the position generates a return in excess of the risk-free rate, the arbitrageurs see an opportunity because the hedged position of the underlying (long asset and short derivative) earns more than the risk-free rate and a risk-free loan undertaken as a borrower incurs a cost equal to the risk-free rate. Therefore, implementing the hedged position and borrowing at the risk-free rate earns a return in excess of the risk-free rate, incurs a cost of the risk-free rate, and has no risk. As a result, an investor can earn excess return at no risk without committing any capital. Arbitrageurs will execute this transaction in large volumes, continuing to exploit the pricing discrepancy until market forces push prices back in line such that both risk-free transactions earn the risk-free rate.

Exhibit 5 Hedging the Underlying with a Derivative (or Vice Versa)



Out of this process, one and only one price can exist for the derivative. Otherwise, there will be an arbitrage opportunity. We typically take the underlying price as given and infer the unique derivative price that prohibits any arbitrage opportunities. Most derivatives pricing models are established on this foundation. We simply assume that no arbitrage opportunities can exist and infer the derivative price that guarantees there are no arbitrage opportunities.

3.3 Arbitrage and Replication

Because a long asset and a short derivative on the asset can be combined to produce a position equivalent to a risk-free bond, it follows that the long asset and a short risk-free asset (meaning to borrow at the risk-free rate) can be combined to produce a long derivative. Alternatively, a short derivative and the short risk-free asset can be combined to produce a short asset position. Exhibit 6 shows this process, referred to as **replication**. Replication is the creation of an asset or portfolio from another asset, portfolio, and/or derivative.

Exhibit 6 Arbitrage, Replication, and Derivatives

Long asset + Short derivative = Long risk-free asset (lending)

Long asset + Short risk-free asset (borrowing) = Long derivative

Short derivative + Short risk-free asset (borrowing) = Short asset

If all assets are correctly priced to prohibit arbitrage, however, the ability to replicate seems useless. Why would one replicate an asset or derivative if there is no cost advantage? Buying a government security to earn the risk-free rate is easier than buying the asset and selling a derivative to produce a risk-free position. At this point, that is certainly a reasonable question. As we progress through this material, however, we will relax the assumption that everything is always correctly priced and we will admit the possibility of occasional arbitrage opportunities. For example, it may be more profitable to hedge a portfolio with a derivative to produce a risk-free rate than to invest in the risk-free asset. In addition, we might find that replication can have lower transaction costs. For example, a derivative on a stock index combined with the risk-free asset [Long derivative (Stock index futures) + Long risk-free asset (Lending) = Long asset (Stock index)] can potentially replicate an index fund at lower transaction costs than buying all the securities in the index. Replication is the essence of arbitrage. The ability to replicate something with something else can be valuable to investors, either through pricing differentials, however temporary, or lower transaction costs.

3.4 Risk Aversion, Risk Neutrality, and Arbitrage-Free Pricing

Most investors are risk averse. They do not accept risk without the expectation of a return commensurate with that risk. Thus, they require risk premiums to justify the risk. One might think that this point implies a method for pricing derivatives based on the application of a risk premium to the expected payoff of the derivative and its risk. As we will describe later, this methodology is not appropriate in the pricing of derivatives.

As previously described, a derivative can be combined with an asset to produce a risk-free position. This fact does not mean that one *should* create such a combination. It merely means that one *can* do so. The derivative price is the price that guarantees the risk-free combination of the derivative and the underlying produces a risk-free return. The derivative price can then be inferred from the characteristics of the underlying, the characteristics of the derivative, and the risk-free rate. The investor's risk aversion is not a factor in determining the derivative price. Because the risk aversion of the investor is not relevant to pricing the derivative, one can just as easily obtain the derivative price by assuming that the investor is risk neutral. That means that the expected payoff of the derivative can be discounted at the risk-free rate rather than the risk-free rate plus a risk premium. Virtually all derivative pricing models ultimately take this form: discounting the expected payoff of the derivative at the risk-free rate.

The entire process of pricing derivatives is not exactly as we have described it at this point. There is an intermediate step, which entails altering the probabilities of the outcomes from the true probabilities to something called risk-neutral probabilities. We will illustrate this process later in this reading. The important point to understand is that while the risk aversion of investors is relevant to pricing assets, it is not relevant to pricing derivatives. As such, derivatives pricing is sometimes called **risk-neutral pricing**. Risk-neutral pricing uses the fact that arbitrage opportunities guarantee that a risk-free portfolio consisting of the underlying and the derivative must earn the risk-free rate. There is only one derivative price that meets that condition. Any mispricing of the derivative will lead to arbitrage transactions that drive the derivative price back to where it should be, the price that eliminates arbitrage opportunities.

The overall process of pricing derivatives by arbitrage and risk neutrality is called **arbitrage-free pricing**. We are effectively determining the price of a derivative by assuming that the market is free of arbitrage opportunities. This notion is also sometimes called the **principle of no arbitrage**. If there are no arbitrage opportunities, combinations of assets and/or derivatives that produce the same results must sell for the same price. The correct derivative price assures us that the market is free of arbitrage opportunities.

3.5 Limits to Arbitrage

As we previously described, there may be reasons to not pick up a coin lying on the ground. Likewise, some small arbitrage profits are never exploited. A bond selling for €1,000 might offer an arbitrage profit by trading a derivative on the bond and a risk-free asset at a total cost of €999, but the profit of €1 might be exceeded by the transaction costs. Such small differentials can easily remain essentially trapped within the bounds of transaction costs. In addition, arbitrage can require capital. Not everyone can borrow virtually unlimited amounts of money at what amounts to a risk-free rate. Moreover, some transactions can require additional capital to maintain positions. The corresponding gains from an offsetting position might not be liquid. Hence, on paper the position is hedged, but in practice, one position has a cash outflow while the other generates gains on paper that are realized only later. Borrowing against those future gains is not always easy.

Moreover, some apparent arbitrage transactions are not completely risk free. As you will learn later, option pricing requires knowledge of the volatility of the underlying asset, which is information that is not easy to obtain and subject to different opinions. Executing an arbitrage can entail risk if one lacks accurate information on the model inputs.

Some arbitrage positions require short-selling assets that can be difficult to short. Some securities are held only by investors who are unwilling to lend the securities and who, by policy, are not arbitrageurs themselves. Some commodities, in particular, can be difficult and costly to sell short. Hence, the arbitrage might exist in only one direction, which keeps the price from becoming seemingly too high or seemingly too low but permitting it to move virtually without limit in the opposite direction.

Arbitrage positions rely on the ultimate realization by other investors of the existence of the mispricing. For some investors, bearing these costs and risks until other investors drive the price back to its appropriate level can be nearly impossible.

The arbitrage principle is the essence of derivative pricing models. Yet, clearly there are limits to the ability of all investors to execute arbitrage transactions. In studying derivative pricing, it is important to accept the no-arbitrage rule as a paradigm, meaning a framework for analysis and understanding. Although no market experts think that arbitrage opportunities never occur, it is a common belief that finding and exploiting them is a challenging and highly competitive process that will not yield frequent success. But it is important that market participants stay alert for and exploit whatever arbitrage opportunities arise. In response, the market functions more efficiently.

EXAMPLE 2

Arbitrage

1. Which of the following *best* describes an arbitrage opportunity? It is an opportunity to:
   - **A** earn a risk premium in the short run.
   - **B** buy an asset at less than its fundamental value.
   - **C** make a profit at no risk with no capital invested.
2. What *most likely* happens when an arbitrage opportunity exists?
   - **A** Investors trade quickly and prices adjust to eliminate the opportunity.
   - **B** Risk premiums increase to compensate traders for the additional risk.
   - **C** Markets cease operations to eliminate the possibility of profit at no risk.
3. Which of the following *best* describes how derivatives are priced?
   - **A** A hedge portfolio is used that eliminates arbitrage opportunities.
   - **B** The payoff of the underlying is adjusted downward by the derivative value.
   - **C** The expected future payoff of the derivative is discounted at the risk-free rate plus a risk premium.
4. An investor who requires no premium to compensate for the assumption of risk is said to be which of the following?
   - **A** Risk seeking
   - **B** Risk averse
   - **C** Risk neutral
5. Which of the following is a limit to arbitrage?
   - **A** Clearinghouses restrict the transactions that can be arbitraged.
   - **B** Pricing models do not show whether to buy or sell the derivative.
   - **C** It may not always be possible to raise sufficient capital to engage in arbitrage.

Solution to 1:

C is correct because it is the only answer that is based on the notion of when an arbitrage opportunity exists: when two identical assets or portfolios sell for different prices. A risk premium earned in the short run can easily have occurred through luck. Buying an asset at less than fair value might not even produce a profit.

Solution to 2:

A is correct. The combined actions of traders push prices back in line to a level at which no arbitrage opportunities exist. Markets certainly do not shut down, and risk premiums do not adjust and, in fact, have no relevance to arbitrage profits.

Solution to 3:

A is correct. A hedge portfolio is formed that eliminates arbitrage opportunities and implies a unique price for the derivative. The other answers are incorrect because the underlying payoff is not adjusted by the derivative value and the discount rate of the derivative does not include a risk premium.

Solution to 4:

C is correct. Risk-seeking investors give away a risk premium because they enjoy taking risk. Risk-averse investors expect a risk premium to compensate for the risk. Risk-neutral investors neither give nor receive a risk premium because they have no feelings about risk.

Solution to 5:

C is correct. It may not always be possible to raise sufficient capital to engage in arbitrage. Clearinghouses do not restrict arbitrage. Pricing models show what the price of the derivative should be.

Thus, comparison with the market price will indicate if the derivative is overpriced and should be sold or if it is underpriced and should be purchased.

4

PRICING AND VALUATION OF FORWARD CONTRACTS: PRICING VS. VALUATION; EXPIRATION; INITIATION

b explain the difference between value and price of forward and futures contracts;

c calculate a forward price of an asset with zero, positive, or negative net cost of carry;

d explain how the value and price of a forward contract are determined at expiration, during the life of the contract, and at initiation;

e describe monetary and nonmonetary benefits and costs associated with holding the underlying asset and explain how they affect the value and price of a forward contract;

i explain the difference between value and price of swaps;

In equity markets, analysis is undertaken with the objective of determining the value, sometimes called the fundamental value, of a stock. When a stock trades in the market for a price that differs from its fundamental value, investors will often buy or sell the stock based on the perceived mispricing. The fundamental value of a stock is typically determined by analyzing the company's financial statements, projecting its earnings and dividends, determining a discount rate based on the risk, and finding the present value of the future dividends. These steps make up the essence of dividend discount models. Other approaches include comparing the book value of a company to its market value, thereby using book value as a proxy for fundamental value, or by application of a price/earnings ratio to projected next-period earnings, or by discounting free cash flow. Each of these approaches purports to estimate the company's fundamental value, leading to the notion that a company is worth something that may or may not correspond to its price in the market.

In derivative markets, the notion of valuation as a representation of fundamental value is still a valid concept, but the terminology can be somewhat different and can lead to some confusion. Options are not a problem in this regard. They can be analyzed to determine their fundamental value, and the market price can be compared with the fundamental value. Any difference can then presumably be exploited via arbitrage. The combined actions of numerous investors should ultimately lead to the market price converging to its fundamental value, subject to the above limits to arbitrage.

The world of forwards, futures, and swaps, however, uses different terminology with respect to price and value. These contracts do not require the outlay of cash at the start the way an option, stock, or bond does. Forwards, futures, and swaps start off with values of zero. Then as the underlying moves, their values become either positive or negative. The forward, futures, or swap price is a concept that represents the fixed price or rate at which the underlying will be purchased at a later date. It is not an amount to be paid at the start. This fixed price or rate is embedded into the contract while the value will fluctuate as market conditions change. But more importantly, the value and price are not at all comparable with each other.

Consider a simple example. Suppose you own a stock priced at \$102. You have a short forward contract to sell the stock at a price of \$100 one year from now. The risk-free rate is 4%. Your position is riskless because you know that one year from now, you will sell the stock for \$100. Thus, you know you will get \$100 one year from now, which has a present value of \$100/(1.04) = \$96.15. Notice the discounting at the risk-free rate, which is appropriate because the position is riskless. Your overall position is that you own an asset worth \$102 and are short a contract worth something, and the two positions combine to have a value of \$96.15. Therefore, the forward contract must have a value of \$96.15 – \$102 = –\$5.85. Your forward contract is thus worth –\$5.85. To the party on the opposite side, it is worth +\$5.85.[7] The price of the forward contract is still \$100, which was set when you created the contract at an earlier date. As you can see, the \$100 forward price is not comparable to the \$5.85 value of the contract.

Although the forward price is fixed, any new forward contract calling for delivery of the same asset at the same time will have a different price. We will cover that point in more detail later. For now, it is important to see that your contract has a price of \$100 but a value of –\$5.85, which are two entirely different orders of magnitude. This information does not imply that the forward contract is mispriced. The value is the amount of wealth represented by owning the forward contract. The price is one of the terms the parties agreed on when they created the contract.[8] This idea applies in the same manner for futures and swaps.

4.1 Pricing and Valuation of Forward Commitments

In this section, we will go into pricing forward commitments in a little more detail. Let us start by establishing that today, at time 0, we create a forward commitment that expires at time T. The value of the underlying today is S_0. At expiration the underlying value is S_T, which is not known at the initiation of the contract.

4.1.1 *Pricing and Valuation of Forward Contracts*

Previously, we noted that price and value are entirely different concepts for forward commitments. We gave an example of a forward contract with a price of \$100 but a value of –\$5.85 to the seller and +\$5.85 to the buyer. In the next subsection, we will delve more deeply into understanding these concepts of pricing and valuation for forward contracts.

---

[7] This concept of the value of the forward contract as it evolves toward expiration is sometimes referred to as its mark-to-market value. The same notion is applicable to swaps. In futures, of course, contracts are automatically marked to market by the clearinghouse, and gains and losses are converted into actual cash flows from one party to the other.

[8] The forward price is more like the exercise price of the option. It is the price the two parties agree will be paid at a future date for the underlying. Of course, the option has the feature that the holder need not ever pay that price, which is the case if the holder chooses not to exercise the option.

4.1.1.1 Pricing and Valuation of Forward Contracts at Expiration Recall that a forward contract specifies that one party agrees to buy the underlying from the other at the expiration date at a price agreed on at the start of the contract. Suppose that you enter into a contract with another party in which you will buy a used car from that party in one year at a price of \$10,000. Then \$10,000 is the forward price. One year later, when the contract expires, you are committed to paying \$10,000 and accepting delivery of the car. Let us say that at that time, you check the used car market and find that an identical car is worth \$10,800. How much is your forward contract worth to you at that time? It obligates you to pay \$10,000 for a car that you would otherwise have to pay \$10,800. Thus, the contract benefits you by \$800, so its value is \$800. If you were on the opposite side of the transaction, its value would be –\$800. If the market price of the car were below \$10,000, the contract would have negative value to you and the mirror image positive value to the seller.

This example leads us to our first important derivative pricing result. The forward price, established at the initiation date of contract is $F_0(T)$. Let us denote the value at expiration of the forward contract as $V_T(T)$. This value is formally stated as

$$V_T(T) = S_T - F_0(T) \tag{1}$$

In words,

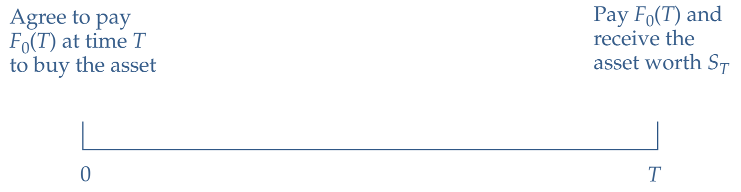
> *The value of a forward contract at expiration is the spot price of the underlying minus the forward price agreed to in the contract.*

In the financial world, we generally define value as the value to the long position, so the above definition is generally correct but would be adjusted if we look at the transaction from the point of view of the short party. In that case, we would multiply the value to the long party by −1 to calculate the value to the short party. Alternatively, the value to the short party is the forward price minus the spot price at expiration.

If a forward contract could be initiated right at the instant of expiration, the forward price would clearly be the spot price. Such a contract would essentially be a spot transaction.

4.1.1.2 Pricing and Valuation at Initiation Date In Exhibit 7, we see the nature of the problem of pricing a forward contract. We are situated at time 0, facing an uncertain future. At the horizon date, time T, the underlying price will be S_T. Of course, at time 0 we do not know what S_T will turn out to be. Yet at time 0, we need to establish the forward price, $F_0(T)$, which is the price we agree to pay at time T to purchase the asset.

Exhibit 7 The Time Horizon of Forward Contracts


Agree to pay $F_0(T)$ at time T to buy the asset
Pay $F_0(T)$ and receive the asset worth S_T
0
T

When a forward contract is initiated, neither party pays anything to the other. It is a valueless contract, neither an asset nor a liability. Therefore, its value at initiation is zero:

$$V_0(T) = 0 \tag{2}$$

The forward price that the parties agree to at the initiation date of the contract is a special price that results in the contract having zero value and prohibiting arbitrage. This is our first important result:

> *Because neither the long nor the short pays anything to the other at the initiation date of a forward contract, the value of a forward contract when initiated is zero.*

If this statement were not true and one party paid a sum of money to the other, the party receiving the money could find another party and engage in the opposite transaction, with no money paid to the other on this second contract. The two transactions would completely offset, thereby eliminating the risk. Yet, the first party would have captured some cash from the second and consequently earned an arbitrage profit because his position is completely hedged. He would walk away with money and never have to worry about paying it back. The forward price is the price the two parties agree on that generates a value of zero at the initiation date. Finding that price is actually quite easy.

Consider a very simple asset price at S_0 today that pays no dividends or interest, nor does it yield any nonfinancial benefits or incur any carrying costs. As described earlier, we can peer into the future, but at best we can make only a forecast of the price of this asset at our horizon date of time T. That forecast was previously referred to as the expected spot price at expiration, $E(S_T)$. On the surface, it might seem that pricing a forward contract would somehow involve a discounting of the expected spot price. As we said earlier, however, that is not how derivatives are priced—they are priced using arbitrage.

Suppose we hold the asset and enter into a forward contract to sell the asset at the price $F_0(T)$. It should be easy to see that we have constructed a risk-free position. We know that the asset, currently worth S_0, will be sold later at $F_0(T)$ and that this price should guarantee a risk-free return. Thus, we should find the following relationship,

$$\frac{F_0(T)}{S_0} = (1 + r)^T \tag{3}$$

We can easily solve for the forward price to obtain

$$F_0(T) = S_0(1 + r)^T \tag{4}$$

Or, in words,

> *The forward price is the spot price compounded at the risk-free rate over the life of the contract.*

There is a nice logic to this relationship. While the spot price is what someone would have to pay today to buy the asset, a forward contract locks in the purchase price at the horizon date. When that date arrives, the investor will own the asset. Instead of buying the asset today, suppose the investor uses the forward contract to guarantee that she will own the asset at the horizon date. By using the forward contract, the investor will not have committed the money, S_0, that would have forgone interest at the rate r for the period 0 to T. Notice how the risk premium on the asset does not directly appear in the pricing relationship. It does appear implicitly, because it determines the spot price paid to buy the asset. Knowing the spot price, however, eliminates the necessity of determining the risk premium. The derivatives market can simply let the spot market derive the risk premium.

As a simple example, let us say the underlying price, S_0, is £50, the risk-free rate, r, is 3%, and the contract expires in three months, meaning that $T = 3/12 = 0.25$. Then the forward price is £50$(1.03)^{0.25}$ = £50.37. Thus, the two parties would agree that the buyer will pay £50.37 to the seller in three months, and the seller will deliver the underlying to the buyer at expiration.

Now suppose the asset generates cash payments and/or benefits and incurs storage costs. As we discussed, the net cost of carry consists of the benefits, denoted as γ (dividends or interest plus convenience yield), minus the costs, denoted as θ, both of which are in present value form. To put these concepts in future value form, we simply compound them at the risk-free rate, $(\gamma - \theta)(1 + r)^T$. Because this is their value at the expiration date of the contract, we can add them to $F_0(T)$ in Equation 3, thereby restating that equation as

$$(1 + r)^T = \frac{F_0(T) + (\gamma - \theta)(1 + r)^T}{S_0}$$

The numerator is how much money we end up with at T. Rearranging, we obtain the forward price as

$$F_0(T) = (S_0 - \gamma + \theta)(1 + r)^T$$

or **(5)**

$$F_0(T) = S_0(1 + r)^T - (\gamma - \theta)(1 + r)^T$$

From Equation 5, we can see that the forward price determined using Equation 4 is reduced by the future value of any benefits and increased by the future value of any costs. In other words,

> *The forward price of an asset with benefits and/or costs is the spot price compounded at the risk-free rate over the life of the contract minus the future value of those benefits and costs.*

Again, the logic is straightforward. To acquire a position in the asset at time T, an investor could buy the asset today and hold it until time T. Alternatively, he could enter into a forward contract, committing him to buying the asset at T at the price $F_0(T)$. He would end up at T holding the asset, but the spot transaction would yield benefits and incur costs, whereas the forward transaction would forgo the benefits but avoid the costs.

Assume the benefits exceed the costs. Then the forward transaction would return less than the spot transaction. The formula adjusts the forward price downward by the expression $-(\gamma - \theta)(1 + r)^T$ to reflect this net loss over the spot transaction. In other words, acquiring the asset in the forward market would be cheaper because it forgoes benefits that exceed the costs. That does not mean the forward strategy is better. It costs less but also produces less. Alternatively, if the costs exceeded the benefits, the forward price would be higher because the forward contract avoids the costs at the expense of the lesser benefits.

Returning to our simple example, suppose the present value of the benefits is γ = £3 and the present value of the costs is θ = £4. The forward price would be £50(1.03)$^{0.25}$ – (£3 – £4)(1.03)$^{0.25}$ = £51.38. The forward price, which was £50.37 without these costs and benefits, is now higher because the carrying costs exceed the benefits.

The value of the contract when initiated is zero provided the forward price conforms to the appropriate pricing formula. To keep the analysis as simple as possible, consider the case in which the asset yields no benefits and incurs no costs. Going long the forward contract or going long the asset produces the same position at T: ownership of the asset. Nonetheless, the strategies are not equivalent. Going long the forward contract enables the investor to avoid having to pay the price of the asset, S_0, so she would collect interest on the money. Thus, the forward strategy would have a value of S_0, reflecting the investment of that much cash invested in risk-free bonds, plus the value of the forward contract. The spot strategy would have a value of S_0, reflecting the investment in the asset. These two strategies must have equal values. Hence, the value of the forward contract must be zero.

Although a forward contract has zero value at the start, it will not have zero value during its life. We now take a look at what happens during the life of the contract.

PRICING AND VALUATION OF FORWARD CONTRACTS: BETWEEN INITIATION AND EXPIRATION; FORWARD RATE AGREEMENTS

5

b explain the difference between value and price of forward and futures contracts;

c calculate a forward price of an asset with zero, positive, or negative net cost of carry;

d explain how the value and price of a forward contract are determined at expiration, during the life of the contract, and at initiation;

e describe monetary and nonmonetary benefits and costs associated with holding the underlying asset and explain how they affect the value and price of a forward contract;

f define a forward rate agreement and describe its uses;

We previously worked an example in which a forward contract established with a price of \$100 later has a value of –\$5.85 to the seller and +\$5.85 to the buyer. Generally we would say the value is \$5.85. We explained that with the spot price at \$102, a party that is long the asset and short the forward contract would guarantee the sale of the asset priced at \$102 at a price of \$100 in one year. The present value of \$100 in one year at 4% is \$96.15. Thus, the party guarantees that his \$102 asset will be effectively sold at a present value of \$96.15, for a present value loss of \$5.85.

In general, we can say that

> *The value of a forward contract is the spot price of the underlying asset minus the present value of the forward price.*

Again, the logic is simple. A forward contract provides a type of synthetic position in the asset, for which we promise to pay the forward price at expiration. Thus, the value of the forward contract is the spot price of the asset minus the present value of the forward price. Let us write out this relationship using $V_t(T)$ as the value of the forward contract at time *t*, which is some point in time after the contract is initiated and before it expires:

$$V_t(T) = S_t - F_0(T)(1 + r)^{-(T-t)} \tag{6}$$

Note that we are working with the spot price at *t*, but the forward price was fixed when the contract was initiated.[9]

Now, recall the problem we worked in which the underlying had a price of £50 and the contract was initiated with a three-month life at a price of £50.37. Move one month later, so that the remaining time is two months: $T - t = 2/12 = 0.167$. Let the underlying price be £52. The value of the contract would be £52 – £50.37$(1.03)^{-0.167}$ = £1.88.

If the asset has a cost of carry, we must make only a small adjustment:

$$V_t(T) = S_t - (\gamma - \theta)(1 + r)^t - F_0(T)(1 + r)^{-(T-t)} \tag{7}$$

Note how we adjust the formula by the net of benefits minus costs. The forward contract forgoes the benefits and avoids the costs of holding the asset. Consequently, we adjust the value downward to reflect the forgone benefits and upward to reflect the avoided costs. Remember that the costs (θ) and benefits (γ) are expressed on a present value basis as of time 0. We need their value at time t. We could compound them from 0 to T and then discount them back to t by the period $T - t$, but a shorter route is to simply compound them from 0 to t. In the problem we previously worked, in which we priced the forward contract when the asset has costs and benefits, the benefits (γ) were £3 and the costs (θ) were £4, giving us a forward price of £51.38. We have now moved one month ahead, so $t = 1/12 = 0.0833$ and $T - t = 2/12 = 0.167$. Hence the value of the forward contract would be £52 – (£3 – £4)$(1.03)^{0.0833}$ – £51.38$(1.03)^{-0.167}$ = £1.88. Notice how the answer is the same as in the case of no costs and benefits, as this effect is also embedded in the original forward price and completely offsets.

---

[9] An alternative approach to valuing a forward contract during its life is to determine the price of a new forward contract that would offset the old one. The discounted difference between the new forward price and the original forward price will lead to the same value.

It is important to note that although we say that Equation 7 holds during the life of the contract at some arbitrary time t, it also holds at the initiation date and at expiration. For the initiation date, we simply change t to 0 in Equation 7. Then we substitute Equation 5 for $F_0(T)$ in Equation 7, obtaining $V_0(T) = 0$, confirming that the value of a forward contract at initiation is zero. At expiration, we let $t = T$ in Equation 7 and obtain the spot price minus the forward price, as presented in Equation 1.[10]

5.1 A Word about Forward Contracts on Interest Rates

Forward contracts in which the underlying is an interest rate are called **forward rate agreements**, or FRAs. These instruments differ slightly from most other forward contracts in that the underlying is not an asset. Changes in interest rates, such as the value of an asset, are unpredictable. Moreover, virtually every company and organization is affected by the uncertainty of interest rates. Hence, FRAs are very useful devices for many companies. FRAs are forward contracts that allow participants to make a known interest payment at a later date and receive in return an unknown interest payment. In that way, a participant whose business will involve borrowing at a future date can hedge against an increase in interest rates by buying an FRA (the long side) and locking in a fixed payment and receiving a random payment that offsets the unknown interest payment it will make on its loan. Note that the FRA seller (the short side) is hedging against a decrease in interest rates. Also, consider that the FRA seller could be a lender wishing to lock in a fixed rate on a loan it will make at a future date.

Even though FRAs do not involve an underlying asset, they can still be combined with an underlying asset to produce a hedged position, thereby leading to fairly straightforward pricing and valuation equations. The math is a little more complex than the math for forwards on assets, but the basic ideas are the same.

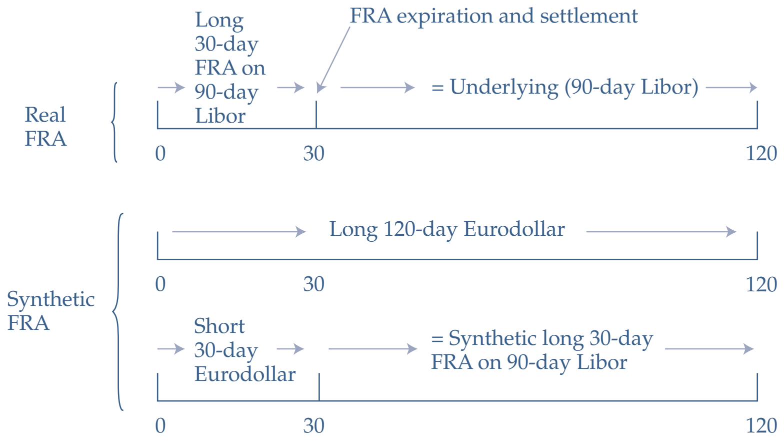
FRAs have often historically been based on Libor, the London Interbank Offered Rate, which represents the rate on a Eurodollar time deposit, a loan in dollars from one London bank to another. Other rates such as Euribor (Euro Interbank Offered Rate) and Tibor (Tokyo Interbank Offered Rate) have also been used.[11] As an example, assume we are interested in going long a 30-day FRA with a fixed rate (the FRA rate) in which the underlying is 90-day Libor. A long position means that in 30 days, we will make a known interest payment and receive an interest payment corresponding to the discounted difference between 90-day Libor on that day and the FRA rate. We can either enter into a 30-day FRA on 90-day Libor or create a synthetic FRA. To do the latter, we would go long a 120-day Eurodollar time deposit and short a 30-day Eurodollar time deposit. Exhibit 8 shows the structure of this strategy. We omit some of the details here, such as how much face value we should take on the two Eurodollar transactions as well as the size of the FRA. Those technical issues are covered in more advanced material. At this time, we focus on the fact that going long over the 120-day period and short over the 30-day period leaves an investor with no exposure over the 30-day period and then converts to a position that starts 30 days from now and matures 90 days later. This synthetic position corresponds to a 30-day FRA on 90-day Libor. Exhibit 8 illustrates this point.[12]

---

[10] You might be wondering whether the cost and benefit terms disappear when $t = T$. With the costs and benefits defined as those incurred over the period t to T, at expiration their value is zero by definition.

[11] Libor is being phased out, as the panel of banks will no longer be required to submit quotations after 2021. In anticipation of this, market participants and regulators have been working to develop alternative reference rates.

Exhibit 8 Real FRA and Synthetic FRA (30-Day FRA on 90-Day Libor)


Long 30-day FRA on 90-day Libor
FRA expiration and settlement
= Underlying (90-day Libor)
Real FRA
0
30
120
Long 120-day Eurodollar
0
30
120
Synthetic FRA
Short 30-day Eurodollar
= Synthetic long 30-day FRA on 90-day Libor
0
30
120

FRAs, and indeed all forward contracts relating to bonds and interest rates, are closely tied to the term structure of interest rates, a concept covered in virtually all treatments of fixed-income securities. Buying a 120-day zero-coupon bond and selling a 30-day zero-coupon bond produces a forward position in a 90-day zero-coupon bond that begins in 30 days. From that forward position, one can infer the forward rate. It would then be seen that the FRA rate *is* the forward rate, even though the derivative itself is not a forward contract on a bond.

EXAMPLE 3

Forward Contract Pricing and Valuation

1 Which of the following *best* describes the difference between the price of a forward contract and its value?

A The forward price is fixed at the start, and the value starts at zero and then changes.

B The price determines the profit to the buyer, and the value determines the profit to the seller.

C The forward contract value is a benchmark against which the price is compared for the purposes of determining whether a trade is advisable.

---

[12] The real FRA we show appears to imply that an investor enters into a Eurodollar transaction in 30 days that matures 90 days later. This is not technically true. The investor does, however, engage in a cash settlement in 30 days that has the same value and economic form as such a transaction. Specifically, settlement at expiration of the FRA is an amount equal to the discounted difference between the underlying 90-day Libor rate on that day and the FRA rate multiplied by a notional principal amount. These details are covered in the Level II and Level III CFA Program curriculum.

2 Which of the following *best* describes the value of the forward contract at expiration? The value is the price of the underlying:

- **A** minus the forward price.
- **B** divided by the forward price.
- **C** minus the compounded forward price.

3 Which of the following factors does *not* affect the forward price?

- **A** The costs of holding the underlying
- **B** Dividends or interest paid by the underlying
- **C** Whether the investor is risk averse, risk seeking, or risk neutral

4 Which of the following *best* describes the forward rate of an FRA?

- **A** The spot rate implied by the term structure
- **B** The forward rate implied by the term structure
- **C** The rate on a zero-coupon bond of maturity equal to that of the forward contract

Solution to 1:

A is correct. The forward price is fixed at the start, whereas the value starts at zero and then changes. Both price and value are relevant in determining the profit for both parties. The forward contract value is not a benchmark for comparison with the price.

Solution to 2:

A is correct because the holder of the contract gains the difference between the price of the underlying and the forward price. That value can, of course, be negative, which will occur if the holder is forced to buy the underlying at a price higher than the market price.

Solution to 3:

C is correct. The costs of holding the underlying, known as carrying costs, and the dividends and interest paid by the underlying are extremely relevant to the forward price. How the investor feels about risk is irrelevant, because the forward price is determined by arbitrage.

Solution to 4:

B is correct. FRAs are based on Libor, and they represent forward rates, not spot rates. Spot rates are needed to determine forward rates, but they are not equal to forward rates. The rate on a zero-coupon bond of maturity equal to that of the forward contract describes a spot rate.

As noted, we are not covering the details of derivative pricing but rather are focusing on the intuition. At this point, we have covered the intuition of pricing forward contracts. We now move to futures contracts.

6 PRICING AND VALUATION OF FUTURES CONTRACTS

b explain the difference between value and price of forward and futures contracts;

g explain why forward and futures prices differ;

Futures contracts differ from forward contracts in that they have standard terms, are traded on a futures exchange, and are more heavily regulated, whereas forward contracts are typically private, customized transactions. Perhaps the most important distinction is that they are marked to market on a daily basis, meaning that the accumulated gains and losses from the previous day's trading session are deducted from the accounts of those holding losing positions and transferred to the accounts of those holding winning positions. This daily settling of gains and losses enables the futures exchange to guarantee that a party that earns a profit from a futures transaction will not have to worry about collecting the money. Thus, futures exchanges provide a credit guarantee, which is facilitated through the use of a clearinghouse. The clearinghouse collects and disburses cash flows from the parties on a daily basis, thereby settling obligations quickly before they accumulate to much larger amounts. There is no absolute assurance that a clearinghouse will not fail, but none has ever done so since the first one was created in the 1920s.

The pattern of cash flows in a futures contract is quite similar to that in a forward contract. Suppose you enter into a forward contract two days before expiration in which you agree to buy an asset at €100, the forward price. Two days later, the asset is selling for €103, and the contract expires. You therefore pay €100 and receive an asset worth €103, for a gain of €3. If the contract were cash settled, instead of involving physical delivery, you would receive €3 in cash, which you could use to defer a portion of the cost of the asset. The net effect is that you are buying the asset for €103, paying €100 plus the €3 profit on the forward contract.

Had you chosen a futures contract, the futures price at expiration would still converge to the spot price of €103. But now it would matter what the futures settlement price was on the next to last day. Let us assume that price was €99. That means on the next to last day, your account would be marked to market for a loss of €1, the price of €100 having fallen to €99. That is, you would be charged €1, with the money passed on to the opposite party. But then on the last day, your position would be marked from €99 to €103, a gain of €4. Your net would be €1 lost on the first day and €4 gained on the second for a total of €3. In both situations you gain €3, but with the forward contract, you gain it all at expiration, whereas with the futures contract, you gain it over two days. With this two-day example, the interest on the interim cash flow would be virtually irrelevant, but over longer periods and with sufficiently high interest rates, the difference in the amount of money you end up with could be noticeable.

The value of a futures contract is the accumulated gain or loss on a futures contract since its previous day's settlement. When that value is paid out in the daily settlement, the futures price is effectively reset to the settlement price and the value goes to zero. The different patterns of cash flows for forwards and futures can lead to differences in the pricing of forwards versus futures. But there are some conditions under which the pricing is the same. It turns out that if interest rates were constant, forwards and futures would have the same prices. The differential will vary with the volatility of interest rates. In addition, if futures prices and interest rates are uncorrelated, forwards and futures prices will be the same. If futures prices are positively correlated with interest rates, futures contracts are more desirable to holders of long positions than are forwards. The reason is because rising prices lead to futures profits that are reinvested in periods of rising interest rates, and falling prices leads to losses that occur in periods of falling interest rates. It is far better to receive cash flows in the interim than all at expiration under such conditions. This condition makes futures more attractive than forwards, and therefore their prices will be higher than forward prices. A negative correlation between futures prices and interest rates leads to the opposite interpretation, with forwards being more desirable than futures to the long position. The more desirable contract will tend to have the higher price.

The practical realities, however, are that the derivatives industry makes virtually no distinction between futures and forward prices.[13] Thus, we will make no distinction between futures and forward pricing, except possibly in noting some subtle issues that may arise from time to time.

EXAMPLE 4

Futures Pricing and Valuation

1 Which of the following *best* describes how futures contract payoffs differ from forward contract payoffs?
   A Forward contract payoffs are larger.
   B They are equal, ignoring the time value of money.
   C Futures contract payoffs are larger if the underlying is a commodity.
2 Which of the following conditions will not make futures and forward prices equivalent?
   A Interest rates are constant.
   B Futures prices are uncorrelated with interest rates.
   C The volatility of the forward price is different from the volatility of the futures price.
3 With respect to the value of a futures contract, which of the following statements is *most* accurate? The value is the:
   A futures price minus the spot price.
   B present value of the expected payoff at expiration.
   C accumulated gain since the previous settlement, which resets to zero upon settlement.

Solution to 1:

B is correct. Forward payoffs occur all at expiration, whereas futures payoffs occur on a day-to-day basis but would equal forward payoffs ignoring interest. Payoffs could differ, so forward payoffs are not always larger. The type of underlying is not relevant to the point of which payoff is larger.

Solution to 2:

C is correct. Constant interest rates or the condition that futures prices are uncorrelated with interest rates will make forward and futures prices equivalent. The volatility of forward and futures prices has no relationship to any difference.

Solution to 3:

C is correct. Value accumulates from the previous settlement and goes to zero when distributed.

---

[13] At the time of this writing, many forwards (and swaps) are being processed through clearinghouses, a response to changes brought about by key legislation in several countries that was adopted following the financial crises of 2008. These OTC instruments are thus being effectively marked to market in a similar manner to the futures contracts described here. The full extent of this evolution of OTC trading through clearinghouses is not yet clear.

PRICING AND VALUATION OF SWAP CONTRACTS

7

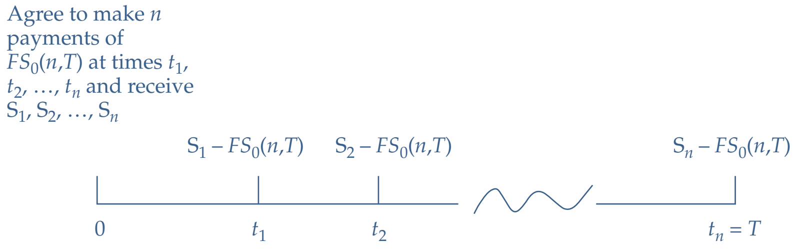
h explain how swap contracts are similar to but different from a series of forward contracts;

i explain the difference between value and price of swaps;

Recall the structure of a forward contract, as depicted in Exhibit 7. The investor is at time 0 and needs to determine the price, $F_0(T)$, that she will agree to pay at time T to purchase the asset. This price is set such that there is no value to the contract at that time. Value can arise later as prices change, but when initiated, the contract has zero value. Neither party pays anything to the other at the start.

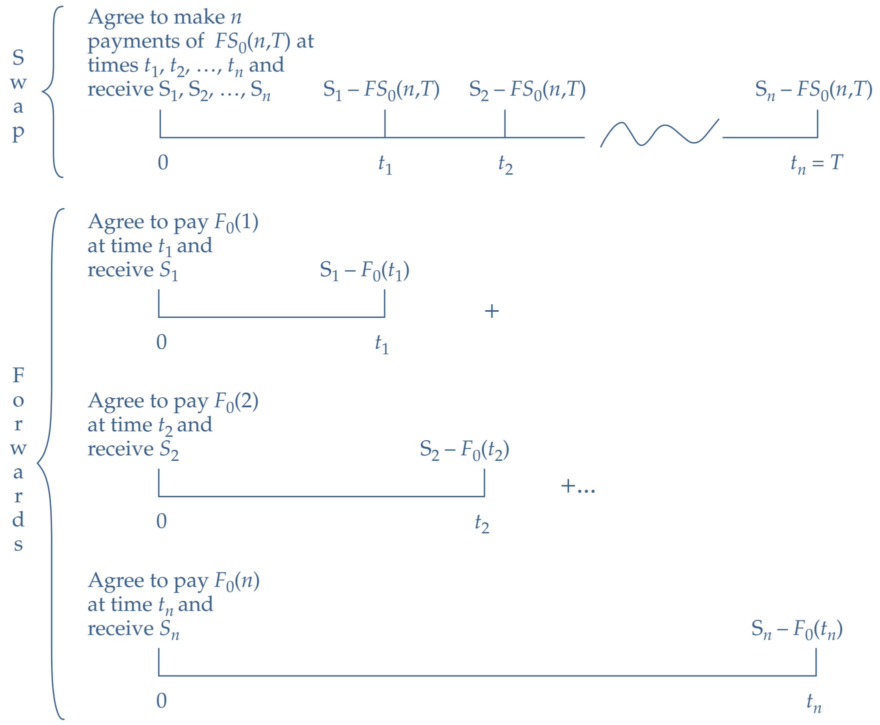
Now consider a swap starting at time 0 and ending at time T. We will let this swap be the type that involves a fixed payment exchanged for a floating payment. The contract specifies that the two parties will make a series of n payments at times that we will designate as 1, 2, …, n, with the last payment occurring at time T. On each of these payment dates, the owner of the swap makes a payment of $FS_0(n, T)$ and receives a payment based on the value of the underlying at the time of each respective payment, $S_1, S_2, \ldots, S_n$. So from the point of view of the buyer, the sequence of cash flows from the swap is $S_1 - FS_0(n, T), S_2 - FS_0(n, T), \ldots, S_n - FS_0(n, T)$. The notation $FS_0(n, T)$ denotes the fixed payment established at time 0 for a swap consisting of n payments with the last payment at time T. We denote the time to each payment as $t_1, t_2, \ldots, t_n$, where $t_n = T$. This structure is shown in Exhibit 9.

Exhibit 9 Structure of Cash Flows in a Swap


Agree to make n payments of $FS_0(n,T)$ at times t_1, $t_2, \ldots, t_n$ and receive $S_1, S_2, \ldots, S_n$
$S_1 - FS_0(n,T)$
$S_2 - FS_0(n,T)$
$S_n - FS_0(n,T)$
0
t_1
t_2
$t_n = T$

Comparing Exhibit 7 with Exhibit 9 reveals some similarities. A swap is in some sense a series of forward contracts, specifically a set of contracts expiring at various times in which one party agrees to make a fixed payment and receive a variable payment. Now consider Exhibit 10, which breaks down a swap into a series of implicit forward contracts, with the expiration of each forward contract corresponding to a swap payment date.

Exhibit 10 A Swap as a Series of Forward Contracts


Swap
Agree to make n payments of FS0(n,T) at times t1, t2, …, tn and receive S1, S2, …, Sn
S1 – FS0(n,T)
S2 – FS0(n,T)
Sn – FS0(n,T)
0
t1
t2
tn = T
Forwards
Agree to pay F0(1) at time t1 and receive S1
S1 – F0(t1)
0
t1
+
Agree to pay F0(2) at time t2 and receive S2
S2 – F0(t2)
0
t2
+…
Agree to pay F0(n) at time tn and receive Sn
Sn – F0(tn)
0
tn

Recall from the material on forward contracts that the forward price is determined by the spot price and the net cost of carry (Equation 5), the latter being partially determined by the length of time of the contract. It should be obvious that a forward contract expiring at time t_1 will not have the same price, $F_0(t_1)$, as a forward contract expiring at time t_2, $F_0(t_2)$, and likewise for all of the implicit remaining forward contracts expiring up through time t_n. The cost of carrying an asset over different time periods will vary by the length of the time periods. In other words, the prices of the implicit forward contracts embedded in a swap will not be equal:

$$F_0(\mathrm{t}_1) \neq F_0(t_2) \neq \ldots \neq F_0(t_n)$$

But for a swap, all the fixed payments are equal. So, how can we equate a swap to a series of forward contracts? It turns out that we can, and in doing so, we recall a valuable point about forward pricing.

Recall that the forward price is the price that produces a zero value of the contract at the start. Zero value is essential if there is no exchange of cash flows from one party to the other. And although no exchange of cash flows is customary, it is not mandatory. The parties could agree on any forward price at the start. If the zero-value forward price were \$30 and the parties agreed on a price of \$28, it should be apparent that the buyer would be getting a great price. The seller, being rational, would require that the buyer compensate him at the start. The seller should be getting \$30 at expiration and instead will get \$28. So the buyer should compensate the seller to the amount of the present value of \$2 at expiration. If the parties agree on a price greater than \$30, similar compensation would have to be paid from seller to buyer.

A forward transaction that starts with a zero value is called an at-market forward. A forward transaction that starts with a nonzero value is called an off-market forward. There is generally no prohibition on the use of off-market forward contracts, so two parties can engage in a series of forward contracts at whatever fixed price they so desire. Assume they agree on the price $FS_0(T)$. That is, each forward contract will be created at the fixed price that corresponds to the fixed price of a swap of the same maturity with payments made at the same dates as the series of forward contracts. That means that some of the forward contracts would have positive values and some would have negative values, but their combined values would equal zero.

Now, it sounds like that price would be hard to find, but it is not. We would not, however, go about finding it by taking random guesses. Doing so would take seemingly forever. Along the way, we would notice that some of these implicit forward contracts would have positive values and some would have negative values. If the positives outweighed the negatives, then the overall swap value would be positive, which is too high. Likewise, we might plug in a number that would produce an overall negative value, with the implicit forward contract values tending to be predominantly negative, which is too low.

Not surprisingly, we can find that price easily by appealing to the principle of arbitrage. We said that the principle of arbitrage will guide us *all the way through* derivative pricing. We will omit the details, but here is the general idea.

Suppose we buy an asset that pays the amounts $S_1, S_2, \ldots, S_n$ at times $t_1, t_2, \ldots, t_n$. These are unknown amounts. A simple example would be a floating-rate bond for which the S values represent the coupons that are unknown at the start but ultimately are determined by the evolution of interest rates. Then suppose we finance the purchase of that asset by borrowing money that we promise to repay with equal fixed payments of $FS_0(T)$. That strategy replicates the swap. As you have already learned, replication is the key to pricing.

Valuation of the swap during its life again appeals to replication and the principle of no arbitrage. We will find a way to reproduce the remaining payments on the swap with other transactions. The value of that strategy is the value of the swap.

To obtain the fixed rate on the swap or to value it later during its life, we will need information from the market for the underlying. As we previously noted, there are derivatives on bonds and interest rates, equities, currencies, and commodities. It is not possible to provide a general and simple statement of how to price swaps that covers all of these cases, but that topic is covered in advanced material.

EXAMPLE 5

Swap Pricing and Valuation

1 A swap is equivalent to a series of:

A forward contracts, each created at the swap price.

B long forward contracts, matched with short futures contracts.

C forward contracts, each created at their appropriate forward prices.

2 If the present value of the payments in a forward contract or swap is not zero, which of the following is most likely to be true?

A The contract cannot legally be created.

B The contract must be replicated by another contract with zero value.

C The party whose stream of payments to be received is greater has to pay the other party the present value difference.

Solution to 1:

A is correct. Each implicit forward contract is said to be off-market, because it is created at the swap price, not the appropriate forward price, which would be the price created in the forward market.

Solution to 2:

C is correct. The party whose stream of payments to be received is greater has to pay the other party the present value difference. Such a contract can legally be created, but the party receiving the greater present value must compensate the other party with a cash payment at the start. Replication is never required.

8

PRICING AND VALUATION OF OPTIONS

j explain the exercise value, time value, and moneyness of an option;

k identify the factors that determine the value of an option and explain how each factor affects the value of an option;

o explain under which circumstances the values of European and American options differ.

Unlike a forward, futures, or swap contract, an option is clearly an asset to the holder and a liability to the seller. The buyer of an option pays a sum of money, called the premium, and receives the right to buy (a call) or sell (a put) the underlying. The seller receives the premium and undertakes a potential obligation because the buyer has the right, but not the obligation, to exercise the option. Options are, therefore, contingent claims. Pricing the option is the same as assigning its value. Some confusion from that terminology may still arise, in that an option could trade in the market for an amount that differs from its value.

As mentioned, there are two general types of options. Calls represent the right to buy, and puts represent the right to sell. There are also two important exercise characteristics of options. American options allow exercise at any time up to the expiration, while European options allow exercise only at expiration. It is important to understand that the terms "American" and "European" have no relationship to where the options are traded. Because the right to exercise can be a complex feature of an option, European options are easier to understand, and we will focus on them first.

We will use the same notation used with forwards. We start by assuming that today is time 0, and the option expires at time T. The underlying is an asset currently priced at S_0, and at time T, its price is S_T. Of course, we do not know S_T until we get to the expiration. The option has an exercise or strike price of X. The symbols we use are as follows:
For calls,

c_0 = value (price) of European call today
c_T = value (price) of European call at expiration
C_0 = value (price) of American call today
C_T = value (price) of American call at expiration

For puts,

p_0 = value (price) of European put today
p_T = value (price) of European put at expiration
P_0 = value (price) of American put today
P_T = value (price) of American put at expiration

8.1 European Option Pricing

Recall that in studying forward contracts earlier in this reading, the first thing we learned is how a forward contract pays off at expiration. Then we backed up and determined how forward contracts are priced and valued prior to expiration. We follow that same approach for options.

8.1.1 *Value of a European Option at Expiration*

Recall that a European call option allows the holder to buy the underlying at expiration by paying the exercise price. Therefore, exercise is justified only if the value of the underlying exceeds the exercise price. Otherwise, the holder would simply let the call expire. So if the call is worth exercising ($S_T > X$), the holder pays X and receives an asset worth S_T. Thus, the option is worth $S_T - X$. If the call is not worth exercising ($S_T \leq X$), the option simply expires and is worth nothing at expiration.[14] Thus, the value of the option at expiration is the greater of either zero or the underlying price at expiration minus the exercise price, which is typically written as

$$c_T = \text{Max}(0, S_T - X) \tag{8}$$

This formula is also sometimes referred to as the **exercise value** or **intrinsic value**. In this reading, we will use the term exercise value.

Taking a simple example, if the exercise price is €40 and the underlying price is at expiration €43, the call is worth c_T = Max(€0,€43 – €40) = Max(€0,€3) = €3. If the underlying price at expiration is €39, the call is worth c_T = Max(0,€39 – €40) = Max(€0,–€1) = €0.

For puts, the holder has the right to sell the underlying at X. If the underlying is worth less than X at expiration ($X > S_T$), the put will be exercised and worth $X - S_T$ because it allowed the holder to avoid the loss in value of the asset of that amount. If the underlying is equal to or worth more than the exercise price at expiration ($S_T \geq X$), the put will simply expire with no value. So, the put is worth the greater of either zero or the exercise price minus the price of the underlying at expiration.

$$p_T = \text{Max}(0, X - S_T) \tag{9}$$

As discussed above, this formula is referred to as the exercise value or intrinsic value, and as noted, we will use the term exercise value.

Using the same example as with the call, if the underlying is €43 at expiration, the put is worth p_T = Max(€0,€40 – €43) = Max(0,–€3) = €0. If the underlying is €39 at expiration, the put is worth p_T = Max(€0,€40 – €39) = Max(€0,€1) = €1.

Thus, the holder of an option looks out into the future and sees these relationships as the payoff possibilities. That does not mean the holder knows what S_T will be, but the holder knows that all of the uncertainty of the option payoff is determined by the behavior of the underlying.

The results of this section can be restated as follows:

> *The value of a European call at expiration is the exercise value, which is the greater of zero or the value of the underlying minus the exercise price.*
>
> *The value of a European put at expiration is the exercise value, which is the greater of zero or the exercise price minus the value of the underlying.*

---

[14] In all the remaining material, we identify conditions at expiration, such as $S_T > X$ and $S_T \leq X$. Here we merged the equality case ($S_T = X$) with the less-than case (<). We could have done it the other way around ($S_T < X$ and $S_T \geq X$), which would have had no effect on our interpretations or any calculations of option value. For convenience, in some situations we will use one specification and in some the other.

To understand option pricing, we have to work our way forward in a gradual manner. The next valuable steps involve using our intuition to identify some characteristics that will influence the value of the option. We might not be able to quantify their effects just yet, but we can rationalize why these factors affect the value of an option.

8.1.2 *Effect of the Value of the Underlying*

The value of the underlying is obviously a critical element in determining the value of an option. It is the uncertainty of the underlying that provides the motivation for using options. It is easy to rationalize the direction of the effect of the underlying.

A call option can be viewed as a mean of acquiring the underlying, whereas a put option can be viewed as a means of selling the underlying. Thus, a call option is logically worth more if the underlying is worth more, and a put option is logically worth more if the underlying is worth less.

The value of the underlying also forms one of the boundaries for calls. The value of a call option cannot exceed the value of the underlying. After all, a call option is only a means of acquiring the underlying. It can never give the holder more benefit than the underlying. Hence, the value of the underlying forms an upper boundary on what a call is worth. The underlying does not provide an upper or lower boundary for puts. That role is played by the exercise price, as we will see in the next section.

To recap what we learned here,

> *The value of a European call option is directly related to the value of the underlying.*
>
> *The value of a European put option is inversely related to the value of the underlying.*

8.1.3 *Effect of the Exercise Price*

The exercise price is a critical factor in determining the value of an option. The exercise price is the hurdle beyond which the underlying must go to justify exercise. For a call, the underlying must rise above the exercise price, and for a put, the underlying must fall below the exercise price, to justify exercise. When the underlying is beyond the exercise price in the appropriate direction (higher for a call, lower for a put), the option is said to be **in the money**. When the underlying is precisely at the exercise price, the option is said to be **at the money**. When the underlying has not reached the exercise price (currently lower for a call, higher for a put), the option is said to be **out of the money**. This characterization of whether the option is in-, at-, or out-of-the-money is referred to as the option's **moneyness**.

For a call option, a lower exercise price has two benefits. One is that there are more values of the underlying at expiration that are above the exercise price, meaning that there are more outcomes in which the call expires in-the-money. The other benefit is that assuming the call expires in-the-money, for any value of the underlying, the option value is greater the lower the exercise price. In other words, at expiration the underlying value S_T will be above the exercise price far more often, the lower is X. And if S_T is indeed higher than X, the payoff of $S_T - X$ is greater, the lower is X.

For puts, the effect is just the opposite. To expire in-the-money, the value of the underlying must fall below the exercise price. The higher the exercise price, the better chance the underlying has of getting below it. Likewise, if the value of the underlying does fall below the exercise price, the higher the exercise price, the greater the payoff. So, if X is higher, S_T will be below it more often, and if S_T is less than X, the payoff of $X - S_T$ is greater, the higher is X for whatever value of S_T occurs.

The exercise price also helps form an upper bound for the value of a European put. If you were holding a European put, the best outcome you could hope for is a zero value of the underlying. For equities, that would mean complete failure and dissolution of the company with shareholders receiving no final payment.[15] In that case, the put would pay $X - S_T$, but with S_T at zero, the put would pay X. If the underlying value goes to zero during the life of the European put, however, the holder cannot collect that payoff until expiration. Nonetheless, the holder would have a risk-free claim on a payoff of X at expiration. Thus, the most the put would be worth is the present value of X, meaning X discounted from expiration to the current day at the risk-free rate.[16] Although the holder cannot collect the payoff by exercising the option, he could sell it for the present value of X.

To recap these results,

> *The value of a European call option is inversely related to the exercise price.*

> *The value of a European put option is directly related to the exercise price.*

8.1.4 *Effect of Time to Expiration*

Logic suggests that longer-term options should be worth more than shorter-term options. That statement is usually true but not always. A call option unquestionably benefits from additional time. For example, the right to buy an asset for $50 is worth a lot more if that right is available for two years instead of one. The additional time provides further opportunity for the underlying to rise above the exercise price. Although that means there is also additional time for the underlying to fall below the exercise price, it hardly matters to the holder of the call because the loss on the downside is limited to the premium paid.

For a European put option, the additional time still provides more opportunity for the underlying price to fall below the exercise price, but with the additional risk of it rising above the exercise price mitigated by the limited loss of the premium if the put expires out-of-the-money. Thus, it sounds as if puts benefit from longer time, but that is not necessarily true. There is a subtle penalty for this additional time. Put option holders are awaiting the sale of the underlying, for which they will receive the exercise price. The longer they have to wait, the lower the present value of the payoff. For some puts, this negative effect can dwarf the positive effect. This situation occurs with a put the longer the time to expiration, the higher the risk-free rate of interest, and the deeper it is in-the-money. The positive effect of time, however, is somewhat more dominant.

Note that we did not mention this effect for calls. For calls, the holder is waiting to pay out money at expiration. More time lowers the value of this possible outlay. Hence, a longer time period helps call option buyers in this regard.

To recap these results,

> *The value of a European call option is directly related to the time to expiration.*

---

[15] You might think this point means that people who buy puts are hoping the company goes bankrupt, a seemingly morbid motivation. Yet, put buyers are often people who own the stock and buy the put for protection. This motivation is no different from owning a house and buying fire insurance. You do not want the house to burn down. If your sole motivation in buying the insurance were to make a profit on the insurance, you would want the house to burn down. This moral hazard problem illustrates why it is difficult, if not impossible, to buy insurance on a house you do not own. Likewise, executives are prohibited from owning puts on their companies' stock. Individual investors can own puts on stocks they do not own, because they cannot drive the stock price down.

[16] For the put holder to truly have a risk-free claim on X at expiration, given zero value of the underlying today, the underlying value must go to zero and have no possibility of any recovery. If there is any possibility of recovery, the underlying value would not go to zero, as is often observed when a legal filing for bankruptcy is undertaken. Many equities do recover. If there were some chance of recovery but the equity value was zero, demand for the stock would be infinite, which would push the price up.

> *The value of a European put option can be either directly or inversely related to the time to expiration. The direct effect is more common, but the inverse effect can prevail with a put the longer the time to expiration, the higher the risk-free rate, and the deeper it is in-the-money.*

8.1.5 *Effect of the Risk-Free Rate of Interest*

We have already alluded to the effect of the risk-free rate. For call options, a longer time to expiration means that the present value of the outlay of the exercise price is lower. In other words, with a longer time to expiration, the call option holder continues to earn interest on the money that could be expended later in paying the exercise price upon exercise of the option. If the option is ultimately never exercised, this factor is irrelevant, but it remains at best a benefit and at worst has no effect. For puts, the opposite argument prevails. A longer time to expiration with a higher interest rate lowers the present value of the receipt of the exercise price upon exercise. Thus, the value today of what the put holder might receive at expiration is lower. If the put is ultimately never exercised, the risk-free rate has no effect. Thus, at best, a higher risk-free rate has no effect on the value of a put. At worst, it decreases the value of the put.

These results are summarized as follows:

> *The value of a European call is directly related to the risk-free interest rate.*
>
> *The value of a European put is inversely related to the risk-free interest rate.*

8.1.6 *Effect of Volatility of the Underlying*

In studying the pricing of equities, we are conditioned to believe that volatility has a negative effect. After all, investors like return and dislike risk. Volatility is certainly an element of risk. Therefore, volatility is bad for investors, right? Well, partially right.

First, not all volatility is bad for investors. Unsystematic volatility should be irrelevant. Investors can hold diversified portfolios. Systematic volatility is clearly undesirable, but do not think that this means that volatility should be completely avoided where possible. If volatility were universally undesirable, no one would take risks. Clearly risks have to be taken to provide opportunity for reward.

With options, volatility of the underlying is, however, universally desirable. The greater the volatility of the underlying, the more an option is worth. This seemingly counterintuitive result is easy to understand with a little explanation.

First, let us make sure we know what volatility really means. In studying asset returns, we typically represent volatility with the standard deviation of the return, which measures the variation from the average return. The S&P 500 Index has an approximate long-run volatility of around 20%. Under the assumption of a normal distribution, a standard deviation of 20% implies that about 68% of the time, the returns will be within plus or minus one standard deviation of the average. About 95% of the time, they will be within plus or minus two standard deviations of the average. About 99% of the time, they will be within plus or minus three standard deviations of the average. When the distribution is non-normal, different interpretations apply, and in some extreme cases, the standard deviation can be nearly impossible to interpret.

Standard deviation is not the only notion of volatility, however, and it is not even needed at this point. You can proceed fairly safely with a measure as simple as the highest possible value minus the lowest, known as the range. The only requirement we need right now is that the concept of volatility reflects dispersion—how high and how low the underlying can go.

So, regardless of how we measure volatility, the following conditions will hold:

1. A call option will have a higher payoff the higher the underlying is at expiration.
2. A call option will have a zero payoff if it expires with the underlying below the exercise price.

If we could impose greater volatility on the underlying, we should be able to see that in Condition 1, the payoff has a better chance of being greater because the underlying has a greater possibility of large positive returns. In Condition 2, however, the zero payoff is unaffected if we impose greater volatility. Expiring more out-of-the-money is not worse than expiring less out-of-the-money, but expiring more in-the-money is better than expiring less-in-the-money.[17]

For puts, we have

1. A put option will have a higher payoff the lower the underlying is at expiration.
2. A put option will have a zero payoff if it expires with the underlying above the exercise price.

If we could impose greater volatility, we would find that it would have a beneficial effect in (1) because a larger positive payoff would have a greater chance of occurring. In (2), the zero payoff is unaffected. The greater of the option expiring more out-of-the-money is irrelevant. Expiring more out-of-the-money is not worse than expiring less out-of-the-money.

Thus, we summarize our results in this section as

> *The value of a European call is directly related to the volatility of the underlying.*
>
> *The value of a European put is directly related to the volatility of the underlying.*

The combined effects of time and volatility give rise to the concept of the time value of an option. The **time value** of an option is the difference between the market price of the option and its intrinsic value. It represents the market valuation of the potential for higher exercise value relative to the potential for lower exercise value given the volatility of the underlying. Time value of an option is not to be confused with the time value of money, which is the notion of money later being worth less than money today as a result of the combined effects of time and interest. Time value results in an option price being greater with volatility and time but declining as expiration approaches. At expiration, no time value remains and the option is worth only its exercise value. As such, an option price is said to decay over time, a process characterized as **time value decay**, which is covered in more advanced material.

8.1.7 *Effect of Payments on the Underlying and the Cost of Carry*

We previously discussed how payments on the underlying and carrying costs enter into the determination of forward prices. They also affect option prices. Payments on the underlying refer to dividends on stocks and interest on bonds. In addition, some commodities offer a convenience yield benefit. Carrying costs include the actual physical costs of maintaining and/or storing an asset.

Let us first consider the effect of benefits. Payments of dividends and interest reduce the value of the underlying. Stocks and bonds fall in value as dividends and interest are paid. These benefits to holders of these securities do not flow to holders of options. For call option holders, this reduction is a negative factor. The price of the underlying is hurt by such payments, and call holders do not get to collect these payments. For put holders, the effect is the opposite. When the value of the underlying is reduced, put holders are helped.

---

[17] Think of an option expiring out-of-the-money as like it being dead. (Indeed, the option is dead.) Being "more dead" is not worse than being "less dead."

Carrying costs have the opposite effect. They raise the effective cost of holding or shorting the asset. Holding call options enables an investor to participate in movements of the underlying without incurring these costs. Holding put options makes it more expensive to participate in movements in the underlying than by short selling because short sellers benefit from carrying costs, which are borne by owners of the asset.

To summarize the results from this section,

> *A European call option is worth less the more benefits that are paid by the underlying and worth more the more costs that are incurred in holding the underlying.*
>
> *A European put option is worth more the more benefits that are paid by the underlying and worth less the more costs that are incurred in holding the underlying.*

9 LOWER LIMITS FOR PRICES OF EUROPEAN OPTIONS

k identify the factors that determine the value of an option and explain how each factor affects the value of an option;

What we have learned so far forms a framework for understanding how European options are priced. Let us now go a step further and establish a minimum price for these options.

First, we need to look at a call option as similar to the purchase of the underlying with a portion of the purchase price financed by borrowing. If the underlying is a stock, this transaction is usually called a margin transaction. Assume that the underlying is worth S_0. Also assume that you borrow cash in the amount of the present value of X, promising to pay X back T periods later at an interest rate of r. Thus, $X/(1 + r)^T$ is the amount borrowed, and X is the amount to be paid back. Now move forward to time T and observe the price of the underlying, S_T. Upon paying back the loan, the overall strategy will be worth $S_T - X$, which can be positive or negative.

Next, consider an alternative strategy of buying a call option expiring at T with an exercise price of X, the same value as the face value of the loan. We know that the option payoffs will be $S_T - X$ if it expires in-the- money ($S_T > X$) and zero if not ($S_T \leq X$). Exhibit 11 compares these two strategies.[18]

---

[18] Note in Exhibit 11, and in others to come, that the inequality ≤ is referred to as out-of-the-money. The case of equality is technically referred to as at-the-money but the verbiage is simplified if we continue to call it out-of-the-money. It is certainly not in-the-money and at-the-money is arguably the same as out-of-the-money. Regardless of one's preference, the equality case can be attached to either of the two outcomes with no effect on our conclusions.

Exhibit 11 Call Option vs. Leveraged (Margin) Transaction

	Outcome at T	
	Call Expires Out-of-the-Money ($S_T \le X$)	**Call Expires In-the-Money ($S_T > X$)**
Call	0	$S_T - X$
Leveraged transaction		
Asset	S_T	S_T
Loan	$-X$	$-X$
Total	$S_T - X$	$S_T - X$

When the call expires in-the-money, both transactions produce identical payoffs. When the call expires out-of-the-money, the call value is zero, but the leveraged transaction is almost surely a loss. Its value $S_T - X$ is negative or zero at best (if S_T is exactly equal to X).

If two strategies are found to produce equivalent results in some outcomes but one produces a better result in all other outcomes, then one strategy dominates the former. Here we see that the call strategy dominates the leveraged strategy. Any strategy that dominates the other can never have a lower value at any time. Why would anyone pay more for one strategy than for another if the former will never produce a better result than the latter? Thus, the value of the call strategy, c_0, has to be worth at least the value of the leveraged transaction, S_0 (the value of the asset), minus $X/(1 + r)^T$ (the value of the loan). Hence, $c_0 \ge S_0 - X/(1 + r)^T$.

The inequality means that this statement provides the lowest price of the call, but there is one more thing we need to do. It can easily be true that $X/(1 + r)^T > S_0$. In that case, we are saying that the lowest value is a negative number, but that statement is meaningless. A call can never be worth less than zero, because its holder cannot be forced to exercise it. Thus, we tend to express this relationship as

$$c_0 \ge \text{Max}\left[0, S_0 - X\big/(1+r)^T\right] \tag{10}$$

which represents the greater of the value of zero or the underlying price minus the present value of the exercise price. This value becomes the lower limit of the call price.

Now consider an analogous result for puts. Suppose we want to profit from a declining price of the underlying. One way to do this is to sell the underlying short. Suppose we do that and invest cash equal to the present value of X into risk-free bonds that pay X at time T. At time T, given a price of the underlying of S_T, the short sale pays off $-S_T$, a reflection of the payment of S_T to cover the short sale. The bonds pay X. Hence, the total payoff is $X - S_T$.

Now, compare that result with the purchase of a put expiring at T with exercise price of X. If the put expires in–the-money ($S_T < X$), it is worth $X - S_T$. If it expires out-of-the-money ($S_T \ge X$), it is worth zero. Exhibit 12 illustrates the comparison of the put with the short sale and bond strategy. We see that for the in-the-money case, the put and short sale and bond strategies match each other. For the out-of-the-money case, however, the put performs better because the short sale and bond strategy pays $X - S_T$. With $S_T \ge X$, this payment amount is negative. With the put dominating the short sale and bond strategy, the put value cannot be less than the value of the short sale and bond strategy, meaning $p_0 \ge X/(1 + r)^T - S_0$. But as with calls, the right-hand side can be negative, and it hardly helps us to say that a put must sell for more than a negative number. A put can never be worth less than zero, because its owner cannot be forced to exercise it. Thus, the overall result is expressed succinctly as

$$p_0 \geq \text{Max}\left[0, X/(1+r)^T - S_0\right] \tag{11}$$

Exhibit 12 Put vs. Short Sale and Bond Purchase

	Outcome at T	
	Put Expires In-the-Money ($S_T < X$)	**Put Expires Out-of-the-Money ($S_T \geq X$)**
Put	$X - S_T$	0
Short sale and bond purchase		
Short sale	$-S_T$	$-S_T$
Bond	X	X
Total	$X - S_T$	$X - S_T$

Let us look at some basic examples. Assume the exercise price is €60, the risk-free rate is 4%, and the expiration is nine months, so $T = 9/12 = 0.75$. Consider two cases:

Underlying: S_0 = €70

Minimum call price = Max[0,€70 – €60/(1.04)$^{0.75}$] = Max(0,€11.74) = €11.74

Minimum put price = Max[0,€60/(1.04)$^{0.75}$ – €70] = Max(0,–€11.74) = €0.00

Underlying: S_0 = €50

Minimum call price = Max[0,€50 – €60/(1.04)$^{0.75}$] = Max(0,–€8.26) = €0.00

Minimum put price = Max[0,€60/(1.04)$^{0.75}$ – €50] = Max(0,€8.26) = €8.26

To recap, in this section we have established lower limits for call and put option values. Formally restating these results in words,

> *The lowest value of a European call is the greater of zero or the value of the underlying minus the present value of the exercise price.*
>
> *The lowest value of a European put is the greater of zero or the present value of the exercise price minus the value of the underlying.*

EXAMPLE 6

Basic Principles of European Option Pricing

1 Which of the following factors does *not* affect the value of a European option?

  A The volatility of the underlying

  B Dividends or interest paid by the underlying

  C The percentage of the investor's assets invested in the option

2 Which of the following statements imply that a European call on a stock is worth more?

A Less time to expiration

B A higher stock price relative to the exercise price

C Larger dividends paid by the stock during the life of the option

3 Why might a European put be worth less the longer the time to expiration?

A The cost of waiting to receive the exercise price is higher.

B The risk of the underlying is lower over a longer period of time.

C The longer time to expiration means that the put is more likely to expire out-of-the-money.

4 The loss in value of an option as it moves closer to expiration is called what?

A Time value decay

B Volatility diminution

C Time value of money

5 How does the minimum value of a call or put option differ from its exercise value?

A The exercise price is adjusted for the time value of money.

B The minimum value reflects the volatility of the underlying.

C The underlying price is adjusted for the time value of money.

Solution to 1:

C is correct. The investor's exposure to the option is not relevant to the price one should pay to buy or ask to sell the option. Volatility and dividends or interest paid by the underlying are highly relevant to the value of the option.

Solution to 2:

B is correct. The higher the stock price and the lower the exercise price, the more valuable is the call. Less time to expiration and larger dividends reduce the value of the call.

Solution to 3:

A is correct. Although the longer time benefits the holder of the option, it also has a cost in that exercise of a longer-term put comes much later. Therefore, the receipt of the exercise price is delayed. Longer time to expiration does not lower the risk of the underlying. The longer time also does not increase the likelihood of the option expiring out-of-the-money.

Solution to 4:

A is correct. An option has time value that decays as the expiration approaches. There is no such concept as volatility diminution. Time value of money relates only to the value of money at one point in time versus another.

Solution to 5:

A is correct. The minimum value formula is the greater of zero or the difference between the underlying price and the present value of the exercise price, whereas the exercise value is the maximum of zero and the appropriate difference between the underlying price and the exercise price. Volatility does not affect the minimum price. It does not make sense to adjust the underlying price for the time value of money for the simple reason that it is already adjusted for the time value of money.

10 PUT-CALL PARITY, PUT-CALL-FORWARD PARITY

l. explain put–call parity for European options;

m. explain put–call–forward parity for European options;

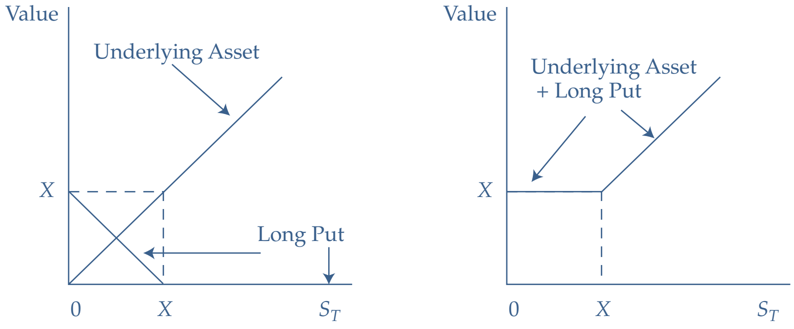
One of the first concepts that a trader learns in options is the parity relationship between puts and calls. Even though the word "parity" means "equivalence," puts and calls are not equivalent. There is, however, a relationship between the call price and the price of its corresponding put, which we refer to as put–call parity.

Suppose Investor A owns an asset that has a current price of S_0. Assume the asset makes no cash payments and has no carrying costs. The end of the holding period is time T, at which point the asset will be worth S_T. Fearing the possibility that S_T will decline, Investor A buys a put option with an exercise price of X, which can be used to sell the asset for X at time T. This put option has a premium of p_0. Combined with the value of the asset, the investor's current position is worth $S_0 + p_0$, which is the investor's money at risk. This strategy of holding the asset and a put is sometimes called a **protective put**.

At expiration, the value of the asset is S_T. The value of the put will be either zero or $X - S_T$. If the asset increases in value such that $S_T \geq X$, then the overall position is worth S_T. The asset has performed well, and the investor will let the put expire. If the asset value declines to the point at which $S_T < X$, the asset is worth S_T, and the put is worth $X - S_T$, for a total of X. In other words, the investor would exercise the put, selling the asset for X, which exceeds the asset's current value of S_T.

This strategy seems like a reasonable and possibly quite attractive investment. Investor A receives the benefit of unlimited upside potential, with the downside performance truncated at X. Exhibit 13 shows the performance of the protective put. The graph on the left illustrates the underlying asset and the put. The graph on the right shows their combined effects.

Exhibit 13 Protective Put (Asset Plus Long Put)

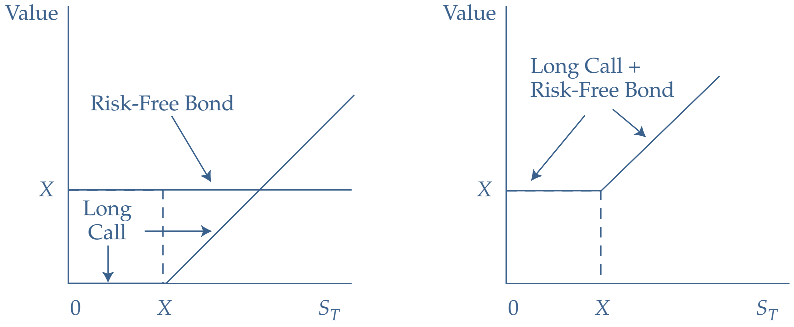


Consider Investor B, an options trader. At time 0, this investor buys a call option on this asset with an exercise price of X that expires at T and a risk-free zero-coupon bond with a face value of X that matures at T. The call costs c_0, and the bond costs the present value of X, which is $X/(1 + r)^T$. Thus, Investor B has invested funds of $c_0 + X/$

$(1 + r)^T$. This strategy is sometimes known as a **fiduciary call**. If the underlying price exceeds the exercise price at expiration, the call will be worth $S_T - X$, and the bond will mature and pay a value of X. These values combine to equal S_T. If the underlying price does not exceed the exercise price at expiration, the call expires worthless and the bond is worth X for a combined value of X.

Exhibit 14 shows the performance of the fiduciary call. The graph on the left shows the call and bond, and the graph on the right shows the combined effects of the two strategies.

Exhibit 14 Fiduciary Call (Long Call Plus Risk-Free Bond)



Comparing Exhibit 13 with Exhibit 14 shows that a protective put and a fiduciary call produce the same result. Exhibit 15 shows this result more directly by identifying the payoffs in the various outcomes. Recall that Investor A committed funds of $S_0 + p_0$, while Investor B committed funds of $c_0 + X/(1 + r)^T$. If both investors receive the same payoffs at time T regardless of the asset price at T, the amounts they invest at time 0 have to be the same. Thus, we require

$$S_0 + p_0 = c_0 + X\big/(1 + r)^T \tag{12}$$

This relationship is known as **put–call parity**.

Exhibit 15 Protective Put vs. Fiduciary Call

	Outcome at *T*	
	Put Expires In-the-Money ($S_T < X$)	**Call Expires In-the-Money ($S_T \geq X$)**
Protective put		
Asset	S_T	S_T
Long put	$X - S_T$	0
Total	X	S_T
Fiduciary call		
Long call	0	$S_T - X$
Risk-free bond	X	X
Total	X	S_T

For a simple example, assume call and put options with an exercise price of ¥100,000 in which the underlying is at ¥90,000 at time 0. The risk-free rate is 2% and the options expire in two months, so $T = 2/12 = 0.167$. To completely fill in the puzzle, we would need to know the put or call price, from which we could obtain the other. For now, let us write this relationship as

$$p_0 - c_0 = X\Big/(1 + r)^T - S_0$$

The right side would be $¥100{,}000/(1.02)^{0.167} - ¥90{,}000 = ¥9{,}670$. Thus, the put price should exceed the call price by ¥9,670. Thus, if the call were priced at ¥5,000, the put price would be ¥14,670. If we knew the put price, we could obtain the call price. Put–call parity does not tell us which price is correct, and it requires knowledge of one price to get the other. Alternatively, it can tell us the difference in the put and call prices.

Put–call parity must hold, at least within transaction costs, or arbitrage opportunities would arise. For example, suppose Investor C observes market prices and finds that the left- hand side of put–call parity, $S_0 + p_0$, is less than the right-hand side, $c_0 + X/(1 + r)^T$. Thus, the put and the stock cost less than the call and the bond. Knowing that there should be equality (parity), Investor C executes an arbitrage transaction, selling the overpriced transactions (the call and the bond) and buying the underpriced transactions (the asset and the put).[19] By selling the higher priced side and buying the lower priced side, Investor C will take in more money than she will pay out, a net inflow of $c_0 + X/(1 + r)^T - (S_0 + p_0)$. At expiration, the long put and long asset will offset the short call and bond, as shown in Exhibit 16.

Exhibit 16 Put–Call Parity Arbitrage

		Outcome at T	
Transaction	**Cash Flow at Time 0**	**Put Expires In-the-Money ($S_T < X$)**	**Call Expires In-the-Money ($S_T \geq X$)**
Buy asset	$-S_0$	S_T	S_T
Buy put	$-p_0$	$X - S_T$	0
Sell call	$+c_0$	0	$-(S_T - X)$
Borrow	$+X/(1 + r)^T$	$-X$	$-X$
Total	$-S_0 - p_0 + c_0 + X/(1 + r)^T > 0$	0	0

In simple terms, if $S_T < X$, the short call expires out-of-the-money and the put is exercised to sell the asset for X. This cash, X, is then used to pay off the loan. The net effect is that no money flows in or out at T. If $S_T \geq X$, the put expires out-of-the money, and the short call is exercised, meaning that Investor C must sell the asset for X. This cash, X, is then used to pay off the loan. Again, no money flows in or out. The net effect is a perfect hedge in which no money is paid out or received at T. But there was money taken in at time 0. Taking in money today and never having to pay it out is an arbitrage profit. Arbitrage opportunities like this, however, will be noticed by many investors who will engage in the same transactions. Prices will adjust until parity is restored, whereby $S_0 + p_0 = c_0 + X/(1 + r)^T$.

---

[19] Selling the bond is equivalent to borrowing, meaning to issue a loan.

Put–call parity provides tremendous insights into option pricing. Recall that we proved that going long the asset and long a put is equivalent to going long a call and long a risk-free bond. We can rearrange the put–call parity equation in the following ways:

$$
\begin{aligned}
S_0 + p_0 &= c_0 + X/(1+r)^T \\
&\Rightarrow \\
p_0 &= c_0 - S_0 + X/(1+r)^T \\
c_0 &= p_0 + S_0 - X/(1+r)^T \\
S_0 &= c_0 - p_0 + X/(1+r)^T \\
X/(1+r)^T &= S_0 + p_0 - c_0
\end{aligned}
$$

By using the symbols and the signs in these versions of put–call parity, we can see several important interpretations. In the equations below, plus signs mean long and minus signs mean short:

$$
\begin{aligned}
p_0 &= c_0 - S_0 + X/(1+r)^T && \Rightarrow && \text{long put} = \text{long call, short asset, long bond} \\
c_0 &= p_0 + S_0 - X/(1+r)^T && \Rightarrow && \text{long call} = \text{long put, long asset, short bond} \\
S_0 &= c_0 - p_0 + X/(1+r)^T && \Rightarrow && \text{long asset} = \text{long call, short put, long bond} \\
X/(1+r)^T &= S_0 + p_0 - c_0 && \Rightarrow && \text{long bond} = \text{long asset, long put, short call}
\end{aligned}
$$

You should be able to convince yourself of any of these points by constructing a table similar to Exhibit 15.[20]

10.1 Put–Call–Forward Parity

Recall that we demonstrated that one could create a risk-free position by going long the asset and selling a forward contract.[21] It follows that one can synthetically create a position in the asset by going long a forward contract and long a risk-free bond. Recall our put–call parity discussion and assume that Investor A creates his protective put in a slightly different manner. Instead of buying the asset, he buys a forward contract and a risk-free bond in which the face value is the forward price. Exhibit 17 shows that this strategy is a synthetic protective put. Because we showed that the fiduciary call is equivalent to the protective put, a fiduciary call has to be equivalent to a protective put with a forward contract. Exhibit 18 demonstrates this point.

Exhibit 17 Protective Put with Forward Contract vs. Protective Put with Asset

	Outcome at T	
	Put Expires In-the-Money ($S_T < X$)	**Put Expires Out-of-the-Money ($S_T \geq X$)**
Protective put with asset		
Asset	S_T	S_T
Long put	$X - S_T$	0
Total	X	S_T

(continued)

---

[20] As a further exercise, you might change the signs of each term in the above and provide the appropriate interpretations.

[21] You might wish to review Exhibit 6.

Exhibit 17 (Continued)

	Outcome at *T*	
	Put Expires In-the-Money ($S_T < X$)	**Put Expires Out-of-the-Money ($S_T \geq X$)**
Protective put with forward contract		
Risk-free bond	$F_0(T)$	$F_0(T)$
Forward contract	$S_T - F_0(T)$	$S_T - F_0(T)$
Long put	$X - S_T$	0
Total	X	S_T

Exhibit 18 Protective Put with Forward Contract vs. Fiduciary Call

	Outcome at *T*	
	Put Expires In-the-Money ($S_T < X$)	**Call Expires In-the-Money ($S_T \geq X$)**
Protective Put with Forward Contract		
Risk-free bond	$F_0(T)$	$F_0(T)$
Forward contract	$S_T - F_0(T)$	$S_T - F_0(T)$
Long put	$X - S_T$	0
Total	X	S_T
Fiduciary Call		
Call	0	$S_T - X$
Risk-free bond	X	X
Total	X	S_T

It follows that the cost of the fiduciary call must equal the cost of the synthetic protective put, giving us what is referred to as **put–call–forward parity**,

$$F_0(T)\big/(1+r)^T + p_0 = c_0 + X\big/(1+r)^T \tag{13}$$

Returning to our put–call parity example, a forward contract on ¥90,000 expiring in two months with a 2% interest rate would have a price of ¥90,000$(1.02)^{0.167}$ = ¥90,298. Rearranging Equation 13, we have

$$p_0 - c_0 = \left[X - F_0(T)\right]\big/(1+r)^T$$

The right-hand side is (¥100,000 – ¥90,298)/$(1.02)^{0.167}$ = ¥9,670, which is the same answer we obtained using the underlying asset rather than the forward contract. Naturally these two models give us the same answer. They are both based on the assumption that no arbitrage is possible within the spot, forward, and options markets.

So far we have learned only how to price options in relation to other options, such as a call versus a put or a call or a put versus a forward. We need a way to price options versus their underlying.

EXAMPLE 7

Put–Call Parity

1 Which of the following statements *best* describes put–call parity?

A The put price always equals the call price.

B The put price equals the call price if the volatility is known.

C The put price plus the underlying price equals the call price plus the present value of the exercise price.

2 From put–call parity, which of the following transactions is risk-free?

A Long asset, long put, short call

B Long call, long put, short asset

C Long asset, long call, short bond

Solution to 1:

C is correct. The put and underlying make up a protective put, while the call and present value of the exercise price make up a fiduciary call. The put price equals the call price for certain combinations of interest rates, times to expiration, and option moneyness, but these are special cases. Volatility has no effect on put–call parity.

Solution to 2:

A is correct. The combination of a long asset, long put, and short call is risk free because its payoffs produce a known cash flow of the value of the exercise price. The other two combinations do not produce risk-free positions. You should work through the payoffs of these three combinations in the form of Exhibit 12.

BINOMIAL VALUATION OF OPTIONS

11

- explain how the value of an option is determined using a one-period binomial model;

Because the option payoff is determined by the underlying, if we know the outcome of the underlying, we know the payoff of the option. That means that the price of the underlying is the only element of uncertainty. Moreover, the uncertainty is not so much the value of the underlying at expiration as it is whether the underlying is above or below the exercise price. If the underlying is above the exercise price at expiration, the payoff is $S_T - X$ for calls and zero for puts. If the underlying is below the exercise price at expiration, the payoff is zero for calls and $X - S_T$ for puts. In other words, the payoff of the option is straightforward and known, as soon as we know whether the option expires in- or out-of-the-money. Note that for forwards, futures, and swaps, there is no such added complexity. The payoff formula is the same regardless of whether the underlying is above or below the hurdle.

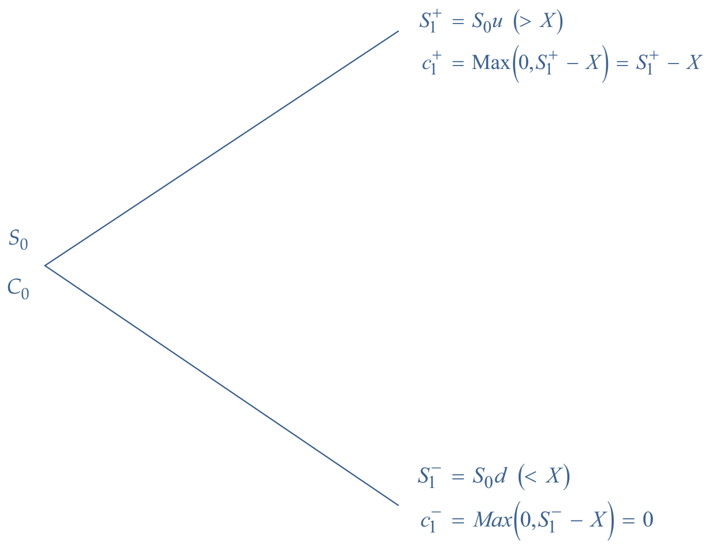
As a result of this characteristic of options, derivation of an option pricing model requires the specification of a model of a random process that describes movements in the underlying. Given the entirely different nature of the payoffs above and below the exercise price, it might seem difficult to derive the option price, even if we could model movements in the underlying. Fortunately, the process is less difficult than it first appears.

At this level of treatment, we will start with a very simple model that allows only two possible movements in the underlying—one going up and one going down from where it is now. This model with two possible outcomes is called the **binomial model**. Start with the underlying at S_0, and let it go up to S_1^+ or down to S_1^-. We cannot arbitrarily set these values at just anything. We will be required to know the values of S_1^+ and S_1^-. That does not mean we know which outcome will occur. It means that we know only what the possibilities are. In doing so, we effectively know the volatility. Assume the probability of the move to S_1^+ is q and the probability of the move to S_1^- is $1 - q$. We specify the returns implied by these moves as up and down factors, u and d, where

$$u = \frac{S_1^+}{S_0},\ d = \frac{S_1^-}{S_0} \tag{14}$$

Now, consider a European call option that expires at time 1 and has an exercise price of X. Let the call prices be c_0 today and c_1^+ and c_1^- at expiration. Exhibit 19 illustrates the model. Our objective is to determine the price of the option today, meaning to determine a formula for c_0. Knowing what we know about arbitrage and the pricing of forward contracts, it would seem we could construct a risk-free portfolio involving this option.

Exhibit 19 The Binomial Option Pricing Model


$S_1^+ = S_0u\ (> X)$
$c_1^+ = \text{Max}(0, S_1^+ - X) = S_1^+ - X$
S_0
C_0
$S_1^- = S_0d\ (< X)$
$c_1^- = Max(0, S_1^- - X) = 0$

Because call options and the underlying move together, one possibility is that buying the underlying and selling a call could create a hedge. Indeed it does, but one unit of each is not the appropriate balance. Let us sell one call and hold h units of the underlying. The value h is unknown at the moment, but we will be able to determine its value. The value today of a combination of h units of the underlying and one short call is

$$V_0 = hS_0 - c_0$$

Think of V_0 as the amount of money invested. Depending on which of the two paths is taken by the underlying, the value of this portfolio at time 1 will be

$$V_1^+ = hS_1^+ - c_1^+ \quad \text{or} \quad V_1^- = hS_1^- - c_1^- \tag{15}$$

If the portfolio were hedged, then V_1^+ would equal V_1^-. We can set V_1^+ and V_1^- equal to each other and solve for the value of h that assures us that the portfolio is hedged:

$$\begin{aligned} &V_1^+ = V_1^- \\ &\quad \Rightarrow hS_1^+ - c_1^+ = hS_1^- - c_1^- \\ &\quad \Rightarrow h = \frac{c_1^+ - c_1^-}{S_1^+ - S_1^-} \end{aligned} \tag{16}$$

The values on the right-hand side are known, so we can easily calculate h. Thus, we can derive the number of units of the underlying that will perfectly hedge one unit of the short call.

We know that a perfectly hedged investment should earn the risk-free rate, r. Thus, the following statement must be true:

$$V_1^+ \text{ (or } V_1^-) = V_0(1 + r)$$

We can substitute the value of V_1^+ or V_1^- from Equation 15 into the above equation. Then we do a little algebra, which is not important to this discussion, and obtain the formula for the option price,

$$c_0 = \frac{\pi c_1^+ + (1 - \pi)c_1^-}{1 + r} \quad \text{where} \quad \pi = \frac{1 + r - d}{u - d} \tag{17}$$

Equation 17 shows that the value of the call today is a weighted average of the next two possible call prices at expiration, where the weights, π and $1 - \pi$, are given by the second formula in Equation 17.

This formula sheds a great deal of light on option pricing. Notice the following:

- The volatility of the underlying, which is reflected in the difference between S_1^+ and S_1^- and affects c_1^+ and c_1^-, is an important factor in determining the value of the option.
- The probabilities of the up and down moves, q and $1 - q$, do not appear in the formula.[22]
- The values π and $1 - \pi$ are similar to probabilities and are often called synthetic or pseudo probabilities. They produce a weighted average of the next two possible call values, a type of expected future value.
- The formula takes the form of an expected future value, the numerator, discounted at the risk-free rate.

---

[22] We introduced them earlier to help make this point, but ultimately they serve no purpose.

On the first point, if volatility increases, the difference between S_1^+ and S_1^- increases, which widens the range between c_1^+ and c_1^-, leading to a higher option value. The upper payoff, c_1^+, will be larger and the lower payoff, c_1^-, will still be zero.[23] On the second point, the actual probabilities of the up and down moves do not matter. This result is because of our ability to construct a hedge and the rule of arbitrage. On the third point, the irrelevance of the actual probabilities is replaced by the relevance of a set of synthetic or pseudo probabilities, π and 1 – π, which are called **risk-neutral probabilities**. On the fourth point, these risk-neutral probabilities are used to find a synthetic expected value, which is then discounted at the risk-free rate. Thus, the option is valued as though investors are risk neutral. As we discussed extensively earlier, that is not the same as assuming that investors are risk neutral.

If the option does not trade at the specified formula, Equation 17, investors can engage in arbitrage transactions. If the option is trading too high relative to the formula, investors can sell the call, buy *h* shares of the underlying, and earn a return in excess of the risk-free rate, while funding the transaction by borrowing at the risk-free rate. The combined actions of arbitrageurs will result in downward pressure on the option price until it converges to the model price. If the option price is too low, buying the call, selling short *h* units of the asset, and investing the proceeds in risk-free bonds will generate risk-free cash that will earn more than the risk-free rate. The combined actions of arbitrageurs doing this will pressure the call price to rise until it reaches the price given by the model.

We will omit the details, but the hedge portfolio can also be constructed with puts.[24] Changing the *c*'s to *p*'s leads to the binomial put option pricing formula,

$$p_0 = \frac{\pi p_1^+ + (1-\pi) p_1^-}{1+r} \tag{18}$$

with the risk-neutral probability π determined by the same formula as for calls, as shown in Equation 17.

Let us construct a simple example. Let S_0 be £40 and the risk-free rate be 5%. The up and down factors are u = 1.20 and d = 0.75. Thus, the next two possible prices of the asset are S_1^+ = £40(1.20) = £48 and S_1^- = £40(0.75) = £30. Consider a call and a put that have exercise prices of £38. Then the next two possible values of the call and put are

$$c_1^+ = \text{Max}(0,£48 - £38) = £10$$

$$c_1^- = \text{Max}(0,£30 - £38) = £0$$

$$p_1^+ = \text{Max}(0,£38 - £48) = £0$$

$$p_1^- = \text{Max}(0,£38 - £30) = £8$$

Next we compute the risk-neutral probability,

$$\pi = \frac{1 + 0.05 - 0.75}{1.20 - 0.75} = 0.667$$

---

[23] Although the lower payoff is zero in this example, that will not always be the case.

[24] A long position in *h* units of the underlying would be hedged with one long put. The formula for *h* is the same as the one given here for calls, with call prices in the numerator instead of put prices.

The values of the call and put are

$$c_0 = \frac{0.667(£10) + (1 - 0.667)£0}{1.05} = £6.35$$

and

$$p_0 = \frac{0.667(£0) + (1 - 0.667)£8}{1.05} = £2.54$$

The binomial model, as we see it here, is extremely simple. In reality, of course, there are more than two possible next-period prices for the underlying. As it turns out, we can extend the number of periods and subdivide an option's life into an increasing number of smaller time periods. In that case, we can obtain a more accurate and realistic model for option pricing, one that is widely used in practice. Given our objective in this reading of understanding the basic ideas behind derivative pricing, the one-period model is sufficient for the time being.

EXAMPLE 8

Binomial Valuation of Options

1 Which of the following terms directly represents the volatility of the underlying in the binomial model?

- **A** The standard deviation of the underlying
- **B** The difference between the up and down factors
- **C** The ratio of the underlying value to the exercise price.

2 Which of the following is *not* a factor in pricing a call option in the binomial model?

- **A** The risk-free rate
- **B** The exercise price
- **C** The probability that the underlying will go up

3 Which of the following *best* describes the binomial option pricing formula?

- **A** The expected payoff is discounted at the risk-free rate plus a risk premium.
- **B** The spot price is compounded at the risk-free rate minus the volatility premium.
- **C** The expected payoff based on risk-neutral probabilities is discounted at the risk-free rate.

Solution to 1:

B is correct. The up and down factors express how high and how low the underlying can go. Standard deviation does not appear directly in the binomial model, although it is implicit. The ratio of the underlying value to the exercise price expresses the moneyness of the option.

Solution to 2:

C is correct. The actual probabilities of the up and down moves are irrelevant to pricing options. The risk-free and exercise price are, of course, highly relevant.

Solution to 3:

C is correct. Risk-neutral probabilities are used, and discounting is at the risk-free rate. There is no risk premium incorporated into option pricing because of the use of arbitrage.

We have now seen how to obtain the price of a European option. Let us now consider what happens if the options are American, meaning they have the right to be exercised early.

12 AMERICAN OPTION PRICING

- explain under which circumstances the values of European and American options differ.

First, we will use upper case letters for American call and put prices: C_0 and P_0. Second, we know that American options possess every characteristic of European options and one additional trait: They can be exercised at any time prior to expiration. Early exercise cannot be required, so the right to exercise early cannot have negative value. Thus, American options cannot sell for less than European options. Thus, we can state the following:

$$\begin{aligned} C_0 &\geq c_0 \\ P_0 &\geq p_0 \end{aligned} \tag{19}$$

Given the price of the underlying at S_0, the early-exercise feature means that we can exercise the option at any time. So, we can claim the value $\text{Max}(0,S_0 - X)$ for calls and $\text{Max}(0,X - S_0)$ for puts. These values establish new minimum prices for American calls and puts,

$$\begin{aligned} C_0 &= \text{Max}\left(0,S_0 - X\right) \\ P_0 &= \text{Max}\left(0,X - S_0\right) \end{aligned} \tag{20}$$

For call options, we previously learned that a European call has a minimum value given by Equation 10, which is restated here:

$$c_0 \geq \text{Max}\left[0,S_0 - X\Big/\left(1 + r\right)^T\right]$$

Comparing $\text{Max}(0,S_0 - X)$ (the minimum for American calls) with $\text{Max}[0,S_0 - X/(1 + r)^T]$ (the minimum for European calls) reveals that the latter is either the same or higher. There are some circumstances in which both minima are zero, some in which the American minimum is zero and the European minimum is positive, and some in which both are positive, in which case $S_0 - X/(1 + r)^T$ is unquestionably more than $S_0 - X$. Given that an American call price cannot be less than a European call price, we have to reestablish the American call minimum as $\text{Max}[0,S_0 - X/(1 + r)^T]$.

For put options, we previously learned that a European put has a minimum value given by Equation 11, which is restated here:

$$p_0 \geq \text{Max}\left[0,X\Big/\left(1 + r\right)^T - S_0\right]$$

Comparing $\text{Max}(0,X - S_0)$ (the minimum for American puts) with $\text{Max}[0,X/(1 + r)^T - S_0]$ (the minimum for European puts) reveals that the former is never less. In some circumstances, they are both zero. In some, $X - S_0$ is positive and $X/(1 + r)$

$^{T} - S_0$ is negative, and in some cases both are positive but $X - S_0$ is unquestionably more than $X/(1 + r)^T - S_0$. Thus, the American put minimum value is the exercise value, which is Max(0,$X - S_0$).

So, now we have new minimum prices for American calls and puts:

$$C_0 \geq \text{Max}\left[0, S_0 - X/(1+r)^T\right]$$
$$P_0 \geq \text{Max}(0, X - S_0) \tag{21}$$

Thus, in the market these options will trade for at least these values.

Let us return to the previous examples for the minimum values. The exercise price is €60, the risk-free rate is 4%, and the expiration is $T = 0.75$. Consider the two cases below:

Underlying: S_0 = €70

- The minimum European call price was previously calculated as €11.74. The exercise value of the American call is Max(0,€70 – €60) = €10. The American call has to sell for at least as much as the European call, so the minimum price of the American call is €11.74.
- The minimum European put price was €0.00. This is also the exercise value of the American put [Max(0,€60 – €70) = €0.00], so the minimum price of the American put is still €0.00

Underlying: S_0 = €50

- The minimum European call price was previously calculated as €0.00. The exercise value of the American call is Max(0,€50 – €60) = €0.00, so €0.00 is still the minimum price of the American call.
- The minimum European put price was previously calculated as €8.26. The exercise value of the American put is Max(0,€60 – €50) = €10. So, €10 is the minimum price of the American put.

The call result leads us to a somewhat surprising conclusion. With the exception of what happens at expiration when American and European calls are effectively the same and both worth the exercise value, an American call is always worth more in the market than exercised. That means that an American call will never be exercised early. This result is probably not intuitive.

Consider a deep in-the-money call. One might think that if the holder expected the underlying to not increase any further, exercise might be justified. Yet, we said the call would sell for more in the market than its exercise value. What is the rationale? If the investor thinks the underlying will not go up any further and thus expects no further gains from the option, why would she prefer the underlying? Would the investor be happier holding the underlying, which she believes is not expected to increase? Moreover, she would tie up more funds exercising to acquire the underlying than if she just held on to the option or, better yet, sold it to another investor.

So far, however, we have left out a possible factor that can affect early exercise. Suppose the underlying is a stock and pays dividends. When a stock goes ex-dividend, its price instantaneously falls. Although we will omit the details, an investor holding a call option may find it worthwhile to exercise the call just before the stock goes ex-dividend. The capture of the dividend, thereby avoiding the ex-dividend drop in the price of the underlying, can make early exercise worthwhile. If the underlying is a bond, coupon interest can also motivate early exercise. But if there are significant carrying costs, the motivation for early exercise is weakened. Storage costs lend a preference for owning the option over owning the underlying.

Because the minimum value of an American put exceeds the minimum value of the European put, there is a much stronger motivation for early exercise. Suppose you owned an American put on a stock that is completely bankrupt, with a zero stock price and no possibility of recovery. You can either wait until expiration and capture its exercise value of $\text{Max}(0,X - S_T) = \text{Max}(0,X - 0) = \text{Max}(0,X) = X$, or you can capture that value by exercising now. Obviously now is better. As it turns out, however, the underlying does not need to go all the way to zero. There is a critical point at which a put is so deep in-the-money that early exercise is justified. This rationale works differently for a call. A deep in-the-money put has a limit to its ultimate value. It can get no deeper than when the underlying goes to zero. For a call, there is no limit to its moneyness because the underlying has no upper limit to its price.

Although dividends and coupon interest encourage early exercise for calls, they discourage early exercise for puts. The loss from the decline in the price of the underlying that is avoided by exercising a call just before the decline works to the benefit of a put holder. Therefore, if a put holder were considering exercising early, he would be better off waiting until right after the dividend or interest were paid. Carrying costs on the underlying, which discourage exercise for calls, encourage exercise for puts.

At this point, we cannot determine the critical prices at which American options are best exercised early. We require more knowledge and experience with option pricing models, which is covered in more advanced material.

EXAMPLE 9

American Option Pricing

1 With respect to American calls, which of the following statements is *most* accurate?

   A American calls should be exercised early if the underlying has reached its expected maximum price.

   B American calls should be exercised early if the underlying has a lower expected return than the risk-free rate.

   C American calls should be exercised early only if there is a dividend or other cash payment on the underlying.

2 The effect of dividends on a stock on early exercise of a put is to:

   A make early exercise less likely.

   B have no effect on early exercise.

   C make early exercise more likely.

Solution to 1:

C is correct. Cash payments on the underlying are the only reason to exercise American calls early. Interest rates, the expected return on the underlying, and any notion of a maximum price is irrelevant. But note that a dividend does not mean that early exercise should automatically be conducted. A dividend is only a necessary condition to justify early exercise for calls.

Solution to 2:

A is correct. Dividends drive down the stock price when the dividend is paid. Thus, all else being equal, a stock paying dividends has a built-in force that drives down the stock price. This characteristic discourages early exercise, because stock price declines are beneficial to holders of puts.

SUMMARY

This reading on derivative pricing provides a foundation for understanding how derivatives are valued and traded. Key points include the following:

- The price of the underlying asset is equal to the expected future price discounted at the risk-free rate, plus a risk premium, plus the present value of any benefits, minus the present value of any costs associated with holding the asset.
- An arbitrage opportunity occurs when two identical assets or combinations of assets sell at different prices, leading to the possibility of buying the cheaper asset and selling the more expensive asset to produce a risk-free return without investing any capital.
- In well-functioning markets, arbitrage opportunities are quickly exploited, and the resulting increased buying of underpriced assets and increased selling of overpriced assets returns prices to equivalence.
- Derivatives are priced by creating a risk-free combination of the underlying and a derivative, leading to a unique derivative price that eliminates any possibility of arbitrage.
- Derivative pricing through arbitrage precludes any need for determining risk premiums or the risk aversion of the party trading the option and is referred to as risk-neutral pricing.
- The value of a forward contract at expiration is the value of the asset minus the forward price.
- The value of a forward contract prior to expiration is the value of the asset minus the present value of the forward price.
- The forward price, established when the contract is initiated, is the price agreed to by the two parties that produces a zero value at the start.
- Costs incurred and benefits received by holding the underlying affect the forward price by raising and lowering it, respectively.
- Futures prices can differ from forward prices because of the effect of interest rates on the interim cash flows from the daily settlement.
- Swaps can be priced as an implicit series of off-market forward contracts, whereby each contract is priced the same, resulting in some contracts being positively valued and some negatively valued but with their combined value equaling zero.
- At expiration, a European call or put is worth its exercise value, which for calls is the greater of zero or the underlying price minus the exercise price and for puts is the greater of zero and the exercise price minus the underlying price.
- European calls and puts are affected by the value of the underlying, the exercise price, the risk-free rate, the time to expiration, the volatility of the underlying, and any costs incurred or benefits received while holding the underlying.
- Option values experience time value decay, which is the loss in value due to the passage of time and the approach of expiration, plus the moneyness and the volatility.
- The minimum value of a European call is the maximum of zero and the underlying price minus the present value of the exercise price.
- The minimum value of a European put is the maximum of zero and the present value of the exercise price minus the price of the underlying.

- European put and call prices are related through put–call parity, which specifies that the put price plus the price of the underlying equals the call price plus the present value of the exercise price.
- European put and call prices are related through put–call–forward parity, which shows that the put price plus the value of a risk-free bond with face value equal to the forward price equals the call price plus the value of a risk-free bond with face value equal to the exercise price.
- The values of European options can be obtained using the binomial model, which specifies two possible prices of the asset one period later and enables the construction of a risk-free hedge consisting of the option and the underlying.
- American call prices can differ from European call prices only if there are cash flows on the underlying, such as dividends or interest; these cash flows are the only reason for early exercise of a call.
- American put prices can differ from European put prices, because the right to exercise early always has value for a put, which is because of a lower limit on the value of the underlying.

PRACTICE PROBLEMS

1 For a risk-averse investor, the price of a risky asset, assuming no additional costs and benefits of holding the asset, is:

A unrelated to the risk-free rate.

B directly related to its level of risk.

C inversely related to its level of risk.

2 An arbitrage opportunity is *least likely* to be exploited when:

A one position is illiquid.

B the price differential between assets is large.

C the investor can execute a transaction in large volumes.

3 An arbitrageur will *most likely* execute a trade when:

A transaction costs are low.

B costs of short-selling are high.

C prices are consistent with the law of one price.

4 An arbitrage transaction generates a net inflow of funds:

A throughout the holding period.

B at the end of the holding period.

C at the start of the holding period.

5 Which of the following combinations replicates a long derivative position?

A A short derivative and a long asset

B A long asset and a short risk-free bond

C A short derivative and a short risk-free bond

6 Most derivatives are priced by:

A assuming that the market offers arbitrage opportunities.

B discounting the expected payoff of the derivative at the risk-free rate.

C applying a risk premium to the expected payoff of the derivative and its risk.

7 The price of a forward contract:

A is the amount paid at initiation.

B is the amount paid at expiration.

C fluctuates over the term of the contract.

8 Assume an asset pays no dividends or interest, and also assume that the asset does not yield any non-financial benefits or incur any carrying cost. At initiation, the price of a forward contract on that asset is:

A lower than the value of the contract.

B equal to the value of the contract.

C greater than the value of the contract.

9 With respect to a forward contract, as market conditions change:

A only the price fluctuates.

B only the value fluctuates.

C both the price and the value fluctuate.

10 The value of a forward contract at expiration is:

© 2019 CFA Institute. All rights reserved.

- **A** positive to the long party if the spot price is higher than the forward price.
- **B** negative to the short party if the forward price is higher than the spot price.
- **C** positive to the short party if the spot price is higher than the forward price.

11 At the initiation of a forward contract on an asset that neither receives benefits nor incurs carrying costs during the term of the contract, the forward price is equal to the:

- **A** spot price.
- **B** future value of the spot price.
- **C** present value of the spot price.

12 Stocks BWQ and ZER are each currently priced at $100 per share. Over the next year, stock BWQ is expected to generate significant benefits whereas stock ZER is not expected to generate any benefits. There are no carrying costs associated with holding either stock over the next year. Compared with ZER, the one-year forward price of BWQ is *most likely*:

- **A** lower.
- **B** the same.
- **C** higher.

13 If the net cost of carry of an asset is positive, then the price of a forward contract on that asset is *most likely*:

- **A** lower than if the net cost of carry was zero.
- **B** the same as if the net cost of carry was zero.
- **C** higher than if the net cost of carry was zero.

14 If the present value of storage costs exceeds the present value of its convenience yield, then the commodity's forward price is *most likely*:

- **A** less than the spot price compounded at the risk-free rate.
- **B** the same as the spot price compounded at the risk-free rate.
- **C** higher than the spot price compounded at the risk-free rate.

15 Which of the following factors *most likely* explains why the spot price of a commodity in short supply can be greater than its forward price?

- **A** Opportunity cost
- **B** Lack of dividends
- **C** Convenience yield

16 When interest rates are constant, futures prices are *most likely*:

- **A** less than forward prices.
- **B** equal to forward prices.
- **C** greater than forward prices.

17 In contrast to a forward contract, a futures contract:

- **A** trades over-the-counter.
- **B** is initiated at a zero value.
- **C** is marked-to-market daily.

18 To the holder of a long position, it is more desirable to own a forward contract than a futures contract when interest rates and futures prices are:

- **A** negatively correlated.
- **B** uncorrelated.
- **C** positively correlated.

19 The value of a swap typically:

A is non-zero at initiation.

B is obtained through replication.

C does not fluctuate over the life of the contract.

20 The price of a swap typically:

A is zero at initiation.

B fluctuates over the life of the contract.

C is obtained through a process of replication.

21 The value of a swap is equal to the present value of the:

A fixed payments from the swap.

B net cash flow payments from the swap.

C underlying at the end of the contract.

22 If no cash is initially exchanged, a swap is comparable to a series of forward contracts when:

A the swap payments are variable.

B the combined value of all the forward contracts is zero.

C all the forward contracts have the same agreed-on price.

23 For a swap in which a series of fixed payments is exchanged for a series of floating payments, the parties to the transaction:

A designate the value of the underlying at contract initiation.

B value the underlying solely on the basis of its market value at the end of the swap.

C value the underlying sequentially at the time of each payment to determine the floating payment.

24 A European call option and a European put option are written on the same underlying, and both options have the same expiration date and exercise price. At expiration, it is possible that both options will have:

A negative values.

B the same value.

C positive values.

25 At expiration, a European put option will be valuable if the exercise price is:

A less than the underlying price.

B equal to the underlying price.

C greater than the underlying price.

26 The value of a European call option at expiration is the greater of zero or the:

A value of the underlying.

B value of the underlying minus the exercise price.

C exercise price minus the value of the underlying.

27 For a European call option with two months until expiration, if the spot price is below the exercise price, the call option will *most likely* have:

A zero time value.

B positive time value.

C positive exercise value.

28 When the price of the underlying is below the exercise price, a put option is:

A in-the-money.

B at-the-money.

C out-of-the-money.

29 If the risk-free rate increases, the value of an in-the-money European put option will *most likely*:

A decrease.

B remain the same.

C increase.

30 The value of a European call option is inversely related to the:

A exercise price.

B time to expiration.

C volatility of the underlying.

31 The table below shows three European call options on the same underlying:

	Time to Expiration	Exercise Price
Option 1	3 months	\$100
Option 2	6 months	\$100
Option 3	6 months	\$105

The option with the highest value is *most likely*:

A Option 1.

B Option 2.

C Option 3.

32 The value of a European put option can be either directly or inversely related to the:

A exercise price.

B time to expiration.

C volatility of the underlying.

33 Prior to expiration, the lowest value of a European put option is the greater of zero or the:

A exercise price minus the value of the underlying.

B present value of the exercise price minus the value of the underlying.

C value of the underlying minus the present value of the exercise price.

34 A European put option on a dividend-paying stock is *most likely* to increase if there is an increase in:

A carrying costs.

B the risk-free rate.

C dividend payments.

35 Based on put-call parity, a trader who combines a long asset, a long put, and a short call will create a synthetic:

A long bond.

B fiduciary call.

C protective put.

36 Which of the following transactions is the equivalent of a synthetic long call position?

A Long asset, long put, short call

B Long asset, long put, short bond

C Short asset, long call, long bond

37 Which of the following is *least likely* to be required by the binomial option pricing model?

A Spot price

B Two possible prices one period later

C Actual probabilities of the up and down moves

38 To determine the price of an option today, the binomial model requires:

A selling one put and buying one offsetting call.

B buying one unit of the underlying and selling one matching call.

C using the risk-free rate to determine the required number of units of the underlying.

39 Assume a call option's strike price is initially equal to the price of its underlying asset. Based on the binomial model, if the volatility of the underlying decreases, the lower of the two potential payoff values of the hedge portfolio:

A decreases.

B remains the same.

C increases.

40 Based on the binomial model, an increase in the actual probability of an upward move in the underlying will result in the option price:

A decreasing.

B remaining the same.

C increasing.

41 If a call option is priced higher than the binomial model predicts, investors can earn a return in excess of the risk-free rate by:

A investing at the risk-free rate, selling a call, and selling the underlying.

B borrowing at the risk-free rate, buying a call, and buying the underlying.

C borrowing at the risk-free rate, selling a call, and buying the underlying.

42 An at-the-money American call option on a stock that pays no dividends has three months remaining until expiration. The market value of the option will *most likely* be:

A less than its exercise value.

B equal to its exercise value.

C greater than its exercise value.

43 At expiration, American call options are worth:

A less than European call options.

B the same as European call options.

C more than European call options.

44 Which of the following circumstances will *most likely* affect the value of an American call option relative to a European call option?

A Dividends are declared

B Expiration date occurs

C The risk-free rate changes

45 Combining a protective put with a forward contract generates equivalent outcomes at expiration to those of a:

A fiduciary call.

B long call combined with a short asset.

C forward contract combined with a risk-free bond.

46 Holding an asset and buying a put on that asset is equivalent to:

- **A** initiating a fiduciary call.
- **B** buying a risk-free zero-coupon bond and selling a call option.
- **C** selling a risk-free zero-coupon bond and buying a call option.

47 If an underlying asset's price is less than a related option's strike price at expiration, a protective put position on that asset versus a fiduciary call position has a value that is:

- **A** lower.
- **B** the same.
- **C** higher.

48 Based on put–call parity, which of the following combinations results in a synthetic long asset position?

- **A** A long call, a short put, and a long bond
- **B** A short call, a long put, and a short bond
- **C** A long call, a short asset, and a long bond

49 For a holder of a European option, put–call–forward parity is based on the assumption that:

- **A** no arbitrage is possible within the spot, forward, and option markets.
- **B** the value of a European put at expiration is the greater of zero or the underlying value minus the exercise price.
- **C** the value of a European call at expiration is the greater of zero or the exercise price minus the value of the underlying.

50 Under put–call–forward parity, which of the following transactions is risk free?

- **A** Short call, long put, long forward contract, long risk-free bond
- **B** Long call, short put, long forward contract, short risk-free bond
- **C** Long call, long put, short forward contract, short risk-free bond

SOLUTIONS

1 C is correct. An asset's current price, S_0, is determined by discounting the expected future price of the asset by r (the risk free rate) plus λ (the risk premium) over the period from 0 to T, as illustrated in the following equation:

$$S_0 = \frac{E(S_T)}{(1 + r + \lambda)^T}$$

Thus, an asset's current price inversely relates to its level of risk via the related risk premium, λ.

A is incorrect because an asset's current price in spot markets is calculated using the risk-free rate plus a risk premium.

B is incorrect because an asset's current price in spot markets is inversely related, not directly related, to its level of risk.

2 A is correct. An illiquid position is a limit to arbitrage because it may be difficult to realize gains of an illiquid offsetting position. A significant opportunity arises from a sufficiently large price differential or a small price differential that can be employed on a very large scale.

3 A is correct. Some arbitrage opportunities represent such small price discrepancies that they are only worth exploiting if the transaction costs are low. An arbitrage opportunity may require short-selling assets at costs that eliminate any profit potential. If the law of one price holds, there is no arbitrage opportunity.

4 C is correct. Arbitrage is a type of transaction undertaken when two assets or portfolios produce identical results but sell for different prices. A trader buys the asset or portfolio with the lower price and sells the asset or portfolio with the higher price, generating a net inflow of funds at the start of the holding period. Because the two assets or portfolios produce identical results, a long position in one and short position in the other means that at the end of the holding period, the payoffs offset. Therefore, there is no money gained or lost at the end of the holding period, so there is no risk.

5 B is correct. A long asset and a short risk-free asset (meaning to borrow at the risk-free rate) can be combined to produce a long derivative position.

A is incorrect because a short derivative and a long asset combine to produce a position equivalent to a long risk-free bond, not a long derivative.

C is incorrect because a short derivative and a short risk-free bond combine to produce a position equivalent to a short asset, not a long derivative.

6 B is correct. Virtually all derivative pricing models discount the expected payoff of the derivative at the risk-free rate.

A is incorrect because derivatives are priced by assuming that the market is free of arbitrage opportunities via the principle of no arbitrage, not by assuming that the market offers them.

C is incorrect because the application of a risk premium to the expected payoff of the derivative and its risk is not appropriate in the pricing of derivatives. An investor's risk premium is not relevant to pricing a derivative.

7 B is correct. The forward price is agreed upon at the start of the contract and is the fixed price at which the underlying will be purchased (or sold) at expiration. Payment is made at expiration. The value of the forward contract may change over time, but the forward price does not change.

8 C is correct. The price of a forward contract is a contractually fixed price, established at initiation, at which the underlying will be purchased (or sold) at expiration. The value of a forward contract at initiation is zero; therefore, the forward price is greater than the value of the forward contract at initiation.

9 B is correct. The value of the forward contract, unlike its price, will adjust as market conditions change. The forward price is fixed at initiation.

10 A is correct. When a forward contract expires, if the spot price is higher than the forward price, the long party profits from paying the lower forward price for the underlying. Therefore, the forward contract has a positive value to the long party and a negative value to the short party. However, if the forward price is higher than the spot price, the short party profits from receiving the higher forward price (the contract value is positive to the short party and negative to the long party).

11 B is correct. At initiation, the forward price is the future value of the spot price (spot price compounded at the risk-free rate over the life of the contract). If the forward price were set to the spot price or the present value of the spot price, it would be possible for one side to earn an arbitrage profit by selling the asset and investing the proceeds until contract expiration.

12 A is correct. The forward price of each stock is found by compounding the spot price by the risk-free rate for the period and then subtracting the future value of any benefits and adding the future value of any costs. In the absence of any benefits or costs, the one-year forward prices of BWQ and ZER should be equal. After subtracting the benefits related to BWQ, the one-year forward price of BWQ is lower than the one-year forward price of ZER.

13 A is correct. An asset's forward price is increased by the future value of any costs and decreased by the future value of any benefits: $F_0(T) = S_0(1 + r)^T - (\gamma - \theta)(1 + r)^T$. If the net cost of carry (benefits less costs) is positive, the forward price is lower than if the net cost of carry was zero.

14 C is correct. When a commodity's storage costs exceed its convenience yield benefits, the net cost of carry (benefits less costs) is negative. Subtracting this negative amount from the spot price compounded at the risk-free rate results in an addition to the compounded spot price. The result is a commodity forward price which is higher than the spot price compounded. The commodity's forward price is less than the spot price compounded when the convenience yield benefits exceed the storage costs and the commodity's forward price is the same as the spot price compounded when the costs equal the benefits.

15 C is correct. The convenience yield is a benefit of holding the asset and generally exists when a commodity is in short supply. The future value of the convenience yield is subtracted from the compounded spot price and reduces the commodity's forward price relative to it spot price. The opportunity cost is the risk-free rate. In the absence of carry costs, the forward price is the spot price compounded at the risk-free rate and will exceed the spot price. Dividends are benefits that reduce the forward price but the lack of dividends has no effect on the spot price relative to the forward price of a commodity in short supply.

16 B is correct. When interest rates are constant, forwards and futures will likely have the same prices. The price differential will vary with the volatility of interest rates. In addition, if futures prices and interest rates are uncorrelated, forward and futures prices will be the same. If futures prices are positively correlated with interest rates, futures contracts are more desirable to holders of long positions than are forwards. This is because rising prices lead to future profits that are reinvested in periods of rising interest rates, and falling prices lead to losses that occur in periods of falling interest rates. If futures prices are negatively correlated with interest rates, futures contracts are less desirable to holders of long positions than are forwards. The more desirable contract will tend to have the higher price.

17 C is correct. Futures contracts are marked-to-market on a daily basis. The accumulated gains and losses from the previous day's trading session are deducted from the accounts of those holding losing positions and transferred to the accounts of those holding winning positions. Futures contracts trade on an exchange, forward contracts are over-the-counter transactions. Typically both forward and futures contracts are initiated at a zero value.

18 A is correct. If futures prices and interest rates are negatively correlated, forwards are more desirable to holders of long positions than are futures. This is because rising prices lead to futures profits that are reinvested in periods of falling interest rates. It is better to receive all of the cash at expiration under such conditions. If futures prices and interest rates are uncorrelated, forward and futures prices will be the same. If futures prices are positively correlated with interest rates, futures contracts are more desirable to holders of long positions than are forwards.

19 B is correct. Valuation of the swap during its life appeals to replication and the principle of arbitrage. Valuation consists of reproducing the remaining payments on the swap with other transactions. The value of that replication strategy is the value of the swap. The swap price is typically set such that the swap contract has a value of zero at initiation. The value of a swap contract will change during the life of the contract as the value of the underlying changes in value.

20 C is correct. Replication is the key to pricing a swap. The swap price is determined at initiation by replication. The value (not the price) of the swap is typically zero at initiation and the fixed swap price is typically determined such that the value of the swap will be zero at initiation.

21 B is correct. The principal of replication articulates that the valuation of a swap is the present value of all the net cash flow payments from the swap, not simply the present value of the fixed payments of the swap or the present value of the underlying at the end of the contract.

22 B is correct. When two parties engage in a series of forward contracts and initially agree on a price of $FS_0(T)$, some of the forward contracts have positive values and some have negative values, but their combined value equals zero.

A is incorrect because for a swap, all payments are fixed and equal, not variable.

C is incorrect because forward prices are determined by the spot price and the net cost of carry, meaning that forward contracts expiring at different times will have different prices, not the same price.

23 C is correct. On each payment date, the swap owner receives a payment based on the value of the underlying at the time of each respective payment.

A is incorrect because in a swap involving a series of fixed payments exchanged for a series of floating payments, each floating payment reflects the value of the underlying at the time of payment, not a designated value at contract initiation.

B is incorrect because in a swap involving a series of fixed payments exchanged for a series of floating payments, each floating payment is based on the value of the underlying at the time of each respective payment, not on the market value at the end of the swap.

24 B is correct. If the underlying has a value equal to the exercise price at expiration, both options will have zero value since they both have the same exercise price. For example, if the exercise price is $25 and at expiration the underlying price is $25, both the call option and the put option will have a value of zero. The value of an option cannot fall below zero. The holder of an option is not obligated to exercise the option; therefore, the options each have a minimum value of zero. If the call has a positive value, the put, by definition, must have a zero value and vice versa. Both cannot have a positive value.

25 C is correct. A European put option will be valuable at expiration if the exercise price is greater than the underlying price. The holder can put (deliver) the underlying and receive the exercise price which is higher than the spot price. A European put option would be worthless if the exercise price was equal to or less than the underlying price.

26 B is correct. The value of a European call option at expiration is the greater of zero or the value of the underlying minus the exercise price.

27 B is correct. A European call option with two months until expiration will typically have positive time value, where time value reflects the value of the uncertainty that arises from the volatility in the underlying. The call option has a zero exercise value if the spot price is below the exercise price. The exercise value of a European call option is $\text{Max}(0,S_t - X)$, where S_t is the current spot price at time t and X is the exercise price.

28 A is correct. When the price of the underlying is below the exercise price for a put, the option is said to be in-the-money. If the price of the underlying is the same as the exercise price, the put is at-the-money and if it is above the exercise price, the put is out-of-the-money.

29 A is correct. An in-the-money European put option decreases in value with an increase in the risk-free rate. A higher risk-free rate reduces the present value of any proceeds received on exercise.

30 A is correct. The value of a European call option is inversely related to the exercise price. A lower exercise price means there are more potential outcomes at which the call expires in-the-money. The option value will be greater the lower the exercise price. For a higher exercise price, the opposite is true. Both the time to expiration and the volatility of the underlying are directly (positively) related to the value of a European call option.

31 B is correct. The value of a European call option is inversely related to the exercise price and directly related to the time to expiration. Option 1 and Option 2 have the same exercise price; however, Option 2 has a longer time to expiration. Consequently, Option 2 would likely have a higher value than Option 1. Option 2 and Option 3 have the same time to expiration; however, Option 2 has a lower exercise price. Thus, Option 2 would likely have a higher value than Option 3.

32 B is correct. The value of a European put option can be either directly or indirectly related to time to expiration. The direct effect is more common, but the inverse effect can prevail the longer the time to expiration, the higher the risk-free rate, and the deeper in-the-money is the put. The value of a European put option is directly related to the exercise price and the volatility of the underlying.

33 B is correct. Prior to expiration, the lowest value of a European put is the greater of zero or the present value of the exercise price minus the value of the underlying.

34 C is correct. Payments, such as dividends, reduce the value of the underlying which increases the value of a European put option. Carrying costs reduce the value of a European put option. An increase in the risk-free interest rate may decrease the value of a European put option.

35 A is correct. A long bond can be synthetically created by combining a long asset, a long put, and a short call. A fiduciary call is created by combining a long call with a risk free bond. A protective put is created by combining a long asset with a long put.

36 B is correct. According to put–call parity, a synthetic call can be constructed by combining a long asset, long put, and short bond positions.

37 C is correct. The actual probabilities of the up and down moves in the underlying do not appear in the binomial option pricing model, only the pseudo or "risk-neutral" probabilities. Both the spot price of the underlying and two possible prices one period later are required by the binomial option pricing model.

38 C is correct. Pricing an option relies on the facts that a perfectly hedged investment earns the risk-free rate and that, based on the binomial option pricing model, the size of the two possible changes in the option price (meaning the potential step up or step down in the option value) after one period are equivalent.

39 B is correct. When the volatility of the underlying decreases, the value of the option also decreases, meaning that the upper payoff value of the hedge portfolio combining them declines. However, the lower payoff value remains at zero.

40 B is correct. The binomial model does not consider the actual probabilities of upward and downward movements in determining the option value. Thus, a change in this probability has no effect on the calculated option price.

41 C is correct. If an option is trading above the value predicted by the binomial model, investors can engage in arbitrage by selling a call, buying shares of the underlying, and funding the transaction by borrowing at the risk-free rate. This will earn a return in excess of the risk-free rate.

42 C is correct. Prior to expiration, an American call option will typically have a value in the market that is greater than its exercise value. Although the American option is at-the-money and therefore has an exercise value of zero, the time value of the call option would likely lead to the option having a positive market value.

43 B is correct. At expiration, the values of American and European call options are effectively the same; both are worth the greater of zero and the exercise value.

44 A is correct. When a dividend is declared, an American call option will have a higher value than a European call option because an American call option holder can exercise early to capture the value of the dividend. At expiration, both types of call options are worth the greater of zero and the exercise value. A change in the risk-free rate does not affect the relative values of American and European call options.

45 A is correct. Put–call forward parity demonstrates that the outcome of a protective put with a forward contract (long put, long risk-free bond, long forward contract) equals the outcome of a fiduciary call (long call, long risk-free bond). The outcome of a protective put with a forward contract is also equal to the outcome of a protective put with asset (long put, long asset).

46 A is correct. Under put–call parity, initiating a fiduciary call (buying a call option on an asset that expires at time T together with a risk-free zero-coupon bond that also expires at time T) is equivalent to holding the same asset and initiating a protective put on it (buying a put option with an exercise price of X that can be used to sell the asset for X at time T).

47 B is correct. On the one hand, buying a call option on an asset and a risk-free bond with the same maturity is known as a fiduciary call. If the fiduciary call expires in the money (meaning that the value of the call, $S_T - X$, is greater than the risk-free bond's price at expiration, X), then the total value of the fiduciary call is $(S_T - X) + X$, or S_T. On the other hand, holding an underlying asset, S_T, and buying a put on that asset is known as a protective put. If the put expires out of the money, meaning that the value of the asset, S_T, is greater than the put's value at expiration, 0, then the total value of the protective put is $S_T - 0$, or S_T. A protective put and a fiduciary call produce the same result.

48 A is correct. One can synthetically create a long asset position by buying a call, shorting a put, and buying a bond.

B is incorrect because combining a short call and a short bond with the right to sell (not buy) another asset via a long put could not result in a new synthetic long asset position.

C is incorrect because combining a long call, a short asset, and a long bond creates a long put, not a synthetic long asset.

49 A is correct. Put–call–forward parity is based on the assumption that no arbitrage is possible within the spot, forward, and option markets.

B is incorrect because the value of a European put at expiration is the greater of either zero or the exercise price minus the value of the underlying, not the greater of zero or the underlying value minus the exercise price. In addition, put–call–forward parity is related to the equality of a fiduciary call and a synthetic protective put or to a protective put and a synthetic fiduciary call, not specifically to the value of a put at expiration.

C is incorrect because the value of a European call at expiration is the greater of either zero or the underlying value minus the exercise price, not the greater of zero or the exercise price minus the value of the underlying. In addition, put–call–forward parity is related to the equality of a fiduciary call and a synthetic protective put or to a protective put and a synthetic fiduciary call, not specifically to the value of a call at expiration.

50 A is correct. Purchasing a long forward contract and a risk-free bond creates a synthetic asset. Combining a long synthetic asset, a long put, and a short call is risk free because its payoffs produce a known cash flow of the value of the exercise price.

Alternative Investments

STUDY SESSION

Study Session 16	Alternative Investments

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to demonstrate a working knowledge of alternative investments, including hedge funds, private capital, real estate, natural resources, and infrastructure. The candidate should be able to describe key attributes and considerations in adding these investments to a portfolio.

Investors often turn to alternative investments for potential diversification benefits and higher returns. As a result, alternative investments now represent meaningful allocations in many institutional and private wealth portfolios. Although the category of "alternative investments" is not always clearly or precisely defined, alternative investments often have a number of characteristics in common. These include lower levels of liquidity, transparency, and disclosure vs. traditional asset classes (equity, fixed income), more complex legal structures, and performance-based compensation arrangements.

© 2021 CFA Institute. All rights reserved.

ALTERNATIVE INVESTMENTS STUDY SESSION

16

Alternative Investments

This study session provides an overview of alternative investments, including hedge funds, private capital, real estate, natural resources, and infrastructure. Each is examined with emphasis on their distinguishing characteristics, potential benefits, and risks. Similarities and differences with traditional investments (stocks, bonds) are also considered.

READING ASSIGNMENTS

Reading 47	Introduction to Alternative Investments by Steve Balaban, CFA, Steven G. Bloom, CFA, David Burkart, CFA, Nasir Hasan, and Barclay T. Leib, CFE, CAIA

© 2021 CFA Institute. All rights reserved.

READING

47

Introduction to Alternative Investments

by Steve Balaban, CFA, Steven G. Bloom, CFA, David Burkart, CFA, Nasir Hasan, and Barclay T. Leib, CFE, CAIA

Steve Balaban, CFA, is at Mink Capital Inc. (Canada). Steven G. Bloom, CFA, is at ARC Fiduciary (USA). David Burkart, CFA, is at Coloma Capital Futures, LLC (USA). Nasir Hasan is at Ernst & Young (UAE). Barclay T. Leib, CFE, CAIA, is at and Spring Advisors LLC (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe types and categories of alternative investments;
☐	**b.** describe characteristics of direct investment, co-investment, and fund investment methods for alternative investments;
☐	**c.** describe investment and compensation structures commonly used in alternative investments;
☐	**d.** explain investment characteristics of hedge funds;
☐	**e.** explain investment characteristics of private capital;
☐	**f.** explain investment characteristics of natural resources;
☐	**g.** explain investment characteristics of real estate;
☐	**h.** explain investment characteristics of infrastructure;
☐	**i.** describe issues in performance appraisal of alternative investments;
☐	**j.** calculate and interpret returns of alternative investments on both before-fee and after-fee bases.

INTRODUCTION

1

a describe types and categories of alternative investments

In this section, we explain what alternative investments are and why assets under management in alternative investments have grown in recent decades. We also explain how alternative investments differ from traditional investments, and we examine

© 2021 CFA Institute. All rights reserved.

their perceived investment merit. We conclude this section with a brief overview of the various categories of alternative investments; these categories will be explored further in later sections.

"Alternative investments" is a label for a disparate group of investments that are distinguished from long-only, publicly traded investments in stocks, bonds, and cash (often referred to as traditional investments). The terms "traditional" and "alternative" should not imply that alternatives are necessarily uncommon or that they are relatively recent additions to the investment universe. Alternative investments include such assets as real estate and commodities, which are arguably two of the oldest types of investments.

Alternative investments also include non-traditional approaches to investing within special vehicles, such as private equity funds and hedge funds. These funds may give the manager flexibility to use derivatives and leverage, to make investments in illiquid assets, and to take short positions. The assets in which these vehicles invest can include traditional assets (stocks, bonds, and cash) as well as less traditional assets. Management of alternative investments is typically active. Alternative investments often have many of the following characteristics:

- Narrow specialization of the investment managers
- Relatively low correlation of returns with those of traditional investments
- Less regulation and less transparency than traditional investments
- Limited historical risk and return data
- Unique legal and tax considerations
- Higher fees, often including performance or incentive fees
- Concentrated portfolios
- Restrictions on redemptions (i.e., "lockups" and "gates")

1.1 Why Investors Consider Alternative Investments

Assets under management in alternative investments have grown rapidly since the mid-1990s. This growth has largely occurred because of interest in these investments from institutions, such as endowments and pension funds, and from family offices seeking diversification and return opportunities. Alternative investments offer broader diversification (because of their lower correlation with traditional asset classes), opportunities for enhanced returns (by increasing the portfolio's risk–return profile), and potentially increased income through higher yields (particularly compared with traditional investments in low–interest rate periods).

The 2019 annual report for the Yale University endowment provides one institutional investor's reasoning behind the attractiveness of investing in alternatives:

> The heavy [75.2%] allocation to nontraditional asset classes stems from the diversifying power they provide to the portfolio as a whole. Alternative assets, by their very nature, tend to be less efficiently priced than traditional marketable securities, providing an opportunity to exploit market inefficiencies through active management. Today's portfolio has significantly higher expected returns and lower volatility than the 1989 portfolio.[1]

This quote neatly illustrates the expected characteristics of alternative investments: diversifying power (low correlations between returns), higher expected returns (positive absolute return), and illiquid and potentially less efficient markets. These links also highlight the importance of having the ability and willingness to take a long-term perspective. Allocating a portion of an endowment portfolio to alternative investments is not unique to Yale. INSEAD, as of April 2020, had allocated 33% of its endowment to private market strategies, including private equity, private debt, and real estate.[2] These examples are not meant to imply that every university endowment fund invests in alternative investments, but many do.

---

[1] Yale University, "Financial Report 2018–2019 Yale University" (2019, p. 18). https://your.yale.edu/sites/default/files/annual-report-2018-2019.pdf (accessed 27 April 2020).

Family offices have also embraced alternative investments. According to the Global Family Office Report 2019 from UBS, over 40% of the average family office portfolio was invested in alternative assets.[3] According to Willis Towers Watson's "Global Pension Assets Study 2020," the typical pension plan had a 23% allocation to alternative investments at the end of 2019.[4]

Alternative investments are not free of risk, of course, and their returns may be correlated with those of other investments, especially in periods of financial crisis. Over a long historical period, the average correlation of returns from alternative investments with those of traditional investments may be low, but in any particular period, the correlation can differ from the average. During periods of economic crisis, correlations among many assets (both alternative and traditional) can increase dramatically.

KNOWLEDGE CHECK

True or false: Alternative investments focus solely on the private markets.

A True

B False

Solution:

B is correct. Although many alternative investments are focused on private markets, there are alternative investments, such as hedge funds, that focus on the public markets.

1.2 Categories of Alternative Investments

Considering the variety of alternative investments, it is not surprising that no consensus exists on a definitive list. There is even considerable debate around categories versus sub-categories. The following list offers one approach to broad category definitions, and each category is described in detail later in this reading.

- **Hedge funds.** Hedge funds are private investment vehicles that manage portfolios of securities and/or derivative positions using a variety of strategies. Although hedge funds may be invested entirely in traditional assets, these vehicles are considered alternative because of their private nature. Hedge funds typically have more leeway to pursue investments and strategies offering the potential for higher returns, whether absolute or compared with a specific market benchmark, but these strategies may increase the risk of investment loss. They may involve long and short positions and may be highly leveraged. Some aim to deliver investment performance that is independent of broader market performance.

---

[2] INSEAD, "Finances and Endowment" (sec. 3), in "INSEAD Annual Report 2018/2019" (2019). https://annual-report.insead.edu/finances-and-endowment (accessed 27 April 2020).

[3] Campden Wealth Limited and UBS, "The Global Family Office Report 2019" (2019, p. 17). www.ubs.com/global/en/wealth-management/uhnw/global-family-office-report/global-family-office-report-2019.html.

[4] Thinking Ahead Institute, "Global Pension Assets Study 2020," Willis Towers Watson (2020).

- **Private Capital**
    - **Private equity.** Investors participate in private equity through direct investments or indirectly through **private equity funds**. Private equity funds generally invest in companies, whether startups or established firms, that are not listed on a public exchange, or they invest in public companies with the intent to take them private. The majority of private equity activity involves leveraged buyouts of established profitable and cash-generating companies with solid customer bases, proven products, and high-quality management. **Venture capital funds**, a specialized form of private equity that typically involves investing in or providing financing to startup or early-stage companies with high growth potential, represent a small portion of the private equity market.
    - **Private debt.** Private debt largely encompasses debt provided to private entities. Forms of private debt include direct lending (private loans with no intermediary), mezzanine loans (private subordinated debt), venture debt (private loans to startup or early-stage companies that may have little or negative cash flow), and distressed debt (debt extended to companies that are "distressed" because of such issues as bankruptcy or other complications with meeting debt obligations).

KNOWLEDGE CHECK

Describe what the majority of private equity activity includes.

Solution:

The majority of private equity activity involves leveraged buyouts of established profitable and cash-generating companies with solid customer bases, proven products, and high-quality management.

- **Real estate.** Real estate investments are made in buildings or land, either directly or indirectly. The growing popularity of securitizations broadened the definition of real estate investing. It now includes private commercial real estate equity (e.g., ownership of an office building), private commercial real estate debt (e.g., directly issued loans or mortgages on commercial property), public real estate equity (e.g., real estate investment trusts, or REITs), and public real estate debt (e.g., mortgage-backed securities).
- **Natural Resources**
    - **Commodities.** Commodity investments may take place in physical commodity products, such as grains, metals, and crude oil, either through owning cash instruments, using derivative products, or investing in businesses engaged in the production of physical commodities. The main vehicles investors use to gain exposure to commodities are commodity futures contracts and funds benchmarked to commodity indexes. Commodity indexes are typically based on various underlying commodity futures.
    - **Agricultural land (or farmland).** Agricultural land is for the cultivation of livestock or plants, and agricultural land investing covers various strategies, including the purchase of farmland in order to lease it back to farmers or receive a stream of income from the growth, harvest, and sale of crops (e.g., corn, cotton, wheat) or livestock (e.g., cattle).

  - **Timberland.** Investing in timberland generally involves investing capital in natural forests or managed tree plantations in order to earn a return when the trees are harvested. Timberland investors often rely on various drivers, such as biological growth, to increase the value of the trees so the wood can be sold at favorable prices in the future.

- **Infrastructure.** Infrastructure assets are capital-intensive, long-lived real assets, such as roads, dams, and schools, that are intended for public use and provide essential services. Infrastructure assets may be financed, owned, and operated by governments, but private sector investment is on the rise. An increasingly common approach to infrastructure investing is a **public–private partnership (PPP)** approach, in which governments and investors each have a stake. Infrastructure investments provide exposure to asset cash flows, but the asset itself is generally part of a long-term concession agreement, ultimately going back to the public authority. Investors may gain exposure to these assets directly or indirectly. Indirect investment vehicles include shares of companies, exchange-traded funds (ETFs), private equity funds, listed funds, and unlisted funds that invest in infrastructure.
- **Other.** Other alternative investments may include tangible assets, such as fine wine, art, antique furniture and automobiles, stamps, coins, and other collectibles, and intangible assets, such as patents and litigation actions.

PRACTICE QUESTIONS FOR LOS A

Question 1

Which of the following is *least likely* to be considered an alternative investment?

A Real estate

B Commodities

C Long-only equity funds

Solution:

C is correct. Long-only equity funds are typically considered traditional investments, and real estate and commodities are typically classified as alternative investments.

Question 2

An investor is seeking an investment that can take long and short positions, may use multi-strategies, and historically exhibits low correlation with a traditional investment portfolio. The investor's goals will be *best* satisfied with an investment in:

A real estate.

B a hedge fund.

C a private equity fund.

Solution:

B is correct. Hedge funds may use a variety of strategies, generally have a low correlation with traditional investments, and may take long and short positions.

Question 3

Relative to traditional investments, alternative investments are *least likely* to be characterized by:

A high levels of transparency.
B limited historical return data.
C significant restrictions on redemptions.

Solution:

A is correct. Alternative investments are characterized as typically having low levels of transparency.

Question 4

Alternative investment funds are typically managed:

A actively.
B to generate positive beta return.
C assuming that markets are efficient.

Solution:

A is correct. There are many approaches to managing alternative investment funds, but typically these funds are actively managed.

Question 5

Compared with traditional investments, alternative investments are *more likely* to have:

A greater use of leverage.
B long-only positions in liquid assets.
C more transparent and reliable risk and return data.

Solution:

A is correct. Investing in alternative investments is often pursued through such special vehicles as hedge funds and private equity funds, which have flexibility to use leverage. Alternative investments include investments in such assets as real estate, which is an illiquid asset, and investments in such special vehicles as private equity and hedge funds, which may make investments in illiquid assets and take short positions. Obtaining information on strategies used and identifying reliable measures of risk and return are challenges of investing in alternatives.

Question 6

The potential benefits of allocating a portion of a portfolio to alternative investments include:

A ease of manager selection.
B improvement in the portfolio's risk–return relationship.
C accessible and reliable measures of risk and return.

Solution:

B is correct. Adding alternative investments to a portfolio may provide diversification benefits because of these investments' less-than-perfect correlation with other assets in the portfolio. As a result, allocating a portion of one's funds to alternatives could potentially result in an improved risk–return relationship.

Challenges to allocating a portion of a portfolio to alternative investments include obtaining reliable measures of risk and return and selecting portfolio managers for the alternative investments.

INVESTMENT METHODS

2

b describe characteristics of direct investment, co-investment, and fund investment methods for alternative investments

This section introduces three methods of investing in alternative investments: direct investing, co-investing, and fund investing. Their advantages and disadvantages are examined, and a brief discussion of due diligence concludes the section.

2.1 Methods of Investing in Alternative Investments

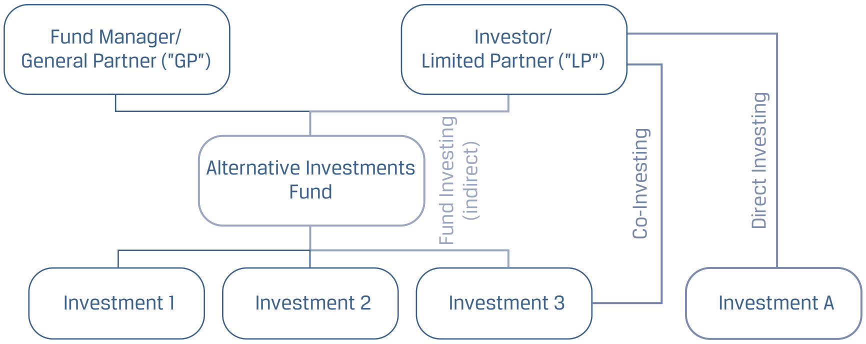
In **fund investing**, the investor contributes capital to a fund, and the fund identifies, selects, and makes investments on the investor's behalf. For the fund's services, the investor is charged a fee based on the amount of the assets being managed, and a performance fee is applied if the fund manager delivers superior results. Fund investing can be viewed as an indirect method of investing in alternative assets. Fund investors have little or no leeway in the sense that their investment decisions are limited to either investing in the fund or not. Furthermore, fund investors are typically unable to affect the fund's underlying investments. Note that fund investing is available for all major alternative investment types, including hedge funds, private capital, real estate, infrastructure, and natural resources.

In **co-investing**, the investor invests in assets *indirectly* through the fund but also possesses rights (known as co-investment rights) to invest *directly* in the same assets. Through co-investing, an investor is able to make an investment *alongside* a fund when the fund identifies deals; the investor is not limited to participating in the deal solely by investing in the fund. Exhibit 1 illustrates the co-investing method: The investor invests in one deal (labeled "Investment #3") indirectly via fund investing while investing an additional amount directly via a co-investment.

Direct investing occurs when an investor makes a direct investment in an asset (labeled "Investment A") without the use of an intermediary. In private equity, this may mean the investor purchases a direct stake in a private company without the use of a special vehicle, such as a fund. Direct investors have great flexibility and control when it comes to choosing their investments, selecting their preferred methods of financing, and planning their approach. The direct method of investing in alternative assets is typically reserved for larger and more sophisticated investors and usually applies to private equity and real estate. Sizable investors, such as major pensions and sovereign wealth funds, however, may also invest directly in infrastructure and natural resources.

Institutional investors typically begin investing in alternative investments via funds. Then, as they gain experience, they can begin to invest via co-investing and direct investing. The largest and most sophisticated direct investors (such as some sovereign wealth funds) compete with fund managers for access to the best investment opportunities.

Exhibit 1 Diagram of Co-Investing in Alternative Assets



KNOWLEDGE CHECK

True or false: In co-investing, the investor is able to invest both directly and indirectly in the same assets.

A True

B False

Solution:

A is correct. In co-investing, the investor is able to invest both directly and indirectly in the same assets.

2.2 Advantages and Disadvantages of Direct Investing, Co-Investing, and Fund Investing

The direct investing, co-investing, and fund investing approaches naturally have distinct advantages and disadvantages.

2.2.1 *Advantages of Fund Investing*

The primary advantages of fund investing include the professional services offered by fund managers, a lower level of investor involvement (compared with the direct and co-investing methods), and access to alternative investments without the prerequisite of advanced expertise. Additionally, diversification benefits stem from the multiple investments found in a single investment vehicle overseen by a fund manager. (Consider instead the time and attention that would be required of a single investor to achieve similar diversification levels by directly investing in and managing multiple alternative assets.) Fund investing in alternative assets demands less participation from the investor than the direct and co-investing approaches, because it is up to the fund manager to identify, select, and manage the fund's investments. The investor may also heavily rely on the fund manager to carry out the due diligence that accompanies the investment process. Finally, specialist fund managers may expand the investment universe for an investor who lacks expertise for the sector in question.

2.2.2 *Disadvantages of Fund Investing*

Fund investing in alternative investments is costly because the investor is required to pay management fees and performance fees that are typically higher than fees for traditional asset classes. The higher fees can be attributed to the increased costs of

(1) running an alternative fund that may have a more complex strategy or require more skills and resources to source, (2) conducting due diligence, and (3) managing opportunities that have limited publicly available information. Alternative managers may also command higher fees if they face little competition. Despite the fact that a fund may provide due diligence expertise when it comes to choosing investments, an investor is still required to conduct thorough due diligence when selecting the right fund in the first place. Selecting the right fund itself is not an easy task because of asymmetry of information.

2.2.3 *Advantages of Co-Investing*

Co-investors, who are essentially engaging in a hybrid of direct investing and fund investing, can learn from the fund's process and leverage the experience they gain to become better at direct investing. Many institutional investors begin with fund investing, because it provides instant diversification while requiring less costly minimum participation levels (lower minimums) and less due diligence. As investors' investment experience and investable assets grow, they may be granted "co-investment rights" with an investment in a fund, giving them a taste of direct investing. It works as follows: Alongside the fund's direct investment, investors co-invest an additional amount into that same investment often without paying management fees on the capital they used for the direct investment (a co-investment, in this case). Compared with fund investing, co-investing allows investors to be more actively involved with the management of their portfolio. Later, as their experience and capital increase, investors may prefer direct investing, bringing the expertise in-house and avoiding fund management fees.

2.2.4 *Disadvantages of Co-Investing*

Co-investors have reduced control over the investment selection process (compared with direct investing) and may be subject to adverse selection bias, where the fund manager makes less attractive investment opportunities available to the co-investor while allocating its own capital to more appealing deals. Co-investing, furthermore, requires more active involvement from the investor, who must evaluate both investment opportunities and the fund manager, which demands more resources, concentration, and expertise. Moreover, co-investors usually have a limited amount of time to make the decision to invest or not. The co-investing approach can prove to be challenging for smaller firms with limited resources and due diligence experience.

2.2.5 *Advantages of Direct Investing*

When an investor chooses direct investing, she avoids paying ongoing management fees to an external manager because the investor conducts the investment process on her own, bypassing the use of a special vehicle, such as the fund. Direct investing allows the investor to build a portfolio of investments to her exact requirements. Direct investing provides the greatest amount of flexibility for the investor and grants the highest level of control over how the asset is managed. For example, an investor who directly purchases an ownership stake in a business typically has the ability to influence important matters, such as selecting the management team and controlling the strategic direction and investment decisions of the company.

2.2.6 *Disadvantages of Direct Investing*

Compared with fund investing and co-investing, direct investing demands more investment expertise and a higher level of financial sophistication. Such experience would otherwise be provided by a fund manager or investment firm during the complex investment and management process. Furthermore, concentration increases risk: The direct investor won't enjoy the ready diversification benefits of fund investing; it would require time and resources for an investor to mirror this advantage through direct investing. Plus, fund managers may enjoy reputational benefits that see them secure participation in attractive investments unavailable to certain direct investors operating on their own behalf.

Exhibit 2 provides a summary of the advantages and disadvantages of the three methods of investing in alternative investments.

Exhibit 2 Summary of Advantages and Disadvantages of Fund Investing, Co-Investing, and Direct Investing

	Advantages	Disadvantages
Fund investing	■ Fund managers offer investment services and expertise ■ Lower level of investor involvement compared with the direct and co-investing methods ■ Access to alternative investments without possessing a high degree of investment expertise ■ Potentially valuable diversification benefits ■ Lower minimum capital requirements	■ Costly management and performance fees ■ Investor must conduct thorough due diligence when selecting the right fund because of the wide dispersion of fund manager returns
Co-investing	■ Investors can learn from the fund's process to become better at direct investing ■ Reduced management fees ■ Allows more active management of the portfolio compared with fund investing and allows for a deeper relationship with the manager	■ Reduced control over the investment selection process compared with direct investing ■ May be subject to adverse selection bias ■ Requires more active involvement compared with fund investing, which can be challenging if resources and due diligence experience are limited
Direct investing	■ Avoids paying ongoing management fees to an external manager ■ Greatest amount of flexibility for the investor ■ Highest level of control over how the asset is managed	■ Requires more investment expertise and a higher level of financial sophistication compared with fund investing and co-investing, resulting in higher internal investment costs ■ Less access to a fund's ready diversification benefits or the fund manager's sourcing network ■ Requires greater levels of due diligence because of the absence of a fund manager ■ Higher minimum capital requirements

KNOWLEDGE CHECK

Identify the advantages of direct investing for an investor.

Solution:

An advantage of direct investing is that the investor can avoid paying ongoing management fees. In addition, a direct investor has great flexibility and control when it comes to choosing his investments, selecting his preferred methods of financing, and planning his approach.

KNOWLEDGE CHECK

In direct investing, an investor puts capital in an asset or business:

A using a special purpose vehicle, such as a fund.

B using a separate business entity, such as a joint venture.

C without the use of an intermediary.

Solution:

C is correct. Direct investing occurs when an investor makes a direct investment in an asset without the use of an intermediary.

2.3 Due Diligence for Fund Investing, Direct Investing, and Co-Investing

One basic question an investor must consider when investing in alternative investments is whether to rely on the expertise of the fund manager or to undertake the investment selection process themselves via direct investing.

For direct investing, the investor has control over the choice of which company to invest in; however, the due diligence process of making an investment in a private company requires considerable expertise. In contrast, funds offer diversified portfolios, due diligence expertise, and the negotiation of favorable redemption terms, but these benefits come with additional fees. And although funds offer due diligence on the underlying investments, responsibility still falls to the investor for conducting due diligence when choosing among funds to invest in.

With co-investing, the investor is currently investing in a fund but is given an opportunity by the fund general partner (co-investment rights) to make an additional investment in a portfolio company (as shown in Exhibit 1). In this case, the investor will conduct direct due diligence on the portfolio company with the support of the general partner. A detailed discussion on due diligence is beyond the scope of this reading (especially for direct investing). However, this section provides a brief overview of the due diligence process.

2.3.1 *Due Diligence for Fund Investing*

Manager selection is a critical factor in portfolio performance. A manager should have a verifiable track record and display a high level of expertise and experience with the asset type. The asset management team should be assigned an appropriate workload and provided sufficient resources. Moreover, it should be rewarded with an effective compensation package to ensure alignment of interest, continuity, motivation, and thoughtful oversight of assets.

Fraud, although infrequent, is always a possibility. The investor should be skeptical of unusually good and overly consistent reported performance. Third-party custody of assets and independent verification of results can help reduce the chance of an investor being defrauded. Diversification across managers is also wise. Finally, separate accounts make theft more difficult because the investor retains custody of the assets and sometimes can select the prime broker or other service providers, binding them to the client's interest.

For an investor considering a new investment, a proper due diligence process should be carried out to ensure that the targeted investment is in compliance with its prospectus and that it will meet her investment strategy, risk and return objectives, and restrictions. Existing investors should monitor results and fund holdings to determine whether a fund has performed in line with expectations and continues to comply with its prospectus.

Exhibit 3 lists key items that should be considered in a typical fund due diligence process.

Exhibit 3 A Typical Due Diligence Process

Organization	■ Experience and quality of management team, compensation, and staffing ■ Analysis of prior and current funds ■ Track record/alignment of interests ■ Reputation and quality of third-party service providers (e.g., lawyers, auditors, prime brokers)
Portfolio management	■ Investment process ■ Target markets/asset types/strategies ■ Sourcing of investments ■ Role of operating partners ■ Underwriting ■ Environmental and engineering review process ■ Integration of asset management/acquisitions/dispositions ■ Disposition process, including its initiation and execution
Operations and controls	■ Reporting and accounting methodology ■ Audited financial statements and other internal controls ■ Valuations—frequency and approach(es) ■ Insurance and contingency plans
Risk management	■ Fund policies and limits ■ Risk management policy ■ Portfolio risk and key risk factors ■ Leverage and currency—risks/constraints/hedging
Legal review	■ Fund structure ■ Registrations ■ Existing/prior litigation
Fund terms	■ Fees (management and performance) and expenses ■ Contractual terms ■ Investment period and fund term and extensions ■ Carried interest ■ Distributions ■ Conflicts ■ Limited partners' rights ■ "Key person" and/or other termination procedures

2.3.2 *Due Diligence for Direct Investing*

Due diligence for direct investing requires the investor to conduct a thorough investigation into the target asset or business, including but not limited to the quality of its management team, the quality of its customers, the competitive landscape, revenue generation, risks, and so on. When considering direct investments in private equity, the investor conducts reference checks and interviews to evaluate the quality of interactions between the business and each of its stakeholders (e.g., customers and suppliers). Due diligence for private debt investing could entail credit analysis of borrowers to assess their ability to service the regular interest and principal payments of debt, and background checks on the owners and management team would round out the effort. In direct real estate, the building's occupancy rate and the quality of its structure and tenants should be analyzed prior to investing. In infrastructure, an investor would perform an assessment of the quality of the assets held and operated (e.g., an airport) and their ability to generate future cash flows. In conducting due diligence, direct investors often supplement their due diligence with analysis prepared by external consultants.

2.3.3 *Due Diligence for Co-Investing*

Given that direct investing is an element of co-investing, aspects of the due diligence process apply to both. In co-investing, investors often rely heavily on the due diligence conducted by the fund manager. Direct investing opportunities, however, are often sourced differently from co-investing opportunities, and so the level of independence sitting behind the due diligence can differ. Consider that direct investing opportunities are usually sourced by the direct investment team at a large pension or sovereign wealth fund, whereas co-investing opportunities are usually provided by the private equity, real estate, or infrastructure fund manager for the investor's consideration. Direct investing due diligence may be more independent than co-investing due diligence because the direct investing team is typically introduced to opportunities by third parties and they have more control over the due diligence process.

KNOWLEDGE CHECK

An investor with limited investment and due diligence experience will *most likely* invest in alternative assets using which method?

A Fund investing

B Co-investing

C Direct investing

Solution:

A is correct. An investor with limited investment and due diligence experience will likely choose fund investing to benefit from the expertise that a fund manager would provide.

PRACTICE QUESTIONS FOR LOS B

Question 1

From the perspective of the investor, the *most* active approach to investing in alternative investments is:

A co-investing.

B fund investing.

C direct investing.

Solution:

C is correct. From the perspective of the investor, direct investing is the most active approach to investing because of the absence of fund managers and the services and expertise they generally provide.

A is incorrect because co-investing includes fund investing, which requires less due diligence compared with direct investing.

B is incorrect because fund investing in alternative assets demands less participation from the investor compared with the direct and co-investing approaches because an investor depends on the fund manager to identify, select, and manage the fund's investments.

Question 2

In comparison to other alternative investment approaches, co-investing is *most likely*:

A more expensive.

B subject to adverse selection bias.

C the most flexible approach for the investor.

Solution:

B is correct. Co-investing may be subject to adverse selection bias. For example, the fund manager may make less attractive investment opportunities available to the co-investor while allocating its own capital to more appealing deals.

A is incorrect because co-investing is likely not more expensive than fund investing since co-investors can co-invest an additional amount alongside the fund directly in a fund investment without paying management fees on the capital that has been directly invested.

C is incorrect because direct investing, not co-investing, provides the greatest amount of flexibility for the investor.

Question 3

Relative to co-investing, direct investing due diligence is *most likely*:

A harder to control.

B more independent.

C equally thorough.

Solution:

B is correct. Direct investing due diligence may be more independent than that of co-investing because the direct investing team is typically introduced to opportunities by third parties rather than fund managers, as is customary in co-investing.

A is incorrect because the direct investing team has more control over the due diligence process compared with co-investing.

C is incorrect because due diligence for direct investing requires the investor to conduct a thorough investigation into the important aspects of a target asset or business, whereas in co-investing, fund managers typically provide investors with access to a data room so they can view the due diligence completed by the fund managers.

Question 4

The investment method that typically requires the greatest amount of or most thorough due diligence from an investor is:

A fund investing.

B co-investing.

C direct investing.

Solution:

C is correct. Due diligence in direct investing will usually be more thorough and more rigid from an investor's perspective because of the absence of a fund manager that would otherwise conduct a large portion of the necessary due diligence.

INVESTMENT AND COMPENSATION STRUCTURES

3

c describe investment and compensation structures commonly used in alternative investments

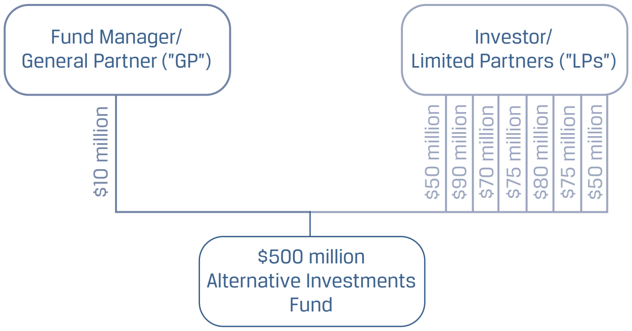
This section explores the investment partnership and compensation structures used in alternative investments. We examine the contractual provisions typically included in partnership agreements and why they exist to protect investors. An overview of the most common investment clauses is provided.

3.1 Partnership Structures

In the world of alternative investments, partnership structures are common. In limited partnerships, the fund manager is the **general partner (GP)** and investors are the **limited partners** (LPs). LPs, who are generally accredited investors (owing to legal restrictions on the fund), are expected to understand and be able to assume the risks associated with the investments, which are less regulated than offerings to the general public. The GP runs the business and theoretically bears unlimited liability for anything that goes wrong. The GP may also run multiple funds at a time.

Limited partners are outside investors who own a fractional interest in the partnership based on the amount of their initial investment and the terms set out in the partnership documentation. LPs commit to future investments, and the upfront cash outflow can be a small portion of their total commitment to the fund. Funds set up as limited partnerships typically have a limit on the number of LPs allowed to invest in the fund. LPs play passive roles and are not involved with the management of the fund (although co-investment rights allow for the LPs to make additional direct investments in the portfolio companies); the operations and decisions of the fund are controlled solely by the GP. Exhibit 4 illustrates the limited partnership structure for a hypothetical $500 million investment fund.

Exhibit 4 Example of a Limited Partnership Structure



The partnership between the GP and LPs is governed by a **limited partnership agreement (LPA)**, a legal document that outlines the rules of the partnership and establishes the framework that ultimately guides the fund's operations throughout its life. LPAs vary in length and complexity and may be dense with provisions and clauses, some of which are discussed later in this section. In addition to LPAs, **side letters** may also be negotiated. Side letters are side agreements created between the GP and a certain number of LPs that exist *outside* the LPA. Some examples of clauses or details that may be included in a side letter include the following:

- Potential additional reporting due to an LP's unique circumstances, such as regulatory or tax requirements
- First right of refusal and other similar clauses to outline potential treatment (regarding fees, co-investment rights, secondary sales, and, potentially, other matters) in comparison to other LPs
- Notice requirements in the event of litigation, insolvency, and related matters
- Most favored nation clauses, such as agreeing that if similar LPs pay lower fees, they will be offered to the LP

Certain structures are commonly adopted for specific alternative investments. For example, infrastructure investors frequently enter into public–private partnerships, which are agreements between the public sector and the private sector to finance, build, and operate public infrastructure, such as hospitals and toll roads. In real estate fund investing, investors may be classified as unitholders, and joint ventures are a partnership structure common in real estate direct investing.

KNOWLEDGE CHECK

Fill in the blank: Investments in limited partnerships are ___________ regulated than offerings to the general public.

Solution:

Investments in limited partnerships are **less** regulated than offerings to the general public.

3.2 Compensation Structures

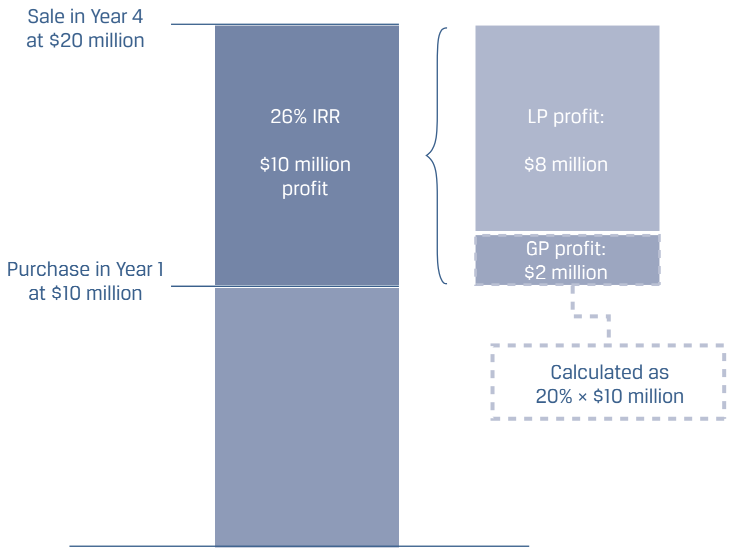
Funds are generally structured with a **management fee** typically ranging from 1% to 2% of assets under management (e.g., for hedge funds) or **committed capital** (e.g., for private equity funds), which is how much money in total that LPs have committed to the fund's future investments. On top of the management fee, a **performance fee** (also referred to as an *incentive fee* or *carried interest*) is applied based on excess returns.

Private equity funds raise committed capital and draw down on those commitments, generally over three to five years, when they have a specific investment to make. Note that the management fee is typically based on committed capital, *not* invested capital; the committed-capital basis for management fees is an important distinction from hedge funds, whose management fees are based on assets under management (AUM). Having committed capital as the basis for management fee calculations reduces the incentive for GPs to deploy the committed capital as quickly as possible to grow their fee base. This allows the GPs to be selective about deploying capital into investment opportunities.

The partnership agreement usually specifies that the performance fee is earned only after the fund achieves a return known as a **hurdle rate**. The hurdle rate is a *minimum* rate of return, typically 8%, that the GP must exceed in order to earn the performance fee. GPs typically receive 20% of the total profit of the private equity fund net of any *hard hurdle rate*, in which case the GP earns fees on annual returns in excess of the hurdle rate, or net of the *soft hurdle rate*, in which case the fee is calculated on the entire annual gross return as long as the set hurdle is exceeded. Hurdle rates are less common for hedge funds but do appear from time to time.

Performance fees are designed to reward GPs for enhanced performance and to motivate investment professionals to work hard and stay involved for years to come. Exhibit 5 illustrates a simple example of how performance fees are calculated using a typical "2 and 20" compensation structure, where management fees are calculated on 2% of AUM or committed capital and performance fees are calculated on 20% of profits. An investment that was purchased in Year 1 for $10 million and sold in Year 4 for $20 million represents a return of 26% on an internal rate of return (IRR) basis. The LPs receive 80% of profits (80% × $10 million profit = $8 million), and the GP receives 20% of profits (20% × $10 million profit = $2 million). Note that this example assumes the hurdle rate has been exceeded and a catch-up clause exists in the partnership agreement; these concepts will be discussed in the next section.

Exhibit 5 Simple Performance Fee Calculation



In addition to both management and performance fees charged to investors, leveraged buyout firms in private equity may charge consulting fees and monitoring fees to the underlying companies.

As for hedge fund managers, they generally accrue an incentive fee on their quarterly performance, but they typically crystalize (realize) their incentive fee once annually. If they make money in the first quarter but then lose the same amount in the second, the investor recoups the Q1 incentive fee and pays only at the end of the year an incentive fee based on total annual returns net of management fees and other expenses.

KNOWLEDGE CHECK

Fill in the blank: A ____________ fee is typically added to a management fee for a fund governed by a limited partnership agreement.

Solution:

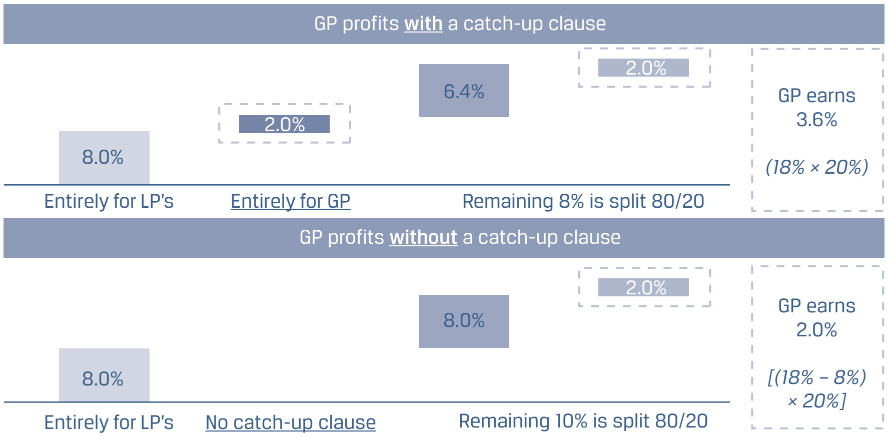
A performance (incentive or carried interest) fee is typically added to a management fee for a fund governed by a limited partnership agreement.

3.3 Common Investment Clauses, Provisions, and Contingencies

For most alternative investment funds, particularly hedge funds and private equity funds, the GP does not earn a performance fee until the LPs have received their initial investment *and* the total return generated on the investment has exceeded a specified hurdle rate. In our example, the LPs receive 80% of the total profit of the fund plus the amount of their initial investment.

A **catch-up clause** may be included in the partnership agreement. Essentially, for a GP who earns a 20% performance fee, a catch-up clause allows the GP to receive 100% of the distributions above the hurdle rate *until* he receives 20% of the profits generated, and then every excess dollar is split 80/20 between the LPs and GP. To illustrate, assume that the GP has earned an 18% IRR on an investment, the hurdle rate is 8%, and the partnership agreement includes a catch-up clause. In this case, the LPs would receive the entirety of the first 8% profit, the GP would receive the entirety of the next 2% profit—because 2% out of 10% amounts to 20% of the profits accounted for so far, as the catch-up clause stipulates—and the remaining 8% would be split 80/20 between the LPs and the GP. Effectively, the LPs earned 14.4% (18% × 80%) and the GP earned 3.6% (18% × 20%). Imagine the same scenario, absent a catch-up clause: The LPs would still receive the entirety of the first 8% profit; however, the remaining 10% would be split 80/20 between the LPs and GP, reducing the GP's return to 2.0% [(18% − 8%) × 20%]. These calculations are presented graphically in Exhibit 6.

Exhibit 6 Simple Catch-Up Clause Illustration



In hedge funds, fee calculations also take into account a **high-water mark**, which reflects the highest value used to calculate an incentive fee. A high-water mark is the *highest* value of the fund investment ever achieved at a performance fee crystallization date, net of fees, by the individual LP. A high-water mark clause states that a hedge fund manager must recuperate declines in value from the high-water mark before performance fees can be charged on newly generated profits. The use of high-water marks protects clients from paying twice for the same performance.

With all alternative investments, investor high-water marks generally carry over into new calendar years, although with hedge funds, an investor will no longer be able to claw back any incentive fees paid for a given calendar year if portfolio losses are subsequently incurred in a later calendar year. Given the generally more illiquid and longer-term nature of their holdings, private equity and real estate investments are more likely to contain clawback clauses for the entire life of the portfolio.

Given that different clients invest at different times, it is possible that not all clients will be at their respective high-water marks at the same time; a client who invests on a dip will enjoy the fund's recovery and pay an incentive fee, whereas a client who invested at the top will need to earn back what was lost before being obliged to pay the incentive fee.

Needless to say, individual capital account fund accounting can become quite complicated when tracking such investment timing differences. Even careful administrators make accounting mistakes when handling high-water mark and incentive fee issues. The astute alternative asset investor will always double-check the incentive fee calculations provided to them. Best-practice accounting for hedge funds is far more complex than the simple issuance of shares (carried at a single net asset value, or NAV), which is more common in the mutual fund industry.

KNOWLEDGE CHECK

The minimum rate of return that a GP must exceed in order to earn an incentive or performance fee is called the:

A high-water mark.

B hurdle rate.

C performance threshold.

Solution:

B is correct. A high-water mark is the highest value used to calculate an incentive fee. "Performance threshold" is not a term that is generally used in the industry.

In alternative investments, a **waterfall** represents the distribution method that defines the order in which allocations are made to LPs and GPs. GPs usually receive a disproportionately larger share of the total profits relative to their initial investment, which incentivizes them to maximize profitability. There are two types of waterfalls: *deal-by-deal* (or *American*) waterfalls and *whole-of-fund* (or *European*) waterfalls. Deal-by-deal waterfalls are more advantageous to the GP because performance fees are collected on a per-deal basis, allowing the GP to get paid before LPs receive both their initial investment *and* their preferred rate of return (i.e., the hurdle rate) on the entire fund. In whole-of-fund waterfalls, all distributions go to the LPs as deals are exited and the GP does not participate in any profits until the LPs receive their initial investment and the hurdle rate has been met. In contrast to deal-by-deal waterfalls, whole-of-fund waterfalls occur at the aggregate fund level (i.e., after *all* investments have been exited) and are more advantageous to the LPs. Exhibits 7 and 8 illustrate deal-by-deal waterfalls and whole-of-fund waterfalls, respectively.

Exhibit 7 Deal-by-Deal (American) Waterfall Example

Investment No.	Year		Amount ($mm)		Profit		
	Invested	Sold	Invested	Sold	$mm	%	GP at 20%
1	1	4	$10	$20	$10	26.0%	$2
2	2	5	$20	$35	$15	20.5%	$3
3	2	7	$40	$80	$40	14.9%	$8
4	3	7	$20	$20	-	-	-
5	3	8	$35	$25	($10)	neg	($2)
6	4	9	$25	$20	($5)	neg	($1)
7	5	9	$30	-	($30)	neg	($6)
8	5	10	$20	-	($20)	neg	($4)
Total	1	10	$200	$200	-	-	-

Exhibit 8 Whole-of-Fund (European) Waterfall Example

Investment No.	Year		Amount ($mm)		Profit		
	Invested	Sold	Invested	Sold	$mm	%	GP at 20%
1	1	4	$10	$20	$10	26.0%	-
2	2	5	$20	$35	$15	20.5%	-
3	2	7	$40	$80	$40	14.9%	-
4	3	7	$20	$20	-	-	-
5	3	8	$35	$25	($10)	neg	-
6	4	9	$25	$20	($5)	neg	-
7	5	9	$30	-	($30)	neg	-
8	5	10	$20	-	($20)	neg	-
Total	1	10	$200	$200	-	-	-

A **clawback** provision reflects the right of LPs to reclaim part of the GP's performance fee. Along either waterfall path, if a GP ever accrues (or actually pays itself) an incentive fee on gains that are not yet fully realized and then subsequently gives back these gains, an investor is typically able to claw back prior incentive fee accruals and payments. Clawback provisions are usually activated when a GP exits successful deals early on but incurs losses on deals later in the fund's life.

PRACTICE QUESTIONS FOR LOS C

Question 1

An alternative investment fund's hurdle rate is a:

A rate unrelated to a catch-up clause.

B tool to protect clients from paying twice for the same performance.

C minimum rate of return the GP must exceed in order to earn a performance fee.

Solution:

C is correct. An alternative investment fund's hurdle rate is a minimum rate of return the GP must exceed in order to earn a performance fee.

A is incorrect because if a catch-up clause is included in the partnership agreement, the catch-up clause permits distributions in relation to the hurdle rate.

B is incorrect because it is a high-water mark (not a hurdle rate) that protects clients from paying twice for the same performance.

Question 2

An investor in a private equity fund is concerned that the general partner can receive incentive fees in excess of the agreed-on incentive fees by making distributions over time based on profits earned rather than making distributions only at exit from investments of the fund. Which of the following is most likely to protect the investor from the general partner receiving excess fees? (2020 Q32)

A A high hurdle rate

B A clawback provision

C A lower capital commitment

Solution:

B is correct. A clawback provision requires the general partner in a private equity fund to return any funds distributed (to the general partner) as incentive fees until the limited partners have received their initial investments *and* the contracted portion of the total profits. A high hurdle rate will result in distributions occurring only after the fund achieves a specified return. A high hurdle rate decreases the likelihood of, but does not prevent, excess distributions. Management fees, not incentive fees, are based on committed capital.

Question 3

Until the committed capital is fully drawn down and invested, the management fee for a private equity fund is based on: (2022 Q33)

A invested capital.
B committed capital.
C assets under management.

Solution:

B is correct. Until the committed capital is fully drawn down and invested, the management fee for a private equity fund is based on committed capital, not invested capital.

Question 4

The distribution method by which profits generated by a fund are allocated between LPs and the GP is called:

A a waterfall.
B an 80/20 split.
C a fair division.

Solution:

A is correct. Although profits are typically split 80/20 between LPs and the GP, the distribution method of profits is not called an "80/20 split." "Fair division" is not a real term that exists in the industry.

Question 5

Fill in the blanks with the correct words: An American waterfall distributes performance fees on a(n) ___________ basis and is more advantageous to the ___________.

A deal-by-deal; LPs
B aggregate fund; LPs
C deal-by-deal; GP

Solution:

C is correct. American waterfalls, also known as deal-by-deal waterfalls, pay performance fees after every deal is completed and are more advantageous to the GP because they get paid sooner (compared with European, or whole-of-fund, waterfalls).

4 HEDGE FUNDS

d explain investment characteristics of hedge funds

4.1 Characteristics of Hedge Funds

In 1949, Alfred Winslow Jones, a sociologist investigating fundamental and technical research in order to forecast the stock market for *Fortune* magazine, set up an investment fund with himself as GP. The fund followed three tenets: (1) Always maintain short positions, (2) always use leverage, and (3) charge only an incentive fee of 20% of profits with no fixed fees. Jones called his portfolio a "hedged" fund (eventually shortened to "hedge" fund) because his short positions offset his long positions in the stock market. Theoretically, the overall portfolio was hedged against major market moves.

Although Jones's original three tenets still have some relevance to today's hedge fund industry, not all hedge funds restrict themselves to equities. Many trade sovereign and corporate debt, commodities, futures contracts, options, and other derivatives. Even real estate investments are found in hedge fund structures. Nor do all hedge funds maintain short positions or use leverage. Instead, many simply exploit niche areas of expertise in a sophisticated manner; hedging and leverage may or may not be part of it.

Fee structures have evolved to include management fees in addition to Jones's incentive fee. The performance fees we discussed previously may vary from his simple 20% while including hurdle rates and other mechanisms.

Investors in modern hedge funds are subject to extended holding periods (known as **lockup periods**) and subsequent **notice periods** before an investment redemption is possible. Such mandatory periods combine with a lighter regulatory compliance burden (in certain areas) than mutual funds face, for instance, allowing hedge funds more flexibility. Most hedge funds do not need to guarantee liquidity on one day's redemption notice as the large majority of mutual funds do. Instead, they may avail themselves of less liquid and unnoticed opportunities, the true valuation of which may at times be opaque.

Hedge funds are generally deemed riskier from an oversight (fraud risk) point of view, but some hedge funds take less absolute market risk in their portfolio construction than registered products available to retail investors might take. The distinction between regulatory risk and illiquidity risk when compared with true market risk is often confused.

Overall, the term "hedge fund" today can be seen more as a reference to a certain "vehicle type"—a private limited partnership or offshore fund open only to high-net-worth and institutional investors from a regulatory perspective—that may contain different types of investments and where incentive fees are allowed to be charged. This situation remains quite distinct from the mutual fund world, where incentive fees are typically not allowed and investment and liquidity restrictions are more tightly defined. Although in recent years, there has been some "blurring of the lines" with the introduction of so-called liquid alternative mutual funds and what are known as UCIT funds in Europe (many of which offer more "hedged" styles and approaches than traditional daily-liquidity products), such offerings are not to be confused with true limited-access and limited-liquidity hedge funds.

A contemporary hedge fund generally has the following characteristics:

- It is a creatively managed portfolio of investments involved in one or more asset classes (equities, credit, fixed income, commodities, futures, foreign exchange, and sometimes even hard assets, such as real estate), sometimes trading in different geographic regions, that is often leveraged, generally takes both long and short positions (when possible), and quite often uses derivatives to express a view or establish a hedge.

- Its goal is to generate high returns, either in an absolute sense or on a risk-adjusted basis relative to its portfolio-level volatility. It may be hard to gauge a hedge fund's performance relative to a traditional index benchmark (after all, low correlation with traditional asset investing is frequently a selling point), and it generally enjoys light investment restrictions that are detailed in each hedge fund's private placement offering memorandum.
- It is set up as a private investment partnership or offshore fund that under certain legal restrictions (which vary by jurisdiction) make the offering open to a limited number of investors willing and able to make a substantive initial investment.

As noted previously, restrictions on **redemptions** are typically imposed. Investors may be required to keep their money in the hedge fund for a minimum period (referred to as a lockup period) before they are allowed to make withdrawals or redeem shares. Investors may be required to give notice of their intent to redeem; the notice period is typically 30–90 days. To redeem shares, investors may be charged a fee, typically payable to the fund itself (rather than the manager) so as not to disadvantage remaining investors in the fund, particularly in circumstances where the redemption takes place during the lockup period. This characteristic is called a *soft lockup*, and it offers a path (albeit an expensive one) to redeem early. Again, such restrictions allow hedge fund managers more flexibility in portfolio construction, giving them leeway to invest in situations in which time may be needed to generate an expected return and that, therefore, are not typically suitable for a mutual fund that offers daily liquidity.

Funds of hedge funds are funds that hold a portfolio of hedge funds. They create a diversified portfolio of hedge funds and invite participation from smaller investors or those who do not have the resources, time, or expertise to choose among hedge fund managers. Also, the managers of funds of hedge funds (commonly shortened to "funds of funds") are required to have some expertise in conducting due diligence on hedge funds and may be able to negotiate better redemption or fee terms than individual investors can. Funds of funds may diversify across fund strategies, investment regions, and management styles.

Funds of funds may charge an additional 1% management fee and an additional 10% incentive fee (1 and 10) on top of the fees charged by the underlying hedge funds held in the fund of funds. At both the hedge fund level and the fund-of-funds level, the incentive fee is typically calculated on profits net of management fees.

Fee structures are important to scrutinize here. Fee layering can sorely reduce the initial gross return of an investment in a hedge fund or fund of funds for the end investor. Each hedge fund into which a fund of funds invests is structured to receive a management fee plus an incentive fee. The result may be that the investor is paying fees more than once for the management of the same assets. Finally, liquidity can be of additional concern for investors in funds of funds in times of crisis, when fund redemptions can hurt performance.

KNOWLEDGE CHECK

Fill in the blank: The goal of hedge fund redemption restrictions is typically to ____________ manager flexibility.

Solution:

The goal of hedge fund redemption restrictions is typically to increase hedge fund manager flexibility.

KNOWLEDGE CHECK

Define funds of hedge funds, and identify their primary purpose.

Solution:

Funds of hedge funds are funds that hold a portfolio of hedge funds. Their primary purpose is to add value in manager selection and due diligence and create a diversified portfolio of hedge funds accessible to smaller investors.

Few investors were worried about high fees as hedge funds were achieving average net-of-fee returns of 10%–20% in the 1990s and early 2000s, but as returns migrated toward 6%–9% and the drag on returns from fee structures became much more relevant on a percentage basis, institutional investors challenged these high fees. In late 2013, the largest pension plan in the United States, CalPERS, announced it would curtail its hedge fund allocations partly for this reason.

According to Eurekahedge, by 2018 the average hedge fund fee level had compressed to 1.3% in management fees and 15.5% in incentive fees—considerably less than the 2 and 20 structure envisions but still quite high relative to traditional mutual funds. Fund-of-funds fee levels similarly have compressed; funds of funds can be found charging, for example, either a 1% flat management fee or a 50 bp management fee and only a 5% incentive fee.

Realization of the value-destruction potential of fee layers conspired with increasingly compressed hedge fund returns (for a variety of reasons) to leave funds of funds far less popular in 2020 than they were between 2000 and 2005. Still, a fund of funds may offer compensating advantages, such as access for smaller investors, diversified hedge fund portfolios, better redemption terms, and due diligence expertise.

The overall popularity of hedge funds is illustrated by AUM and net asset flows. Hedge Fund Research, Inc. (HFR, or HFRI in named indexes), reported that AUM grew from approximately $39 billion in 1990 to $491 billion in 2000 and to $3.32 trillion as of December 2019—a clear decade-spanning boom in investment, even if growth has slowed somewhat in the last five years.

Exhibit 9 compares the returns, risk, and performance measures for the HFRI Fund of Funds Composite Index, the MSCI ACWI Index, and the Bloomberg Barclays Global Aggregate Index. The HFRI Fund of Funds Composite Index is an equally weighted performance index of funds of hedge funds included in the HFR Database. Hedge fund indexes suffer from issues related to self-reporting and **survivorship bias**. Survivorship bias relates to the inclusion of only current investment funds in a database. As such, the returns of funds that are no longer available in the marketplace (have been liquidated) are excluded from the database. However, the HFRI Fund of Funds Composite Index reflects the actual performance of portfolios of hedge funds. The measures shown here may reflect a lower reported return because of the added layer of fees, but they likely represent a fairer, more conservative, and more accurate estimate of average hedge fund performance than HFR's composite index of individual funds. The returns are likely just mildly biased toward equity long–short funds since these are always a substantial portion of funds of funds' allocation mix.

As shown in Exhibit 9, over the 25-year period between 1990 and 2014, hedge funds enjoyed higher returns than either stocks or bonds and a standard deviation almost identical to that of bonds. On the basis of the Sortino ratio, hedge funds do not appear as attractive as bonds if returns are adjusted for downside deviation, as reflected in the relative Sortino measures shown in Exhibit 9. The worst **drawdown**, or period of largest cumulative negative returns for hedge funds and global equities, began in 2007 (when each peaked) and ended in 2009.

Notwithstanding this period of stress, when accounting for hedge fund returns' modest overall correlation with global stock returns (0.56) and negligible correlation with global bond returns (0.07) over this 25-year period, hedge funds have certainly offered added value to institutional investors as a portfolio diversification agent.

Exhibit 9 Historical Risk–Return Characteristics of Hedge Funds and Other Investments, 1990–2014

	FoF	Global Stocks	Global Bonds
Annualized return	7.2%	6.9%	6.3%
Annualized volatility	6.0%	16.5%	5.8%
Sharpe ratio	0.63	0.21	0.49
Sortino ratio	0.74	0.43	1.09
Percentage of positive months	69.3%	61.3%	62.7%
Best month	6.8%	11.9%	6.2%
Worst month	−7.5%	−19.8%	−3.8%
Worst drawdown	−22.2%	−54.6%	−10.1%
FoF correlation (avg. monthly)		0.56%	0.07%

Sources: Fund-of-funds (FoF) data are from the HFRI Fund of Funds Composite Index; global stock data are from the MSCI ACWI Index; global bond data are from the Bloomberg Barclays Global Aggregate Index.

Notably, as shown in Exhibit 10, for the subsequent five-year period between 2015 and 2019, the absolute return of funds of funds relative in particular to global equities has declined, while their performance correlation with equity markets actually increased. This trend has made hedge fund allocations arguably less useful and also somewhat less popular. But some allocators have continued to find value to maintain or actually increase their allocation to a mix of hedge funds as a bond market substitute in their overall portfolio building.

Exhibit 10 Historical Risk-Return Characteristics of Hedge Funds and Other Investments, 2015–2019

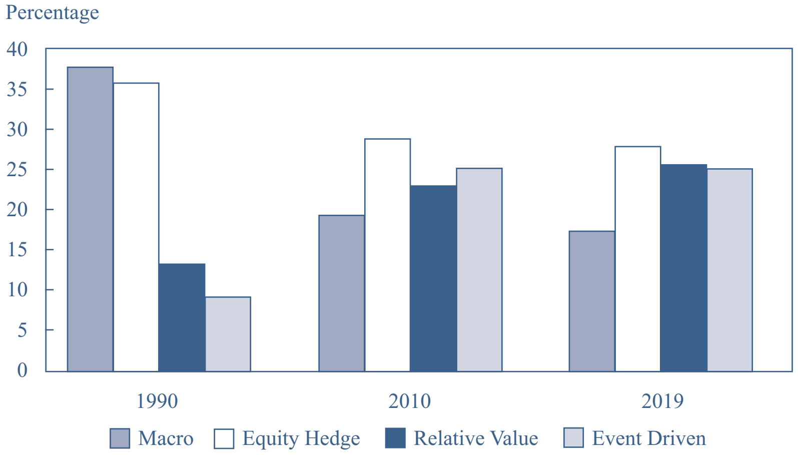
	FoF	Global Stocks	Global Bonds
Annualized return	2.48%	10.41%	2.42%
Annualized volatility	1.08%	11.42%	1.33%
FoF Correlation (avg. monthly)		0.86%	0.03%

Sources: Fund-of-funds (FoF) data are from the HFRI Fund of Funds Composite Index; global stock data are from the MSCI ACWI Index; global bond data are from the Bloomberg Barclays Global Aggregate Index (USD Unhedged). The risk-free rate used in the Sharpe ratio calculation is from the average of one-month T-bill rates over the period.

4.2 Hedge Fund Strategies

Hedge funds are typically classified by strategy, although categorizations vary. Many classifying organizations focus on the most common strategies, but others classify on the basis of different criteria, such as the underlying assets in which the funds are invested. Classifications change over time as new strategies, often based on new products and market opportunities, are introduced. Classifying hedge funds is important so that investors can review aggregate performance data, select strategies with which to build a portfolio of funds, and select or construct appropriate performance benchmarks. HFR classifies hedge funds under four broad strategy categories: event-driven, relative value, macro, and equity hedge. As of March 2019, hedge fund assets stood at approximately $3.18 trillion. Exhibit 11 shows the approximate percentage of hedge fund AUM by strategy, according to HFR, for 1990, 2010, and 2019. Based on this AUM measure, event-driven and relative value strategies have grown in popularity during the last 30 years, while macro and equity hedge funds have declined.

Exhibit 11 Percentage of AUM by Strategy



Source: HFR data, as of March 2019.

4.2.1 *Equity Hedge Strategies*

Equity hedge strategies can be thought of as the original hedge fund category. They are focused on public equity markets and take long and short positions in equity and equity derivative securities. Most equity hedge strategies use a “bottom-up” security-specific approach—company-level analysis, followed by overall industry analysis, followed by overall market analysis—with relatively balanced long and short exposures. A “top-down” approach, in contrast, entails global macro analysis, followed by sector/regional analysis, followed by individual company analysis or any market-timing approach. Some strategies use individual equities only, whereas others use index ETF securities for portfolio-balancing purposes. Examples of equity hedge strategies include the following:

- **Market neutral.** These strategies use quantitative (technical) and fundamental analysis to identify under- and overvalued equity securities. The hedge fund takes long positions in securities it has identified as undervalued and short positions in securities it has identified as overvalued. The hedge fund tries to maintain a net position that is neutral with respect to market risk and other risk factors (size, industry, momentum, value, etc.). Ideally, the portfolio has an overall beta of approximately zero. The intent is to profit from the movements of individual securities while remaining largely indifferent to market risk. To achieve a meaningful return, market-neutral portfolios may require the application of leverage. They are generally stable, low-return portfolios but are exposed to occasional spurts of volatility, at which point sudden leverage reduction may be required.

- **Fundamental L/S growth.** These strategies use fundamental analysis to identify companies expected to exhibit high growth and capital appreciation. The hedge fund takes long positions in these companies while shorting companies with business models that are being disintermediated or are under downward pressure—where continued revenue growth is deemed unlikely. Long–short growth managers tend to end up long biased overall.
- **Fundamental value.** These strategies use fundamental analysis to identify companies that are deemed undervalued and unloved for any number of corporate-performance or sector-driven reasons, but the manager identifies a path to corporate rejuvenation. The hedge fund takes long positions in these companies and sometimes hedges the portfolio by shorting index ETFs or more growth-oriented companies deemed overvalued. Such managers also tend to be long biased and have a positive factor bias to value as well as to small-cap factors since many underappreciated value situations tend to reside in the small-cap sector.
- **Short biased.** These strategies use quantitative (technical) and fundamental analysis to focus the bulk of their portfolio on shorting overvalued equity securities (against limited long-side exposures or none at all). Many of these funds are more forensic in their fundamental analysis, and some turn activist by trying to expose previously unrecognized accounting or business flaws, thereby also improving their own portfolio’s chance for profits. The degree to which these funds adjust their short exposure over time varies by manager. Such managers tend to be contrarian by nature, and the funds can be useful additions to larger portfolios in terms of weathering periods of market stress. Short-biased managers, however, have had a difficult time overall posting meaningful long-term returns during the past 30 years of generally positive market conditions.
- **Sector specific.** These strategies exploit manager expertise in a particular sector and use quantitative (technical) and fundamental analysis to identify opportunities. Technology, biotech/life sciences, and financial services are common areas of focus. Given our era of increased democratization in the dissemination of news, added sector-specialist expertise may be a prerequisite for truly differentiated security selection. The more complex a sector (as in biotechnology, with its binary outcomes on drug trials) or the more opaqueness in accounting practices (as with financial or insurance stocks), the more value sector-specific managers bring.

4.2.2 *Event-Driven Strategies*

Event-driven strategies seek to profit from defined catalyst events, typically those that involve changes in corporate structure, such as an acquisition or restructuring. This strategy is considered to be underpinned by bottom-up security-specific analysis, as opposed to top-down analysis. Investments may include long and short positions in common and preferred stocks, debt securities, and options. Further subdivisions of this category by HFR include the following:

- **Merger arbitrage.** Generally, these strategies involve going long (buying) the stock of the company being acquired at a discount to its announced takeover price and going short (selling) the stock of the acquiring company when the merger or acquisition is announced. The manager may expect to profit once the initial deal spread narrows to the closing value of the transaction after it is fully consummated. The spread exists because there is always some uncertainty over whether the deal will actually close in the face of legal and regulatory hurdles. Equally, a sudden change in market conditions could damage the perceived attractiveness of the merger to the acquirer, who could step away. Shorting the acquirer is also a way to express the risk that the acquirer has overpaid for the acquisition or will suffer from an increased debt load following the merger. The primary risk in merger arbitrage is that the announced merger or acquisition fails to take place and the deal spread re-widens to pre-merger levels before the hedge fund manager is able to unwind its position. Since the expected risk and return on a merger arbitrage strategy typically starts off being quite modest, managers regularly use leverage to amplify merger spreads into financially worthwhile total return targets. Unfortunately, this application of leverage also increases the magnitude of losses should real-world circumstances move against the manager's bet.

- **Distressed/restructuring.** These strategies focus on the securities of companies that are either in bankruptcy or perceived to be near bankruptcy. One strategy sees hedge funds simply purchasing fixed-income securities that are trading at a significant discount to par but are still senior enough to be deemed "money good" (backed by enough corporate assets to be fully repayable at par or at least at a significant premium to the available bond purchase price) in a bankruptcy situation. Alternatively, a hedge fund may purchase the so-called fulcrum debt instrument that is expected to convert into new equity in the case of a restructuring or bankruptcy.
- **Special situations.** These strategies focus on opportunities to get involved in the equity of companies that are engaged in restructuring activities other than mergers, acquisitions, or bankruptcy. These activities include security issuance or repurchase, special capital distributions, rescue finance, and asset sales/spin-offs.
- **Activist.** The term "activist" is short for "activist shareholder." Here, managers secure sufficient equity holdings to allow them to influence a company's policies or direction. The hedge fund manager thus tries to create his or her own catalyst, influencing the investment's ultimate destiny by creating a desired corporate outcome. For example, the activist hedge fund may advocate for divestitures, restructuring, capital distributions to shareholders, or changes in management and company strategy that will affect their equity holdings. Such hedge funds are distinct from private equity because they operate primarily in the public equity market.
- These event-driven strategies tend to be long biased. Although merger arbitrage is the least long biased among them, the strategy still tends to suffer when market conditions weaken because the risk that mergers will fail increases.

4.2.3 *Relative Value Strategies*

Relative value funds seek to profit from a pricing discrepancy, an unusual short-term relationship, between related securities. The expectation is that the pricing discrepancy will be resolved over time. Examples of relative value strategies include the following:

- **Convertible bond arbitrage.** This conceptually market-neutral investment strategy seeks to exploit a perceived mispricing between a convertible bond and its component parts—namely, the underlying bond and the embedded stock option—relative to the pricing of a reference equity into which the bond may someday convert. The strategy typically involves buying convertible debt securities and simultaneously selling a certain amount of the same issuer's common stock. Residual bankruptcy risks can be further hedged using equity put options or credit default swap derivative hedges on the credit of the issuer.

- **Fixed income (general).** These strategies focus on the relative value within the fixed-income markets, with an emphasis on sovereign debt and, at times, the relative pricing of investment-grade corporate debt. Strategies may incorporate long–short trades between two different issuers, between corporate and government issuers, between different parts of the same issuer's capital structure, or between different parts of an issuer's yield curve. Currency dynamics and government yield-curve considerations may also come into play.
- **Fixed income (asset backed, mortgage backed, and high yield).** These strategies focus on the relative value of a variety of higher-yielding securities, including asset-backed securities (ABS), mortgage-backed securities (MBS), and high-yield loans and bonds. The hedge fund seeks both to earn an attractive highly secured coupon and to take advantage of relative security mispricings. At times, some of these securities may be further parsed into structured note products with unique return attributes relative to interest rate movements and general credit spread changes. Opaque **mark-to-market** pricing and illiquidity issues are significant considerations. "Mark-to-market" refers to the current expected fair market value for which a given security would likely be available for purchase or sale if traded in current market conditions, but in the case of structured products, absent hard broker/dealer quotes, this is often a "mark-to-model" type of calculation that is less reliable than an actual bid–offer spread seen between active market participants.
- **Volatility.** These strategies typically use options to go long or short market volatility, either in a specific asset class or across asset classes. For example, a short-volatility strategy involves selling options to earn the premiums and benefit from calm markets but can experience significant losses during unexpected periods of market stress. A long-volatility strategy tends to suffer the cost of small premiums in anticipation of larger market moves where positions may benefit from both directionality and the ability to rebalance option-based exposure through a variety of methods.
- **Multi-strategy.** These strategies trade relative value within and across asset classes or instruments. Rather than focusing on one type of trade (e.g., convertible arbitrage), a single basis for a trade (e.g., volatility), or a particular asset class (e.g., fixed income), this strategy instead looks for investment opportunities wherever they might exist, often with different pods of managers executing their own unique market approaches. The goal of a multi-strategy manager is to initially deploy (and later redeploy) capital efficiently and quickly across a variety of strategy areas as market conditions change.

4.2.4 *Macro and CTA Strategies*

Macro hedge funds emphasize a top-down approach to identify economic trends. Trades are made on the basis of expected movements in economic variables. Generally, these funds trade opportunistically in fixed-income, equity, currency, and commodity markets. Macro hedge funds use long and short positions to profit from a view on the overall direction of the market as it is influenced by major economic trends and events. Macro managers were generally very successful during the 1980s and 1990s, but return profiles have weakened over the past two decades as markets have arguably become more closely managed by central bankers. Because macro managers generally enjoy periods of higher volatility compared with lower volatility, the active moves by national authorities, such as central banks, to smooth out economic shocks likely shrink managers' investment sphere.

Managed futures funds are actively managed funds making diversified directional investments primarily in the futures markets on the basis of a variety of technical and fundamental strategies. Managed futures funds are also known as commodity trading advisers (CTAs) because they historically focused on commodity futures. However, CTAs may include investments in a variety of futures, including commodities, equities, fixed income, and foreign exchange. CTAs generally use models that measure trends and momentum over different time horizons. Investing in CTAs can be useful for portfolio diversification, particularly in times of strong trending market conditions and especially during periods of acute market stress when other fundamental strategies may be expected to perform poorly. However, mean-reverting markets, which may cause false momentum breakout signals, can lead to uncomfortable and occasionally extended drawdown periods before strong trends emerge for the CTA. To the extent that many CTAs have migrated to trade more and more financial products (such as stock index futures and bond futures), the reliability of CTA diversification benefits has diminished.

4.3 Hedge Funds and Diversification Benefits

Hedge funds were generally a niche business throughout much of the 1980s and 1990s, supported mostly by high-net-worth investors. Traditional long–short equity managers shared the spotlight with successful macro-oriented managers, bringing to mind George Soros's Quantum Fund or Julian Robertson's Tiger Management, which took advantage of global currency, equity, and interest rate imbalances between countries.

But when hedge funds generally performed well during the dot-com bubble unwinding of 2000–2002—particularly compared with traditional long-only investment products—endowments, foundations, and pensions started to allocate more money to them. Initially investing in funds of funds, after the 2008 period of market stress, they increasingly made direct allocations. They did so in spite of the high fees charged by hedge funds, and this trend stemmed from an overall effort to achieve better diversification and risk mitigation. This latter era of hedge funds can be characterized by a search for absolute and uncorrelated risk-adjusted returns rather than for outsized upside performance. The institutionalization of the hedge fund industry changed its very nature: Most hedge funds failed to keep pace with the positive equity and bond market advances of 2009–2019, but they maintained a place in institutional asset allocations because of their risk-diversification properties.

PRACTICE QUESTIONS FOR LOS D

Question 1

Which approach is *most commonly* used by equity hedge strategies?

A Top down

B Bottom up

C Market timing

Solution:

B is correct. Most equity hedge strategies use a bottom-up strategy.

A is incorrect because most equity hedge strategies use a bottom-up (not top-down) strategy

C is incorrect because most equity hedge strategies use a bottom-up (not market-timing) strategy.

Question 2

An investor may prefer a single hedge fund to a fund of funds if she seeks: (2022 Q8)

A due diligence expertise.

B better redemption terms.

C a less complex fee structure.

Solution:

C is correct. Hedge funds of funds have multi-layered fee structures, whereas the fee structure for a single hedge fund is less complex. Funds of funds presumably have some expertise in conducting due diligence on hedge funds and may be able to negotiate more favorable redemption terms than an individual investor in a single hedge fund could.

Question 3

Hedge funds are similar to private equity funds in that both:

A are typically structured as partnerships.

B assess management fees based on assets under management.

C do not earn an incentive fee until the initial investment is repaid.

Solution:

A is correct. Private equity funds and hedge funds are typically structured as partnerships where investors are limited partners and the fund is the general partner. The management fee for private equity funds is based on committed capital, whereas for hedge funds, the management fees are based on assets under management. For most private equity funds, the general partner does not earn an incentive fee until the limited partners have received their initial investment back.

Question 4

Both event-driven and macro hedge fund strategies use:

A long–short positions.

B a top-down approach.

C long-term market cycles.

Solution:

A is correct. Long–short positions are used by both types of hedge funds to potentially profit from anticipated market or security moves. Event-driven strategies use a bottom-up approach and seek to profit from a catalyst event typically involving a corporate action, such as an acquisition or a restructuring. Macro strategies seek to profit from expected movements in evolving economic variables.

Question 5

Hedge fund losses are *most likely* to be magnified by a:

A margin call.

B lockup period.

C redemption notice period.

Solution:

A is correct. Margin calls can magnify losses. To meet the margin call, the hedge fund manager may be forced to liquidate a losing position in a security, which, depending on the position size, could exert further price pressure on the security, resulting in further losses. Restrictions on redemptions, such as lockup and notice periods, may allow the manager to close positions in a more orderly manner and minimize forced-sale liquidations of losing positions.

Question 6

An equity hedge fund following a fundamental growth strategy uses fundamental analysis to identify companies that are *most likely* to: (2022 Q16)

A be undervalued.

B be either undervalued or overvalued.

C experience high growth and capital appreciation.

Solution:

C is correct. Fundamental growth strategies take long positions in companies identified, using fundamental analysis, to have high growth and capital appreciation. Fundamental value strategies use fundamental analysis to identify undervalued companies. Market-neutral strategies use quantitative and fundamental analysis to identify under- and overvalued companies.

PRIVATE CAPITAL

5

e explain investment characteristics of private capital

5.1 Overview of Private Capital

Private capital is the broad term for funding provided to companies that is not sourced from the public markets, such as proceeds raised from the sale of equities, bonds, and other securities on exchanges, nor from traditional institutional providers, such as a government or bank. When capital is raised from sources other than public markets and traditional institutions and it is provided in the form of an equity investment, it is called private equity. If similarly sourced capital is instead extended to companies through a loan or other form of debt, it is referred to as private debt. Private capital looks at the entire capital structure, comprising private equity and private debt.

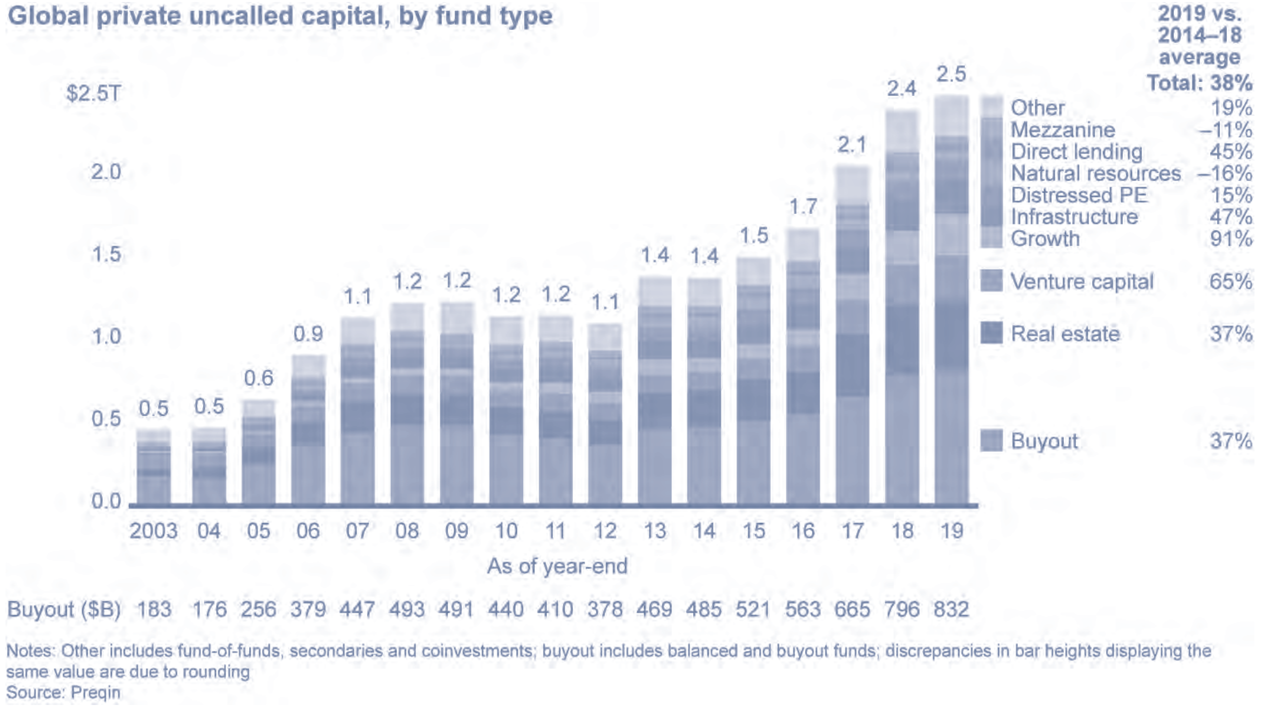
The private capital space largely consists of private investment funds and entities that invest in the equity or debt securities of companies, real estate, or other assets that are privately held. Many private investment firms have private equity and private debt arms; however, these teams typically refrain from investing in the same assets or businesses to avoid overexposure to a single investment and to avoid the conflict of interest that arises from sitting on both sides of the creditors' table. Private investment firms are typically referred to as "private equity firms" even when they have private debt arms. Although private equity is the largest component of private capital, using "private equity" as a generic term could be less accurate and possibly misleading because other forms of private alternative finance have grown considerably in size and popularity.

5.2 Description: Private Equity

Private equity refers to investment in privately owned companies or in public companies with the intent to take them private. As business conditions and the availability of financing change, the focus of private equity firms may change. A firm may manage many private equity funds, each composed of several investments, and the company the firm invests in is often called a **portfolio company** because it will become part of a private equity fund portfolio. According to one possible categorization, we can name leveraged buyouts, venture capital, and growth capital as the primary private equity strategies.

Private equity activity has grown over time, but it is cyclical. Exhibit 12 shows the portion of committed capital that has not yet been called; the industry term for it is "dry powder." Note that detailed information on private equity activity is not always readily available.

Exhibit 12 Global Private Equity: Uncalled Capital by Fund Type (US$ trillions)



Notes: Other includes fund-of-funds, secondaries and coinvestments; buyout includes balanced and buyout funds; discrepancies in bar heights displaying the same value are due to rounding
Source: Preqin

Source: Bain & Company, "Global Private Equity Report 2020" (2020).

5.2.1 *Leveraged Buyouts*

Leveraged buyouts (LBOs) or highly leveraged transactions arise when private equity firms establish buyout funds (or LBO funds) to acquire public companies or established private companies, with a significant percentage of the purchase price financed through debt. The target company's assets typically serve as collateral for the debt, and the target company's cash flows are expected to be sufficient to service the debt. The debt becomes part of the target company's capital structure if the buyout goes through. After the buyout, the target company becomes or remains a privately owned company. LBOs are sometimes referred to as "going-private" transactions because after the acquisition of a publicly traded company, the target company's equity is substantially no longer publicly traded. (When the target company is private, it is not a going-private transaction.) The LBO may also be of a specific type. In **management buyouts** (MBOs), the current management team is involved in the acquisition, and in **management buy-ins** (MBIs), the current management team is being replaced and the acquiring team will be involved in managing the company. LBO managers seek to add value by improving company operations, boosting revenue, and ultimately increasing profits and cash flows. Cash-flow growth, in order of contribution, comes from organic revenue growth, cost reductions and restructuring, acquisitions, and then all other sources. The financial returns in this category, however, depend to a large extent on the use of leverage. If debt financing is unavailable or costly, LBOs are less likely to take place.

5.2.2 *Venture Capital*

Venture capital (VC) entails investing in or providing financing to private companies with high growth potential. Typically these are start-ups or young companies, but venture capital can be provided at a variety of stages, ranging from the inception of an idea for a company to the point when the company is about to launch an IPO (initial public offering) or be acquired. The investment return required varies on the basis of the company's stage of development. Investors in early-stage companies will demand higher expected returns relative to later-stage investors because the earlier the stage of development, the higher the risk.

Venture capitalists, like all private equity managers, are not passive investors. They are actively involved with the companies in which they invest.

VC funds typically invest in companies and receive an equity interest but may also provide financing in the form of debt (commonly, convertible debt).

Formative-stage financing is for a company that is still in the process of being formed. Its steps are as follows:

a *Pre-seed capital*, or *angel investing*, is capital provided at the idea stage. Funds may be used to transform the idea into a business plan and to assess market potential. The amount of financing at this stage is typically small and provided by individuals, often friends and family, rather than by VC funds.

b *Seed-stage financing*, or *seed capital*, generally supports product development and marketing efforts, including market research. This is the first stage at which VC funds usually invest.

c *Early-stage financing* (early-stage VC), or *start-up stage financing*, is provided to companies that are moving toward operation but have not yet started commercial production or sales, both of which early-stage financing may be injected to initiate.

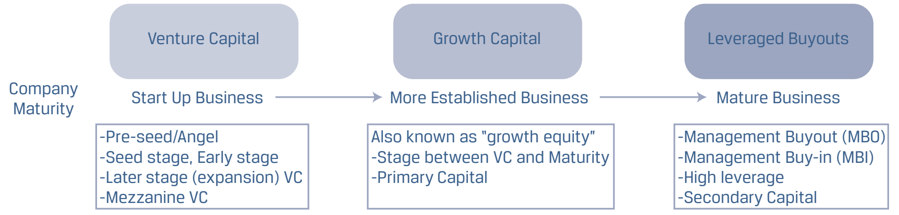
Later-stage financing (expansion VC) is provided after commercial production and sales have begun but before an IPO takes place. Funds may be used to support initial growth, a major expansion (such as a physical plant upgrade), product improvements, or a major marketing campaign.

Mezzanine-stage financing (mezzanine venture capital) prepares a company to go public. It represents the bridge financing needed to fund a private firm until it can complete an IPO or be sold. The term mezzanine-stage financing is used because this financing is provided at the stage between being a private company and being a public company. The focus is on when the financing occurs rather than the financing mechanism itself.

Formative-stage financing is generally carried out by providing ordinary or convertible preferred share sales to the investor or investors, likely the VC fund, while management retains control of the company. Later-stage financing generally involves management selling control of the company to the VC investor; financing is provided through equity and debt, although the fund may also use convertible bonds or convertible preferred shares. The VC fund offers debt financing not for reasons of income generation but, rather, for reasons of recovery and the control of assets in a bankruptcy situation. Simply put, debt financing affords the VC fund more protection than equity does.

To make an investment, a venture capitalist needs to be convinced that the portfolio company's management team is competent and is armed with a solid business plan and strong prospects for growth. Because these investments are not made in mature businesses with years of operational and financial performance history, estimating company valuations on the basis of future prospects is highly uncertain. Accurate estimation is more difficult than in LBO investing, which targets mature, underperforming public companies. Certainty around valuation increases as the portfolio company matures and moves into later-stage financing, but even then, there remains more certainty around LBO investments. Exhibit 13 take us through the growth stages of a company and the types of financing it may receive at each stage:

Exhibit 13 The Private Equity Stage Continuum


Venture Capital
Growth Capital
Leveraged Buyouts
Company Maturity
Start Up Business
More Established Business
Mature Business
-Pre-seed/Angel
-Seed stage, Early stage
-Later stage (expansion) VC
-Mezzanine VC
Also known as "growth equity"
-Stage between VC and Maturity
-Primary Capital
-Management Buyout (MBO)
-Management Buy-in (MBI)
-High leverage
-Secondary Capital

Source: Private Equity Primer, "What Is Private Equity" (12 February 2016). www.pe-primer.com/dealsherpapress/2016/2/12/bbcati52gcwdzrl34pts0x2tl783do (accessed June 18, 2020).

5.2.3 *Other Private Equity Strategies*

Among several other specialties, some private equity firms specialize in *growth capital*, also known as growth equity or minority equity investing. Growth capital generally refers to minority equity investments, whereby the firm takes a less-than-controlling interest in more mature companies that are looking for capital to expand or restructure operations, enter new markets, or finance major acquisitions. Many times, minority equity investing is initiated and sought by the management of the investee company, which is interested in realizing earnings from selling a portion of its shares before the company can go public but still seeks to retain control and participation in the success of the company. Although this scenario occurs most commonly with private companies, publicly quoted companies can seek private equity capital through PIPEs (private investments in public equities). Other private equity strategies secure returns by investing in companies in specific industries.

KNOWLEDGE CHECK

Identify two of the four stages of the private equity continuum and the company growth stage each finances.

Solution:

The stages of the entire PE continuum and their associated growth stage financing are as follows:

1 Venture capital focuses on start-up and seed-stage businesses.

2 Growth equity focuses on more established businesses.

3 Recapitalizations focus on more mature businesses, either healthy or distressed.

4 Buyouts/LBOs focuses on mature businesses.

5.2.4 *Exit Strategies*

Private equity firms seek to improve new or underperforming businesses and then exit them at higher valuations, buying and holding companies for an average of five years. The time to exit, however, can range from less than six months to more than 10 years. Before deciding on an exit strategy, private equity managers take into account the dynamics of the industry in which the portfolio company competes, the overall economic cycle, interest rates, and company performance. Managers pursue the following common exit strategies:

- **Trade sale.** This refers to the sale of a company to a strategic buyer, such as a competitor. A trade sale can be conducted by auction or by private negotiation. Benefits include (a) an immediate cash exit for the private equity fund, (b) the potential willingness of strategic buyers to pay more because they anticipate synergies with their own business, (c) fast and simple execution, (d) lower transaction costs than for an IPO, and (e) lower levels of disclosure and higher confidentiality than for an IPO because private equity firms generally deal with only one other party. Disadvantages of trade sales include (a) possible opposition by management (management may wish to avoid being purchased by a competitor for job security reasons), (b) lower attractiveness to employees of the portfolio company than for an IPO (an IPO allows for the monetizing of shares), (c) a limited universe of potential trade buyers, and (d) a potentially lower price for the sale than would be achieved from an IPO.
- **IPO.** In an initial public offering, the portfolio company sells its shares, including some or all of those held by the private equity firm, to public investors. Advantages of an IPO exit include (a) the potential for the highest price, (b) management approval (because management will be retained), (c) publicity for the private equity firm, and (d) future upside potential (because the private equity firm may choose to remain a large shareholder). Disadvantages of an IPO exit include (a) the high transaction costs paid to investment banks and lawyers, (b) long lead times, (c) the risk of stock market volatility (including short-term focus of some investors), (d) onerous disclosure requirements, (e) a potential lockup period (which requires the private equity firm to retain an equity position for a specified period after the IPO), and (f) the fact that IPOs are usually appropriate only for larger companies with attractive growth profiles.
- **Recapitalization.** Recapitalization in the context of private equity describes the steps a firm takes to increase or introduce leverage to its portfolio company and pay itself a dividend out of the new capital structure. A recapitalization is not a true exit strategy, because the private equity firm typically maintains control; however, it does allow the private equity investor to extract money from the company to pay its investors and improve IRR. A recapitalization may be a prelude to a later exit. However, LP investors should be aware that a recapitalization can be used as a method for the GP to manipulate fund IRRs.

- **Secondary sales.** This approach represents a sale of the company to another private equity firm or another group of investors. With the growth of "dry powder" (Exhibit 12), we have seen an increase in the proportion of secondary sales exits.
- **Write-off/liquidation.** A write-off occurs when a transaction has not gone well: The private equity firm revises the value of its investment downward or liquidates the portfolio company before moving on to other projects.

The foregoing exit strategies may be pursued individually or in concert or may be used for a partial exit strategy. For example, private equity funds may sell a portion of a portfolio company to a competitor via a trade sale and then complete a secondary sale to another private equity firm for the remaining portion. Company shares may also be distributed to fund investors, although such a move is unusual.

5.3 Description: Private Debt

Private debt primarily refers to the various forms of debt provided by investors to private entities. In the past decade, the expansion of the private debt market has been largely driven by private lending funds filling the gap between borrowing demand and reduced lending supply from traditional lenders in the face of tightened regulations following the 2008 financial crisis. We can organize the primary methods of private debt investing into four categories: direct lending, mezzanine loans, venture debt, and distressed debt. The broad array of debt strategies offers not only diversification benefits but also exposure to other investment spheres, such as real estate and infrastructure.

5.3.1 *Direct Lending*

Private debt investors get involved in direct lending by providing capital directly to borrowers and subsequently receiving interest, the original principal, and possibly other payments in exchange for their investment. As with typical bank loans, payments are usually received on a fixed schedule, and the loan itself typically is senior and secured and has covenants in place to protect the lender/investor. The loan is provided by a small number of investors to private and sometimes public entities, and it differs from traditional debt instruments, such as bonds, that are issued to many participants and can be publicly traded in the market.

Direct lending primarily involves private debt firms (or private equity firms with private debt arms) establishing funds with money raised from investors looking for higher-yielding debt. Fund managers will then seek opportunities to deploy that capital, such as providing a loan to a mid-market corporation or extending debt to another private equity fund that is seeking financing for acquisitions. In general, private debt funds provide debt, at higher interest rates, to entities that require capital but lack favorable alternatives to traditional bank lenders, who themselves may be uninterested or unable to provide debt to these borrowers. As in private equity, private debt fund managers conduct thorough due diligence before selecting the fund's investments.

In direct lending, many firms may also provide debt in the form of a *leveraged loan*, which is a loan that is itself levered. In other words, private debt firms that invest in leveraged loans will borrow money to finance the debt that the firm then extends to another borrower. By using leverage, a private debt firm can enhance the return on its loan portfolio.

5.3.2 *Mezzanine Debt*

In private debt, mezzanine debt refers to private credit that is subordinated to senior secured debt but is senior to equity in the borrower's capital structure. Mezzanine debt makes a pool of additional capital available to borrowers beyond senior secured debt, and it is often used to finance LBOs, recapitalizations, corporate acquisitions, and similar transactions. Because of its typically junior ranking and the fact that it is usually unsecured, mezzanine debt is riskier than senior secured debt; to compensate investors for this heightened risk profile, investors commonly demand higher interest rates and may require options for equity participation. Mezzanine debt often comes with additional features, such as warrants or conversion rights, which provide equity participation to lenders/investors, meaning they have the option of converting their debt into equity or purchasing the equity of the underlying borrower under certain circumstances.

5.3.3 *Venture Debt*

Venture debt is private debt funding provided to start-up or early-stage companies with venture capital backing that may be generating little or negative cash flow. Entrepreneurs may seek venture debt, which often takes the form of a line of credit or term loan, as a way to obtain additional financing without further diluting shareholder ownership in their business. Venture debt can complement existing equity financing, allowing shareholders to maintain ownership and control over the company for a longer period of time. Similar to mezzanine debt, venture debt may carry additional features that compensate the investor/lender for the increased risk of default or for the fact that start-up and early-stage companies often lack substantial assets that can be pledged as collateral for the debt. One such feature could grant the lender rights to purchase equity in the borrowing company under certain circumstances.

5.3.4 *Distressed Debt*

Involvement in distressed debt typically entails buying the debt of mature companies with financial difficulty. These companies may be in bankruptcy proceedings, have defaulted on debt, or seem likely to default on debt. Some investors identify companies with a temporary cash-flow problem but a good business plan that will help the company survive and, in the end, flourish. These investors buy the company's debt in expectation of both the company and its debt increasing in value. Turnaround investors buy debt and plan to be more active in the management and direction of the company. They seek distressed companies to restructure and revive.

5.3.5 *Other Private Debt Strategies*

Private debt firms may have specialties other than the aforementioned strategies, one of which is investing in or issuing collateralized loan obligations (CLOs), which are leveraged structured vehicles that are collateralized by a portfolio of loans covering a diverse range of tranches, issuers, and industries. A CLO manager extends several loans to corporations—usually to firms involved in LBOs, corporate acquisitions, or similar transactions—pools these loans together, and then divides that pool into various tranches of debt and equity that differ in seniority and security. The CLO manager then sells each tranche to different investors according to their risk profile; the most senior portion of the CLO will be the least risky, and the most junior portion of the CLO (i.e., equity) will be the riskiest.

Another type of debt that could be directly extended to borrowers is **unitranche debt**. Unitranche debt consists of a hybrid or blended loan structure that combines different tranches of secured and unsecured debt into a single loan with a single, blended interest rate. Since unitranche debt is a blend of secured and unsecured debt, the interest rate on this type of loan will generally fall in between the interest rates often demanded on secured and unsecured debt, and the unitranche loan itself will usually be structured between senior and subordinated debt in terms of priority ranking.

Some private debt firms invest in real estate debt or infrastructure debt. *Real estate debt* refers to loans and other forms of debt provided for real estate financing, where a specified real estate asset or property serves as collateral. *Infrastructure debt* encompasses the many forms of debt used to finance the construction, operation, and maintenance of infrastructure assets.

Private debt firms may also provide *specialty loans*, where debt is extended to niche borrowers in specific situations. For example, litigation finance is the practice of a specialist funding company providing debt to clients, usually plaintiffs in litigation, for their legal fees and expenses in exchange for a portion of any case winnings.

KNOWLEDGE CHECK

Which of the following is not considered a strategy in private debt investing?

A Direct lending

B Recapitalization

C Mezzanine debt

Solution:

B is correct. Recapitalization is when a private equity firm increases leverage or introduces it to the company and pays itself a dividend.

KNOWLEDGE CHECK

Which of the following forms of debt are likely to have additional features, such as warrants or conversion rights?

A Mezzanine debt

B Venture debt

C Both A and B

Solution:

C is correct. Both mezzanine and venture debt are likely to have additional features, such as warrants and conversion rights, to compensate debt holders for increased risk, including the risk of default.

5.4 Risk/Return of Private Equity

The higher-return opportunities that private equity funds may provide relative to traditional investments are a function of their ability to invest in private companies, their influence on portfolio companies' management and operations, and their use of leverage. Investing in private equity, including venture capital, is riskier than investing in common stocks. Investors require a higher return for accepting higher risk, including illiquidity and leverage risks.

Exhibit 14 shows the mean annual returns for the Cambridge Associates US Private Equity index, the NASDAQ index, and the S&P 500 Index for a variety of periods ending 31 December 2019. Because private equity returns are based on IRR, however, it is difficult to identify reliable benchmarks for comparing returns to investments.

Some investors may use the public market equivalent (PME) to match the timing of cash flows with the public market. Other investors may use an index of publicly traded private equity firms as a benchmark for private equity returns.

Exhibit 14 Comparison of Mean Annual Returns for US Private Equity and US Stocks

Index	1 Year	5 Years	10 Years	20 Years	25 Years
US private equity	13.95	12.00	14.35	11.08	13.24
NASDAQ	35.23	13.63	14.74	4.03	10.43
S&P 500	31.49	11.70	13.56	6.06	10.22

Source: US Private Equity index and selected benchmark statistics, 31 December 2019, Cambridge Associates.

Published private equity indexes may be an unreliable measure of performance. Measuring historical private equity performance is challenging; as with hedge funds, private equity return indexes rely on self-reporting and are subject to survivorship, backfill, and other biases, which typically lead to an overstatement of returns. Moreover, prior to 2009, in the absence of a liquidity event, private equity firms did not necessarily mark their investments to market. This failure to mark to market leads to an understatement of measures of volatility and correlations with other investments. Thus, data adjustments are required to more reliably measure the benefits of private equity investing. Investors should require a higher return for accepting a higher risk, including illiquidity and leverage risks.

5.5 Risk/Return of Private Debt

Private debt investments may provide higher-yielding opportunities to fixed-income investors seeking increased returns relative to traditional bonds. Private debt funds may generate higher returns by taking opportunistic positions based on market inefficiencies. Indeed, private lending funds stepped into the financing gap left by traditional lenders following the 2008 financial crisis. Investors in private debt could realize higher returns from the illiquidity premium, which is the excess return investors require to compensate for lack of liquidity, and may benefit from increased diversification in their portfolios.

The potential for higher returns is connected to higher levels of risk. Private debt investments vary in risk and return, with senior private debt providing a steadier yield and moderate risk and mezzanine private debt carrying higher growth potential, equity upside and higher risk. As a whole, however, investing in private debt is riskier than investing in traditional bonds. Investors should be aware of these risks, which include illiquidity and heightened default risk when loans are extended to riskier entities or borrowers in riskier situations.

Exhibit 15 shows annualized returns and standard deviations for private debt funds (2004–2016). Based on these figures, private debt funds appear to offer an attractive risk–return trade-off.

Exhibit 15 Historical Annualized Returns and Standard Deviation of Returns for Private Debt Funds, 2004–2016

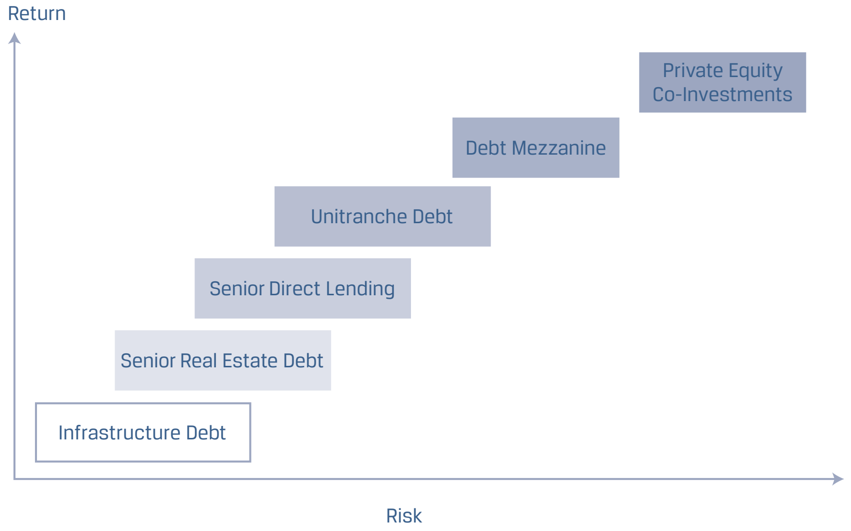
Private Debt Fund	Annualized Return	Standard Deviation
Mezzanine	7.4%	2.5%
Distressed	7.9%	6.7%
Direct lending only:		
All direct lending	9.2%	2.0%
Direct lending (excl. mezzanine)	11.1%	3.4%

Source: S. Munday, W. Hu, T. True, and J. Zhang, "Performance of Private Credit Funds: A First Look," Institute for Private Capital (7 May 2018, pp. 32–3).

Although there is no widely used benchmark for private debt, various indexes may serve investors' purposes as long as they understand the pros and cons of each index. The ICE Bank of America Merrill Lynch Global High Yield Index (HYI) tracks publicly traded, non-investment-grade corporate bonds of large, international corporations; the S&P/LSTA U.S. Leveraged Loan Index (LLI) covers syndicated and over-the-counter traded, non-investment-grade loans for large US corporations; the S&P Net Total Return BDC Index tracks publicly traded business-development companies (BDCs); and the Cliffwater Direct Lending Index (CDLI) is based on quarterly SEC filings covering more than 60 private and public BDCs.

Investments in private equity and private debt vary in terms of risk and return because of various factors, including but not limited to ranking in an entity's capital structure. Exhibit 16 contains a graph that plots private equity and private debt categories by their risk and return levels. (Mirroring the risk–return pathway for traditional equity and debt investing, Exhibit 16 plots private equity and private debt using categories that are broadly similar to how we have defined them. Note the trade-off as investors select between junior and senior debt and between equity and debt.)

Exhibit 16 Private Capital Risk and Return Levels by Category



Source: Based on graph from Leon Sinclair, “The Rise of Private Debt,” IHS Markit (7 August 2017).

KNOWLEDGE CHECK

Fill in the blank by choosing from among the following options:

____________ is the type of private debt expected to have the greatest excess return potential.

A Unitranche

B Mezzanine

C Infrastructure

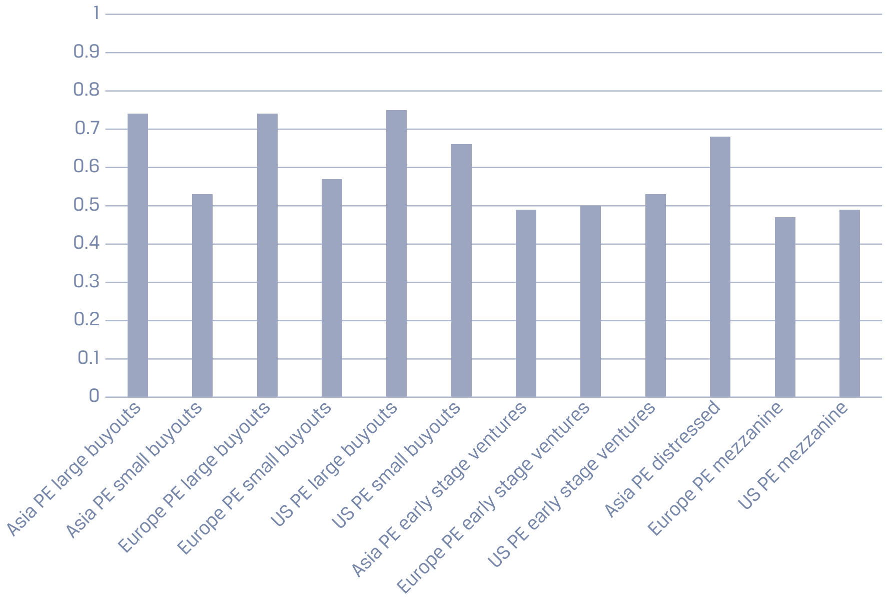
Solution:

B is correct. As a junior form of subordinated debt, mezzanine private debt private debt offers higher growth potential, equity upside, and higher risk, with the comparatively highest returns. Infrastructure debt is senior and poses the lowest risk. Unitranche debt is less risky than subordinated debt but riskier than infrastructure debt and is a blend of secured and unsecured debt; its interest rate generally falls in between the interest rates often demanded on secured and unsecured debt, and the loan itself is usually structured between senior and subordinated debt.

5.6 Diversification Benefits of Investing in Private Capital

Investments in private capital funds can add a moderate diversification benefit to a portfolio of publicly traded stocks and bonds, with correlations with public market indexes varying from 0.47 to 0.75, as shown in Exhibit 17. And if investors identify skillful fund managers, they may benefit from excess returns given the additional leverage, market, and liquidity risks.

Exhibit 17 Private Capital's Average Correlations with Public Market Indexes, 2013–2018



Source: MSCI. www.msci.com/www/blog-posts/did-private-capital-deliver-/01697376382 (accessed 20 July 2020).

Private equity investments may offer vintage diversification because capital is not deployed at a single point in time but invested over several years. Private debt investments, which offer more options than bonds and public forms of traditional fixed income, can also serve diversification efforts. Ultimately, investors should prudently avoid concentration risk, aiming to diversify across different managers, industries, and geographies.

PRACTICE QUESTIONS FOR LOS E

Question 1

A collateralized loan obligation specialist is *most likely* to:

A sell its debt at a single interest rate.

B cater to niche borrowers in specific situations.

C rely on diverse risk profiles to complete deals.

Solution:

C is correct. A CLO manager will extend several loans to corporations (usually to firms involved in LBOs, corporate acquisitions, or other similar types of transactions), pool these loans, and then divide that pool into various tranches of debt and equity that range in seniority and security. The CLO manager will then sell each tranche to different investors according to their risk profiles; the most senior portion of the CLO will be the least risky, and the most junior portion of the CLO (i.e., equity) will be the riskiest.

A is incorrect because with the different CLO tranches having distinct risks varying with their seniority and security, they will be priced over a range of interest rates. In contrast, unitranche debt combines different tranches of secured and unsecured debt into a single loan with a single, blended interest rate.

B is incorrect because debt extended to niche borrowers in specific situations is more commonly offered through specialty loans. For example, in litigation finance, a specialist funding company provides debt to a client to finance the borrower's legal fees and expenses in exchange for a portion of any case winnings.

Question 2

Private capital is:

A accurately described by the generic term "private equity."

B a source of diversification benefits from both debt and equity.

C predisposed to invest in both the debt and equity of a client's firm.

Solution:

B is correct. Investments in private capital funds can add diversity to a portfolio composed of publicly traded stocks and bonds because they have less-than-perfect correlation with those investments. There is also the potential to offer further diversification within the private capital asset class. For example, private equity investments may also offer vintage diversification since capital is not deployed at a single point in time but is invested over several years. Private debt provides investors with the opportunity to diversify the fixed-income portion of their portfolios since private debt investments offer more options than bonds and other public forms of traditional fixed income.

A is incorrect because although private equity is considered by many to be the largest component of private capital, using "private equity" as a generic term could be less accurate and possibly misleading since other private forms of alternative finance have grown considerably in size and popularity.

C is incorrect because although many private investment firms often have private equity and private debt arms, these teams typically won't invest in the same assets or businesses to avoid overexposure to a single investment.

Question 3

The first stage of financing at which a venture capital fund *most likely* invests is the: (2020 Q13)

A seed stage.

B mezzanine stage.

C angel investing stage.

Solution:

A is correct. The seed stage supports market research and product development and is generally the first stage at which venture capital funds invest. The seed stage follows the angel investing stage. In the angel investing stage, funds are typically provided by individuals (often friends or family), rather than a venture capital fund, to assess an idea's potential and to transform the idea into a plan. Mezzanine-stage financing is provided by venture capital funds to prepare the portfolio company for its IPO.

Question 4

A private equity fund desiring to realize an immediate and complete cash exit from a portfolio company is *most likely* to pursue: (2020 Q18)

A an IPO.

B a trade sale.

C a recapitalization.

Solution:

B is correct. Private equity funds can realize an immediate cash exit in a trade sale. Using this strategy, the portfolio company is typically sold to a strategic buyer.

Question 5

Angel investing capital is typically provided in which stage of financing? (2020 Q24)

A Later stage

B Formative stage

C Mezzanine stage

Solution:

B is correct. Formative-stage financing occurs when the company is still in the process of being formed and encompasses several financing steps. Angel investing capital is typically raised in this early stage of financing.

Question 6

Private equity funds are *most likely* to use:

A merger arbitrage strategies.

B leveraged buyouts.

C market-neutral strategies.

Solution:

B is correct. The majority of private equity activity involves leveraged buyouts. Merger arbitrage and market neutral are strategies used by hedge funds.

6 NATURAL RESOURCES

f explain investment characteristics of natural resources

6.1 Overview of Natural Resources

We can define natural resources as a unique asset category comprising commodities and raw land used for farming and timber. Some investors also include real estate assets and infrastructure in this category to form a “real assets” asset class, which can include inflation-protected securities and miscellaneous investments designed to protect against an increase in consumer prices. Others include timberland and farmland in their real estate portfolios. Regardless, the particular investment characteristics of natural resources, the specialized knowledge that their investors require, and their inclusion in more portfolios today than a decade ago call for a separate examination of the sector.

Commodities are broadly familiar to both professional investors and the general public: Crude oil, soybeans, copper, and gold are all commodities. Commodities are considered either "hard" (those that are mined, such as copper, or extracted, such as oil) or "soft" (those that are grown over a period of time, such as livestock, grains, and cash crops, such as coffee). Timberland investment involves ownership of raw land and the harvesting of its trees for lumber, thus generating an income stream and the potential for capital gain; it has formed part of large institutional portfolios for decades. Farmland as an investment is a more recent phenomenon, with only a few dedicated funds involved. With population growth, weather, and water management becoming more topical, however, investors may turn to these sustainable land assets to address such concerns in their portfolio.

An examination of natural resources investment requires an important clarification up front: The aspects discussed here are associated with investing in the physical land and products that come from that land—petroleum, metals, grains—not from the companies that produce those products. Until about 20 years ago, the only commonly available investment vehicles related to this asset class were, indeed, financial instruments, such as stocks and bonds. Nowadays, the wide variety of direct investments via ETFs, limited partnerships, REITS, swaps, and futures opens the door for almost everyone to invest in these assets directly.

This section covers both commodities and land (farmland and timberland) in each subsection, leaving the broader real estate category for the next section.

6.2 Characteristics of Natural Resources

Here, we offer an overview and describe properties that separate these investments from the more commonly known asset classes addressed earlier in this reading.

6.2.1 *Commodities*

Commodities are physical products that can be standardized on quality, location, and delivery for the purpose of investing. Returns on commodity investments are primarily based on changes in price rather than on income from interest, dividends, or rent. In fact, holding commodities (i.e., the physical products) incurs transportation and storage costs. Therefore, trading in physical commodities is primarily limited to a smaller group of entities that are part of the physical supply chain. Most commodity investors do not trade actual physical commodities but, rather, trade commodity derivatives. The underlying asset of the commodity derivative may be a single commodity or an index of commodities.

Because the prices of commodity derivatives are, to a significant extent, a function of the underlying commodity prices, it is important to understand the physical supply chain and general supply–demand dynamics. In fact, the price volatility associated with commodity derivative indexes is highly correlated with the underlying physical goods. This fact should make intuitive sense because the supply chain participants use futures to hedge their forward purchases and sales of the physical commodities. Imagine how uncertain food prices would be if a soybean farmer could not hedge his crop risk. Investors, sometimes referred to as speculators, trade commodity derivatives in search of profit based largely on changes or expected changes in the price of the underlying commodities. Such non-hedging investors include retail and institutional investors, hedge funds, proprietary desks within financial institutions, and trading desks operating in the physical supply chain. (Yes, commodity producers and consumers both hedge *and* speculate on commodity prices.)

Commodity sectors include precious and base (i.e., industrial) metals, energy products, and agricultural products. Exhibit 18 offers examples of each type. The relative importance, amount, and price of individual commodities evolve with society's preferences and needs. For example, the increasing industrialization of China, India, and other emerging markets has driven strong global demand for commodities. Developing markets need increasing amounts of oil, steel, and other materials to support manufacturing, infrastructure development, and the consumption demands of their populations. Emerging technologies, such as advanced cell phones and electric vehicles, create demand for new materials and destroy demand for old resources. The supply and demand for specific commodities evolves over time.

Exhibit 18 Examples of Commodities

Sector	Sample Commodities
Energy	Oil, natural gas, electricity, coal
Base metals	Copper, aluminum, zinc, lead, tin, nickel
Precious metals	Gold, silver, platinum
Agriculture	Grains, livestock, coffee
Other	Carbon credits, freight, forest products

Commodities may be further classified on the basis of physical location and grade or quality, for example. There are many grades and delivery locations for crude oil, and there are many grades and delivery locations for wheat. Commodity derivative contracts, therefore, specify quantity, quality, maturity date, and delivery location.

KNOWLEDGE CHECK

Which of the following can be considered as an investment in the commodities asset class?

A A set of rare antique coins, some of which are made of silver and gold

B A lease on an oil tanker shipping oil between Saudi Arabia and China

C Ownership of a battery factory that uses industrial metals

D A metric tonne of coffee packed in jute bags in a warehouse in Brazil

Solution:

D is correct, assuming the investor owns the commodity directly, not via a derivatives contract. A describes an investment in collectables (whose value has more to do with the rarity of the collectible rather than the actual value of the silver or gold). B describes a financial contract—effectively, a lease or bond. C is also a financial contract that pays on the use of commodities but depends on many other factors (technology, marketing, rent, and employees) for returns.

In order to be transparent, investable, and replicable, commodity indexes typically use the price of the futures contracts rather than the prices of the underlying commodities. As a result, the performance of a commodity index can differ from the performance of the physical commodities. Different commodity indexes are composed of different commodities and weight their component commodities differently. Thus, exposures to specific commodities and commodity sectors vary. However, the low correlation between commodities and other asset classes (e.g., stocks and bonds) means that an investor can achieve improved portfolio diversification regardless of the index she chooses.

Futures are the basis for the vast majority of commodity investment, and each contract is for the future delivery or receipt of a set amount of the commodity in question—for example, 1,000 barrels of oil or 10 metric tonnes of cocoa. This means that the price can be formalized in the following form:

$$\text{Futures price} \approx \text{Spot price}(1 + r) + \text{Storage costs} - \text{Convenience yield},$$

where r is the period's short-term risk-free interest rate. The collateral posted for the contract should also be included if a total return calculation is required. The storage and interest costs together are sometimes referred to as the "cost of carry" or the "carry." We recognize that the buyer of the futures contract has no immediate access to the commodity and, therefore, cannot benefit from it, and so the futures price is adjusted for the loss of convenience, a value that varies. For example, the convenience yield from possessing heating oil in in Finland during the winter is higher than the convenience yield from possessing heating oil in Finland during the summer.

Futures prices may be higher or lower than spot prices depending on the convenience yield. When futures prices are higher than the spot price, the commodity forward curve is upward sloping, and the prices are referred to as being in **contango**. Contango generally occurs when there is little or no convenience yield. When futures prices are lower than the spot price, the commodity forward curve is downward sloping, and the prices are referred to as being in **backwardation**. Backwardation occurs when the convenience yield is high. As a rule of thumb, a contango scenario generally lowers the return of the long-only investor, and a backwardation scenario enhances it.

The pricing of derivatives and the theories of commodity pricing are covered in other readings in the CFA Program curriculum.

6.2.2 *Timberland and Farmland*

Real estate property ownership is represented by a title and may reflect access to air rights, mineral rights, and surface rights in addition to building and land-use rights. For the purposes of this section, we are discussing land owned or leased for the benefit of the returns it generates from crops and timber. Given that these resources consume carbon as part of the plant life cycle, their value comes not just from the harvest but also from the offset to human activity. Water rights are also part of the direct and implied value of these properties; conservation easements may create value by supporting traditions and nature conservation. As interest in investments that adhere to environmental, social, and governance (ESG) considerations grows, arable land may fit the criteria.

Sustained interest in these investments stems from their global nature (everyone eats, everyone requires shelter), the income generated from selling crops, inflation protection from holding land, and the degree of insulation they offer from financial market volatility. US farmland, for example, enjoyed positive returns during the two periods after World War II when US GDP declined significantly (1973-1975 and 2007-2009) and during the three periods after World War I (1915-1920, 1940-1951, and 1967-1981) when the United States experienced higher-than-normal inflation. Timberland has been part of institutional and ultra-high-net-worth portfolios for decades, typically trading in larger units of land. Farmland can be found in much smaller sizes—perhaps tens of or a few hundred acres. Many farms are still family owned, as is 98% of US farmland. One of the main challenges of these investments cited by industry participants is their long market cycle, particularly in new-growth forest and crops that are picked, such as fruit.

Timberland offers an income stream based on the sale of trees, wood, and other timber products and has been not highly correlated with other asset classes. Timberland can be thought of as both a factory and a warehouse. Timber (trees) can be grown and easily stored by simply not harvesting them. This characteristic offers the flexibility of harvesting more trees when timber prices are up and delaying harvests when prices are down. The three primary return drivers are biological growth, changes in spot prices and futures prices of lumber (cut wood), and changes in the price of the underlying land.

Farmland is often perceived to provide a hedge against inflation. Similar to timberland, the returns include an income component related to harvest quantities and agricultural commodity prices. Farmland consists mainly of row crops that are planted and harvested (i.e., more than one round of planting and harvesting can occur in a year depending on the crop and the climate) and permanent crops that grow on trees (e.g., nuts) or vines (e.g., grapes). Unlike timberland, farm products must be harvested when ripe, so there is little flexibility in production. Therefore, commodity futures contracts can be combined with farmland holdings to generate an overall hedged return. Recall that a farm is inherently "long" the crop and, therefore, will sell futures that require delivery at the time of the harvest. Farmland may also be used as pastureland for livestock. Similar to timberland, farmland has three primary return drivers: harvest quantities, commodity prices (e.g., the price of corn), and land price changes.

KNOWLEDGE CHECK

Alexandra is considering buying a tract of farmland for long-term capital appreciation and current income. Which of the following factors play a role in evaluating the attractiveness of the investment?

A The land's rights to preferential water access

B The land's proximity to a large, privately held nature reserve

C The land's soil chemical composition allowing the growth of a variety of crops

D All of the above

Solution:

The answer is D. A is important to ensure that crops can be grown with a reasonable chance of success. B allows for the land to be rented to the stewards of the nature reserve and to lay fallow instead of growing food for human consumption. Nature preserve foundations are interested in the land around them because (1) runoff and pesticide use affect the plants and animals in the preserve, (2) migration of animals in the preserve may incur onto the farmland, and (3) foundations often look to add to their holdings, allowing for a future sale of the land. C allows for the possibility of being able to react to market preferences and more stable productivity and thus more reliable income.

KNOWLEDGE CHECK

Large institutional investors consider timberland investments because:

A The small parcel sizes permit fine tuning of their holdings across geography and wood types.

B The optionality around harvesting gives investors the choice between cutting trees for lumber and current income or letting them grow another year for future gain.

C The short return history allows for many alpha opportunities by knowledgeable active managers.

D Clear-cutting trees and destroying nature is appreciated by ESG investors, which are becoming a larger portion of the investment universe.

Solution:

B is the correct answer. A is incorrect because the parcel sizes are generally large, especially compared with farmland. C is incorrect because the return history is relatively long, not short. D states the opposite reason why ESG investors may be interested in timberland investments—for the opportunity to create "conservation zones."

6.3 Risk/Return of Natural Resources

Although we present commodities, farmland, and timberland in the same section, they do have different return drivers and cycles. Commodities are priced on a second-by-second basis on public exchanges, whereas land generally has an infrequent pricing mechanism and may include estimates as opposed to actual transactions. Keeping these market structure differences in mind will help investors consider their relative benefits and challenges.

6.3.1 *Risk/Return: Commodities*

The arguments for investing in commodities include the potential for returns, portfolio diversification, and inflation protection. Investors may choose commodities if they believe prices will increase in the short or intermediate term. Since commodity prices directly feed into inflation index level calculations, commodities serve as a real hedge against inflation risk even when they yield little or no real return. Commodity futures contracts may offer investors a liquidity premium or other trading opportunities, creating the prospect for a real return greater than zero. Note that commodity investments, especially when combined with leverage, exhibit high volatility (see Exhibit 19), which has led to many well-publicized losses among commodity players.

Exhibit 19 Historical Returns of Commodities, 1990–Q1 2020 (quarterly data)

	Global Stocks	Global Bonds	Commodities
Annualized return	4.8%	5.3%	−1.7%
Annualized standard deviation	16.2%	6.0%	23.8%
Worst calendar year	−43.5%	−5.2%	−46.5%
	(2008)	(1999)	(2008)
Best calendar year	31.6%	19.7%	49.7%
	(2003)	(1995)	(2000)

Sources: Global stocks = MSCI ACWI; global bonds = Bloomberg Barclays Global Aggregate Index; commodities = S&P GSCI Total Return.

Before investors dismiss commodities as a source for returns, let us examine the period before the Global Financial Crisis (GFC) in 2008–2009. If we look at the world economy before central banks began injecting trillions of dollars, euros, yuan, yen, and so on, in liquidity and capital market purchases, we can see that commodities held their own in terms of returns. In fact, commodities well outperformed global stocks. Therefore, asset allocators need to ask whether conditions or sentiment can return to the pre-GFC situation when inflation was a factor and government bonds yielded mid- to high single digits.

Exhibit 20 Historical Returns of Commodities, 1990–Q4 2007 (quarterly data)

	Global Stocks	Global Bonds	Commodities
Annualized return	5.2%	7.0%	6.8%
Annualized standard deviation	17.0%	6.3%	18.3%
Worst calendar year	–20.5%	–5.2%	–35.7%
	(2002)	(1999)	(1998)
Best calendar year	31.6%	19.7%	49.7%
	(2003)	(1995)	(2000)

Sources: Global stocks = MSCI ACWI; global bonds = Bloomberg Barclays Global Aggregate Index; commodities = S&P GSCI Total Return.

Commodity spot prices are a function of supply and demand, the costs of production and storage, the value to users, and global economic conditions. Supplies of physical commodities are determined by production and inventory levels and secondarily by the actions of non-hedging investors. Demand for commodities is determined by the needs of end users and secondarily by the actions of non-hedging investors. Investor actions can both dampen and stimulate commodity price movements, at least in the short term.

Producers cannot alter commodity supply levels quickly because extended lead times are often needed to affect production levels. For example, agricultural output may be altered by planting more crops and changing farming techniques, but at least one growing cycle is required before there is an effect. And for agricultural products, at least one factor that is outside the producer's control—the weather—will significantly affect output. Building the necessary infrastructure for increased oil and mining production may take many years: It is a matter not just of developing the mine itself but also of the necessary transportation and smelting components. For commodities, suppliers' inability to quickly respond to changes in demand levels may result in supply levels that are too low in times of economic growth and too high when the economy slows. And despite advancing technology, the cost of new supply may grow over time. For example, new energy and mineral exploration is frequently more expensive because the easy discoveries tend to be exploited first. If fixed production costs are high, producers are unlikely to produce more than what is needed to meet anticipated demand, and they are unlikely to maintain more than modest levels of inventory, leading to the risk of shortages and price spikes.

Overall demand levels are influenced by global manufacturing dynamics and economic growth. Manufacturing needs can change in a period of months as orders and inventories vary. Investors seek to anticipate these changes by watching inventories closely and monitoring economic conditions, including those relating to government policy and growth forecasts. When demand levels and investors' orders to buy and sell during a given period change quickly, the resulting mismatch of supply and demand may lead to price volatility.

6.3.2 *Risk/Return: Timberland and Farmland*

Turning to land, Exhibit 21 provides a comparison of returns on US timber and farmland. The National Council of Real Estate Investment Fiduciaries (NCREIF) constructs a variety of appraisal-based indexes for property, timber, and farmland. Over the 1990–Q1 2020 period, farmland had the highest annualized return and timber had the highest standard deviation.

Exhibit 21 Historical Returns of US Real Estate Indexes, 1990–Q1 2020, Quarterly Data

	NCREIF Data	
	Timber	**Farmland**
Annualized return	9.1%	11.0%
Annualized standard deviation	7.0%	5.9%
Worst calendar year	–5.2% (2001)	2.0% (2001)
Best calendar year	37.3% (1992)	33.9% (2005)

Although the data in Exhibit 21 make farmland appear to be a very attractive investment, they do not tell the whole story. Liquidity is very low and the risk of negative cash flow high because fixed costs are relatively high (remember that the land must be cared for and crops need fertilizer, seed, and so on) and revenue is highly variable based on the weather. The risks of timber and farmland are similar to those of other real estate investments in raw land, but we highlight weather as a unique and more exogenous risk for these assets compared to traditional commercial and residential real estate properties. Drought and flooding can dramatically decrease the harvest yields for crops and thus the income stream expected by investors. The second primary risk is the international competitive landscape. Although real estate is often considered a local investment, productive land generates commodities that are globally traded and consumed. For example, there have been interruptions in world trade, and growing agricultural competition has resulted in declining grain prices. Therefore, looking ahead, these returns seem unlikely to be repeated in the next 30 years. Timber and farmland investments should consider the international context as a major risk factor.

KNOWLEDGE CHECK

Describe a risk of farmland that distinguishes it from real estate investment in raw land.

Solution:

There are two significant risks differentiating farmland from raw land investment:

1. Weather is a more unique and exogenous risk for farmland, with drought or flooding dramatically decreasing many crop yields and thus the expected income stream.
2. Productive land generates globally traded and consumed commodities. This international competitive landscape can result in interruptions in world trade, growing foreign agricultural competition, and resulting declines in crop prices.

6.4 Diversification Benefits of Natural Resources

Although they often entail higher transaction costs and higher informational hurdles, alternative assets generally offer diversification as a major benefit. Where that diversification comes from needs to be understood in order for managers to better set investor expectations.

6.4.1 *Diversification Benefits: Commodities*

Commodity investing may be attractive to investors not only for the potential profits but also because of the following perceptions: (1) Commodities are effective hedges against inflation, which is to say that commodity prices have historically been positively correlated with inflation, because they are an input to prices, and (2) commodities are effective for portfolio diversification, which is to say that commodity returns have historically had a low correlation with other investment returns. Institutional investors, particularly endowments, foundations, and increasingly corporate and public pension funds and sovereign wealth funds, are allocating more of their portfolios to commodities and commodity derivatives.

The portfolio diversification argument is based on the observation that commodities historically have behaved differently from stocks and bonds during the business cycle. Exhibit 22 shows the quarterly correlation between selected commodity, global equity, and global bond indexes. In the period from 1990 through Q1 2020, commodities exhibited a low correlation with traditional assets; the correlations of commodities with global stocks and global bonds were 0.37 and 0.05, respectively. The correlations of stocks, bonds, and commodities are expected to be positive because each of the assets has some exposure to the global business cycle. Note that the selected commodity index, the S&P GSCI (Goldman Sachs Commodity Index), is heavily weighted toward the energy sector and that each underlying commodity may exhibit unique behavior.

Exhibit 22 Historical Commodity Return Correlations, 1990–Q1 2020 (quarterly data)

	Global Stocks	Global Bonds	Commodities	US CPI
Global stocks	1	0.09	0.37	0.21
Global bonds		1	0.05	–0.03
Commodities			1	0.64

Sources: Global stocks = MSCI ACWI; global bonds = Bloomberg Barclays Global Aggregate Index; commodities = S&P GSCI Total Return.

The argument for commodities as a hedge against inflation is related to the fact that some commodity prices are a component of inflation calculations. Commodities, especially energy and food, affect consumers' cost of living. The positive correlation between quarterly commodity price changes and quarterly changes in the US CPI of 0.64 supports this assertion. In contrast, the quarterly return correlations between the US CPI and global stocks and global bonds are close to zero. The volatility of commodity prices, especially energy and food, is much higher than that of reported consumer inflation. Note that consumer inflation is computed from many products used by consumers, including housing, that change more slowly than commodity prices, and inflation calculations use statistical smoothing techniques and behavioral assumptions. However, even in the recent period of low inflation (or even negative inflation in many countries), investing in commodities outperformed global stocks and bonds on average during the same calendar year if inflation was relatively high (i.e., US CPI greater than 2%; see Exhibit 23).

Exhibit 23 Historical Asset Class Returns When US CPI Was >2% versus <2%, 1990–2019 (quarterly data)

	Global Stocks	Global Bonds	Commodities
Higher inflation	+10.9%	+7.1%	+15.6%
Lower inflation	+0.3%	+2.7%	−20.0%

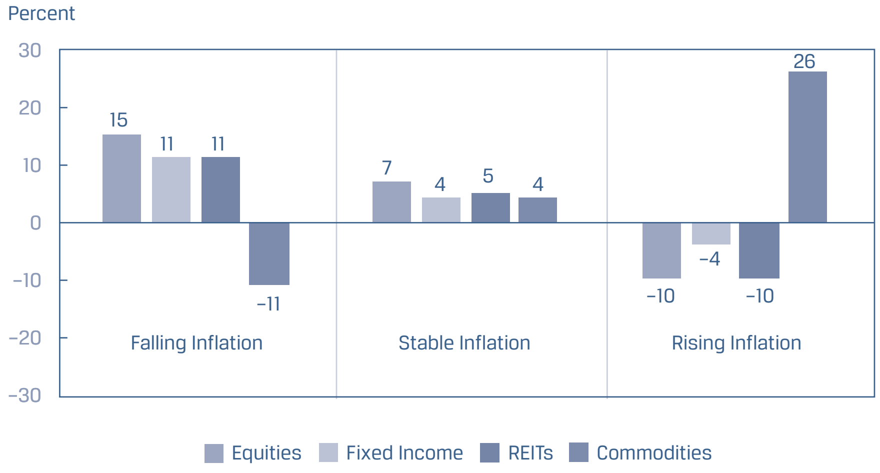
Sources: Global stocks = MSCI ACWI; global bonds = Bloomberg Barclays Global Aggregate Index; commodities = S&P GSCI Total Return.

In addition, Exhibit 23 affirms the sensitivity of commodities to inflation as measured by its relatively high correlation.

Note that changes in commodity prices will coincidentally affect inflation because those changes are included in inflation measures. However, in looking at the weight of food and energy in inflation calculations, they are a relatively small portion. A review of the US CPI calculations indicates an approximate 13% weight of the commodity price impacts. Furthermore, this number is overstated because these price changes as measured by CPI include all the other costs to consumers, such as gasoline taxes, supermarket rent and labor costs, and brand marketing. Therefore, even though there is a direct impact from changes in commodity costs on consumer inflation, it is marginal for the purposes of this discussion.

Exhibit 24 outlines a number of common assets in various regimes of inflation with certain sector callouts that highlight their impact. One example is that mining equities have performed notably better than overall equities in periods of high inflation because of the operational leverage inherent in mining businesses (high fixed costs, low operating costs). This example again highlights the diversification opportunities that can exist between commodities and traditional asset classes.

Exhibit 24



Source: Wellington Management Company LLP

6.4.2 *Diversification Benefits: Timberland and Farmland*

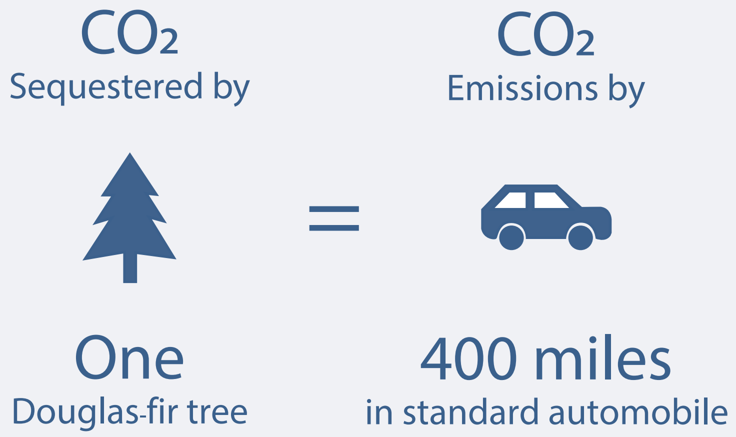
Investors look at diversification in many forms; the numbers tell part of the story, but there are qualitative considerations as well. For example, investors can adhere to ESG principles of responsible and sustainable investing by including timberland and farmland in their portfolios. The following example provides a case study on timberland assets and how investing in timberland can help mitigate climate change.

TIMBERLAND CASE STUDY: CAMPBELL GLOBAL

Investing Responsibly in Timberland Assets: A Climate-Conscious Case Study

Campbell Global (CG) is a global investment manager focused on forest and natural resources investments. Based in Portland, Oregon, with offices in 14 US states and New Zealand, the firm has nearly four decades of experience in sustainable value creation. CG is committed to managing its forests in a manner that promotes the best long-term interests of its clients, while also striving to address economic and ESG considerations.

Figure 1



Source: Campbell Global, LLC

In addition to their economic value, forests serve as vast carbon sinks, with trees removing CO_2 from the atmosphere and using it as carbon storage. In one year, a single Douglas fir tree, a common commercial timber species in the US Pacific Northwest, stores the CO_2 equivalent of driving 400 miles in a standard automobile.[5] Globally, it is estimated that the Earth's forests absorb as much as 30% of human-induced CO_2 emissions.[6]

Sustainably harvested wood products and materials also store atmospheric CO_2 long after they have been removed from a forest, with one cubic meter of wood capable of storing nearly a metric ton of CO_2[7]. In addition to carbon sequestration, forests also provide other benefits, such as clean water and wildlife habitat, recreational opportunities, and a source of living-wage jobs in rural communities. These attributes positively align with the UN's Sustainable Development Goals and contribute to advancing the UN's mission for a sustainable future for all. In light of these considerations, there is increasing awareness that well-managed forests are a critical component of any global climate change strategy.

CG uses scenario analyses to identify climate-related risks beginning at a broad country-level scale, narrowing down to a specific property, and then testing the impact of various risks to site suitability now and into the future. Factors analyzed to gauge climate risks include precipitation patterns, temperature fluctuations, the severity of weather events, the presence of pests or disease, and the annual average growth rates for commercial tree species. While many climate-related risks in forestry are mitigated through active management, during this iterative process CG analyzes both the potential positive and negative impacts associated with these risks to assess potential changes in net asset value. The following table illustrates climate risks evaluated, their impact on the forest, and our ability to mitigate the risks.

Climate Risk	**Implication**	**CG Mitigants**
Change in temperature	Increased fire danger	Property-specific fire plans; re-evaluate target regions/country for investment
Change in precipitation patterns	Changes in tree species range; increased drought and related fire risk	Vegetation suitability modeling and genetic tree improvement; re-evaluate target regions/country for investment
Frequency of extreme weather events	Loss of standing timber from wind events	Re-evaluate target regions; property-specific response plans; geographically diverse portfolio construction

(continued)

---

[5] Estimated carbon storage of a 20-inch Douglas fir using the National Tree Benefit Calculator.

[6] Y. Pan et al., "A Large and Persistent Carbon Sink in the World's Forests," *Science* (19 August 2011): 988–93.

[7] Oregon Forest Resources Institute, "Forest to Frame" (2017).

Climate Risk	Implication	CG Mitigants
Presence of pests or disease	Early onset and increased frequency of individual tree mortality	Immediate treatment, which may include removal of effected trees to prevent further spread of pests or disease in the forest
Change in growth	Increased or decreased growth rates	Effects will vary by region, may influence planting stock decisions; re-evaluate forest growth model assumptions

Specific examples of how CG identifies opportunities and challenges in its investment process related to climate change include the following:

- Identifying afforestation opportunities that mitigate climate change by sequestering CO_2 emissions from the atmosphere in trees and soil, while offering many important co-benefits for communities, biodiversity, and soil and water quality.
- Protecting existing carbon stocks by minimizing impacts to carbon stored on the forest floor through tailored forest management practices.
- Enhancing forest carbon sequestration by replanting areas as soon as possible so the new forest will quickly begin removing CO_2 from the atmosphere.

The ability to quantify, evaluate, and report the year-over-year changes in the carbon footprint of a forest has the potential to influence the impacts an organization has on the environment, leading to increased transparency and more-informed business decisions. Incorporating climate change factors in its investment process not only mitigates climate-related risks; it also promotes and enhances the natural solutions forests provide. Understanding and measuring the comprehensive carbon stores of forests may lead to business decisions that improve carbon sequestration, a critical factor in addressing climate change.

Exhibit 25 shows the quarterly correlation between timber, farmland, and selected global equity and global bond indexes. In the 30-year period from 1990 through 2020, timber and farmland exhibited low correlation with traditional assets; the correlation of timber with global stocks and global bonds was 0.04 and 0.07, while the correlation of farmland with global stocks and global bonds was 0.15 and –0.04, respectively.

Exhibit 25 Historical Commodity Return Correlations, 1990–Q1 2020

NCREIF Data	Timber	Farmland
Global stocks	0.04	0.15
Global bonds	0.07	–0.04

Sources: Global stocks = MSCI ACWI; global bonds = Bloomberg Barclays Global Aggregate Index. Quarterly basis.

The United States has experienced three periods of elevated inflation as measured by the US Consumer Price Index since the early 1900s: 1915–1920, 1940–1951, and 1967–1981. In all three periods, farmland provided investors with meaningful inflation protection: +9.9% annualized farmland compound return versus +13.5% annualized inflation, +7.9% versus +5.9%, and +12.0% versus +7.6%, respectively.

Real estate return data pitfalls will be discussed in more detail in the next section, but note the inherent smoothing in quarterly versus monthly or daily data and the relative infrequency of transactions that make return projections problematic. After all, can one reliably or consistently compare a plot of prime farmland with a more marginal neighbor? How about the price of one piece of land that has generous water rights with another that is water restricted? As is the case with other alternative benchmarks (e.g., hedge fund indexes), productive land indexes may offer a good overall understanding but limited specific value.

KNOWLEDGE CHECK

Is the following statement true or false?

For the most appropriate instrument to invest in natural resources today, retail investors should focus on the stocks and bonds of companies producing in this sector.

A True, statement to support this option on why it’s true.

B False, statement to support this option on why it’s false.

Solution:

B is correct; the statement is false. Up to about 20 years ago, the only commonly available investment vehicles related to this asset class were indeed financial instruments (stocks and bonds). Rather than investing in the physical land and the products that come from it, investors concentrated on the companies that produced natural resources. Nowadays, however, the wide variety of direct investments via ETFs, limited partnerships, REITS, swaps, and futures opens the door for almost everyone to participate in these assets directly.

6.5 Instruments

To achieve the exposures they are after, investors require suitable instruments. Fortunately, there are many publicly traded and indexed choices, and there is still room for specialized value-added alternative mangers.

6.5.1 *Instruments: Commodities*

The majority of commodity investing is implemented through derivatives, and commodity futures are a popular choice. Physical commodities often generate unwelcome price opacity, tax obligations, and costs arising from storage, brokerage, and transportation, all of which may increase the attractiveness of standardized futures traded on transparent exchanges. Commodity derivatives include futures, forwards, options, and swaps. These contracts may be traded on exchanges or over the counter (OTC).

Futures contracts are obligations to buy or sell a specific amount of a given commodity at a fixed price, location, and date in the future. Futures contracts are exchange traded, are marked to market daily, and may or may not be settled with the delivery or receipt of the physical commodity at the end of the contract. This delivery obligation became dramatically important during the global financial crisis in 2008 and in the 2020 COVID-19 pandemic as demand collapsed, oil producers could not find buyers for their petroleum, and global storage filled suddenly. Even commodity-related ETFs were affected, forcing some to close and impose large losses on their investors. For futures contracts, counterparty risk is managed by the exchange and clearing broker. Commodity exposure can be achieved through means other than direct investment in commodities or commodity derivatives, including the following:

- Exchange-traded products (ETPs, either funds or notes) may be suitable for investors who are restricted to equity shares or are seeking the simplicity of trading through a standard brokerage account. ETPs may invest in commodities or commodity futures. For example, the SPDR Gold Shares exchange-traded fund attempts to track the price of physical gold by holding bullion in vaults. It owned just under \$49 billion in gold bullion as of March 2020. ETPs may use leverage and may be long or short (also known as "inverse"). Similar to mutual funds or unit trusts, ETPs charge fees that are included in their expense ratios.

- Commodity exposure can also be obtained by investing in managed futures, also known as CTAs. However, in order to obtain pure commodity exposure, one would need to choose to invest in a managed futures fund focused solely on commodities because modern CTAs often invest in a variety of futures, including commodities, equities, fixed income, and foreign exchange. Such a fund might concentrate on a specific commodity, such as grains or livestock, or it may be broadly diversified to cover all commodities. Commodity-focused managed futures funds are unique (versus global macro) because there is a constant price tension between suppliers and consumers: High prices cripple demand (and thus lead to lower prices), and low prices shut in supply (and thus raise prices). This situation creates a unique balance that is not present in traditional asset classes, such as stocks and bonds.
- Funds that specialize in specific commodity sectors exist. For example, private energy partnerships, similar to private equity funds, are a popular way for institutions to gain exposure to the energy sector. Management fees can range from 1% to 3% of committed capital, with a lockup period of 10 years and extensions of 1- and 2-year periods. Publicly available energy mutual funds and unit trusts typically focus on the oil and gas sector and often act as fixed-income investments, paying out dividends from rents or capital gains. They may focus on upstream (drilling), midstream (refineries), or downstream (chemicals). Management fees for these funds are in line with those of other public equity managers and range from 0.4% to 1%.

6.5.2 *Instruments: Timberland and Farmland*

The primary investment vehicles for timber and farmland are investment funds, whether offered on the public markets, such as real estate investment trusts (REITs) in the United States, or administered privately through limited partnerships. Larger investors can consider direct investments if there are particular assets that have appeal. For example, Middle Eastern sovereign wealth funds have made investments in farmland in Africa and Southeast Asia. Owning physical farmland opens the door to a wider variety of foodstuffs: spices, nuts, fruits and vegetables—a much broader array than the corn, soy, and wheat offered by futures investment. However, there is limited price transparency or information to guide investment decisions without the assistance of sector specialists. The illiquidity of direct farm and timberland investments is also limiting.

PRACTICE QUESTIONS FOR LOS F

Question 1

A significant challenge to investing in timber is *most likely* its:

A high correlation with other asset classes.

B dependence on an international competitive context.

C return volatility compounded by financial market exposure.

Solution:

B is correct. A primary risk of timber is the international competitive landscape. Timber is a globally sold and consumed commodity subject to world trade interruptions. So the international context can be considered one of its major risk factors.

A is incorrect because timberland offers an income stream based on the sale of trees, wood, and other timber products that has not been highly correlated with other asset classes.

C is incorrect because investors are interested in timber because of its global nature (everyone requires shelter), the current income generated from the sale of the crop, inflation protection from holding the land, and its safe haven characteristics (it offers some insulation from financial market volatility).

Question 2

A characteristic of farmland strongly distinguishing it from timberland is its:

A commodity price-driven returns.

B inherent rigidity of production for output.

C value as an offset to other human activities.

Solution:

B is correct. Unlike timberland products, farm products must be harvested when ripe, so there is little flexibility in the production process. In contrast, timber (trees) can be grown and easily "stored" by simply not harvesting. This feature offers the flexibility of harvesting more trees when timber prices are up and delaying harvests when prices are down.

A is incorrect because just as a primary return driver for timberland is change in commodity price (of lumber from cut wood) in either the spot or futures price, farmland's returns are driven by agricultural commodity prices, with commodity futures contracts potentially combined with farmland holdings to generate an overall hedged return.

C is incorrect because for both farmland and timberland owned or leased for the benefit of the bounty each generates in the form of crops and more broadly timber, since these resources consume carbon as part of the plant life cycle, the considered value comes not just from the harvest but also from the offset to other human activities.

Question 3

Which of the following statements about commodity investing is invalid?

A Few commodity investors trade actual physical commodities.

B Commodity producers and consumers both hedge and speculate.

C Commodity indexes are based on the price of physical commodities.

Solution:

C is correct. Commodity indexes typically use the price of futures contracts on the commodities included in them rather than the prices of the physical commodities themselves in order to be transparent, investable, and replicable.

A is incorrect because trading in physical commodities is primarily limited to a smaller group of entities that are part of the physical supply chain. Thus, most commodity investors do not trade actual physical commodities but, rather, trade commodity derivatives.

B is incorrect because although supply chain participants use futures to hedge their forward purchases and sales of the physical commodities, those commodity producers and consumers nonetheless both hedge and speculate on commodity prices.

Question 4

An investor seeks a current income stream as a component of total return and desires an investment that historically has low correlation with other asset classes. The investment *most likely* to achieve the investor's goals is: (2020 Q10)

A timberland.

B collectibles.

C commodities.

Solution:

4. A is correct. Timberland offers an income stream based on the sale of timber products as a component of total return and has historically generated returns not highly correlated with other asset classes.

Question 5

If a commodity's forward curve is downward sloping and there is little or no convenience yield, the market is said to be in:

A backwardation.

B contango.

C equilibrium.

Solution:

5. B is correct. Contango is a condition in the futures markets in which the spot price is lower than the futures price, the forward curve is upward sloping, and there is little or no convenience yield. Backwardation is the opposite condition in the futures markets, where the spot price exceeds the futures price, the forward curve is downward sloping, and the convenience yield is high. Equilibrium is an economic term where supply is equal to demand.

7 REAL ESTATE

g explain investment characteristics of real estate

7.1 Overview of the Real Estate Market

Individuals and institutions buy real property for their own use, as an investment, or both. The residential, or housing, market is made up of individual single-family detached homes and multi-family attached units, which share at least one wall with another unit (e.g., condominiums, townhouses), owned by the resident. Commercial real estate includes primarily office buildings, retail shopping centers, and warehouses. In contrast to the owner-occupied market, rental properties are leased to tenants. A lease is a contract that conveys the use of the property from the owner (the landlord or *lessor*) to the tenant (*lessee*) for a predetermined period in exchange for compensation (rent). When residential real estate, be it single or multi-unit, is owned with the intention to let, lease, or rent the property in order to generate income, it is classified as commercial (i.e., income-producing) real estate.

Real estate investing is typically thought of as either direct or indirect ownership (equity investing) in real estate property, such as land and buildings. However, it also includes lending (debt investing) against real estate property. Loans secured by real estate are called mortgages. Investors can access real estate debt by lending, buying mortgages, or purchasing mortgage-backed securities with the property ultimately serving as collateral. The key reasons for investing in real estate include the following potential benefits:

- Competitive long-term total returns driven by both income generation and capital appreciation
- Multiple-year leases with fixed rents for some property types potentially providing stable income over many economic cycles
- Historically low correlations with other asset classes
- Inflation hedge if leases provide regular contractual rent step-ups or can be frequently marked to market

A title or deed represents real estate property ownership and covers building and land-use rights along with air, mineral, and surface rights. Titles can be purchased, leased, sold, mortgaged, or transferred together or separately, in whole or in part. Title searches are a crucial part of buyer and lender due diligence, ensuring the seller/borrower owns the property without any liens or other claims against the asset, such as from other owners, lenders, or investors or from the government for unpaid taxes.

Residential real estate is by far the largest sector of the real estate market by value and size. Savills World Research estimated in July 2018 that residential real estate accounted for more than 75% of global real estate values. Although the average value of a home is less than the average value of an office building, the aggregate space required to house people is much larger than the space required to accommodate office use and retail shopping. Homeownership rates are at least 80% in many of the largest countries by population, including China, India, Mexico, Japan, and Russia, and are high in other large countries, such as Brazil, at more than 74%, and the United States, with at least 64% homeownership. Germany is a notable exception, at approximately 52%, according to OECD.org.

The size of the professionally managed, institutional-quality global real estate market increased to US$9.6 trillion at the end of 2019, as shown in Exhibit 26, from US$6.8 trillion in 2013, representing a compound average annual growth rate of 5.8%. It should come as no surprise that the largest and most developed countries have the highest-valued real estate markets. Institutional-quality real estate generally consists of high-quality properties owned by institutional investors and high-net-worth individuals. It is typically of higher value than most individuals can afford. Smaller and out-of-date properties typically do not qualify as institutional quality.

Exhibit 26 Size of Professionally Managed Global Real Estate Market, 2018 (US$ billions)

Country or Region	Size	% of Total
United States	3,418.1	35.8%
Japan	881.4	9.2%
United Kingdom	745.5	7.8%

(continued)

Exhibit 26 (Continued)

Country or Region	Size	% of Total
China	592.2	6.2%
Germany	580.1	6.1%
France	441.2	4.6%
Hong Kong SAR	378.3	4.0%
Canada	361.0	3.8%
Australia	306.8	3.2%
Rest of world	1,848.4	19.3%
Total	9,553.0	100%

Source: MSCI Real Estate.

Real estate is different from other asset classes: the large capital investment required; the illiquidity; the fact that no two properties are identical in terms of location, tenant credit mix, lease term, age, and market demographics; and the necessarily fixed location of the asset. All these have important investment implications.

Furthermore, price discovery in the private market is opaque, historical prices may not reflect market conditions, transaction costs are high, and transaction activity may be limited in certain markets. It may be difficult for small investors to establish a diversified portfolio of wholly owned properties. Private market indexes are not investable, and property typically requires professional operational management.

Numerous other considerations will give investors pause before investing independently. Involvement in unfamiliar real estate markets calls for specialized knowledge of country, regional, and local market dynamics and the practices, regulations, and taxation that apply. Government regulations dictate what can be built or modified, with further complications coming from each locality's zoning and permit approval process, rules around ownership transfer, and possible rent controls. Local supply and demand dynamics may override wider market trends, complicating analysis. Many countries have limits on foreign investors with respect to property type and location, and they may require higher withholding taxes on foreign-owned property.

KNOWLEDGE CHECK

True or false: The largest sector of the real estate market by value and size is commercial real estate.

A True, statement to support this option on why it's true.

B False, statement to support this option on whey it's false.

Solution:

B is correct; the statement is false. Residential real estate is by far the largest sector of the real estate market by value and size. The residential debt market greatly exceeds commercial property debt because of the larger total value of residential properties combined with property owners' greater ability to use leverage—up to 80% of the property's value or more in some cases. In addition, home mortgages are subsidized in some markets, including government guarantees.

7.2 Characteristics: Forms of Real Estate Ownership

Real estate investing takes a variety of forms best illustrated with the four-quadrant model that arranges the quadrants by type of capital (debt or equity) and source of capital (private or public markets). Investors can choose to own all or part of a property's equity, either unlevered (100% equity) or with a levered approach that relies on debt as well as equity to finance the purchase. Alternatively, investors can gain exposure to real estate through debt ownership, either as a lender/originator or as a purchaser of debt instruments.

Banks and insurance companies are among the largest originators/owners of mortgage debt. The residential debt market greatly exceeds commercial property debt because the total value of residential properties is larger and property owners have a greater ability to use leverage—securing loans for up to 80% of the property's value or more in some cases. In addition, home mortgages are subsidized in some markets and enjoy government guarantees.

Exhibit 27 presents examples of the basic forms of real estate in each of the quadrants: public or private structures, debt or equity ownership.

Exhibit 27 Basic Forms of Real Estate Investments and Examples

	Debt	Equity
Private	■ Mortgages ■ Construction lending ■ Mezzanine debt	■ Direct ownership of real estate: ownership through sole ownership, joint ventures, separate accounts, or real estate limited partnerships ■ Indirect ownership via real estate funds ■ Private REITs
Public	■ MBS (residential and commercial) ■ Collateralized mortgage obligations ■ Mortgage REITs ■ ETFs that own securitized mortgage debt	■ Shares in real estate operating and development corporations ■ Listed REIT shares ■ Mutual funds ■ Index funds ■ ETFs

Within the basic forms, there can be many variations.

7.2.1 *Direct Real Estate Investing*

Direct private investing involves purchasing a property and originating debt for one's own account. Ownership can be free and clear, whereby the title to the property is transferred to the owner(s) unencumbered by any financing liens, such as from outstanding mortgages. Owners can also borrow from mortgage lenders to fund the equity acquisition. Debt investors can also use leverage. Initial purchase costs associated with direct ownership may include legal expenses, survey costs, engineering/environmental studies, and valuation (appraisal) fees. In addition, ongoing maintenance and refurbishment charges are also incurred. Additional debt closing costs are incurred when owners take out loans to fund their investments.

There are several benefits to owning real estate directly. One is control. Only the owner decides when to buy or sell, when and how much to spend on capital projects (subject to lease terms and regulatory requirements), whom to select as tenants based on credit quality preference and tenant mix, and what types of lease terms to offer. Another benefit relates to taxes. Property investors in the largest countries as ranked by the size of the institutional-quality real estate market can use non-cash property depreciation expenses and interest expense, with some limitations, to reduce taxable income and lower their income tax bills. In fact, real estate investors in a country that permits accelerated depreciation and interest expense deductions can reduce taxable income below zero in the early years of asset ownership, and losses can be carried forward to offset future income. Thus, a property investment can be cash-flow positive while generating accounting losses and deferring tax payments. If the tax losses do not reverse during the life of the asset, depreciation-recapture taxes can be triggered when the property is sold.

Major disadvantages to investing directly include the extensive time required to manage the property, the importance of local real estate market expertise to success, and the large capital requirements. Smaller investors bear additional risks from portfolio concentration.

The owner may choose to handle all aspects of investing in and operating the property, including property selection, asset management, property management, leasing, and administration. To address the complexity of owning and managing commercial real estate, investors hire advisers to perform any number of functions, including identifying investments, negotiating terms, performing due diligence, conducting operations, asset management, and eventual disposal. Asset management focuses on maximizing property returns by deciding when and how much to invest in the property and when to sell the property.

Many investors prefer to hire advisers or managers to manage the investors' direct real estate investment in what is called a *separate account*, which contains only the single investor's equity or perhaps a nominal stake from the adviser to help align interests. A separate account allows the investor to control the timing and value of acquisitions and dispositions and perhaps to make operating decisions as well. There is no commingling of capital with other investors.

Sometimes investors will form joint ventures with other investors to access real estate. Joint ventures are especially common when one party can contribute something of value that the other lacks, such as land, capital, development expertise, debt due diligence, or entrepreneurial talent. Joint ventures can be structured as general partnerships or, more commonly, as limited liability companies.

Numerous parties can pool their resources to acquire property, to develop or redevelop assets, or to lend capital. The equity stakes that investors contribute to pooled investment vehicles vary on the basis of their contribution of real assets, capital, entrepreneurial talent, and services. When structured as a limited partnership, there must by a general partner to manage the partnership and accept unlimited liability.

7.2.2 *Indirect Real Estate Investing*

Indirect investing provides access to the underlying real estate assets through a variety of pooled investment vehicles, which can be public or private, such as limited partnerships, mutual funds, corporate shares, REITs, and ETFs. Equity REITs own real estate equity, mortgage REITs own real estate mortgages and MBS, and hybrid REITs own both. Institutional investors and some high-net-worth investors can invest in MBS. Intermediaries facilitate the raising and pooling of capital and creating investable structures.

7.2.3 *Mortgages*

Mortgages represent passive investments in which the lender expects to receive a predefined stream of payments throughout the finite life of the mortgage. Mortgages may require full amortization or partial amortization with balloon payments due at maturity or may be interest only (IO). Borrowers can often choose among fixed-rate, floating-rate, and adjustable-rate options. Some of the floating- and variable-rate mortgages may have limitations, or caps, on how much the rate can change over a given period. For example, a borrower could apply for a mortgage that carries a fixed rate for five years, after which the rate resets to a predetermined spread over the lender's benchmark rate, with any rate change limited to 200 bps during the first year of the reset. Residential mortgages in developed markets usually have 15-, 25-, or 30-year maturities. Commercial property loans may be similarly long for long-term property owners. Other owners will take out 5-, 7-, or 10-year loans to correspond with the owners' anticipated holding period. Selling the property before the mortgage is due may result in prepayment penalties for the commercial borrower. If the borrower defaults on the loan, the lender may seek to take possession of the property. Investments may take place in the form of "whole" loans based on specific properties, typically by way of direct investment through private markets, or through participation in a pool of mortgage loans, typically by way of indirect investment in real estate through publicly traded securities, such as MBS.

REITs and partnerships carry fees and administrative overhead for managing the assets embedded in their valuations. Fee structures for private REITs can be similar to those for private equity funds, with investment management fees based on either committed capital or invested capital. These fees typically range from 1% to 2% of capital per annum. Funds also charge performance-based fees.

7.2.4 *Private Fund Investing Styles*

Capital committed to or invested by *real estate private equity funds* exceeded $900 billion in mid-2018. Most real estate private equity funds are structured as *infinite-life open-end funds*, which, like mutual funds, allow investors to contribute or redeem capital throughout the life of the fund. Investor subscriptions and redemptions are commonly accepted by the GP on a quarterly basis and are subject to the manager's ability to acquire or sell real estate or match investor redemption requests with pending investor subscriptions. Open-end funds generally offer exposure to *core* real estate, which is characterized by well-leased, high-quality institutional real estate in the best markets. Investors expect core real estate to deliver stable returns, primarily from income. Returns are typically driven by real estate beta, although managers can add value through better market selection, property management, and execution.

In addition to investing in core real estate, *core-plus* strategies will also accept slightly higher risks derived from *non-core* markets and sectors or properties with slightly more leasing risk. Non-core properties include large sectors with different risk profiles, such as hotels and nursing homes. Assets in secondary and tertiary markets and such niches as student housing, self-storage, and data centers are also considered non-core.

Investors seeking higher returns may also accept development, redevelopment, repositioning, and leasing risk. *Finite-life closed-end* funds are more commonly used for alpha- and beta-generating value-add and opportunistic investment styles. *Value-add* investments may require modest redevelopment or upgrades, the leasing of vacant space, or repositioning the underlying properties to earn a higher return than core properties. *Opportunistic* investing accepts the much higher risks of development, major redevelopment, repurposing of assets, taking on large vacancies, and speculating on significant improvement in market conditions. Closed-end fund managers will sell assets to realize the value created by management and lock in the investment's IRR. There are exceptions to the fund style/life/structure relationships—that is, core/infinite life/open end or opportunistic/finite life/closed end. On occasion, you will see core, closed-end, finite-life funds or open-end, infinite-life, value-add funds.

7.2.5 *REITs*

Real estate investment trusts are the preferred investment vehicles for owning income-producing real estate of both private and public investors. The main appeal of the REIT structure is the elimination of double corporate taxation: Corporations pay taxes on income, and then the dividend distributions of after-tax earnings are subsequently taxed at the shareholder's personal tax rate. REITs can avoid corporate income taxation by distributing dividends equal to 90%–100% of their taxable net rental income.

In the United States, taxable real estate corporations can elect REIT status by meeting a number of important criteria. Among the many rules, the company must invest at least 75% of its assets in real estate, cash, or US Treasuries and derive at least 75% of its income from rents on real property or real estate mortgages.

REIT and REIT-like structures in most countries are similar to those in the United States. Whereas REITs are a popular vehicle in the United States, other countries have structures separate from corporations for REITs that still permit dividend deductibility from the entities' income taxes. Primary differences relate to whether REITs can be internally managed, operated, and advised or are required to hire external managers and advisers, how much income the REIT must distribute (90%–100% is the standard), and whether the REIT is a corporation as in the United States or an independent structure. REITs that are self-managed and self-administered typically benefit from a better alignment of interests between management and shareholders than externally advised or managed companies.

Companies that have other sources of non-qualifying income or wish to retain earnings to reinvest in the business will choose an operating company structure. Note that the income taxes due for companies that can deduct depreciation and interest expense from operating income in calculating pretax income may not be high enough to accept the restrictions that come with REIT and REIT-like structures.

Publicly traded REITs address many of the disadvantages related to private real estate investing. Listed REITs provide investors with much greater liquidity, lower trading costs, and better transparency. Management is employed internally by the REIT in most countries rather than brought in as a separate organization on contract, which results in better alignment of interests with investors. Importantly, when an investor wants to modify its real estate holdings, the investor need only buy or sell REIT shares on listed markets instead of buying or selling real estate directly. The REIT is not forced to sell the company's underlying real estate the way open-end funds are when there are mass redemptions.

KNOWLEDGE CHECK

Fill in the blank: Compared with private real estate investing, publicly traded real estate investment trusts provide much greater ___________, lower ___________, and better ___________.

Choices of words: transparency, liquidity, trading costs, variety, happiness

Solution:

Compared with private real estate investing, publicly traded real estate investment trusts provide much greater liquidity, lower trading costs, and better transparency.

7.3 Characteristics: Real Estate Investment Categories

The majority of real estate property may be classified as either commercial or residential. In this reading, residential properties are defined narrowly to include only owner-occupied, single residences (often referred to as single-family residential property). Residential properties owned with the intention to let, lease, or rent them are classified as commercial. Commercial properties also include office, retail, industrial and warehouse, and hospitality (e.g., hotel and motel) properties. Commercial properties may also have mixed uses. Commercial properties generate returns from income (e.g., rent) and capital appreciation. Several factors will affect opportunities for capital appreciation, including development strategies, market conditions, and property-specific features.

7.3.1 *Residential Property*

For many individuals and families, real estate investment takes the form of direct equity investment (i.e., ownership) in a residence with the intent to occupy it. In other words, a home is purchased. Given the price of homes, most purchasers cannot pay 100% cash up front and must borrow funds to make the purchase. Most lenders require an equity contribution of at least 10%–20% of the property purchase price in countries with well-developed mortgage markets. Any appreciation (depreciation) in the value of the home increases (decreases) the owner's equity in the home and is magnified by mortgage leverage. In countries without well-developed mortgage markets, homebuyers must save for a much longer period and may have to ask family for assistance in order to raise enough money to buy a home.

Home loans may be held on the originator's balance sheet or securitized and offered to the financial markets. Securitization provides indirect debt investment opportunities in residential property to other investors via securitized debt products, such as residential mortgage-backed securities (RMBS).

7.3.2 *Commercial Real Estate*

Commercial property has traditionally been considered an appropriate direct investment, whether through equity or debt, for institutional funds and high-net-worth individuals with long time horizons and limited liquidity needs. This perception was primarily the result of the complexity, size, and relative illiquidity of the investments. Direct equity investing (i.e., ownership) is further complicated because commercial property requires active day-to-day management. The success of the equity investment is a function of a variety of factors, including how well the property is managed, general economic and specific real estate market conditions, and the extent and terms of any debt financing.

In order to provide direct debt financing, the lender (investor) will conduct financial analysis to determine the borrower's creditworthiness, to ensure the property can generate sufficient cash flow to service the debt, to estimate the property's value, and to evaluate economic conditions. The estimate of the property value is critical because the size of the loan relative to the property value, the loan-to-value ratio, determines the amount of risk held by the lender versus the borrower (equityholder). The borrower's equity in the property is an indicator of commitment to the success of the project, and it provides the lender with a cushion because the property is generally the sole collateral for the loan.

7.3.3 *REIT Investing*

As of 2019, REITs were listed on stock exchanges in 39 countries, and their combined market capitalization exceeded $1.6 trillion.[8] The risk and return characteristics of REITs depend on the type of investment they make. Mortgage REITs, which invest primarily in mortgages, are similar to fixed-income investments. Equity REITs, which invest primarily in commercial or residential properties and use leverage, are similar to direct equity investments in leveraged real estate.

---

[8] www.reit.com/investing/global-real-estate-investment (accessed 12 August 2019).

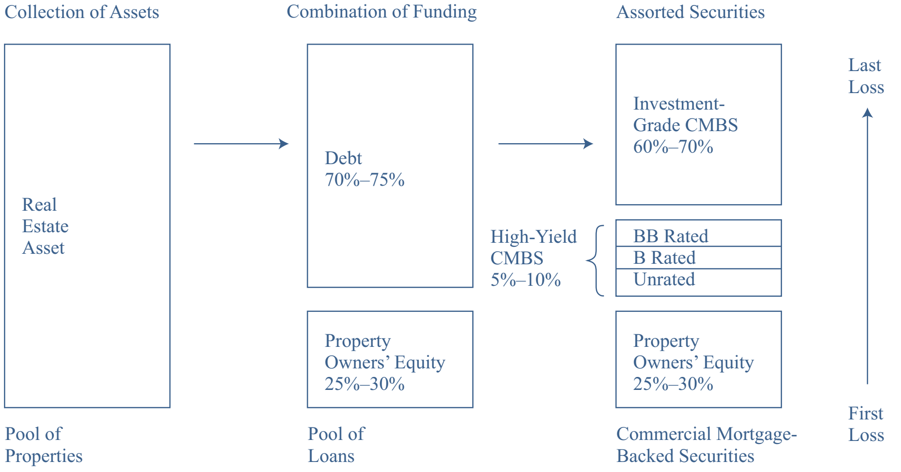
The business strategy for equity REITs is simple: Maximize property occupancy rates and rents in order to maximize income and dividends. Equity REITs, like other public companies, must report earnings per share based on net income as defined by generally accepted accounting principles (GAAP) or IFRS. Many report non-traditional measures, such as net asset value or variations of gross cash flow, to better estimate a company's dividend-paying ability, because non-cash depreciation expenses can be high for asset-intensive businesses.

7.3.4 *Mortgage-Backed Securities*

MBS structuring is based on the asset-backed securitization model of transforming illiquid assets (mortgages) into liquid securities and transferring risk from asset owners (banks, finance companies) to investors. An MBS issuer/originator forms a special purpose vehicle (SPV) to buy mortgages from lenders and other mortgage owners and use them to create a diversified mortgage pool. The MBS issuer assigns the incoming stream of mortgage interest income and principal payments to individual security tranches, which will be sold to investors. Each tranche is assigned a priority distribution ranking.

Exhibit 28 illustrates the basic process for creating commercial mortgage-backed securities (CMBS). On the right-hand side of the exhibit, the ranking of losses indicates the priority of claims against the real estate property. Risk-averse investors, primarily insurance companies, prefer the lowest-risk tranches, which are the first to receive interest and principal. Investors who choose the senior notes with the highest credit rating expect to earn a low return consistent with the senior tranche's low risk profile. Investors seeking the highest returns, which carry the highest risk, will invest in the lowest-rated, most junior securities, which are the last to receive interest and principal distributions. If mortgage defaults and losses are high, the lowest-ranked tranches bear the cost of the shortfall. The most junior tranche is referred to as the first-loss tranche.

Exhibit 28 CMBS Security Structure


Collection of Assets
Combination of Funding
Assorted Securities
Real Estate Asset
Debt 70%–75%
Investment-Grade CMBS 60%–70%
Last Loss
High-Yield CMBS 5%–10%
BB Rated
B Rated
Unrated
Property Owners' Equity 25%–30%
Property Owners' Equity 25%–30%
First Loss
Pool of Properties
Pool of Loans
Commercial Mortgage-Backed Securities

MBS may be issued privately or publicly, and the securities are typically offered in $1,000 increments in the United States. MBS valuations are influenced by the underlying borrowers' behavior. When interest rates decline, fixed-income security prices usually rise, but borrowers are also likely to refinance their loans at a faster pace than before, resulting in the faster amortization of each MBS tranche, which leaves MBS investors to reinvest their principal at the lower rates. Conversely, when rates rise, not only do fixed-income instrument prices decline, but also property owner prepayments can also slow, lengthening the duration of most MBS tranches and contributing to further price weakness. These unusual traits are well understood by investors. MBS are often included in broad fixed-income indexes and in indexes that are used to indicate the performance of real estate investments.

7.4 Risk and Return Characteristics

7.4.1 *Real Estate Indexes*

There are a variety of indexes globally that are designed to measure total and component real estate returns for listed securities and non-listed investment vehicles. Listed REIT indexes are straightforward in that, like other listed equity indexes, security pricing is reported by the major exchanges. Total returns are calculated assuming dividends are reinvested in the index. Importantly, listed REIT indexes are investable when their constituent stocks are freely traded. The more frequently the shares trade, the more reliable the index. Indexes containing small-cap, closely held, and low-float real estate stocks may not be as representative of investor returns, and REIT indexes in general are not necessarily representative of the entire real estate universe because of different geographic concentrations, property types (residential, commercial, etc.), and the asset quality of the publicly traded REIT company portfolios. The National Association of Real Estate Investment Trusts (Nareit) in the United States, the European Public Real Estate Association (EPRA), and the Asian Public Real Estate Association (APREA) publish listed real estate company indexes for their respective regions and contribute to global indexes. Most of the leading for-profit index providers also offer a variety of REIT and real estate indexes.

As shown in Exhibit 29, REITs are the predominant structure among listed real estate companies in the United States. Outside the United States, there is an equal balance of REITs and non-REITs, as represented by the FTSE EPRA Nareit index series. Some other countries in which REITs make up most of the index include Japan, the United Kingdom, Australia, and Singapore. In contrast, non-REIT property companies make up most of the real estate indexes in Hong Kong SAR, Germany, and Sweden.

Exhibit 29 **REITs Make Up 50% of the Property Index Outside the United States**

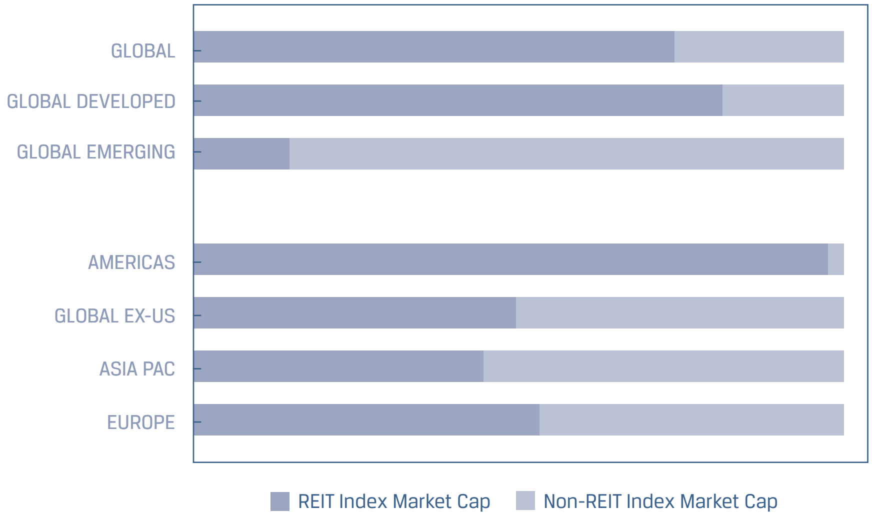


Exhibit 30 displays global and regional listed REIT returns. The table shows some disparity among regional returns and supports the importance of country-specific, regional, and local knowledge.

Exhibit 30 Historical Returns of Global REITs (Through 31 December 2019)

FTSE EPRA Nareit Developed Index Series, USD

Period	Global	North America	Asia Pacific	Europe
3 years	9.31%	7.76%	10.24%	13.05%
5 years	6.53%	6.63%	5.69%	7.40%
10 years	9.25%	11.39%	6.58%	8.66%
20 years	9.23%	11.06%	7.05%	9.38%
25 years	8.92%	11.13%	7.00%	8.88%
3-year volatility	8.24%	10.17%	9.05%	10.51%
10-year volatility	11.81%	13.05%	13.22%	15.16%
Dividend yield (Dec. 19)	3.83%	4.00%	3.60%	3.64%

Sources: FTSE, EPRA, NAREIT.

A variety of indexes measure private investment performance, and they may report on private fund performance, underlying property values, or property operating performance. Funds that track listed REIT indexes make public REIT performance investable, but the same instruments are not available for private indexes. Real estate values are reported by property owners, advisers, and fund managers, and many private investors do not report results, giving rise to selection bias. The private fund and property indexes vary by property selection, valuation methodology, and longevity. Although in many cases private index property values are based on third-party property appraisal and, therefore, offer a measure of independence from owner/manager estimates, appraisals are nonetheless backward looking, subject to the biases of the appraisers, and likely to offer a smoothed-out picture of actual market volatility.

Industry associations, practitioners, and academics have developed *repeat sales indexes* that are transaction based rather than appraisal based. When repeat sales take place, the changing property prices are measured and used to construct the index. These indexes also suffer from a sample selection bias because the properties that sell in each period vary and may not be representative of the larger market. Also, the properties that change hands are not a random sample and may be biased toward those that have increased or decreased in value, depending on economic conditions. The higher the number of sales, the more reliable and relevant the index.

In the United States, NCREIF, in conjunction with the Pension Real Estate Association (PREA), collects fund and property data, maintains numerous indexes, and disseminates the results. The NCREIF Property Index (NPI), which has data going back to 1977, measures property-level performance each quarter. NCREIF also produces open-end diversified core equity (OEDCE) fund, open-end (OE) fund, closed-end value-add (CEVA) fund, and transaction-based repeat sales indexes. The European Association for Investors in Non-Listed Real Estate (INREV) and the Asian Association for Investors in Non-Listed Real Estate (ANREV) produce similar indexes for their respective regions. These three organizations also contribute data from their respective non-listed real estate vehicle indexes, which are rolled up into the quarterly Global Real Estate Fund Index (GREFI).

A wide range of indexes for public stocks and private real estate funds and properties are published by such organizations as Cambridge Associates, Bloomberg, FTSE, MSCI, Real Capital Analytics, Savills, and S&P Dow Jones Indices.

Although investors can benefit from this variety and may well identify an index that is relevant to them, they should be aware of how these indexes are constructed and the inherent limitations that result. Investors should also be aware that real estate's apparent low volatility and low correlation with other asset classes may result from these limitations.

7.4.2 *Real Estate Investment Risks*

Real estate investments, like any investment, may fail to perform in accordance with expectations. Property values are variable in the face of national and global economic conditions, local conditions, and interest rate levels. There are risks around the ability of the fund management teams to select, finance, and manage real properties and to account for changes in government regulation. Management of the underlying properties themselves includes handling rentals and leases, controlling expenses, directing maintenance and improvements, and ultimately disposing of the property. Expenses may increase because of circumstances beyond management's control or be covered by insurance. Returns to both debt and equity investors in real estate depend to a large extent on the ability of the owners or their agents to successfully operate the underlying properties.

Investments in distressed properties and in property development are subject to greater risks than investments in properties that are in sound financial condition or enjoy stable operations. Property development is subject to special risks, including regulatory issues, construction delays, and cost overruns. Environmental regulation is one regulatory hurdle, as is the failure to receive zoning, occupancy, and other approvals and permits. Because the development and disposition period of such projects can be very lengthy, economic conditions may change. All these moving pieces may conspire to increase construction time or delay successful leases, which increases construction costs and reduces the level of rents relative to initial expectations. And there is also financing risk: Long-term financing with acceptable terms simply might not be available, forcing real estate acquisitions and developments to be financed with lines of credit or other forms of temporary financing. Financing problems with one property may delay or limit further development of the owner's other projects.

It is important to recognize that the vast majority of direct investors, private funds, and public companies pursue leverage to increase returns for their investors. Leverage magnifies the effects of both gains and losses for equity investors and increases the risks that the real estate owner or investor will be left with insufficient funds to make interest payments or repay the debt at maturity. In that case, the lender or debt investor will receive less than the entire interest and principal balances. As the loan-to-value ratio increases and interest coverage weakens, the probability of default increases.

KNOWLEDGE CHECK

Fill in the blanks: Adding real estate to a portfolio has been demonstrated to increase portfolio ___________ and reduce portfolio ___________.

Solution:

Adding real estate to a portfolio has been demonstrated to increase portfolio diversification and reduce portfolio risk.

7.5 Diversification Benefits

Many investors find the real estate sector attractive for providing high, steady current income. Although broad stock market gains are derived mostly from long-term capital appreciation, more than half of the returns from private core and listed real estate are derived from income. Throughout a market cycle, income is a more consistent source of return than capital appreciation, and hence, it is lower risk. The source of that consistency stems from the underlying medium- to long-term property leases: The longer the lease the more predictable the income. And the better the credit quality of a property's tenants, the more reliable the rents.

Are REITs equity, debt, or something else? In line with our categorization of all the assets in this reading, many investors view REITs as alternative investments in the broader real estate category, and others see REITs as a fixed-income/equity hybrid, adding exposure simply for the incremental yield pickup over both high-grade bonds and equities. Others still have a different focus: Some institutional investors use their public equity department to analyze their REIT investments rather than engaging their real estate or alternative asset teams.

Real estate's moderate correlation with other asset classes adds to its attractiveness. Exhibit 31 lists the correlation of REITs with equity and debt securities during the 20-year period ended March 2020. The market capitalization of most REITs is on the smaller side, with the largest among them being notable exceptions.

Exhibit 31 REIT Allocation Can Improve Portfolio Diversification

Index	Correlation with the FTSE Nareit All Equity REITs Index
S&P 500	0.57
Russell 2000 Index	0.64
ICE BofAML Corp./Gov't.	0.19
ICE BofAML MBS	0.09
Bloomberg Barclays US Agg./Credit/Corp. High Yield	0.61

Sources: Nareit, using data from FactSet and Intercontinental Exchange (ICE).

There are periods when equity REIT correlations with other securities are elevated, and they are at their highest during steep market downturns, such as the 2007–08 financial crisis. These correlations remained high during the post-crisis recovery, which lifted the value of most asset classes.

Nevertheless, the diversification benefit persists, and including real estate has been demonstrated to reduce portfolio risk. Oxford Economics studied the expected performance of listed European real estate as an asset class, comparing it with equities, fixed income, and commodities.[9] "A substantial allocation to listed real estate," the authors concluded, "does enhance the risk-return characteristics of a multi-asset portfolio." They went on to recommend larger allocations to real estate for European investors.

---

[9] Listed Real Estate in A Multi-Asset Portfolio: A European Perspective, pg,. 2, Oxford Economics, EPRA, September 2019.

Exhibit 32 contains a summary of the performance and correlations of major asset classes and shows that European investors would benefit from real estate allocations.

Exhibit 32 Total Returns by Asset Class, Jan. 1999–May 2019

	Total Return (CAGR)	Standard Deviation	Sharpe Ratio	Correlation with Listed Real Estate (Annual Returns)
Small-cap equities	8.2%	22.9%	0.47	0.60
Large-cap equities	2.2	21.3	0.21	0.55
Government bonds	3.8	3.9	1.02	0.15
Investment-grade corporate bonds	4.3	5.5	0.84	0.41
High-yield corporate bonds	7.1	18.0	0.45	0.44
Diversified commodities	1.8	18.9	0.19	0.14
Listed real estate	8.4	22.7	0.49	1.00

Source: Data sources are as follows: small-cap equities, S&P Europe SmallCap index; large-cap equities, STOXX Europe 50 Index; government bonds, Bloomberg Barclays Pan-European Aggregate Government A Index; investment-grade corporate bonds, Bloomberg Barclays Pan-European Aggregate Corporate Index; high-yield corporate bonds, Bloomberg Barclays Pan-European High Yield Total Return Index; diversified commodities, Bloomberg Commodity Index; listed real estate, FTSE EPRA Nareit Developed Europe Index.

Whether listed real estate behaves like stocks or private real estate is a matter of ongoing debate. We know that listed REITs are priced continuously whereas private real estate is appraised perhaps once a year, and there are other indications that their correlation is weak. Listed equity investors will discount future cash flows, and appraisers place heavy emphasis on current market conditions and recent trends.

CEM Benchmarking, based in Toronto, has performed several studies looking at the role of listed and unlisted real estate in the portfolio of defined contribution pension funds. CEM Benchmarking also concludes that real estate improves the risk–return profile of funds diversified across equities, fixed income, and alternatives. CEM analyzed the performance of each asset class over 12 years, from 2005 to 2016, across €2 trillion of AUM at the end of 2016. The data covered 36% of European pension assets and provided a comprehensive view about how real estate performs relative to the pension funds' other assets. Exhibit 33 presents the asset correlations for Dutch and UK pension funds, which are among the deepest pension fund markets with real estate concentrations.

Exhibit 33 Correlations between Aggregate Asset Class Net Returns for Dutch and UK Pension Funds

Correlation Matrix	Public Equity	Private Equity	Fixed Income	Hedge Funds	Listed Real Estate	Unlisted Real Estate	Unlisted Infrastructure
Public equity	1.00						
Private equity	0.93	1.00					
Fixed income	0.12	0.06	1.00				
Hedge funds	0.91	0.77	0.16	1.00			

(continued)

Exhibit 33 (Continued)

Correlation Matrix	Public Equity	Private Equity	Fixed Income	Hedge Funds	Listed Real Estate	Unlisted Real Estate	Unlisted Infrastructure
Listed real estate	0.86	0.83	0.28	0.81	1.00		
Unlisted real estate	0.76	0.86	0.37	0.64	0.88	1.00	
Unlisted infrastructure	0.73	0.67	0.56	0.62	0.66	0.67	1.00

Notes: The period spans 2005–2016. Unlisted real estate and private equity are adjusted for lagged and smoothed reporting. Actual, as-reported correlations would appear higher and standard deviations, lower.

CEM Benchmarking conducted another study addressing some of the differences between listed and unlisted real estate performance by comparing the returns, volatility, correlations, and Sharpe ratios for various private real estate investing styles. The study covered more than 200 US public and private defined benefit pension funds over 20 years through 2017. Unlisted real estate's correlation with listed real estate was 0.91, which supports the view that listed REITs behave more like stocks in the short term and more like real estate in the long term. The outperformance of listed real estate was primarily due to fees and carry charged by external managers in the unlisted markets.

PRACTICE QUESTIONS FOR LOS G

Question 1

The majority of real estate property may be classified as either:

A debt or equity.

B commercial or residential.

C direct ownership or indirect ownership.

Solution:

B is correct. The majority of real estate property may be classified as either commercial or residential.

Question 2

Which of the following relates to a benefit when owning real estate directly?

A Taxes

B Capital requirements

C Portfolio concentration

Solution:

A is correct. When owning real estate directly, there is a benefit related to taxes. The owner can use property non-cash depreciation expenses to reduce taxable income and lower the current income tax bill. In fact, accelerated depreciation and interest expense can reduce taxable income below zero in the early years of asset ownership, and losses can be carried forward to offset future income. Thus, a property investment can be cash-flow positive while generating accounting losses and deferring tax payments. If the tax losses do not reverse during the life of the asset, depreciation-recapture taxes can be triggered when the property is sold.

B is incorrect because the large capital requirement is a major disadvantage of investing directly in real estate.

C is incorrect because a disadvantage for smaller investors who own real estate directly is that they bear the risk of portfolio concentration.

Question 3

Which of the following statements is true regarding mortgage-backed securities?

A Insurance companies prefer the first-loss tranche.

B When interest rates rise, prepayments will likely accelerate.

C When interest rates fall, the low-risk senior tranche will amortize more quickly.

Solution:

C is correct. When interest rates decline, borrowers are likely to refinance their loans at a faster pace than before, resulting in faster amortization of each MBS tranche, including the senior tranche, which is the lowest-risk tranche.

A is incorrect because risk-averse investors, primarily insurance companies, prefer the lowest-risk tranches, which are the first to receive interest and principal. The junior-most tranche is referred to as the first-loss tranche. It is the highest-risk tranche and is the last to receive interest and principal distributions.

B is incorrect because when interest rates rise, prepayments will likely slow down, lengthening the duration of most MBS tranches. Prepayments will likely increase when interest rates decline, because borrowers are likely to refinance their loans at a faster pace.

Question 4

Which of the following statements is true for REITs?

A According to GAAP, equity REITs are exempt from reporting earnings per share.

B Though equity REIT correlations with other asset classes are typically moderate, they are highest during steep market downturns.

C The REIT corporation pays taxes on income, and the REIT shareholder pays taxes on the REIT's dividend distribution of after-tax earnings.

Solution:

B is correct. Real estate investments, including REITs, provide important portfolio benefits due to moderate correlation with other asset classes. However, there are periods when equity REIT correlations with other securities are high, and their correlations are highest during steep market downturns.

A is incorrect because equity REITs, like other public companies, must report earnings per share based on net income as defined by GAAP or IFRS.

C is incorrect because REITs can avoid this double taxation. A REIT can avoid corporate income taxation by distributing dividends equal to 90%–100% of its taxable net rental income. This ability to avoid double taxation is the main appeal of the REIT structure.

Question 5

What is the most significant drawback of a repeat sales index to measure returns to real estate? (2020 Q14)

A Sample selection bias

B Understatement of volatility

C Reliance on subjective appraisals

Solution:

A is correct. A repeat sales index uses the changes in price of repeat sales properties to construct the index. Sample selection bias is a significant drawback because the properties that sell in each period vary and may not be representative of the overall market the index is meant to cover. The properties that transact are not a random sample and may be biased toward properties that changed in value. Understated volatility and reliance on subjective appraisals by experts are drawbacks of an appraisal index.

Question 6

As the loan-to-value ratio increases for a real estate investment, risk *most likely* increases for: (2020 Q19)

A debt investors only.

B equity investors only.

C both debt and equity investors.

Solution:

C is correct. The higher the loan-to-value ratio, the higher leverage is for a real estate investment, which increases the risk to both debt and equity investors.

8 INFRASTRUCTURE

h explain investment characteristics of infrastructure

8.1 Introduction and Overview

The assets underlying infrastructure investments are real, capital intensive, and long lived. These assets are intended for public use, and they provide essential services—for instance, airports, health care facilities, and power plants. Most infrastructure assets are financed, owned, and operated by governments. A 2017 World Bank/PPIAF study of infrastructure spending in developing countries reported that 83% came from public investment,[10] but more and more infrastructure is being financed privately, with the increasing use by local, regional, and national governments of public–private partnerships. PPPs are typically defined as a long-term contractual relationship between the public sector and the private sector for the purpose of having the private sector deliver a project or service traditionally provided by the public sector.

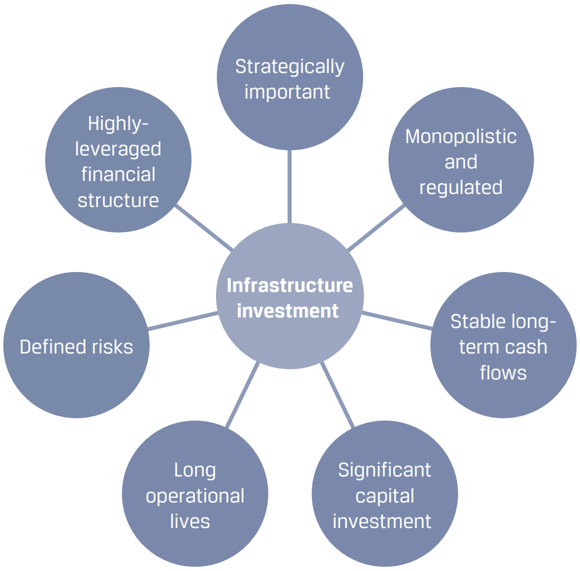
---

[10] World Bank Group and Public-Private Infrastructure Advisory Facility, “Who Sponsors Infrastructure Projects? Disentangling Public and Private Contributions” (2019).

Allocations to infrastructure investments have increased not only because of increasing interest by investors (demand-side growth) but also because of governments providing more investment opportunities (supply-side growth) as they expand the financing of infrastructure assets and the privatization of services. By the end of 2019, according to Preqin, global infrastructure funds had raised more than US$98 billion for infrastructure projects, a year-over-year increase of approximately 15% and up from US$30 billion in 2012, with assets under management of US$582 billion (as of the end of June 2019),[11] reflecting what is commonly considered a growth market. Indeed, Global Infrastructure Hub (a G–20 initiative) and Oxford Economics forecast an infrastructure investment gap of US$15 trillion in 2040, based on the US$79 trillion presently forecast to be spent. (A further $3.5 trillion is required by 2030 to meet the United Nations' Sustainable Development Goals for electricity and water).[12]

Infrastructure investors may intend to lease the assets back to the government, to sell newly constructed assets to the government, or to hold and operate the assets until they reach operational maturity or perhaps for even longer. Exhibit 34 shows the characteristics of a typical infrastructure investment.

Exhibit 34 Typical Infrastructure Investment Characteristics


Strategically important
Highly-leveraged financial structure
Monopolistic and regulated
Infrastructure investment
Defined risks
Stable long-term cash flows
Long operational lives
Significant capital investment

From an investment perspective, if the assets are being held and operated, the relatively inelastic demand for the assets and services is advantageous; regulation and the high costs of the assets create high barriers to entry, which give the provider of the services a strong competitive position. (Regulation, conversely, deters adverse monopolistic behavior from the public perspective.) Maintenance, asset replacement, and operating costs are factored into the pricing of the services.

Investors expect these assets to generate stable long-term cash flows that adjust for economic growth and inflation. Well-defined contractual structures that allocate the risk and responsibilities of asset delivery, service provision, and legal and financial obligations contribute to such stability.

---

[11] Preqin, "2020 Preqin Global Infrastructure Report" (4 February 2020).

[12] Global Infrastructure Hub and Oxford Economics, "Global Infrastructure Outlook" (2017).

These structural aspects allow for relatively high levels of leverage, which is particularly important for funding given the typically high up-front capital investment. Leverage enhances investor returns, and with debt commonly issued on a non-recourse basis, the exposure of investors to their investment is capped.

An increasingly important theme of infrastructure investing is the incorporation of ESG criteria, whether through the application of technology (e.g., renewable energy sources), consideration of the project's impact on the surrounding environment, or how the criteria inform the business relationships and governance approach. All key infrastructure stakeholders are increasingly supporting this theme: the government/public procuring authority, by incorporating environmental and social impact assessments in the planning and permit processes; financiers, through, for instance, the application of the Equator Principles as a condition for advancing loan facilities; and investors, who consciously use ESG criteria to choose investments. The Equator Principles represent a risk management and due diligence framework that has been adopted by financial institutions globally to take account of environmental and social risk in projects.

8.2 Description

8.2.1 *Categories of Infrastructure Investments*

Infrastructure investments are frequently categorized on the basis of the underlying assets. The broadest categorization organizes investments into economic and social infrastructure assets.

Economic infrastructure assets support economic activity and include three broad types of assets: transportation assets, information and communication technology (ICT) assets, and utility and energy assets:

- *Transportation assets* include roads, bridges, tunnels, airports, seaports, and heavy and light/urban railway systems. Income will usually be linked to demand based on traffic, airport, and seaport charges, tolls, and rail fares and hence is deemed to carry market risk.
- *ICT assets* include infrastructure that stores, broadcasts, and transmits information or data, such as telecommunication towers and data centers.
- *Utility and energy assets* generate power and produce potable water; they transmit, store, and distribute gas, water, and electricity; and they treat solid waste.

Utility assets increasingly encompass environmentally sustainable development, with a greater focus on renewable technologies, including solar, wind, and waste-to-energy power generation. (The increasing rejection of less environmentally viable fuel sources, such as coal, is part of the story.) Other energy assets may encompass downstream oil and gas infrastructure, such as pipelines and LNG terminals, and also natural resource assets, such as mining assets. The income earned from utility assets may also carry demand risk as buyers' energy and natural resources needs fluctuate. Alternatively, utilities can institute "take-or-pay" arrangements that lock buyers into minimum purchases whether they need the supply or not. Buyers usually have recourse if the utility falls short on performance, delivering supplies that are late or of inferior quality.

Social infrastructure assets are directed toward human activities and include such assets as educational, health care, social housing, and correctional facilities, with the focus on providing, operating, and maintaining the asset infrastructure. The relevant services that are administered through those facilities, meanwhile, be they medical or schooling, are usually provided separately by the public authority or by a private service provider contracted by the public authority. In some countries, this model has been extended to other public infrastructure, such as courthouses and government and municipal buildings. Income from social infrastructure is typically derived from a type of lease payment that depends on assuring availability (often referred to as availability payments) and on managing and maintaining the asset according to predefined standards. For instance, an availability payment may be reduced or voided if a hospital operating room is not available because of an electromechanical fault.

Infrastructure investments may also be categorized by the underlying asset's stage of development. Investing in infrastructure assets *that are to be constructed* is generally referred to as **greenfield investment**. The intent may be to lease or sell the assets to the government after construction or to either hold and operate the assets over the long term or hold and operate them for a shorter period until operational maturity and then sell them to new investors, thus ensuring capital appreciation that reflects the construction and commissioning risk. Greenfield investors typically invest alongside strategic investors or developers who specialize in developing the underlying assets. Investing in *existing* infrastructure assets may be referred to as **brownfield investment**. Perhaps the assets are owned by a government that wants to privatize them, lease them out, or sell them and lease them back; or they may be owned by greenfield investors seeking to realize the value of their investment through a sale. Typically, some of the assets' financial and operating history is available. Brownfield investors may include strategic investors who specialize in the operation of the underlying assets, but particularly in the case of privatizations, there will be financial investors involved who focus on the long-term stable returns.

Finally, infrastructure may be categorized by location. Infrastructure investments are available globally, and the geographic location of the underlying assets will inform the political and macroeconomic risks, particularly in light of the government's relationship to the assets.

Risks and expected returns may differ on the basis of the underlying asset's category, stage of development, and geographic location. How the investment is actualized—its form—also affects the risks and expected returns.

KNOWLEDGE CHECK

Describe one of the three ways by which an infrastructure investment is typically categorized.

Solution:

Infrastructure assets are frequently characterized on the basis of (1) underlying assets, (2) the underlying asset's state of development, and (3) the geographic location on the underlying assets.

The first category comprises economic assets supporting economic activity, such as transportation and energy/utility assets. It also includes social infrastructure assets, which enable public services directed toward human activities, such as education, health care, and housing.

The second category refers to the relevant stages of the infrastructure asset's life cycle. These include a greenfield investment for a newly created asset and a brownfield investment for an established asset.

The third category highlights the specific place (local, regional, national) associated with the government entity directly involved with the assets.

8.2.2 *Forms of Infrastructure Investments*

As is the case with real estate investments, infrastructure investments take a variety of forms. The choice affects liquidity, cash flow, and income streams. Investing directly in the underlying assets provides control and the opportunity to capture full value. It entails a large investment, however, and results in concentration and liquidity risks while the assets have to be managed and operated. Hence, direct investment often happens in consortia, with strategic partners who are better placed to manage certain risks (e.g., operational) or with other financial/institutional investors to limit individual concentration risk. However, most investors invest indirectly.

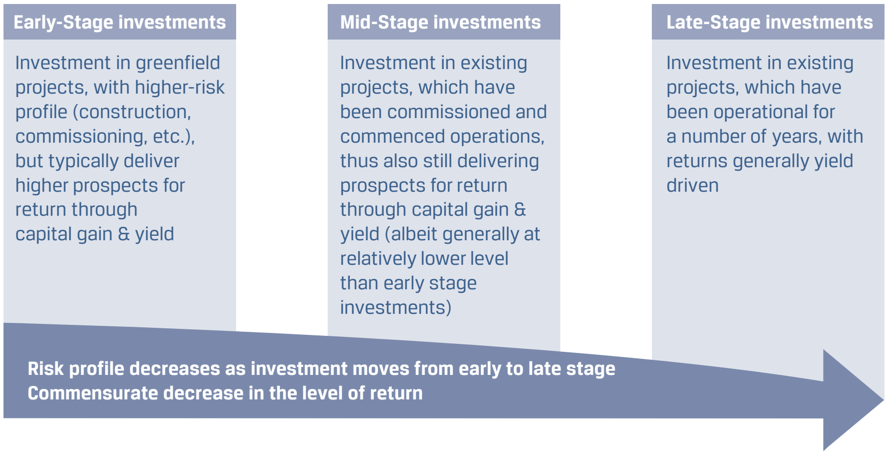
Indirect investment vehicles include infrastructure funds (which are similar in structure to private equity funds and may be closed end or open end), infrastructure ETFs, and company shares. Different asset class strategies, including debt and mezzanine investment, are offered, but the vast majority of investors are focused on equity instruments. According to IJInvestor, equity is the preferred instrument for 77% of fund investors, with 10% looking for mixed strategies and 9% looking for debt and mezzanine strategies.[13]

Investors concerned about liquidity and diversification may choose publicly traded infrastructure securities or master limited partnerships. Publicly traded infrastructure securities provide the benefits not only of liquidity but also of reasonable fees, transparent governance, market prices, and the ability to diversify among underlying assets. An investor should be aware, however, that publicly traded infrastructure securities represent a small segment of the infrastructure investment universe and tend to be clustered in certain asset categories. Master limited partnerships (MLPs) trade on exchanges and are pass-through entities, similar to REITs. They also share with REITs applicable regulations that vary among countries, and income is passed through to the investors for taxation; MLPs generally distribute most free cash flow to their investors. Typically, the GP manages the partnership, receives a fee, and holds a small partnership interest, with LPs owning the remaining majority partnership interest.

8.3 Risk and Return Characteristics

The lowest-risk infrastructure investments have more stable cash flows and higher dividend payout ratios. They typically enjoy fewer growth opportunities, however, and lower expected returns. For example, an MLP with an investment in a brownfield asset that is being leased back to a government represents a low-risk infrastructure investment, as does an investment in assets with a history of steady cash flows, such as certain established toll roads. An investment in a fund that is building a new power plant without any operating history (a greenfield investment) is riskier.

Exhibit 35 Typical Infrastructure Investment Risk–Return Profile



Early-Stage investments	Mid-Stage investments	Late-Stage investments
Investment in greenfield projects, with higher-risk profile (construction, commissioning, etc.), but typically deliver higher prospects for return through capital gain & yield	Investment in existing projects, which have been commissioned and commenced operations, thus also still delivering prospects for return through capital gain & yield (albeit generally at relatively lower level than early stage investments)	Investment in existing projects, which have been operational for a number of years, with returns generally yield driven

Risk profile decreases as investment moves from early to late stage
Commensurate decrease in the level of return

---

[13] IJInvestor, “IJInvestor Funds & Investors Report - Q1 2020” (17 April).

There is a risk that infrastructure revenues diverge meaningfully from expectations, leverage creates financing risk, and operational risk and construction risk are ever present. As for regulatory risk, because the projects involve essential services, governments usually regulate the investments with strictures on the sale of the underlying assets, on operations and service quality, and on prices and profit margins. Global infrastructure investing introduces additional uncertainty, such as currency risk, political risk, and profit-repatriation risk. Hence, the fund manager must ensure these risks are appropriately managed and mitigated through insurance, financial instruments, and whatever other mechanisms are available.

Exhibit 36 Typical Infrastructure Risk Management

PERFORMANCE RISKS	**Demand/ volume risk**	Enter into take-or-pay arrangements, where payments are based upon the availability rather than the use of an asset
	Operational risk	Enter into operation and maintenance contracts with reputable and experienced operators who are incentivized to meet or exceed the minimum contractual performance standards through an appropriate regime
	Construction risk	Enter into fixed-price date-certain contracts with reputable and experienced contractors to minimize construction cost overruns and delays
STRUCTURAL RISKS	**Financing/ interest rate risk**	Enter into interest rate swaps
	Regulatory risk	If regulation is by contract (e.g., PPP agreement), ensure clear, unambiguous provisions that govern such risks; if regulation is defined and enforced by a regulatory body, then ensure due diligence is undertaken on that body
	Political risk	Purchase political risk insurance for protection, as appropriate, against events such as expropriation, political violence, sovereign debt default, etc.
	Currency risk	Ensure PPP agreement incorporates appropriate mechanisms to permit an adjustment to tariffs when there are material foreign exchange fluctuations
	Tax/profit repatriation risks	Ensure PPP agreement incorporates appropriate mechanisms to permit an adjustment to tariffs and/or other compensation when there are significant changes in law/regulations affecting tax and profit repatriation

Returns, of course, depend on such factors as investment type. Exhibit 37 provides an illustrative range of returns for a private infrastructure fund according to different risk profiles, noting that returns may vary at an individual investment/asset level depending on its specific characteristics.

Exhibit 37 Private Infrastructure Fund Illustrative Target Returns

Higher-Risk Profile	Medium-Risk Profile	Lower-Risk Profile
Greenfield projects without guarantees of demand upon completion—e.g., variable electricity prices, uncertain traffic on roads and through ports	Mostly brownfield assets (with some capital expenditure requirements) and some greenfield assets (with limited construction and demand risk)	Brownfield assets with mitigated risks—e.g., fully constructed with contracted/regulated revenues
Located in OECD countries and emerging markets	Located primarily in OECD countries	Located in the most stable OECD countries
High weighting to capital appreciation	Mix of yield and capital appreciation	High weighting to current yield
Target equity returns of 14%+	**Target equity returns of 10%–12%**	**Target equity returns of 6%–8%**

Note: Target equity returns are net of fees.
Source: Cambridge Associates, "Digging In: Assessing the Private Infrastructure Opportunity Today," Research Note (April 2017).

Most infrastructure funds gravitate toward the medium- and lower-risk profiles. Rolling one-year-horizon returns have been around 10% annually since 2016.[14]

In summary, infrastructure investments generally provide investors with stable, long-term capital growth and cash distributions, according to the risks assumed, and these risks tend to be relatively well defined. Preqin maintains a return series of private funds investing in infrastructure deals. Standard & Poor's, FTSE Group, and other firms also publish indexes of publicly traded infrastructure companies.

KNOWLEDGE CHECK

Fill in the blank by choosing from among the following options:

___________ is an infrastructure investment characteristic *most likely* valuable to investors aiming to sell newly constructed assets to the government.

A "Strategically important"

B "Monopolistic and regulated"

C "Significant capital investment"

Solution:

A is the correct answer because the related priority will probably increase the demand of the public buyer to effectively provide essential services to its citizens. B would be more advantageous for investors holding and operating an asset and charging fees to the buyer, with the inelastic demand supporting pricing and regulations increasing the barriers to entry that improve the competitive position. C is more of a challenge than a benefit to the investor, requiring greater funding capability and potentially higher financial risks.

8.4 Diversification Benefits

We have established that investors expect infrastructure assets to generate stable long-term cash flows that adjust for economic growth and inflation. Investors may also expect capital appreciation, depending on the type and timing of their investment.

---

[14] "2020 Preqin Global Infrastructure Report."

Investing in infrastructure may allow for the addition of an income stream, for further portfolio diversification by adding an asset class with typically low correlation with existing investments, and for some protection against inflation. Low exposure to short-term GDP growth issues may also be a factor.

Infrastructure has proven to be relatively resilient to recent swings in the equity markets. According to Preqin, from December 2017 to December 2018 as equity returns fell from 22.8% to –9.6% and hedge fund returns fell from 12.2% to –3.05%, infrastructure returns dropped only from 11.4% to 9.6%.[15] Such resilience is a key factor for many institutional investors with an eye on portfolio construction.

Exhibit 38 provides an illustrative example of the low correlation of infrastructure—and in particular, unlisted Australian infrastructure—with other asset classes over the 10-year period through June 2019.

Exhibit 38 Quarterly Return Correlations for Selected Asset Classes, 2009–2019

	Global Treasury Bonds	Global Equities	Australian Equities	Global REITs	Listed Infrastructure	Unlisted Australian Infrastructure
Global Treasury bonds	1.00					
Global equities	–0.49	1.00				
Australian equities	–0.36	0.89	1.00			
Global REITs	0.18	0.64	0.70	1.00		
Listed infrastructure	–0.12	0.81	0.75	0.68	1.00	
Unlisted Australian infrastructure	–0.08	0.06	–0.05	–0.21	0.13	1.00

Notes: All data are for the period 31 March 2009–30 June 2019, except for the FTSE EPRA NAREIT Developed Rental Net index (Global REITs), where data series began in May 2009.

- Global Treasury bonds: Bloomberg Barclays Global Treasury GDP Index
- Global equities: MSCI World Net Accumulation Index
- Australian equities: S&P/ASX 200
- Global REITs: FTSE EPRA NAREIT Developed Rental Net index
- Listed infrastructure: Dow Jones Brookfield Global Infrastructure Net Accumulation Index
- Unlisted infrastructure: MSCI Australian Unlisted Infrastructure Index

Source: John Julian, "The Pros of Infrastructure Investment in a Lower-for-Longer Environment," AMP Capital (October 2019). www.ampcapital.com/au/en/insights-hub/articles/2019/october/the-pros-of-infrastructure-investment-in-a-lower-for-longer-environment.

Because these investments are placed in long-lived assets, infrastructure may better match the longer-term liability structure of certain investors, such as pension funds, superannuation schemes, and life insurance companies. It also suits the longer-term horizon of sovereign wealth funds, which tend to make the largest allocations to this asset class—around 5%–6% of total AUM, according to Preqin.[16]

---

[15] "2020 Preqin Global Infrastructure Report."

[16] "2020 Preqin Global Infrastructure Report."

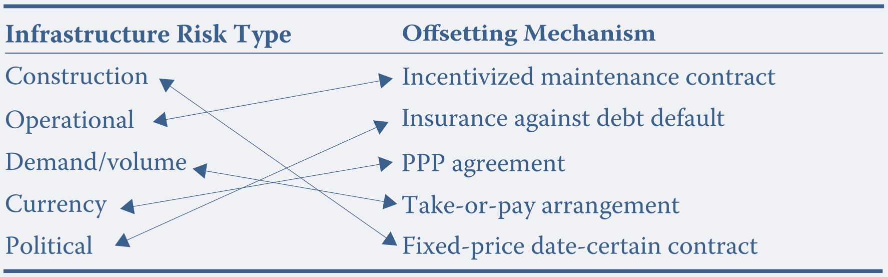
KNOWLEDGE CHECK

To highlight typical risk management approaches on infrastructure projects, draw connecting lines to match a particular risk type with its appropriate mitigating mechanism:

Infrastructure Risk Type	Offsetting Mechanism
Construction	Incentivized maintenance contract
Operational	Insurance against debt default
Demand/volume	PPP agreement
Currency	Take-or-pay arrangement
Political	Fixed-price date-certain contract

Solution:

The correct matches are as follows:


Infrastructure Risk Type
Offsetting Mechanism
Construction
Operational
Demand/volume
Currency
Political
Incentivized maintenance contract
Insurance against debt default
PPP agreement
Take-or-pay arrangement
Fixed-price date-certain contract

PRACTICE QUESTIONS FOR LOS H

Question 1

Compared with direct investment in infrastructure, publicly traded infrastructure securities are characterized by:

A higher concentration risk.

B more transparent governance.

C greater control over the infrastructure assets.

Solution

B is correct. Publicly traded infrastructure securities, which include shares of companies, exchange-traded funds, and listed funds that invest in infrastructure, provide the benefits of transparent governance, liquidity, reasonable fees, market prices, and the ability to diversify among underlying assets. Direct investment in infrastructure involves a large capital investment in any single project, resulting in high concentration risks. Direct investment in infrastructure provides control over the assets and the opportunity to capture the assets' full value.

Question 2

Which of the following forms of infrastructure investment is the most liquid?

A An unlisted infrastructure mutual fund

B A direct investment in a greenfield project

C An exchange-traded MLP

Solution:

C is correct. A publicly traded infrastructure security, such as an exchange-traded MLP, provides the benefit of liquidity.

Question 3

An investor chooses to invest in a brownfield, rather than a greenfield, infrastructure project. The investor is *most likely* motivated by:

A growth opportunities.

B predictable cash flows.

C higher expected returns.

Solution:

B is correct. A brownfield investment is an investment in an existing infrastructure asset, which is more likely to have a history of steady cash flows compared with that of a greenfield investment. Growth opportunities and returns are expected to be lower for brownfield investments, which are less risky than greenfield investments.

Question 4

The privatization of an existing hospital is best described as:

A a greenfield investment.

B a brownfield investment.

C an economic infrastructure investment.

Solution:

B is correct. Investing in an existing infrastructure asset with the intent to privatize, lease, or sell and lease back the asset is referred to as a brownfield investment. An economic infrastructure asset supports economic activity and includes such assets as transportation and utility assets. Hospitals are social infrastructure assets, which are focused on human activities.

Question 5

Risks in infrastructure investing are *most likely* greatest when the project involves:

A construction of infrastructure assets.

B investment in existing infrastructure assets.

C investing in assets that will be leased back to a government.

Solution:

A is correct. Infrastructure projects involving construction have more risk than investments in existing assets with a demonstrated cash flow or investments in assets that are expected to generate regular cash flows.

ISSUES IN PERFORMANCE APPRAISAL

9

i. describe issues in performance appraisal of alternative investments

9.1 Overview of Performance Appraisal for Alternative Investments

Portfolio managers invest in one of two basic ways to achieve returns: passively or actively. Passive investments focus on index or asset class coverage. Portfolios comprising real estate, commodity, and infrastructure instruments may be passively managed to provide exposure to these alternative asset classes. However, alternative investments are generally actively managed, aiming to achieve added portfolio benefits, but typically at a higher net cost for such active management.

Investors frequently look to alternative investments for diversification and a chance to earn relatively high returns on a risk-adjusted basis. They also value low correlation and a more risk-neutral source of alpha.

Evaluating an alternative investment can be a subtle qualitative exercise, one that depends on the initial objectives of the investor, as opposed to a purely quantitative, one-size-fits-all exercise. Much of the nuance revolves around not just the total net return created by an alternative investment but also the path and volatility (drawdown risk) required to create the total return and how an alternative investment fits into and benefits a larger portfolio of assets—in other words, its portfolio-level correlation benefit.

However, there is often a naive attraction to alternative investments based on their expected returns, which neglects to consider the atypical risks they present—risks we can examine on both a standalone and portfolio basis:

- Limited transparency
- Often low portfolio-level liquidity
- The active use of leverage (and, at times, derivatives)
- General strategy and product complexity
- Mark-to-market issues for a portfolio with niche specialized products
- Limited redemption availability and portfolio pressures from eventual redemption activities
- The general challenge of manager differentiation and diversification
- The fee drag associated with alternative investments, which can be non-trivial

In this section, we examine some of these issues in the context of applying traditional performance measurement tools to alternative investments, which is sometimes done incorrectly. In Section 10, we will delve into the impact of fee structures on alternative asset evaluation.

9.2 Common Approaches to Performance Appraisal and Application Challenges

9.2.1 *Sharpe Ratio*

The Sharpe ratio is probably the first basic intuitive metric that some people use to evaluate an alternative investment. It is often prominently displayed in marketing materials. The single biggest flaw, however, in a dependence on the Sharpe ratio is probably the underlying assumption of normally distributed returns.

In light of our list of atypical risks—leverage, illiquidity, mark-to-market smoothing, poor portfolio transparency—return profiles of alternative investments tend to be asymmetric and skewed, making the Sharpe ratio a less-than-ideal performance measure. For non-normal return distributions with significant skewness (fat tails in one direction or the other) and kurtosis (a measure of whether the data are heavy tailed or light tailed relative to a normal distribution), volatility is not a perfect measure of dispersion. Although still widely cited, the Sharpe ratio may not, therefore, be a good risk-adjusted performance measurement to rely on.

Given that caveat and taken by itself as a simple starting point, an attractive Sharpe ratio for an alternative strategy might be deemed to be anywhere between 1.0 and 2.0. Getting twice the return per unit of risk is generally quite attractive and hard to achieve; getting only a single unit of return for a unit of risk (a Sharpe ratio of 1.0) might be deemed acceptable but not overwhelmingly compelling, unless the strategy in question also has negative correlation attributes compared with other strategy areas in a portfolio, as with macro or CTA managers. In comparison, as of the end of 2019, the long-only S&P 500 equity index delivered an average annual return of around 10% over the past 32 years but with a relatively high annual standard deviation around the mean, 15%, yielding a Sharpe ratio close to 0.66. Alternative asset managers generally endeavor to offer a far superior risk–reward profile.

Ironically, if a Sharpe ratio for an alternative manager is too high—say, 3.0–4.0—extra caution in the evaluation of that manager may be warranted. These types of Sharpe ratios are typically only available if some degree of *return smoothing* is taking place. Illiquid securities that are hard to value and coupon return streams that might one day suddenly come to an end (sometimes along with a sudden loss of principal, as with a structured bond product that extinguishes itself or a short option strategy) can produce very high Sharpe ratios for a period of time that may deceive and then ultimately disappoint naive investors.

The ultimate test for a portfolio is really upon liquidation, but investment value at that point is frequently at odds with how accounting standards require alternative asset positions to be carried on the books. Furthermore, neither expected liquidation value nor reported value may account for the tail-risk aspects in an unfortunate macro environment.

The nature of the positions that a portfolio holds must be considered: Is the Sharpe ratio even the relevant metric?

9.2.2 *Sortino Ratio*

A second, arguably more valuable, quantitative measure of performance is return relative to downside volatility—the Sortino ratio. However, the Sharpe and Sortino risk measures alone still do not take into account the correlation of alternative assets with the traditional assets that their inclusion in the portfolio may be intended to hedge.

We have established that alternative investments' diversifying potential is part of the motivation for investing in them: Investors perceive an opportunity to improve the risk–return relationship in the portfolio context. Given the historical return, volatility, and correlation profiles of alternative investments, combining a portfolio of alternative investments with a portfolio of traditional investments should improve the overall portfolio's risk–return profile. Doing so can increase the risk-adjusted return of the overall portfolio because of potentially higher returns to the portfolio and a less-than-perfect correlation with traditional investments.

It is key for investors to consider how an alternative manager's investment choices will correlate with traditional assets and how the assets in question affect drawdown options and liquidity.

9.2.3 *Treynor Ratio and Correlation Behavior*

Alternative assets are generally deemed more valuable when they have a lower beta to traditional assets, and this is where the Treynor ratio proves useful: It is a measure of the excess average return of an investment relative to its beta to a relevant benchmark, such as a broad equity index.

The lower the beta of the alternative asset, the higher the Treynor ratio will be, and all else being equal, when comparing alternative investment possibilities, an asset with a higher Treynor ratio will be deemed more attractive than an asset with a low Treynor ratio.

The main limitation of the Treynor ratio is that it is based on historical beta data that may change in the future. The ratio also becomes less meaningful if the beta of the alternative asset to its systematic benchmark is negative, which is certainly possible for some alternative assets.

The Treynor ratio is the first ratio we have covered that encompasses the return of an alternative asset relative to its expected systematic risk. Examining an alternative asset's correlation coefficient with the overall market and with the other assets in the portfolio will also be instructive. All else being equal, less correlated alternative assets are generally deemed more attractive for portfolio diversification purposes than more correlated assets. It is nevertheless important to note that the expected diversification benefits from alternative investments are not always realized, even when they are most needed. Correlations between risky investments increase during periods of market stress and can approach 1.0 during financial crises.

For now, one easy proxy for measuring return relative to risk for alternative investments—and a far simpler calculation to observe and calculate—is to analyze the performance record for average return relative to the worst drawdown loss. A maximum drawdown (MDD) is the maximum observed loss from a peak to a trough of a portfolio, before a new peak is attained. The **Calmar ratio** is a comparison of the average annual compounded return to this maximum drawdown risk. The higher (lower) the Calmar ratio, the better (worse) an alternative asset performed on a risk-adjusted basis over a specified period of time. The Calmar ratio is typically calculated using the prior three years of performance, and it thus adjusts over time. Variations of the ratio exist: The **MAR ratio** uses a full investment history and the average drawdown. Both ratios help address the left-tailed return profile that sometimes characterizes alternative assets.

Other ratios are applied to specific alternative strategies and require more granular disclosure from investment managers. Some quantitative managers refer to the percentage of profitable trades (or "batting average") and their "slugging percentage," the magnitude of the gains from winning trades divided by the losses from losing trades. A manager who makes outsized gains on winning trades while incurring small losses on losing trades is worth considering.

KNOWLEDGE CHECK

True or false: The Sharpe ratio measures the amount of risk per unit of return.

A True, statement to support this option on why it's true.

B False, statement to support this option on why it's false.

Solution:

B is correct; the statement is false. The Sharpe, Sortino, and Treynor ratios are risk-adjusted performance measurements. They are measures of return per unit of risk.

KNOWLEDGE CHECK

Is the following statement true or false? The Sharpe and Sortino ratios share the same denominator.

A True, statement to support this option on why it's true.

B False, statement to support this option on why it's false.

Solution:

B is correct; the statement is false. The Sharpe and Sortino ratios share the same numerator—average annualized return net of the risk free-rate. Their denominators are different. The denominator for the Sharpe ratio is standard deviation of returns, and the denominator for the Sortino ratio is downside deviation of returns—a semi-deviation measure of volatility only during periods of loss for an alternative investment.

KNOWLEDGE CHECK

Which of the following performance measures uses beta as the risk measure?

A Sharpe ratio

B Treynor ratio

C Sortino ratio

Solution:

The correct answer is B. Both Sharpe Ratio and Sortino Ratio use standard deviation as a risk measure. Treynor ratio uses beta as a risk measure.

9.3 Private Equity and Real Estate Performance Evaluation

Unlike liquid hedge fund strategies, alternative investments that involve private equity and real estate tend to have variable cash flows.

Private equity investments generally involve an initial capital commitment, but actual capital flows often lag that commitment because capital "calls" are staggered over substantive periods of time. Private equity returns are frequently described in terms of the J-curve effect: a substantive initial capital commitment promise, followed by high initial fee drag (calculated on total committed capital, not the capital actually called), followed by the identification of longer-term growth or turnaround opportunities, and an eventual positive return (the investor hopes) when the staggered returns of the fund are realized. The component investments mature and are sold at various times, and the partnership finally closes—but usually only after six to eight years (often with two-year extensions allowed at the discretion of the general partner or after a vote of approval by a majority of limited partners). The line representing the return changes from downward sloping to positive and then to energetically upward sloping later in the investment's life (if all goes well).

The real estate pathway starts with initial property purchases, followed at times by substantive cash outlays for improvements, followed by instances of accounting depreciation that can influence after-tax performance, followed typically by the receipt of rents and then an eventual property sale (often at a long-term tax-advantaged tax rate).

The measurement of success in both instances depends far more on the timing and magnitude of cash flows in and out of the investments, and these are often hard to standardize, let alone always properly anticipate. Given the long time horizon, the application of different tax treatments can have a non-trivial impact on after-tax investment returns.

As a general rule, the best way to start evaluating such investments is with the IRR of the respective cash flows into an investment and the timing thereof, versus the magnitude and the timing of the cash flows returned by the investment (inclusive of tax benefits).

In an independent, fixed-life private equity fund, the decisions to raise money, take money in the form of capital calls, and distribute proceeds are all at the discretion of the private equity manager. Timing of cash flows is an important part of the investment decision process. The private equity manager should thus be rewarded or penalized for the results of those timing decisions, and the calculation of an IRR is a key metric for doing so.

Although the determination of an IRR involves certain assumptions about a financing rate to use for outgoing cash flows (typically a weighted average cost of capital) and a reinvestment rate to use for incoming cash flows (which must be assumed and may or may not actually be earned), the IRR is the key metric used to assess longer-term alternative investments in the private equity and real estate worlds.

Because of this complexity, a shortcut methodology often used by both private equity and real estate managers involves simply citing a **multiple of invested capital (MOIC)**, or money multiple, on total paid-in capital. Here, one simply measures the total value of all distributions and residual asset values (assets that may still be awaiting their ultimate sale) relative to an initial total investment. Although this shortcut valuation completely ignores the timing of cash flows, it is easier to calculate, and it is intuitively easier to understand when someone says they received two or three times their initial investment. But how long it takes to realize this value does matter. A 2× return on one's initial investment would be phenomenal if the return were collected over two years but far less compelling if it took 15 years to realize.

In general, because private equity and real estate investing involve longer holding periods, there is less emphasis on evaluating them in terms of shorter-term portfolio correlation benefits. After a private equity fund has fully drawn in its monetary commitments, interim accounting values for a private equity partnership become less critical for a period of time because no incoming or outgoing cash flows may immediately hinge on such valuations. During this "middle period" in the life of a private equity fund, accounting values may not always be particularly reflective of the future potential realizations (and hence the expected returns) of the fund. It is not that the value of the investments is not actually rising and falling in the face of economic influences; rather, accounting conventions simply leave longer-lived investments marked at their initial cost for some time or make only modest adjustments to carrying value until clearer impairments or realization events take place.

Ironically, many investors gravitate to private equity *because* they are not easily marked to market and can deliver somewhat smoothed returns over time—with less short-term mark-to-market angst. Although most private equity managers are conservative in their marks (awaiting actual realization events), a lagged mark-to-market process can at times offer a false sense of success, diminishing short-term investment worry but subsequently delivering the occasional back-end-loaded investment disappointment. Private equity interim valuations and real estate appraisals can certainly be inaccurate or skewed at times despite the best efforts of auditors to present a fair valuation.

The lagging impact of shorter-term economic events on the interim accounting valuations of these strategies makes them appear more resilient and less correlated than they really are. A more realistic picture may emerge when premature portfolio liquidations are forced on managers. The lack of transparency around such investments and the slowness to mark them to market can be incorrectly construed by investors as an overall lack of volatility.

Along this imperfect path and in an effort to benchmark how longer-term investments may be faring, private equity and real estate managers are generally judged by where they fall in terms of a *quartile ranking*, which depicts their performance against a cohort of peer investment vehicles constructed with similar investment attributes and funded around the same time, or what is often referred to as the same vintage year.

Real estate managers are also often judged by the **cap rate** being earned on their properties, which is simply the annual rent actually being earned (net of any vacancies) divided by the price originally paid for the property. This approach ignores the true values of properties should they need to be sold.

The large latitude around carried valuations for both private equity and real estate strategies makes any application of shorter-term risk metrics highly inappropriate.

Instead, the probability of permanent impairment on an investment is likely the best risk measure. Along these lines, an admittedly backward-looking metric of some interest and potential value that many private equity managers refer to is their historic **capital loss ratio**, defined as the percentage of capital in deals that have been realized below cost, net of any recovered proceeds, divided by total invested capital. Because markets have been generally buoyant for the past 30 years, investors can overestimate the true resilience of private equity and real estate. Having money tied up in longer-lived investments, for 7 to 10 years, should conceptually earn an illiquidity premium but may easily result in return disappointment given the wrong economic environment.

9.4 Hedge Funds: Leverage, Illiquidity, and Redemption Terms

9.4.1 *Leverage*

Hedge funds may use leverage to obtain higher returns on their investments. Leverage has the effect of magnifying gains and losses because it allows for taking a larger position relative to the capital committed. Hedge funds leverage their portfolios by using derivatives or borrowing capital from **prime brokers**, negotiating with them to establish margin requirements, interest, and fees in advance of trading. The hedge fund deposits cash or other collateral into a margin account with the prime broker, and the prime broker essentially lends the hedge fund the shares, bonds, or derivatives to make additional investments. The margin account represents the hedge fund's net equity in its positions. The minimum margin required depends on the riskiness of the investment portfolio and the creditworthiness of the hedge fund.

Leverage is a large part of the reason that some hedge funds either earn larger-than-normal returns or suffer significant losses. If the margin account or the hedge fund's equity in a position declines below a certain level, the lender initiates a margin call and requests that the hedge fund put up more collateral. An inability to meet margin calls can have the effect of magnifying or locking in losses because the hedge fund may have to liquidate (close) the losing position. This liquidation can lead to further losses if the order size is sufficiently large to move the security's market price before the fund can sufficiently eliminate the position.

Under normal conditions, the application of leverage may be necessary for yielding meaningful returns from given quantitative, arbitrage, or relative value strategies. But with added leverage comes increased risk. For example, in August 2007, a sudden cascade of problems arose from the application of leverage: so-called unattractive stocks rallied sharply over a 10-day period of short covering while stocks deemed fundamentally "attractive" declined as quantitative managers were forced to sell them. Many managers positioning themselves as market neutral and "safe" quickly lost 20%–25% of their portfolio value. The tail-risk aspect of such strategies was revealed to be the leverage and the “crowding” phenomenon of so many market participants holding similar positions that needed to be downsized when the market moved. Leverage can cause left-tailed events when mechanical downsizing and crowd psychology impacts come into play. Market watchers said that the source of the 2007 upheaval was no more than the butterfly effect of one levered quantitative manager after another deciding to de-lever at the same time.

The application of leverage to a strategy may not be revealed by studying the Sharpe ratio, the Sortino ratio, or another financial performance metric. Instead, it may underpin the strategy undetected. Understanding the impact of leverage on a portfolio is less about what the track record has been and more about *how* that track record was created. Analysts evaluating the alternative asset space should, therefore, scrutinize the returns of any alternative asset that relies on high leverage.

9.4.2 *Illiquidity and Potential Redemption Pressures*

A second qualitative issue for many hedge funds is the manner in which portfolios are marked to market. This issue is less important for traditional long–short managers trading only publicly traded equities, but it still can cause problems.

Consider the long–short equity manager involved in thinly traded small-capitalization stocks that see only a few thousand shares traded daily. Perhaps the manager was able at one point to source a block of stock from a retiring firm founder in order to establish the long exposure. An outside fund administrator uses a daily quoted price to value these shares, but the manager knows there is little chance of actually liquidating all the shares at that price.

This problem will be worse for a convertible bond manager, a credit-oriented manager, or a structured products manager who faces wider bid–offer spreads and deals in some securities that are particularly illiquid or that trade only “by appointment” (in other words, securities that trade so infrequently an appointment is almost necessary).

Proper valuations are important for calculating performance and meeting potential redemptions without incurring undue transaction costs for liquidating exposures. The frequency with which alternative assets are valued and how they are valued vary among funds. Hedge funds are generally valued by the manager internally, and these valuations are confirmed by an outside administrator on a daily or perhaps weekly basis; performance is reported to investors by the administrator on a monthly or quarterly basis. The valuation may use market values or estimated values of the underlying positions. When market prices or quotes are used for valuation, funds may differ in which price or quote they use (bid price, ask price, average quote, or median quote). A common practice is to use an average of the bid and the ask. A more conservative and theoretically more accurate approach is to use bid prices for long positions and ask prices for short positions because these are more realistic prices at which the positions could be closed.

In some instances, the underlying positions may be in highly illiquid or even non-traded investments, and since such securities may have no reliable market values, it becomes necessary to estimate values. Starting in 2006, with later amendments and modifications, GAAP accounting rules in the United States created a methodology that involves the categorization of investments into three buckets: Level 1, 2, and 3 asset pricing. Level 1 assets involve situations where an exchange-traded, publicly traded price is available and is mandated to be used for valuation purposes. When such pricing is not available, outside broker quotes, or Level 2 values, may be relied on. As a final recourse, when such broker quotes are deemed either unavailable or unreliable, assets may be computed using only internal models—a process referred to as Level 3 asset pricing. No matter the model used by a manager in such circumstances, it should be independently tested, benchmarked, and calibrated to industry-accepted standards to ensure a consistency of approach. Because of the potential for conflicts of interest when applying estimates of value, hedge funds must develop procedures for in-house valuation, communicate these procedures to clients, and adhere to them.

Notwithstanding best practice, the very nature of assets that can be valued only on a "mark-to-model" basis can and should be a concern for the alternative asset investor. A model may reflect an imperfect theoretical valuation and not a true liquidation value. The illiquid nature of these assets means that estimates, rather than observable transaction prices, may well have factored into any valuation. As a result, returns may be smoothed or overstated and the volatility of returns, understated. Any investor relying on a risk metric such as a Sharpe ratio for such illiquid strategy situations is basically involved in self-deception as to the true drawdown risk of the strategy. As a generalized statement, any investment vehicle that is heavily involved with Level 3–priced assets deserves increased scrutiny and due diligence.

EXAMPLE 1

Hedge Fund Valuation

A hedge fund with a market-neutral strategy restricts its investment universe to domestic publicly traded equity securities that are actively traded on an exchange or between over-the-counter brokers. In calculating net asset value, the fund is most likely to use which of the following to value underlying positions?

A Exchange last-trade pricing and/or averaged quotes of any available over-the-counter bid–offer spreads

B Average quotes adjusted for liquidity

C Bid price for shorts and ask price for longs

Solution:

A is correct. The fund is most likely to use exchange-traded last-trade pricing (Level 1 pricing) or averaged quotes of publicly available bid–offer spreads (Level 2 pricing). The securities are actively traded, so no liquidity adjustment is required.

If the fund uses bid–ask prices, it will use ask prices for shorts and bid prices for longs; these are the prices at which the positions are closed.

Understanding and evaluating "tail events"—low-probability, high-severity instances of stress—is an important yet extraordinarily difficult aspect of the risk management process, particularly when dealing with illiquid securities. Stress testing and scenario analysis are often used as a complement to other risk measures to develop a better understanding of a portfolio's potential loss under both normal and stressed market conditions. Stress testing involves estimating losses under extremely unfavorable conditions.

Another factor that can lock in or magnify losses for hedge funds is investor redemptions. Redemptions frequently occur when a hedge fund is performing poorly. Redemptions may require the hedge fund manager to liquidate some positions and potentially receive particularly disadvantageous prices when forced to do so by redemption pressures, while also incurring transaction costs. Funds sometimes charge redemption fees (typically payable to the remaining investors) to discourage redemption and to offset the transaction costs for remaining investors. Notice periods provide an opportunity for the hedge fund manager to liquidate a position in an orderly fashion without magnifying the losses. Lockup periods—time periods when investors cannot withdraw their capital—provide the hedge fund manager the required time to implement and potentially realize a strategy's expected results. If the hedge fund receives a drawdown request shortly after a new investment, the lockup period forces the investors who made the request to stay in the fund for a period of time rather than be allowed to immediately withdraw. A hedge fund's ability to demand a long lockup period while raising a significant amount of investment capital depends a great deal on the reputation of either the firm or the hedge fund manager. Funds of hedge funds may offer more redemption flexibility than is afforded to direct investors in hedge funds because of special redemption arrangements with the underlying hedge fund managers, the maintenance of added cash reserves, access to temporary bridge-loan financing, or the simple avoidance of less liquid hedge fund strategies.

Ideally, redemption terms should be designed to match the expected liquidity of the assets being invested in, but even with careful planning, an initial drawdown can turn into something far more serious when it involves illiquid and obscure assets. These left-tailed loss events are not easily modeled for hedge funds.

EXAMPLE 2

Effect of Redemption

A European credit hedge fund has a very short redemption notice period—one week—because the fund's managers believe it invests in highly liquid asset classes and is market neutral. The fund has a small number of holdings that represent a significant portion of the outstanding issue of each holding. The fund's lockup period has expired. Unfortunately, in one particular month, because of the downgrades of two large holdings, the hedge fund has a drawdown (decline in NAV) of more than 10%. The declines in value of the two holdings result in margin calls from their prime broker, and the drawdown results in requests to redeem 50% of total partnership interests. The combined requests are *most likely* to:

A force the hedge fund to liquidate or unwind 50% of its positions in an orderly fashion throughout the week.

B have little effect on the prices received when liquidating the positions because the hedge fund has a week before the partnership interests are redeemed.

C result in a forced liquidation, which is likely to further drive down prices and result in ongoing pressures on the hedge fund as it tries to convince nervous investors to remain in the fund.

Solution:

C is correct. One week may not be enough time to unwind such a large portion of the fund's positions in an orderly fashion that also does not further drive down prices. A downgrade is not likely to have a temporary effect, so even if other non-losing positions are liquidated to meet the redemption requests, it is unlikely that the two large holdings will return to previous or higher values in short order. Also, the hedge fund may have a week to satisfy the requests for redemptions, but the margin call must be met immediately. Overall, sudden redemptions at the fund level can have a cascading negative impact on a fund.

The previous discussion applies mostly to liquid alternative asset strategies, principally hedge funds. In the world of private equity and real estate alternative assets, other methodologies are used to measure relative performance. Here, we find more issues with lagged and smoothed pricing.

It is important to note that although the ratios we have considered are among the best performance and risk tools available, they can still lead to inappropriate conclusions, as shown in the following example.

EXAMPLE 3

Steamboat Structured Products LLC is a manager that specializes in the purchase and sale of residential mortgage-backed securities, sometimes hedged with other put option and short equity index exposures. For the most part, its strategy is geared to earn an attractive monthly mortgage payment on as low of a loan-to-value ratio of a secured property as possible. Steamboat managers are generally very good at sourcing such investments.

However, Steamboat's managers discover that in the wake of the 2008 financial crisis, many small pieces of "odd-lot" RMBS paper are being sold by financial institutions across the United States, sometimes with face values of just \$200,000–\$400,000. This paper is often offered at a 10% or 15% discount to the price a more sizable and significant "round-lot" block of \$1 million–\$2 million of the same type of paper might trade.

Steamboat managers know that their outside administrator will place only a single round-lot valuation on each security identifier (often referred to as a CUSIP) in their portfolio, and thus they can't resist starting to buy as many of these odd-lot offerings as possible. When they do, their administrator does indeed immediately mark them higher, which, as Steamboat continues this practice month after month, creates a lovely track record of constant "trading profits" on top of their natural coupon income. To the greatest extent possible, Steamboat hopes to create odd-lot "matchers" to eventually aggregate its exposures into larger tradable blocks. But what happens instead is that the firm ends up with 2,000 small-position line items in its book and a glorious track record.

This is a story of an illiquid asset class that is naturally hard to trade and mark. Steamboat's Sharpe ratio, Sortino ratio, and Calmar ratio all look stronger than they really should be, mostly because of accounting conventions. If Steamboat were ever forced to sell its portfolio, the odd-lot discount earned as trading revenue would largely disappear and be replaced by the true liquidation values that Steamboat would find in the market from others for its accumulated odd-lot portfolio.

Performance ratio analysis must, therefore, be discounted when dealing with illiquid securities as described here. Although not necessarily fraudulent, standard accounting practices can be purposefully gamed.

PRACTICE QUESTIONS FOR LOS I

Question 1

The Sharpe ratio is a less-than-ideal performance measure for alternative investments because:

A it uses a semi-deviation measure of volatility.

B returns of alternative assets are not normally distributed.

C alternative assets exhibit low correlation with traditional asset classes.

Solution:

B is correct. The Sharpe ratio assumes normally distributed returns. However, alternative assets tend to have non-normal return distributions with significant skewness (fat tails in one direction or the other) and kurtosis (sharper peak than a normal distribution has, with fatter tails). Therefore, the Sharpe ratio may not be a good risk-adjusted performance measure to rely on for alternative investments.

A is incorrect because the Sharpe ratio does not use a semi-deviation measure of volatility; it uses standard deviation. The Sortino ratio uses a semi-deviation measure of volatility. Further, the use of semi-deviation instead of standard deviation actually makes the Sortino ratio a more attractive measure of alternative asset performance than the Sharpe ratio.

C is incorrect because correlation does not enter into the calculation of the Sharpe ratio. However, it is true that alternative assets can have low correlations with other asset classes. In contrast to the Sharpe ratio, the Treynor ratio incorporates the beta of the alternative asset relative to a benchmark, which is conceptually similar to correlation.

Question 2

Which of the following is true regarding private equity performance calculations?

A The money multiple calculation relies on the amount and timing of cash flows.

B The IRR calculation involves the assumption of two rates.

C Because private equity funds have low volatility, accounting conventions allow them to use a lagged mark-to-market process.

Solution:

B is correct. The determination of an IRR involves certain assumptions about a financing rate to use for outgoing cash flows (typically a weighted average cost of capital) and a reinvestment rate assumption to make on incoming cash flows (which must be assumed and may or may not actually be earned).

A is incorrect because the money multiple calculation completely ignores the timing of cash flows.

C is incorrect because it is somewhat of a reversal of cause and effect: Private equity (PE) funds can appear to have low volatility because of the lag in their mark-to-market process. It's not that PE investments don't actually rise and fall behind the scenes with economic influences, but accounting conventions may simply leave longer-lived investments marked at their initial cost for some time or with only modest adjustments to such carrying value until known impairments or realization events begin to transpire. Also, because PE funds are not easily marked to market, their returns can appear somewhat smoothed, making them appear more resilient and less correlated with other assets than they really are. The slowness to re-mark them can unfortunately be confused by investors as an overall lack of volatility.

Question 3

Which is *not* true of mark-to-model valuations?

A Return volatility may be understated.

B Returns may be smooth and overstated.

C A calibrated model will produce a reliable liquidation value.

Solution:

C is correct. It is not true that a calibrated model will produce a reliable liquidation value in a mark-to-model valuation. The need to use a model for valuation arises when an asset is so illiquid that there are not reliable market values available. A model may reflect only an imperfect theoretical valuation, not a true liquidation value, should liquidation become necessary. The illiquid nature of alternative assets means that estimates, rather than observable transaction prices, may have been used for valuation purposes.

A and B are not correct because they are both true statements.

Question 4

An analyst wanting to assess the downside risk of an alternative investment is *least likely* to use the investment's: (2020 Q34)

A Sortino ratio.

B value at risk (VaR).

C standard deviation of returns.

Solution:

C is correct. Downside risk measures focus on the left side of the return distribution curve, where losses occur. The standard deviation of returns assumes that returns are normally distributed. Many alternative investments do not exhibit close-to-normal distributions of returns, which is a crucial assumption for the validity of a standard deviation as a comprehensive risk measure. Assuming normal probability distributions when calculating these measures will lead to an underestimation of downside risk for a negatively skewed distribution. Both the Sortino ratio and the VaR measure are measures of downside risk.

CALCULATING FEES AND RETURNS

10

j calculate and interpret returns of alternative investments on both before-fee and after-fee bases

Now that we have considered alternative assets, their risks and rewards, their investment characteristics, and headline strategies, performing a few calculations will take us close to the real-world evaluations made by investment managers and the investors who hire them.

10.1 Alternative Asset Fee Structures and Terms

EXAMPLE 4

Incentive Fees Relative to Waterfall Types

A PE fund invests $15 million in a nascent luxury yacht manufacturer and $17 million in a new casino venture. The yacht manufacturer generates a $9 million profit when the company is acquired by a larger competitor, but the casino venture turns out to be a flop when its state licensing is eventually denied and it generates a $10 million loss. If the manager's carried interest incentive fee is

20% of the profits, what would this incentive be with a European-style waterfall whole-of-fund approach, and what would it be if the incentive is paid on an American-style waterfall deal-by-deal basis (assuming no clawback applies)?

Solution:

In aggregate, the fund lost money (+\$9 million – \$10 million = –\$1 million), so with a European-style whole-of-fund waterfall and assuming the time period for the gain and the loss are the same, there is no incentive fee. With an American-style waterfall, the GP could still earn 20% × \$9 million = \$1.8 million on the yacht company, thereby further compounding the loss to the ultimate investor to –\$2.8 million net of fees.

	Aggregate	**Yacht Company**	**Casino Venture**
Investment	\$32 m	\$15 m	\$17 m
Profit/Loss	-\$1 m	\$9 m	-\$10 m
Incentive by Deal		20% × \$9m = \$1.8m	\$0 m
Total Incentive	\$0 m	\$1.8 m	

If the gain and loss in this example transpired sequentially over different years, perhaps with the yacht company gain occurring first and then the Casino venture loss later on, the outcome for a European-style waterfall would typically result in an initial accrued incentive fee for the yacht manufacturer gain, but if there is a clawback provision in place, then there would be a clawback of that fee for the investor in the subsequent year when the casino venture loss is eventually realized, still resulting in no overall incentive fee. Waterfall language and clawback provisions on fees are very important to study and understand in offering memorandums, and these terms can vary widely.

10.2 Custom Fee Arrangements

Although "2 and 20" and "1 and 10" are commonly quoted fee structures for hedge funds and funds of funds, respectively, many variations exist.

1. *Fees based on liquidity terms and asset size:* Hedge funds may charge different rates depending on the liquidity terms that an investor is willing to accept (longer lockups are generally associated with lower fees), and hedge fund managers may discount their fees for larger investors or for placement agents who introduced these investors. Different investors in the same fund may well face different fee structures. Hedge fund managers negotiate terms, including fees and notice and lockup periods, with individual investors via side letters, which are special amendments to a standard offering memorandum's terms and conditions. For perspective, management fees for large LPs could range from 0.5% to 1.5%, with incentive fees reduced to 10%–15%, depending on the mandate. Such reductions can be meaningful in terms of net realized returns. However, some smaller hedge funds with strong performance (and capacity constraints) are able to maintain higher fees and may even turn down business from larger investors rather than agree to a lower fee.
2. *Founders' shares*: As a way to entice early participation in start-up and emerging hedge funds, managers have increasingly offered incentives known as founders' class shares. Founders' shares entitle investors to a lower fee structure, such as 1.5 and 10 rather than 2 and 20, and are typically available to be applied only to the first \$100 million in assets, although cutoff thresholds vary. Another option is to reduce the fees for early founders' share investors once the fund achieves a

critical mass of assets or performance targets. Both paths act as an incentive to spur investors to make faster investment commitments than might otherwise be the case.

3 *"Either/or" fees*: As a pushback against high hedge fund fees, institutional investors, such as the Teacher Retirement System of Texas, have recently requested a new fee model that some hedge fund managers have started to accept in return for substantive allocations. Managers agree *either* to charge a 1% management fee (simply to cover expenses during down years) *or* to receive a 30% incentive fee above a mutually agreed-on annual hurdle (to incentivize and reward managers during up years), whichever is greater. If a manager were to go without profits for a year or two, the 1% management fee becomes effectively an advance against an eventual 30% incentive-fee year (thereby reducing that future-year incentive fee by the prior years' advanced management fees. Although a far cry from the traditional 2 and 20, such novel fee structures, which are designed to reward performance and the delivery of true alpha above a benchmark, are likely to become even more in demand in the institutional hedge fund industry. Hedge funds will likely also endeavor to charge high-net-worth investors (with smaller commitment sizes) more traditional fees.

The following example demonstrates fee structures and their effect on the resulting returns to investors.

EXAMPLE 5

Fee and Return Calculations

AWJ Capital is a hedge fund with $100 million of initial investment capital. It charges a 2% management fee based on year-end AUM and a 20% incentive fee. In its first year, AWJ Capital has a 30% return. Assume management fees are calculated using end-of-period valuation.

1 What are the fees earned by AWJ if the incentive and management fees are calculated independently? What is an investor's effective return given this fee structure?

2 What are the fees earned by AWJ assuming that the incentive fee is calculated from the return net of the management fee? What is an investor's net return given this fee structure?

3 If the fee structure specifies a hurdle rate of 5% and the incentive fee is based on returns in excess of the hurdle rate, what are the fees earned by AWJ assuming the performance fee is calculated net of the management fee? What is an investor's net return given this fee structure?

4 In the second year, the fund value declines to $110 million. The fee structure is as specified for Question 1 but also includes the use of a high-water mark (computed net of fees). What are the fees earned by AWJ in the second year? What is an investor's net return for the second year given this fee structure?

5 In the third year, the fund value increases to $128 million. The fee structure is as specified in Questions 1 and 4. What are the fees earned by AWJ in the third year? What is an investor's net return for the third year given this fee structure?

Solution to 1:

AWJ fees

$130 million × 2% = $2.6 million management fee.

($130 – $100) million × 20% = $6 million incentive fee.

Total fees to AWJ Capital = $8.6 million.

Investor return: ($130 – $100 – $8.6) million/$100 million = 21.40%.

Solution to 2:

$130 million × 2% = $2.6 million management fee.

($130 – $100 – $2.6) million × 20% = $5.48 million incentive fee.

Total fees to AWJ Capital = $8.08 million.

Investor return: ($130 – $100 – $8.08) million/$100 million = 21.92%.

Solution to 3:

$130 million × 2% = $2.6 million management fee.

($130 – $100 – $5 – $2.6) million × 20% = $4.48 million incentive fee.

Total fees to AWJ Capital = $7.08 million.

Investor return: ($130 – $100 – $7.08) million/$100 million = 22.92%.

Solution to 4:

$110 million × 2% = $2.2 million management fee.

No incentive fee because the fund has declined in value.

Total fees to AWJ Capital = $2.2 million.

Investor return: ($110 – $2.2 – $121.4) million/$121.4 million = –11.20%.

The beginning capital position in the second year for the investors is ($130 – $8.6) million = $121.4 million. The ending capital position at the end of the second year is ($110 – $2.2) million = $107.8 million.

Solution to 5:

$128 million × 2% = $2.56 million management fee.

($128 – $121.4) million × 20% = $1.32 million incentive fee.

The $121.4 million represents the high-water mark established at the end of Year 1.

Total fees to AWJ Capital = $3.88 million.

Investor return: ($128 – $3.88 – $107.8) million/$107.8 million = 15.14%.

The ending capital position at the end of Year 3 is $124.12 million. This amount is the new high-water mark.

As the previous example illustrates, the return to an LP investor in a fund may differ significantly from the quoted return for the fund as a whole, which generally reflects the return that a “Day 1” investor who made no capital movements would have earned. Hedge fund databases and indexes generally report performance net of fees. If fee structures vary, however, the actual net-of-fees returns earned by various investors often may vary from the one included in a given database or index.

The multi-layered fee structure of funds of hedge funds has the effect of further diluting returns to the investor, but as discussed, this disadvantage may be balanced by positive features, such as access to a diversified portfolio and to hedge funds that may otherwise be closed to direct investments, as well as expertise in due diligence in hedge fund selection. There is thus both added cost and added value.

Generally, over time many funds of funds have earned a reputation for being "fast" money because their managers tend to be the first to redeem their investment when a hedge fund performs poorly. They may also have negotiated more favorable redemption terms—a shorter lockup or notice period, for example. Because of the overall compression of hedge fund returns and issues of excessive fee layering, many fund-of-funds managers have been pressured to drop their incentive fees and simply charge a flat management fee.

KNOWLEDGE CHECK

Fill in the blank: Fund offering documents typically offer terms that include a ___________, whereby incentive fees will accrue, and subsequently be paid, only on new earnings above and beyond the recoupment of any prior losses.

Solution:

Fund offering documents typically offer terms that include a high-water mark, whereby incentive fees will accrue, and subsequently be paid, only on new earnings above and beyond the recoupment of any prior losses.

KNOWLEDGE CHECK

True or false: Advantages of funds-of-hedge funds include: due diligence in selecting individual hedge funds, access to hedge funds that may be closed to direct investments, and dilution of returns to the investor.

A True, statement to support this option on why it's true.

B False, statement to support this option on why it's false.

Solution:

B is correct; the statement is false. Although these three attributes are indeed true of funds of hedge funds, the dilution of returns to the investor is a disadvantage for the investor, not an advantage. The "due diligence" and "access" attributes are advantages.

EXAMPLE 6

Comparison of Returns: Investment Directly into a Hedge Fund or through a Fund of Hedge Funds

An investor is contemplating investing €100 million in either the ABC Hedge Fund (ABC HF) or the XYZ Fund of Funds (XYZ FOF). XYZ FOF has a "1 and 10" fee structure and invests 10% of its AUM in ABC HF. ABC HF has a standard "2 and 20" fee structure with no hurdle rate. Management fees are calculated on an annual basis on AUM at the beginning of the year. For simplicity, assume that management fees and incentive fees are calculated independently. ABC HF has a 20% return for the year before management and incentive fees.

1 Calculate the return to the investor from investing directly in ABC HF.
2 Calculate the return to the investor from investing in XYZ FOF. Assume that the other investments in the XYZ FOF portfolio generate the same return before management fees as those of ABC HF and that XYZ FOF has the same fee structure as ABC HF.
3 Why would the investor choose to invest in a fund of funds instead of a hedge fund given the effect of the "double fee" demonstrated in the answers to Questions 1 and 2?

Solution to 1:

ABC HF has a profit before fees on a €100 million investment of €20 million (= €100 million × 20%). The management fee is €2 million (= €100 million × 2%), and the incentive fee is €4 million (= €20 million × 20%). The return to the investor is 14% [= (20 – 2 – 4)/100].

Solution to 2:

XYZ FOF earns a 14% return or €14 million profit after fees on €100 million invested with hedge funds. XYZ FOF charges the investor a management fee of €1 million (= €100 million × 1%) and an incentive fee of €1.4 million (= €14 million × 10%). The return to the investor is 11.6% [= (14 – 1 – 1.4)/100].

Solution to 3:

This scenario assumes that returns are the same for all underlying hedge funds. In practice, this result will not likely be the case, and XYZ FOF may provide due diligence expertise and potentially valuable diversification.

10.3 Alignment of Interests and Survivorship Bias

The alternative asset business is attractive to portfolio managers because of how significant the fees can be if the fund performs well and AUM are significant. But as discussed previously, high fees destroy value and reduce the attractiveness of alternative investing. If a hedge fund manager can stay in business for just four to five years with acceptable returns, incentive fee allocations to the general partner can be substantive.

Comparatively, the commitment of a private equity or real estate manager to stay in business for 6–10 years (often with offering memorandum language that first attempts to return all capital to investors, then attempts in many instances to pay a preferred minimum return to investors, and then only belatedly allows a manager to crystalize, or realize, its own incentive fees) is arguably a more aligned overall incentive structure. However, the overall time commitment for the investor is obviously of a much longer duration. The investor runs a larger risk of being stuck in something that might be a disappointment, as opposed to an ability to simply move on, as with hedge funds. Landing a private equity commitment basically guarantees a manager a very attractive management fee runway for an extended period of time.

So, there remain these types of trade-offs that still make the alternative investing world overall quite lucrative regardless of the specific fund or fee format. Because of this generally lucrative incentive structure—particularly for hedge funds and their ability to crystalize incentive fees (even on gains that may still be unrealized) on an annualized basis—many new hedge funds launched in the late 1990s and early 2000s. Not all hedge funds, however, remain in business for very long. One study suggests that more than a quarter of all hedge funds fail within the first three years because of performance problems that result in investor defections and the ultimate failure to generate sufficient revenue to cover the fund's operating costs. This is one reason survivorship bias is a major problem with hedge fund indexes. *Backfill bias* is another problem: Certain surviving hedge funds may be added to databases and various hedge fund indexes only after they are initially successful and start to report their returns. Because of survivorship and backfill biases, hedge fund indexes may not reflect actual hedge fund performance but, rather, reflect only the performance of hedge funds that are performing well and thus "survived."

EXAMPLE 7

1 Clawbacks Due to Return Timing Differences

The Granite Rock Fund makes investments in leveraged-buyout Company A and start-up Company B, each for $10 million. One year later, the leveraged-buyout company returns a $14 million profit, and two years later, the start-up company turns into a complete bust, deemed to be worth zero.

If the GP's carried interest incentive fee is 20% of aggregate profits and there is a clawback provision, how much carried interest will the GP initially accrue and ultimately receive?

Solution:

From leveraged-buyout Company A, the GP would initially accrue a 20% of $14 million profit at the end of the first year, equal to $2.8 million. Typically, this amount would be held in escrow for the benefit of the GP but not actually paid.

But then the GP loses $10 million of the initial $14 million gain, so the aggregate whole-of-fund gain at the end of the second year would be only $4 million; this amount times 20% would result in only an $800,000 incentive fee. The general partner would then have to return $2 million of the previously accrued incentive fees to the capital accounts of LP investors because of the clawback provision.

2 Soft and Hard Hurdles

A real estate investment fund has a $100 million initial drawdown structure in its first year and fully draws this capital to purchase a property. The fund has a soft hurdle preferred return to investors of 8% per annum and an 80%/20% carried interest incentive split thereafter. At the end of year two, the property is sold for a total of $160 million.

What are the correct distributions to the LPs and to the GP? And how would these have been different if the real estate investment fund had a hard hurdle of 8% per annum instead of a soft hurdle?

Solution:

One needs to construct a waterfall of cash flows.

First, the LPs would be due their $100 million initial investment.

Then, they would be due $16 million (8% preferred return on initial capital for two years).

The soft hurdle has been met, and the GP is ultimately due 20% of $60 million, or $12 million, which would be paid to the GP next as a catch-up to the achieved hurdle return.

The residual amount would be $160 million – $100 million – $16 million – $12 million = $32 million. This amount would then be split 80% to the LPs and 20% to the GP, or $25.6 million and $6.4 million, respectively.

So, the total payout with a soft annual hurdle of 8% of the $160 million would end up with the following waterfall:

	LP	GP
Return of Capital	$100 m	
8% Prfd per Annum	$16 m	
GP Catch-Up 20%		$12 m
80%/20% Split	$25.6 m	$6.4 m
Total Payout	$141.6 m	18.4 m

If the fund instead had a hard hurdle rate, only the amount above the $100 return of capital and $16 million preferred return would be subject to the 20% carried interest incentive to the GP: 20% × $44 million = $8.8 million, quite a bit less than the carried interest payment with the soft hurdle. The LPs would be due the balance of $35.2 million ($44 million – $8.8 million incentive). This would result in the following total payout:

	LP	GP
Return of Capital	$100 m	
8% Prfd per Annum	$16 m	
80%/20% Split Above Hurdle	$35.2 m	$8.8 m
Total Payout	$151.2 m	8.8 m

PRACTICE QUESTIONS FOR LOS J

Question 1

United Capital is a hedge fund with $250 million of initial capital. United charges a 2% management fee based on assets under management at year end and a 20% incentive fee based on returns in excess of an 8% hurdle rate. In its first year, United appreciates 16%. Assume management fees are calculated using end-of-period valuation. The investor's net return assuming the performance fee is calculated net of the management fee is *closest* to: (2020 Q26)

A 11.58%.

B 12.54%.

C 12.80%.

Solution:

B is correct. The net investor return is 12.54%, calculated as follows:

End-of-year capital = $250 million × 1.16 = $290 million.

Management fee = $290 million × 2% = $5.8 million.

Hurdle amount = 8% of $250 million = $20 million.

Incentive fee = ($290 – $250 – $20 – $5.8) million × 20% = $2.84 million.

Total fees to United Capital = ($5.8 + $2.84) million = $8.64 million.

Investor net return: ($290 – $250 – $8.64)/$250 = 12.54%.

Question 2

Capricorn Fund of Funds invests GBP100 million in each of Alpha Hedge Fund and ABC Hedge Fund. Capricorn Fund of Funds has a "1 and 10" fee structure. Management fees and incentive fees are calculated independently at the end of each year. After one year, net of their respective management and incentive fees, Capricorn's investment in Alpha is valued at GBP80 million and Capricorn's investment in ABC is valued at GBP140 million. The annual return to an investor in Capricorn Fund of Funds, net of fees assessed at the fund-of-funds level, is *closest* to: (2020 Q27)

A 7.9%.

B 8.0%.

C 8.1%.

Solution:

A is correct because the net investor return is 7.9%, calculated as follows:

First, note that "1 and 10" refers to a 1% management fee and a 10% incentive fee.

End-of-year capital = GBP140 million + GBP80 million = GBP220 million.

Management fee = GBP220 million × 1% = GBP2.2 million.

Incentive fee = (GBP220 – GBP200) million × 10% = GBP2 million.

Total fees to Capricorn = (GBP2.2 + GBP2) million = GBP4.2 million.

Investor net return: (GBP220 – GBP200 – GBP4.2)/GBP200 = 7.9%.

Question 3

The following information applies to Rotunda Advisers, a hedge fund:

- $288 million in AUM as of prior year end
- 2% management fee (based on year-end AUM)
- 20% incentive fee calculated:
  - net of management fee
  - using a 5% soft hurdle rate
  - using a high-water mark (high-water mark is $357 million)
- Current-year fund gross return is 25%.

The total fee earned by Rotunda in the current year is *closest* to: (2020 Q28)

A $7.20 million.

B $20.16 million.

C $21.60 million.

Solution:

A is correct. Although the gross return of Rotunda results in a $360 million gross NAV, the deduction of the $7.2 million incentive fee brings NAV to $352.8 million, which is below the prior high-water mark. Rotunda earns a management fee of $7.20 million but does not earn an incentive fee because the year-end fund value net of management fee does not exceed the prior high-water mark of $357 million. Since Rotunda is still also below the prior-year high-water mark, the hurdle rate of return is also basically irrelevant in this fee calculation.

The specifics of this calculation are as follows:

End-of-year AUM = Prior year-end AUM × (1 + Fund return) = \$288 million × 1.25 = \$360 million.

\$360 million × 2% = \$7.20 million management fee.

\$360 million – \$7.2 million = \$352.8 million AUM net of management fee.

The year-end AUM net of fees do not exceed the \$357 million high-water mark. Therefore, no incentive fee is earned.

Question 4

A hedge fund has the following fee structure:

Annual management fee based on year-end AUM	2%
Incentive fee	20%
Hurdle rate before incentive fee collection starts	4%
Current high-water mark	\$610 million

The fund has a value of \$583.1 million at the beginning of the year. After one year, it has a value of \$642 million before fees. The net percentage return to an investor for this year is *closest* to: (2020 Q29)

A 6.72%.

B 6.80%.

C 7.64%.

Solution:

C is correct. The management fee for the year is

\$642 × 0.02 = \$12.84 million.

Because the ending gross value of the fund of \$642 million exceeds the high-water mark of \$610 million, the hedge fund can collect an incentive fee on gains above this high-water mark but net of the hurdle rate of return. The incentive fee calculation becomes

{\$642 – [\$610 × (1 + 0.04)]} × 0.20 = \$1.52 million.

The net return to the investor for the year is

[(\$642 – \$12.84 – \$1.52)/\$583.1] – 1 ≈ 0.07638 ≈ 7.64%.

SUMMARY

This reading provides a comprehensive introduction to alternative investments. Some key points of the reading are as follows:

- Alternative investments are supplemental strategies to traditional long-only positions in stocks, bonds, and cash. Alternative investments include investments in five main categories: hedge funds, private capital, natural resources, real estate, and infrastructure.
- Alternative investment strategies are typically active, return-seeking strategies that also often have risk characteristics different from those of traditional long-only investments.

- Characteristics common to many alternative investments, when compared with traditional investments, include the following: lower liquidity, less regulation, lower transparency, higher fees, and limited and potentially problematic historical risk and return data.
- Alternative investments often have complex legal and tax considerations and may be highly leveraged.
- Alternative investments are attractive to investors because of the potential for portfolio diversification resulting in a higher risk-adjusted return for the portfolio.
- Investors can access alternative invests in three ways:
  - Fund investment (such as a in a PE fund)
  - Direct investment into a company or project (such as infrastructure or real estate)
  - Co-investment into a portfolio company of a fund
- Investors conduct due diligence prior to investing in alternative investments. The due diligence approach depends on the investment method (direct, co-investing, or fund investing).
- Operational, financial, counterparty, and liquidity risks may be key considerations for those investing in alternative investments. These risks can be analyzed during the due diligence process. It is critical to perform fund due diligence to assess whether (a) the manager can effectively pursue the proposed investment strategy; (b) the appropriate organizational structure and policies for managing investments, operations, risk, and compliance are in place; and (c) the fund terms appear reasonable.
- Many alternative investments, such as hedge and private equity funds, use a partnership structure with a general partner that manages the business and limited partners (investors) who own fractional interests in the partnership.
- The general partner typically receives a management fee based on assets under management or committed capital (the former is common to hedge funds, and the latter is common to private equity funds) and an incentive fee based on realized profits.
- Hurdle rates, high-water marks, lockup and notice periods, and clawback provisions are often specified in the LPA.
- The fee structure affects the returns to investors (limited partners), with a waterfall representing the distribution method under which allocations are made to LPs and GPs. Waterfalls can be on a whole-of-fund basis (European) or deal-by-deal basis (American).
- Hedge funds are typically classified by strategy. One such classification includes four broad categories of strategies: equity hedge (e.g., market neutral), event driven (e.g., merger arbitrage), relative value (e.g., convertible bond arbitrage), macro and CTA strategies (e.g., commodity trading advisers).
- Funds-of-hedge-funds are funds that create a diversified portfolio of hedge funds. These vehicles are attractive to smaller investors that don't have the resources to select individual hedge funds and build a portfolio of them.
- Private Capital is a broad term for funding provided to companies that is sourced from neither the public equity nor debt markets. Capital that is provided in the form of equity investments is called private equity, whereas capital that is provided as a loan or other form of debt is called private debt.

- Private equity refers to investment in privately owned companies or in public companies with the intent to take them private. Key private equity investment strategies include leveraged buyouts (e.g., MBOs and MBIs) and venture capital. Primary exit strategies include trade sale, IPO, and recapitalization.
- Private debt refers to various forms of debt provided by investors to private entities. Key private debt strategies include direct lending, mezzanine debt, and venture debt. Private debt also includes specialized strategies, such as CLOs, unitranche debt, real estate debt, and infrastructure debt.
- Natural resources include commodities (hard and soft), agricultural land (farmland), and timberland.
- Commodity investments may involve investing in actual physical commodities or in producers of commodities, but more typically, these investments are made using commodity derivatives (futures or swaps). One can also invest in commodities via a CTA (see hedge funds)
- Returns to commodity investing are based on changes in price and do not include an income stream, such as dividends, interest, or rent (apart from income earned on the collateral). However, timberland offers an income stream based on the sale of trees, wood, and other products. Timberland can be thought of as both a factory and a warehouse. Plus, timberland is a sustainable investment that mitigates climate-related risks.
- Farmland, like timberland, has an income component related to harvest quantities and agricultural commodity prices. However, farmland doesn't have the production flexibility of timberland, because farm products must be harvested when ripe.
- Real estate includes two major sectors: residential and commercial. Residential real estate is the largest sector, making up some 75% of the market globally. Commercial real estate primarily includes office buildings, shopping centers, and warehouses. Real estate property has some unique features compared with other asset classes, including heterogeneity (no two properties are identical) and fixed location.
- Real estate investments can be direct or indirect, in the public market (e.g., REITs) or private transactions, and in equity or debt.
- The assets underlying infrastructure investments are real, capital intensive, and long lived. These assets are intended for public use, and they provide essential services. Examples include airports, health care facilities, and power plants. Funding is often done on a public–private partnership basis.
- Social infrastructure assets are directed toward human activities and include such assets as educational, health care, social housing, and correctional facilities, with the focus on providing, operating, and maintaining the asset infrastructure.
- Infrastructure investments may also be categorized by the underlying asset's stage of development. Investing in infrastructure assets *that are to be constructed* is generally referred to as greenfield investment. Investing in *existing* infrastructure assets may be referred to as brownfield investment.
- Conducting performance appraisal on alternative investments can be challenging because these investments are often characterized by asymmetric risk–return profiles, limited portfolio transparency, illiquidity, product complexity, and complex fee structures.

- Traditional risk and return measures (such as mean return, standard deviation of returns, and beta) may provide an inadequate picture of alternative investments' risk and return characteristics. Moreover, these measures may be unreliable or not representative of specific investments.
- A variety of ratios can be calculated in order to review the performance of alternative investments, including the Sharpe ratio, Sortino ratio, Treynor ratio, Calmar ratio, and MAR ratio. In addition, batting average and slugging percentage can also be used. The IRR calculation is often used to evaluate private equity investments, and the cap rate is often used to evaluate real estate investments.
- Redemption rules and lockup periods can bring special challenges to performance appraisal of alternative investments.
- When comparing the performance of alternative investments versus an index, the analyst must be aware that indexes for alternative investments may be subject to a variety of biases, including survivorship and backfill biases.
- Analysts need to be aware of any custom fee arrangements in place that will affect the calculation of fees and performance. These can include such arrangements as fee discounts based on custom liquidity terms or significant asset size; special share classes, such as "founders' shares"; and a departure from the typical management fee + performance fee structure in favor of "either/or" fees

Portfolio Management

STUDY SESSION

Study Session 17	Portfolio Management (1)
Study Session 18	Portfolio Management (2)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to explain and demonstrate the use of fundamentals of portfolio and risk management, including return and risk measurement, and portfolio planning and construction.

© 2021 CFA Institute. All rights reserved.

PORTFOLIO MANAGEMENT

STUDY SESSION

17

Portfolio Management (1)

This study session introduces the concept of a portfolio approach to investments. The needs of individual and institutional investors are each examined, along with the range of available investment solutions. The three main steps in the portfolio management process (planning, execution, and feedback) are outlined. Common measures of portfolio risk and return and the introduction of modern portfolio theory—a quantitative framework for asset pricing and portfolio selection—then follow.

READING ASSIGNMENTS

Reading 48	Portfolio Management: An Overview by Owen M. Concannon, CFA, Robert M. Conroy, DBA, CFA, Alistair Byrne, PhD, CFA, and Vahan Janjigian, PhD, CFA
Reading 49	Portfolio Risk and Return: Part I by Vijay Singal, PhD, CFA
Reading 50	Portfolio Risk and Return: Part II by Vijay Singal, PhD, CFA

© 2021 CFA Institute. All rights reserved.

READING

48

Portfolio Management: An Overview

by Owen M. Concannon, CFA, Robert M. Conroy, DBA, CFA, Alistair Byrne, PhD, CFA, and Vahan Janjigian, PhD, CFA

Owen M. Concannon, CFA, is at Neuberger Berman (USA). Robert M. Conroy, DBA, CFA, is at the Darden School of Business, University of Virginia (USA). Alistair Byrne, PhD, CFA, is at State Street Global Advisors (United Kingdom). Vahan Janjigian, PhD, CFA, is at Greenwich Wealth Management, LLC (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the portfolio approach to investing;
☐	**b.** describe the steps in the portfolio management process;
☐	**c.** describe types of investors and distinctive characteristics and needs of each;
☐	**d.** describe defined contribution and defined benefit pension plans;
☐	**e.** describe aspects of the asset management industry;
☐	**f.** describe mutual funds and compare them with other pooled investment products.

INTRODUCTION

1

This reading provides an overview of portfolio management and the asset management industry, including types of investors and investment plans and products. A portfolio approach is important to investors in achieving their financial objectives. We outline the steps in the portfolio management process in managing a client's investment portfolio. We next compare the financial needs of different types of investors: individual and institutional investors. We then describe both defined contribution and defined benefit pension plans. The asset management[1] industry, which serves as a critical link between providers and seekers of investment capital around the world, is broadly discussed. Finally, we describe mutual funds and other types of pooled investment products offered by asset managers.

---

[1] Note that both "investment management" and "asset management" are commonly used throughout the CFA Program curriculum. The terms are often used interchangeably in practice.

© 2019 CFA Institute. All rights reserved.

2 PORTFOLIO PERSPECTIVE: DIVERSIFICATION AND RISK REDUCTION

a describe the portfolio approach to investing;

One of the biggest challenges faced by individuals and institutions is to decide how to invest for future needs. For individuals, the goal might be to fund retirement needs. For such institutions as insurance companies, the goal is to fund future liabilities in the form of insurance claims, whereas endowments seek to provide income to meet the ongoing needs of such institutions as universities. Regardless of the ultimate goal, all face the same set of challenges that extend beyond just the choice of what asset classes to invest in. They ultimately center on formulating basic principles that determine how to think about investing. One important question is: Should we invest in individual securities, evaluating each in isolation, or should we take a portfolio approach? By "portfolio approach," we mean evaluating individual securities in relation to their contribution to the investment characteristics of the whole portfolio. In the following section, we illustrate a number of reasons why a diversified portfolio perspective is important.

2.1 Historical Example of Portfolio Diversification: Avoiding Disaster

Portfolio diversification helps investors avoid disastrous investment outcomes. This benefit is most convincingly illustrated by examining what may happen when individuals have *not* diversified.

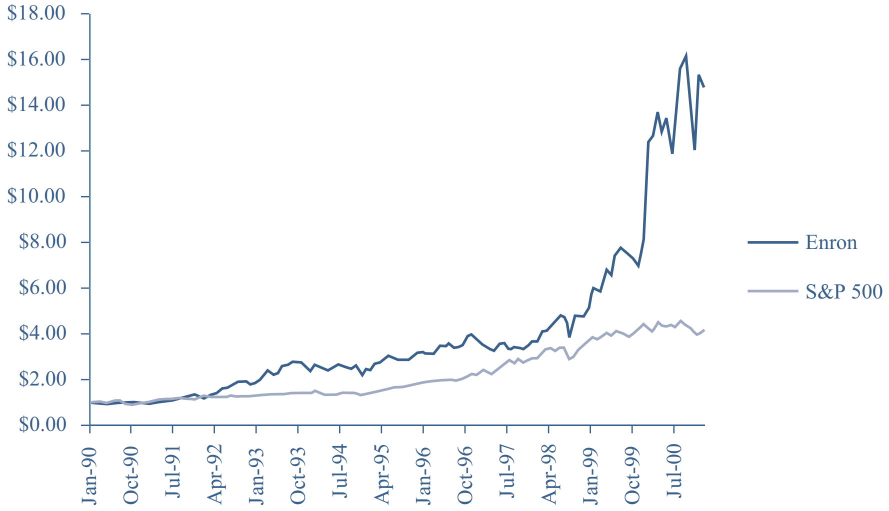
We are usually not able to observe how individuals manage their personal investments. However, in the case of US 401(k) individual retirement portfolios,[2] it is possible to see the results of individuals' investment decisions. When we examine their retirement portfolios, we find that some individual participants make sub-optimal investment decisions.

During the 1990s, Enron Corporation was one of the most admired corporations in the United States. A position in Enron shares returned over 27 percent per year from 1990 to September 2000, compared to 13 percent for the S&P 500 Index for the same time period.

---

[2] In the United States, 401(k) plans are employer-sponsored individual retirement savings plans. They allow individuals to save a portion of their current income and defer taxation until the time when the savings and earnings are withdrawn. In some cases, the sponsoring firm will also make matching contributions in the form of cash or shares. Individuals within certain limits have control of the invested funds and consequently can express their preferences as to which assets to invest in.

Exhibit 1 Value of US$1 Invested from January 1990 to September 2000 Enron vs. S&P 500 Composite Index (01/01/1990 = US$1.00)



Source: Thomson Reuters Datastream.

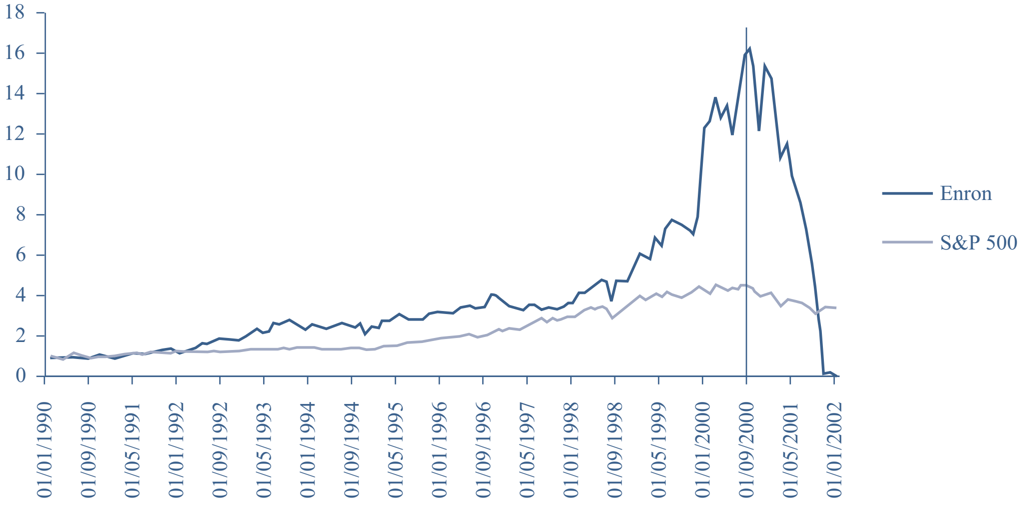
During this time period, thousands of Enron employees participated in the company's 401(k) retirement plan. The plan allowed employees to set aside some of their earnings in a tax-deferred account. Enron participated by matching the employees' contributions. Enron made the match by depositing required amounts in the form of Enron shares. Enron restricted the sale of its contributed shares until an employee turned 50 years old. In January 2001, the employees' 401(k) retirement accounts were valued at over US$2 billion, of which US$1.3 billion (or 62 percent) was in Enron shares. Although Enron restricted the sale of shares it contributed, less than US$150 million of the total of US$1.3 billion in shares had this restriction. The implication was that Enron employees continued to hold large amounts of Enron shares even though they were free to sell them and invest the proceeds in other assets.

A typical individual was Roger Bruce,[3] a 67-year-old Enron retiree who held all of his US$2 million in retirement funds in Enron shares. Between January 2001 and January 2002, Enron's share price fell from about US$90 per share to zero.

---

[3] Singletary (2001).

Exhibit 2 Value of US$1 Invested from January 1990 to January 2002 Enron vs. S&P 500 Composite Index (1/1/1990 = US$1.00)



Source: Thomson Reuters Datastream.

Employees and retirees who had invested all or most of their retirement savings in Enron shares, just like Mr. Bruce, experienced financial ruin. The hard lesson that the Enron employees learned from this experience was to "not put all your eggs in one basket."[4] Unfortunately, the typical Enron employee did have most of his or her eggs in one basket. Most employees' wages and financial assets were dependent on Enron's continued viability; hence, any financial distress on Enron would have a material impact on an employee's financial health. The bankruptcy of Enron resulted in the closing of its operations, the dismissal of thousands of employees, and its shares becoming worthless. Hence, the failure of Enron was disastrous to the typical Enron employee.

Enron employees were not the only ones to be victims of over-investment in a single company's shares. In the defined contribution retirement plans at Owens Corning, Northern Telecom, Corning, and ADC Telecommunications, employees all held more than 25 percent of their assets in the company's shares during a time (March 2000 to December 2001) in which the share prices in these companies fell by almost 90 percent. The good news in this story is that the employees participating in employer-matched 401(k) plans since 2001 have significantly reduced their holdings of their employers' shares.

Thus, by taking a diversified portfolio approach, investors can spread away some of the risk. Rational investors are concerned about the risk–return trade-off of their investments. The portfolio approach provides investors with a way to reduce the risk associated with their wealth without necessarily decreasing their expected rate of return.

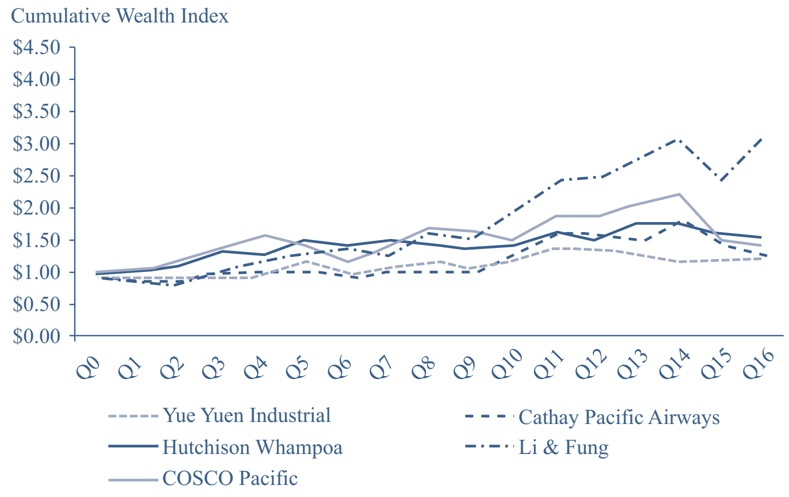
2.2 Portfolios: Reduce Risk

In addition to avoiding a potential disaster associated with over investing in a single security, portfolios also generally offer equivalent expected returns with lower overall **volatility** of returns—as represented by a measure such as standard deviation. Consider this simple example: Suppose you wish to make an investment in companies listed on the Hong Kong Stock Exchange (HKSE) and you start with a sample of five companies.[5] The cumulative returns for 16 fiscal quarters are shown in Exhibit 3.

---

[4] This expression, which most likely originated in England in the 1700s, has a timeless sense of wisdom.

Exhibit 3 Cumulative Wealth Index of Sample of Shares Listed on HKSE (initial amount= US$1.00)



Source: Thomson Reuters Datastream.

The individual quarterly returns for each of the five shares are shown in Exhibit 4. The annualized means and annualized standard deviations for each are also shown.[6]

Exhibit 4 Quarterly Returns (in Percent) for Sample of HKSE Listed Shares over 16 Fiscal Quarters

	Yue Yuen Industrial	Cathay Pacific Airways	Hutchison Whampoa	Li & Fung	COSCO Pacific	Equally Weighted Portfolio
Q1	−11.1%	−2.3%	0.6%	−13.2%	−1.1%	−5.4%
Q2	−0.5	−5.4	10.8	1.7	21.0	5.5
Q3	5.7	6.8	19.1	13.8	15.5	12.2
Q4	5.3	4.6	−2.1	16.9	12.4	7.4
Q5	17.2	2.4	12.6	14.5	−7.9	7.8
Q6	−17.6	−10.4	−0.9	4.4	−16.7	−8.2
Q7	12.6	7.4	4.2	−10.9	15.4	5.7
Q8	7.5	−0.4	−3.6	29.2	21.9	10.9
Q9	−7.9	1.3	−5.1	−2.0	−1.6	−3.1
Q10	8.2	27.5	0.1	26.0	−10.1	10.3
Q11	18.3	24.3	16.5	22.8	25.7	21.5
Q12	0.1	−2.6	−6.7	−0.4	0.3	−1.8

(continued)

---

[5] A sample of five companies from a similar industry group was arbitrarily selected for illustration purposes.

[6] Mean quarterly returns are annualized by multiplying the quarterly mean by 4. Quarterly standard deviations are annualized by taking the quarterly standard deviation and multiplying it by 2.

Exhibit 4 (Continued)

	Yue Yuen Industrial	Cathay Pacific Airways	Hutchison Whampoa	Li & Fung	COSCO Pacific	Equally Weighted Portfolio
Q13	–6.2	–4.2	16.7	11.9	11.1	5.8
Q14	–8.0	17.9	–1.8	12.4	8.4	5.8
Q15	3.5	–20.1	–8.5	–20.3	–31.5	–15.4
Q16	2.1	–11.8	–2.6	24.2	–6.1	1.2
Mean annual return	7.3%	8.7%	12.3%	32.8%	14.2%	15.1%
Annual standard deviation	20.2%	25.4%	18.1%	29.5%	31.3%	17.9%
Diversification ratio						71.9%

Source: Thomson Reuters Datastream.

Suppose you want to invest in one of these five securities next year. There is a wide variety of risk–return trade-offs for the five shares selected. If you believe that the future will replicate the past, then choosing Li & Fung would be a good choice. For the prior four years, Li & Fung provided the best trade-off between return and risk. In other words, it provided the most return per unit of risk. However, if there is no reason to believe that the future will replicate the past, it is more likely that the risk and return on the one security selected will be more like selecting one randomly. When we randomly selected one security each quarter, we found an average annualized return of 15.1 percent and an average annualized standard deviation of 24.9 percent, which would now become your expected return and standard deviation, respectively.

Alternatively, you could invest in an equally weighted portfolio of the five shares, which means that you would invest the same dollar amount in each security for each quarter. The quarterly returns on the equally weighted portfolio are just the average of the returns of the individual shares. As reported in Exhibit 4, the equally weighted portfolio has an average return of 15.1 percent and a standard deviation of 17.9 percent. As expected, the equally weighted portfolio's return is the same as the return on the randomly selected security. However, the same does not hold true for the portfolio standard deviation. That is, the standard deviation of an equally weighted portfolio is not simply the average of the standard deviations of the individual shares. In a more advanced reading we will demonstrate in greater mathematical detail how such a portfolio offers a lower standard deviation of return than the average of its individual components due to the correlations or interactions between the individual securities.

Because the mean return is the same, a simple measure of the value of diversification is calculated as the ratio of the standard deviation of the equally weighted portfolio to the standard deviation of the randomly selected security. This ratio may be referred to as the **diversification ratio**. In this case, the equally weighted portfolio's standard deviation is approximately 72 percent of the average standard deviation of the 5 stocks (24.9%). The diversification ratio of the portfolio's standard deviation to the individual asset's standard deviation measures the risk reduction benefits of a simple portfolio construction method, equal weighting. Even though the companies were chosen from a similar industry grouping, we see significant risk reduction. An even greater portfolio effect (i.e., lower diversification ratio) could have been realized if we had chosen companies from completely different industries.

This example illustrates one of the critical ideas about portfolios: Portfolios affect risk more than returns. In the prior section portfolios helped avoid the effects of downside risk associated with investing in a single company's shares. In this section we extended the notion of risk reduction through portfolios to illustrate why individuals and institutions should hold portfolios.

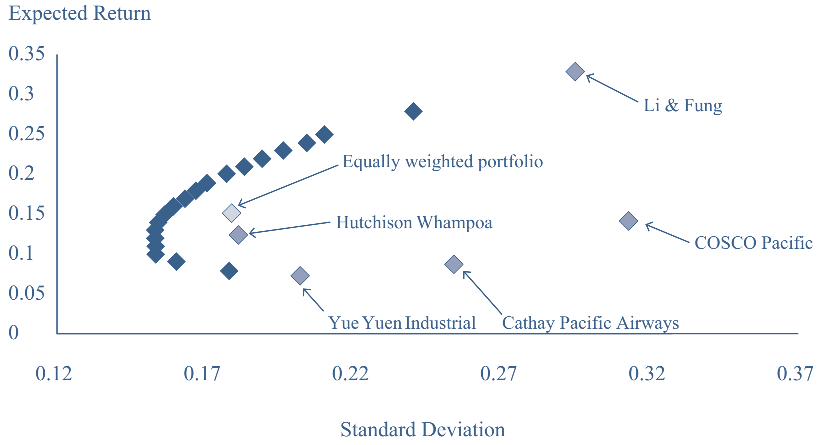
PORTFOLIO PERSPECTIVE: RISK-RETURN TRADE-OFF, DOWNSIDE PROTECTION, MODERN PORTFOLIO THEORY

3

a describe the portfolio approach to investing;

In the previous section we compared an equally weighted portfolio to the selection of a single security. In this section we examine additional combinations of the same set of shares and observe the trade-offs between portfolio volatility of returns and expected return (for short, their risk–return trade-offs). If we select the portfolios with the best combination of risk and return (taking historical statistics as our expectations for the future), we produce the set of portfolios shown in Exhibit 5.

Exhibit 5 Optimal Portfolios for Sample of HKSE Listed Shares



Source: Thomson Reuters Datastream.

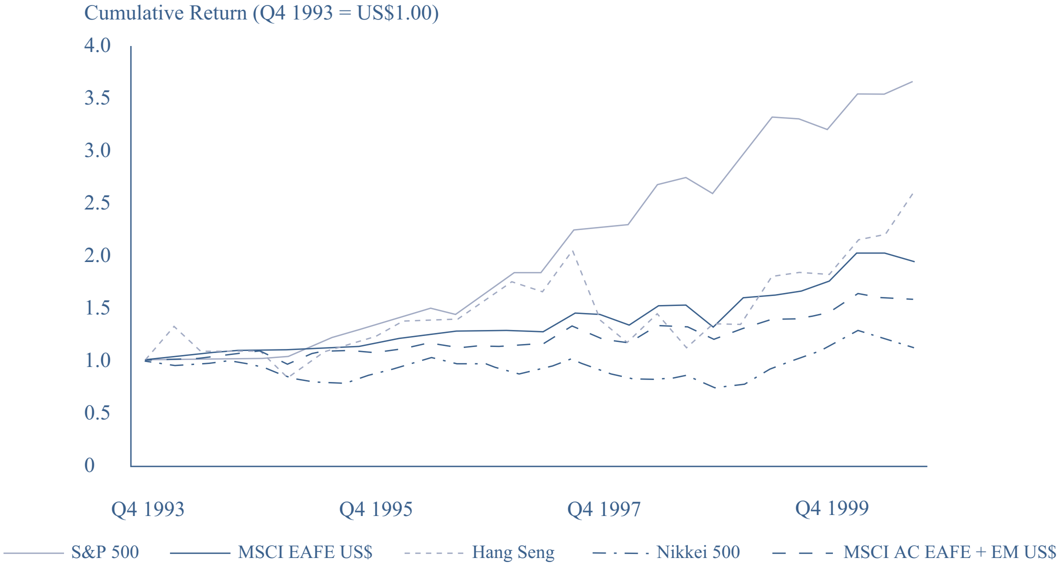
In addition to illustrating that the diversified portfolio approach reduces risk, Exhibit 5 also shows that the composition of the portfolio matters. For example, an equally weighted portfolio (20 percent of the portfolio in each security) of the five shares has an expected return of 15.1 percent and a standard deviation of 17.9 percent. Alternatively, a portfolio with 25 percent in Yue Yuen Industrial (Holdings), 3 percent in Cathay Pacific, 52 percent in Hutchison Whampoa, 20 percent in Li & Fung, and 0 percent in COSCO Pacific produces a portfolio with an expected return of 15.1 percent and a standard deviation of 15.6 percent. Compared to a simple equally weighted portfolio, this provides an improved trade-off between risk and return because a lower level of risk was achieved for the same level of return.

3.1 Historical Portfolio Example: Not Necessarily Downside Protection

A major reason that portfolios can effectively reduce risk is that combining securities whose returns do not move together provides diversification. Sometimes a subset of assets will go up in value at the same time that another will go down in value. The fact that these may offset each other creates the potential diversification benefit we attribute to portfolios. However, an important issue is that the co-movement or correlation pattern of the securities' returns in the portfolio can change in a manner unfavorable to the investor. We use historical return data from a set of global indexes to show the impact of changing co-movement patterns.

When we examine the returns of a set of global equity indexes over the last 15 years, we observe a reduction in the diversification benefit due to a change in the pattern of co-movements of returns. Exhibits 6 and 7 show the cumulative returns for a set of five global indexes[7] for two different time periods. Comparing the first time period, from Q4 1993 through Q3 2000 (as shown in Exhibit 6), with the last time period, from Q1 2006 through Q1 2009 (as shown in Exhibit 7), we show that the degree to which these global equity indexes moved together increased over time.

Exhibit 6 Returns to Global Equity Indexes Q4 1993–Q3 2000

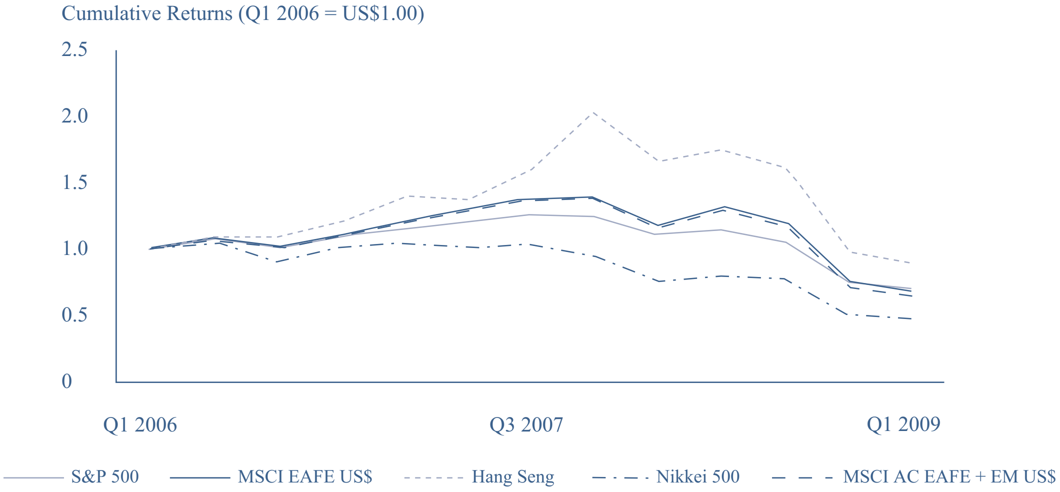


Source: Thomson Reuters Datastream.

---

[7] The S&P 500, Hang Seng, and Nikkei 500 are broad-based composite equity indexes designed to measure the performance of equities in the United States, Hong Kong SAR, and Japan. MSCI stands for Morgan Stanley Capital International. EAFE refers to developed markets in Europe, Australasia, and the Far East. AC indicates all countries, and EM is emerging markets. All index returns are in US dollars.

Exhibit 7 Returns to Global Equity Indexes Q1 2006–Q1 2009



Source: Thomson Reuters Datastream.

The latter part of the second time period, from Q4 2007 to Q1 2009, was a period of dramatic declines in global share prices. Exhibit 8 shows the mean annual returns and standard deviation of returns for this time period.

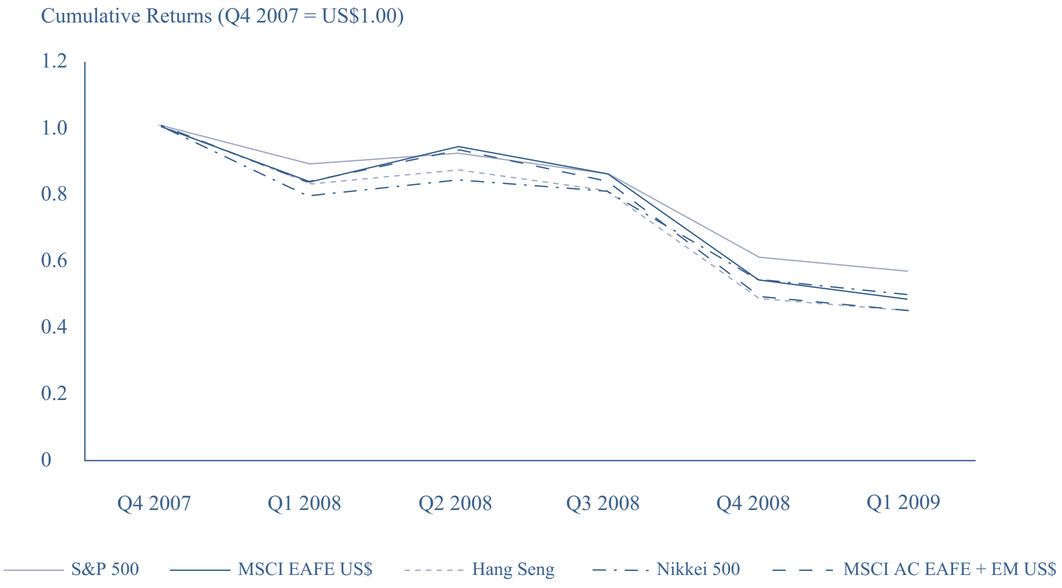
Exhibit 8 Returns to Global Equity Indexes

	Q4 1993–Q3 2000		Q1 2006–Q1 2009		Q4 2007–Q1 2009	
Global Index	**Mean**	**Stand. Dev.**	**Mean**	**Stand. Dev.**	**Mean**	**Stand. Dev.**
S&P 500	20.5%	13.9%	–6.3%	21.1%	–40.6%	23.6%
MSCI EAFE US$	10.9	14.2	–3.5	29.4	–48.0	35.9
Hang Seng	20.4	35.0	5.1	34.2	–53.8	34.0
Nikkei 500	3.3	18.0	–13.8	27.6	–48.0	30.0
MSCI AC EAFE + EM US$	7.6	13.2	–4.9	30.9	–52.0	37.5
Randomly selected index	12.6%	18.9%	–4.7%	28.6%	–48.5%	32.2%
Equally weighted portfolio	12.6%	14.2%	–4.7%	27.4%	–48.5%	32.0%
Diversification ratio		75.1%		95.8%		99.4%

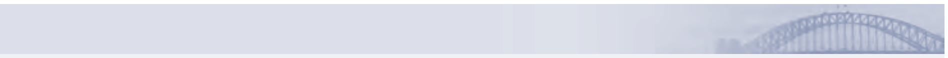
Source: Thomson Reuters Datastream.

During the period Q4 2007 through Q1 2009, the average return for the equally weighted portfolio, including dividends, was –48.5 percent. Other than reducing the risk of earning the return of the worst performing market, the diversification benefits were small. Exhibit 9 shows the cumulative quarterly returns of each of the five indexes over this time period. All of the indexes declined in unison. The lesson is that although portfolio diversification generally does reduce risk, it does not necessarily provide the same level of risk reduction during times of severe market turmoil as it does when the economy and markets are operating ‘normally’. In fact, if the economy or markets fail totally (which has happened numerous times around the world), then diversification is a false promise. In the face of a worldwide contagion, diversification was ineffective, as illustrated at the end of 2008.

Exhibit 9 Return to Global Equity Indexes Q4 2007–Q1 2009



Source: Thomson Reuters Datastream.



Portfolios are *most likely* to provide:

A risk reduction.

B risk elimination.

C downside protection.

Solution:

A is correct. Combining assets into a portfolio should reduce the portfolio's volatility. However, the portfolio approach does not necessarily provide downside protection or eliminate all risk.

3.2 Portfolios: Modern Portfolio Theory

The concept of diversification has been around for a long time and has a great deal of intuitive appeal. However, the actual theory underlying this basic concept and its application to investments only emerged in 1952 with the publication of Harry Markowitz's classic article on portfolio selection.[8] The article provided the foundation for what is now known as **modern portfolio theory** (MPT). The main conclusion of MPT is that investors should not only hold portfolios but should also focus on how individual securities in the portfolios are related to one another. In addition to the diversification benefits of portfolios to investors, the work of William Sharpe (1964), John Lintner (1965), and Jack Treynor (1961) demonstrated the role that portfolios play in determining the appropriate individual asset risk premium (i.e., the return in excess of the risk-free return expected by investors as compensation for the asset's risk). According to capital market theory, the priced risk of an individual security is affected by holding it in a well-diversified portfolio. The early research provided the insight that an asset's risk should be measured in relation to the remaining systematic or non-diversifiable risk, which should be the only risk that affects the asset's price. This view of risk is the basis of the capital asset pricing model, or CAPM, which is discussed in greater detail in other readings. Although MPT has limitations, the concepts and intuitions illustrated in the theory continue to be the foundation of knowledge for portfolio managers.

---

[8] Markowitz (1952).

STEPS IN THE PORTFOLIO MANAGEMENT PROCESS

4

b describe the steps in the portfolio management process;

In the previous section we discussed a portfolio approach to investing. When establishing and managing a client's investment portfolio, certain critical steps are followed in the process. We describe these steps in this section.

- The Planning Step
  - Understanding the client's needs
  - Preparation of an investment policy statement (IPS)
- The Execution Step
  - Asset allocation
  - Security analysis
  - Portfolio construction
- The Feedback Step
  - Portfolio monitoring and rebalancing
  - Performance measurement and reporting

4.1 Step One: The Planning Step

The first step in the investment process is to understand the client's needs (objectives and constraints) and develop an **investment policy statement** (IPS). A portfolio manager is unlikely to achieve appropriate results for a client without a prior understanding of the client's needs. The IPS is a written planning document that describes the client's investment objectives and the constraints that apply to the client's portfolio. The IPS may state a benchmark—such as a particular rate of return or the performance of a particular market index—that can be used in the feedback stage to assess the performance of the investments and whether objectives have been met. The IPS should be reviewed and updated regularly (for example, either every three years or when a major change in a client's objectives, constraints, or circumstances occurs).

4.2 Step Two: The Execution Step

The next step is for the portfolio manager to construct a suitable portfolio based on the IPS of the client. The portfolio execution step consists of first deciding on a target asset allocation, which determines the weighting of asset classes to be included in the portfolio. This step is followed by the analysis, selection, and purchase of individual investment securities.

4.2.1 *Asset Allocation*

The next step in the process is to assess the risk and return characteristics of the available investments. The analyst forms economic and capital market expectations that can be used to form a proposed allocation of asset classes suitable for the client. Decisions that need to be made in the **asset allocation** of the portfolio include the distribution between equities, fixed-income securities, and cash; sub-asset classes, such as corporate and government bonds; and geographical weightings within asset classes. Alternative assets—such as real estate, commodities, hedge funds, and private equity—may also be included.

Economists and market strategists may set the top down view on economic conditions and broad market trends. The returns on various asset classes are likely to be affected by economic conditions; for example, equities may do well when economic growth has been unexpectedly strong whereas bonds may do poorly if inflation increases. The economists and strategists will attempt to forecast these conditions.

Top down—A **top-down analysis** begins with consideration of macroeconomic conditions. Based on the current and forecasted economic environment, analysts evaluate markets and industries with the purpose of investing in those that are expected to perform well. Finally, specific companies within these industries are considered for investment.

Bottom up—Rather than emphasizing economic cycles or industry analysis, a **bottom-up analysis** focuses on company-specific circumstances, such as management quality and business prospects. It is less concerned with broad economic trends than is the case for top-down analysis, but instead focuses on company specifics.

4.2.2 *Security Analysis*

The top-down view can be combined with the bottom-up insights of security analysts who are responsible for identifying attractive investments in particular market sectors. They will use their detailed knowledge of the companies and industries they cover to assess the expected level and risk of the cash flows that each security will produce. This knowledge allows the analysts to assign a valuation to the security and identify preferred investments.

4.2.3 *Portfolio Construction*

The portfolio manager will then construct the portfolio, taking account of the target asset allocation, security analysis, and the client's requirements as set out in the IPS. A key objective will be to achieve the benefits of diversification (i.e., to avoid putting all the eggs in one basket). Decisions need to be taken on asset class weightings, sector weightings within an asset class, and the selection and weighting of individual securities or assets. The relative importance of these decisions on portfolio performance depends at least in part on the investment strategy selected; for example, consider an investor that actively adjusts asset sector weights in relation to forecasts of sector performance and one who does not. Although all decisions have an effect on portfolio performance, the asset allocation decision is commonly viewed as having the greatest impact.

Exhibit 10 shows the broad portfolio weights of the endowment funds of Yale University and the University of Virginia as of June 2017. As you can see, the portfolios have a heavy emphasis on such alternative assets as hedge funds, private equity, and real estate—Yale University particularly so.

Exhibit 10 Endowment Portfolio Weights, June 2017

Asset Class	Yale University Endowment	University of Virginia Endowment
Public equity	19.1%	26.7%
Fixed income	4.6	9.1
Private equity	14.2	15.7
Real assets (e.g., real estate)	18.7	12.1
Absolute return (e.g., hedge funds)	25.1	19.6
Cash	1.2	2.3
Other	17.2	14.5
Portfolio value	US$27.2bn	US$8.6bn

Sources: "2017 Yale Endowment Annual Report" (p. 2): www.yale.edu/investments/Yale_Endowment_17.pdf; "University of Virginia Investment Management Company Annual Report 2017" (p. 26): http://uvm-web.eservices.virginia.edu/public/reports/FinancialStatements_2017.pdf.

Risk management is an important part of the portfolio construction process. The client's risk tolerance will be set out in the IPS, and the portfolio manager must make sure the portfolio is consistent with it. As noted above, the manager will take a diversified portfolio perspective: What is important is not the risk of any single investment, but rather how all the investments perform as a portfolio.

The endowments shown above are relatively risk tolerant investors. Contrast the asset allocation of the endowment funds with the portfolio mix of the insurance companies shown in Exhibit 11. You will notice that the majority of the insurance assets are invested in fixed-income investments, typically of high quality. Note that the Yale University portfolio has less than 5 percent invested in fixed income, with the remainder invested in such growth assets as equity, real estate, and hedge funds. This allocation is in sharp contrast to the Massachusetts Mutual Life Insurance Company (MassMutual) portfolio, which is 80 percent invested in bonds, mortgages, loans, and cash—reflecting the differing risk tolerance and constraints (life insurers face regulatory constraints on their investments).

Exhibit 11 MassMutual Portfolio, December 2017[9]

Asset Classes	Portfolio %
Bonds	56%
Preferred and common shares	9
Mortgages	14
Real estate	1
Policy loans	8
Partnerships	5
Other assets	5
Cash	2

Source: "MassMutual Financial Group 2017 Annual Report" (p. 8): www.massmutual.com/mmfg/docs/annual_report/index.html.

The portfolio construction phase also involves trading. Once the portfolio manager has decided which securities to buy and in what amounts, the securities must be purchased. In many investment firms, the portfolio manager will pass the trades to a buy-side trader—a colleague who specializes in securities trading—who will contact a stockbroker or dealer to have the trades executed.

4.3 Step Three: The Feedback Step

Finally, the feedback step assists the portfolio manager in rebalancing the portfolio due to a change in, for example, market conditions or the circumstances of the client.

4.3.1 *Portfolio Monitoring and Rebalancing*

Once the portfolio has been constructed, it needs to be monitored and reviewed and the composition revised as the security analysis changes because of changes in security prices and changes in fundamental factors. When security and asset weightings have drifted from the intended levels as a result of market movements, some rebalancing may be required. The portfolio may also need to be revised if it becomes apparent that the client's needs or circumstances have changed.

4.3.2 *Performance Evaluation and Reporting*

Finally, the performance of the portfolio must be evaluated, which will include assessing whether the client's objectives have been met. For example, the investor will wish to know whether the return requirement has been achieved and how the portfolio has performed relative to any benchmark that has been set. Analysis of performance may suggest that the client's objectives need to be reviewed and perhaps changes made to the IPS. As we will discuss in the next section, there are numerous investment products that clients can use to meet their investment needs. Many of these products are diversified portfolios that an investor can purchase.

---

[9] Asset class definitions: Bonds—Debt instruments of corporations and governments as well as various types of mortgage- and asset-backed securities; Preferred and Common Shares—Investments in preferred and common equities; Mortgages—Mortgage loans secured by various types of commercial property as well as residential mortgage whole loan pools; Real Estate—Investments in real estate; Policy Loans—Loans by policyholders that are secured by insurance and annuity contracts; Partnerships—Investments in partnerships and limited liability companies; Cash—Cash, short-term investments, receivables for securities, and derivatives. Cash equivalents have short maturities (less than one year) or are highly liquid and able to be readily sold.

TYPES OF INVESTORS

5

c describe types of investors and distinctive characteristics and needs of each;

d describe defined contribution and defined benefit pension plans;

The portfolio management process described in the previous section may apply to different types of investment clients. Such clients are broadly divided among individual (or retail) and institutional investors. Each of these segments has distinctive characteristics and needs, as discussed in the following sub-sections.

5.1 Individual Investors

Individual investors have a variety of motives for investing and constructing portfolios. Short-term goals can include providing for children's education, saving for a major purchase (such as a vehicle or a house), or starting a business. The retirement goal—investing to provide for an income in retirement—is a major part of the investment planning of most individuals. Many employees of public and private companies invest for retirement through **defined contribution pension plans** (DC plans). DC plans are retirement plans in the employee's name usually funded by both the employee and the employer. Examples include 401(k) plans in the United States, group personal pension schemes in the United Kingdom, and superannuation plans in Australia. With DC plans, individuals will invest part of their wages while working, expecting to draw on the accumulated funds to provide income during retirement or to transfer some of their wealth to their heirs. The key to a DC plan is that the employee accepts the investment and inflation risk and is responsible for ensuring that there are enough assets in the plan to meet their needs upon retirement.

Some individuals will be investing for growth and will therefore seek assets that have the potential for capital gains. Others, such as retirees, may need to draw an income from their assets and may therefore choose to invest more in fixed-income and dividend-paying shares. The investment needs of individuals will depend in part on their broader financial circumstances, such as their employment prospects and whether or not they own their own residence. They may also need to consider such issues as building up a cash reserve and the purchase of appropriate insurance policies before undertaking longer-term investments.

Asset managers serving individual investors typically distribute their products directly to investors or through intermediaries such as financial advisers and/or retirement plan providers. The distribution network for individual investors varies globally. In the United States, financial advisers are independent or employed by national or regional broker–dealers, banks, and trust companies. Additionally, many asset managers distribute investment strategies to investors through major online brokerage and custodial firms.

In Europe, retail investment product distribution is fragmented and, in turn, varies by country/region. In continental Europe, for example, distribution is primarily driven through financial advisers affiliated with retail and private banks. In the United Kingdom, products are sold through independent advisers as well as through advisers representing a bank or insurance group. Retail distribution in Switzerland and in the Nordic countries is driven mainly through large regional and private banks. In contrast to the United States and Europe, in many Asian markets retail distribution is dominated by large regional retail banks and global banks with private banking divisions.

Globally, many wealth management firms and asset managers target high-net-worth investors. These clients often require more customized investment solutions alongside tax and estate planning services.

5.2 Institutional Investors

Institutional investors primarily include defined benefit pension plans, endowments and foundations, banks, insurance companies, investment companies, and sovereign wealth funds. Each of these has unique goals, asset allocation preferences, and investment strategy needs.

5.2.1 *Defined Benefit Pension Plans*

Pension plans are typically categorized as either defined contribution (DC) or defined benefit (DB). We previously described DC plans, which relate to individual investors. **Defined benefit pension plans** (DB plans) are company-sponsored plans that offer employees a predefined benefit on retirement. The future benefit is defined because the DB plan requires the plan sponsor to specify the obligation stated in terms of the retirement income benefits owed to participants. Generally, employers are responsible for the contributions made to a DB plan and bear the risk associated with adequately funding the benefits offered to employees. Plans are committed to paying pensions to members, and the assets of these plans are there to fund those payments. Plan managers need to ensure that sufficient assets will be available to pay pension benefits as they come due. The plan may have an indefinitely long time horizon if new plan members are being admitted or a finite time horizon if the plan has been closed to new members. In some cases, the plan managers attempt to match the fund's assets to its liabilities by, for example, investing in bonds that will produce cash flows corresponding to expected future pension payments. There may be many different investment philosophies for pension plans, depending on funded status and other variables.

An ongoing trend is that plan sponsors increasingly favor DC plans over DB plans because DC plans typically have lower costs/risk to the company. As a result, DB plans have been losing market share of pension assets to DC plans. Nevertheless, DB plans, both public and private, remain sizable sources of investment funds for asset managers. As Exhibit 12 shows, global pension assets totaled more than US$41 trillion by the end of 2017. The United States, United Kingdom, and Japan represent the three largest pension markets in the world, comprising more than 76% of global pension assets.

Exhibit 12 Global Pension Assets (as of year-end 2017)

Country/Region	Total Assets (US$ billions)
United States	25,411
United Kingdom	3,111
Japan	3,054
Australia	1,924
Canada	1,769
Netherlands	1,598
Switzerland	906
South Korea	725
Germany	472
Brazil	269
South Africa	258
Finland	233
Malaysia	227
Chile	205
Mexico	177

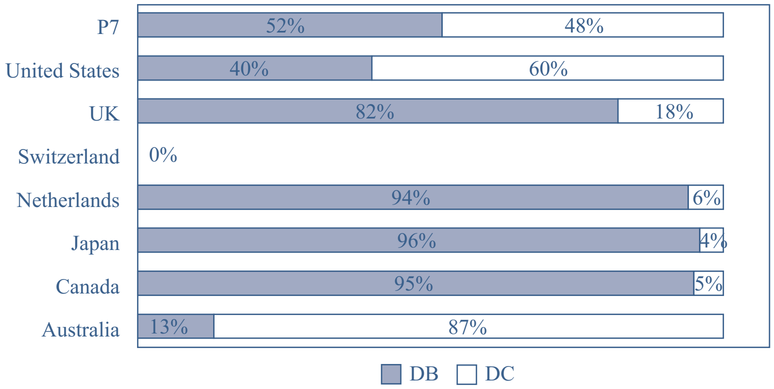
Exhibit 12 (Continued)

Country/Region	Total Assets (US$ billions)
Italy	184
France	167
Chinese mainland	177
Hong Kong SAR	164
Ireland	157
India	120
Spain	44
Total	**41,355**

Note: Column does not sum precisely because of rounding.
Source: Willis Towers Watson.

By geography, the United States and Australia have a higher proportion of pension assets in DC plans, whereas Canada, Japan, the Netherlands, and the United Kingdom remain weighted toward DB plans (see Exhibit 13).

Exhibit 13 Pension Plan Type by Geography



Notes: "P7" represents the combination of the seven countries listed. No data were available for Switzerland for this study.
Sources: Willis Towers Watson and secondary sources.

5.2.2 *Endowments and Foundations*

Endowments are funds of non-profit institutions that help the institutions provide designated services. In contrast, foundations are grant-making entities. Endowments and foundations collectively represent an estimated US$1.6 trillion in assets in the United States, which is the primary market for endowments and foundations.

Endowments and foundations typically allocate a sizable portion of their assets in alternative investments (Exhibit 14). This large allocation to alternative investments primarily reflects the typically long time horizon of endowments and foundations, as well the popularity of endowment-specific asset allocation models developed by Yale University's endowment managers David Swensen and Dean Takahashi.

Exhibit 14 Asset Allocations for US College and University Endowments and Affiliated Foundations (as of 30 June 2017, dollar weighted)

Asset Class	Percentage Allocation
Domestic equity	15
Fixed income	7
Foreign equity	20
Alternatives	54
Cash	4

Source: National Association of College and University Budget Officers and Commonfund Institute.

A typical investment objective of an endowment or a foundation is to maintain the real (inflation-adjusted) capital value of the fund while generating income to fund the objectives of the institution. Most foundations and endowments are established with the intent of having perpetual lives. Exhibit 15 describes the Yale University endowment's approach to balancing short-term spending needs with ensuring that future generations also benefit from the endowment, and it also shows the Wellcome Trust's approach. The investment approach undertaken considers the objectives and constraints of the institution (for example, no tobacco investments for a medical endowment).

Exhibit 15 Spending Rules

The following examples of spending rules are excerpts from the Yale University endowment (in the United States) and from the Wellcome Trust (in the United Kingdom).

Yale University Endowment

> The spending rule is at the heart of fiscal discipline for an endowed institution. Spending policies define an institution's compromise between the conflicting goals of providing substantial support for current operations and preserving purchasing power of Endowment assets. The spending rule must be clearly defined and consistently applied for the concept of budget balance to have meaning.
>
> The Endowment spending policy, which allocates Endowment earnings to operations, balances the competing objectives of providing a stable flow of income to the operating budget and protecting the real value of the Endowment over time. The spending policy manages the trade-of between these two objectives by combining a long-term spending rate target with a smoothing rule, which adjusts spending in any given year gradually in response to changes in Endowment market value.
>
> The target spending rate approved by the Yale Corporation currently stands at 5.25%. According to the smoothing rule, Endowment spending in a given year sums to 80% of the previous year's spending and 20% of the targeted long-term spending rate applied to the fiscal year-end market value two years prior.

Exhibit 15 (Continued)

Source: 2017 Yale Endowment Annual Report (p.18) [http://investments.yale.edu/endowment-update/]

Wellcome Trust

> Our overall investment objective is to generate 4.5% percent real return over the long term.
>
> This is to provide for real increases in annual expenditure while reserving the Trust's capital base to balance the needs of current and future beneficiaries.
>
> We use this absolute return strategy because it aligns asset allocation with funding requirements and provides a competitive framework in which to judge individual investments.

Source: Wellcome Trust website (https://wellcome.ac.uk/about-us/investments)

5.2.3 *Banks*

Banks are financial intermediaries that accept deposits and lend money. Banks often have excess reserves that are invested in relatively conservative and very short-duration fixed-income investments, with a goal of earning an excess return above interest obligations due to depositors. Liquidity is a paramount concern for banks that stand ready to meet depositor requests for withdrawals. Many large banks have asset management divisions that offer retail and institutional products to their clients.

5.2.4 *Insurance Companies*

Insurance companies receive premiums for the policies they write, and they need to invest these premiums in a manner that will allow them to pay claims.

Insurance companies can be segmented into two broad types: life insurers and property and casualty (P&C) insurers. Insurance premiums from policyholders comprise an insurance company's general account. To pay claims to policyholders, regulatory guidelines maintain that an insurance company's general account is typically invested conservatively in a diverse allocation of fixed-income securities. General account portfolio allocations differ among life, P&C, and other specialty insurers (e.g., reinsurance) because of both the varying duration of liabilities and the unique liquidity considerations across insurance type.[10] In contrast to the general account, an insurer's surplus account is the difference between its assets and liabilities. An insurer's surplus account typically targets a higher return than the general account and thus often invests in less-conservative asset classes, such as public and private equities, real estate, infrastructure, and hedge funds.

Many insurance companies have in-house portfolio management teams responsible for managing general account assets. Some insurance companies offer portfolio management services and products in addition to their insurance offerings. An increasing trend among insurers (particularly in the United States) is outsourcing some of the portfolio management responsibilities—primarily sophisticated alternative asset classes—to unaffiliated asset managers. Several insurers manage investments for third-party clients, often through separately branded subsidiaries.

---

[10] For example, life insurers tend to invest in longer-term assets (e.g., 30-year government and corporate bonds) relative to P&C insurers because of the longer-term nature of their liabilities.

5.2.5 *Sovereign Wealth Funds*

Sovereign wealth funds (SWFs) are state-owned investment funds or entities that invest in financial or real assets. SWFs do not typically manage specific liability obligations, such as pensions, and have varying investment horizons and objectives based on funding the government's goals (for example, budget stabilization or future development projects). SWF assets more than doubled from 2007 to March 2018, totaling more than US$7.6 trillion.[11] Exhibit 16 lists the 10 largest SWFs in the world. The largest SWFs tend to be concentrated in Asia and in natural resource-rich places.

Exhibit 16 Largest Sovereign Wealth Funds (as of August 2018, in US$ billions)

Place	Sovereign Wealth Fund (Inception Year)	Assets
Norway	Government Pension Fund—Global (1990)	1,058
Chinese Mainland	China Investment Corporation (2007)	941
UAE – Abu Dhabi	Abu Dhabi Investment Authority (1976)	683
Kuwait	Kuwait Investment Authority (1953)	592
Hong Kong SAR	Hong Kong Monetary Authority Investment Portfolio (1993)	523
Saudi Arabia	SAMA Foreign Holdings (1952)	516
Chinese Mainland	SAFE Investment Company (1997)	441
Singapore	Government of Singapore Investment Authority (1981)	390
Singapore	Temasek Holdings (1974)	375
Saudi Arabia	Public Investment Fund (2008)	360
Total SWF Assets under Management		**8,109**

Source: SWF Institute (www.swfinstitute.org).

Investment needs vary across client groups. With some groups of clients, generalizations are possible. In other groups, needs vary by client. Exhibit 17 summarizes needs within each group.

Exhibit 17 Summary of Investment Needs by Client Type

Client	Time Horizon	Risk Tolerance	Income Needs	Liquidity Needs
Individual investors	Varies by individual	Varies by individual	Varies by individual	Varies by individual
Defined benefit pension plans	Typically long term	Typically quite high	High for mature funds; low for growing funds	Varies by maturity of the plan

---

[11] SWFI, "Sovereign Wealth Fund Rankings" (https://www.swfinstitute.org/sovereign-wealth-fund-rankings/; retrieved October 2018).

Exhibit 17 (Continued)

Client	Time Horizon	Risk Tolerance	Income Needs	Liquidity Needs
Endowments and foundations	Very long term	Typically high	To meet spending commitments	Typically quite low
Banks	Short term	Quite low	To pay interest on deposits and operational expenses	High to meet repayment of deposits
Insurance companies	Short term for property and casualty; long term for life insurance companies	Typically quite low	Typically low	High to meet claims
Investment companies	Varies by fund	Varies by fund	Varies by fund	High to meet redemptions
Sovereign wealth funds	Varies by fund	Varies by fund	Varies by fund	Varies by fund

THE ASSET MANAGEMENT INDUSTRY

6

e describe aspects of the asset management industry;

The portfolio management process and investor types are broad components of the asset management industry, which is an integral component of the global financial services sector. At the end of 2017, the industry managed more than US$79 trillion of assets owned by a broad range of institutional and individual investors (Exhibit 18).[12] Although nearly 80% of the world's professionally managed assets are in North America and Europe, the fastest-growing markets are in Asia and Latin America.

Exhibit 18 Global Assets under Management (AUM) by Region (year-end 2017)

	Market Size (US$ trillions)	Market Share (%)
North America	37.4	47%
Europe	22.2	28
Japan and Australia	6.2	8
Chinese mainland	4.2	5
Asia (excluding Japan, Australia, and Chinese mainland)	3.5	4
Latin America	1.8	2
Middle East and Africa	1.4	2
Total Global AUM	79.2	100%

Notes: Total Global AUM in this exhibit represents assets professionally managed in exchange for a fee. The total of US$79.2 trillion includes certain offshore assets that are not represented in the specific regional categories above.

(continued)

---

[12] http://image-src.bcg.com/Images/BCG-Seizing-the-Analytics-Advantage-June-2018-R-3_tcm9-194512.pdf (accessed on 6 September 2018).

Exhibit 18 (Continued)

Source: Boston Consulting Group.

The asset management industry is highly competitive. The universe of firms in the industry is broad, ranging from "pure-play" independent asset managers to diversified commercial banks, insurance companies, and brokerages that offer asset management services in addition to their core business activities. Given the increasingly global nature of the industry, many asset managers have investment research and distribution offices around the world. An asset manager is commonly referred to as a **buy-side firm** given that it uses (buys) the services of sell-side firms. A **sell-side firm** is a broker/dealer that sells securities and provides independent investment research and recommendations to their clients (i.e., buy-side firms).

Asset managers offer a broad range of strategies. Specialist asset managers may focus on a specific asset class (e.g., emerging market equities) or style (e.g., quantitative investing), while "full service" managers typically offer a wide variety of asset classes and styles. Another type of asset manager firm is a "multi-boutique," in which a holding company owns several asset management firms that typically have specialized investment strategies. The multi-boutique structure allows individual asset management firms to retain their own unique investment cultures—and often equity ownership stakes—while also benefiting from the centralized, shared services of the holding company (e.g., technology, sales and marketing, operations, and legal services).

6.1 Active versus Passive Management

Asset managers may offer either active or passive management. As of year-end 2017, active management considerably exceeded passive management in terms of global assets under management and industry revenue (Exhibit 19), although passive management has demonstrated significant growth.

Exhibit 19 Global Asset Management Industry Assets and Revenue (as of year-end 2017)

Category	Assets (US$ trillions)	Revenue (US$ billions)	Market Share by Assets (%)	Market Share by Revenue (%)
Actively Managed	**64**	**258**	**80%**	**94%**
Alternatives	12	117	15	43
Active Specialties	15	55	19	20
Multi Asset Class	11	27	14	10
Core	26	59	33	21
Passively Managed	**16**	**17**	**20%**	**6%**
Total	**80**	**275**	**100%**	**100%**

Note: Some columns may not sum precisely because of rounding.
Source: Boston Consulting Group.

Through fundamental research, quantitative research, or a combination of both, active asset managers generally attempt to outperform either predetermined performance benchmarks, such as the S&P 500, or, for multi-asset class portfolios, a combination of benchmarks. In contrast to active managers, passive managers attempt to replicate the returns of a market index. Despite the rise of passive management in asset share, its share of industry revenue remains small given the low management fees relative to active management. As Exhibit 19 illustrates, passive management represents a fifth of global assets under management but only 6% of industry revenue.

Asset managers are increasingly offering other strategies beyond traditional market-cap-weighted exposures. Some of these other strategies, commonly known as **smart beta**, are based on such factors as size, value, momentum, or dividend characteristics. Smart beta involves the use of simple, transparent, rules-based strategies as a basis for investment decisions. Typically, smart beta strategies feature somewhat higher management fees and higher portfolio turnover relative to passive market-cap weighted strategies.

6.2 Traditional versus Alternative Asset Managers

Asset managers are typically categorized as either “traditional” or “alternative.” Traditional managers generally focus on long-only equity, fixed-income, and multi-asset investment strategies, generating most of their revenues from asset-based management fees. Alternative asset managers, however, focus on hedge fund, private equity, and venture capital strategies, among others, while generating revenue from both management and performance fees (or “carried interest”). As Exhibit 19 demonstrates, alternative asset managers have a relatively low proportion of total global assets under management but generate a disproportionately high total of industry revenue.

Increasingly, the line between traditional and alternative managers has blurred. Many traditional managers have introduced higher-margin alternative products to clients. Concurrently, alternative managers seeking to reduce the revenue volatility associated with performance fees have increasingly offered retail versions of their institutional alternative strategies (typically referred to as “liquid alternatives”) as well as long-only investment strategies. These liquid alternatives are often offered through highly regulated pooled investment products (e.g., mutual funds) and typically feature less leverage, no performance fees, and more liquid holdings than typical alternative products.

6.3 Ownership Structure

The ownership structure of an asset manager can play an important role in retaining and incentivizing key personnel. Portfolio managers who have personal capital invested in their firms or investment strategies are often viewed favorably by potential investors because of perceived alignment of management and client interests.

The majority of asset management firms are privately owned, typically by individuals who either established their firms or play key roles in their firms’ management. Privately owned firms are typically structured as limited liability companies or limited partnerships.

While less common than privately owned managers, publicly traded asset managers have substantial assets under management. A prevalent ownership form in the industry is represented by asset management divisions of large, diversified financial services companies that offer asset management alongside insurance and banking services.

6.4 Asset Management Industry Trends

The asset management industry is evolving and continues to be shaped by socioeconomic trends, shifting investor demands, advances in technology, and the expansion of global capital markets. Three key trends that we discuss in this section include the growth of passive investing, "big data" in the investment process, and the emergence of robo-advisers in the wealth management industry.

6.4.1 *Growth of Passive Investing*

As we saw in Exhibit 19, passively managed assets comprised nearly a fifth of global assets under management at the end of 2017. Management of passive assets is concentrated among a reasonably small group of asset managers and tends to be concentrated in equity strategies. As shown in Exhibit 20, the top three managers account for 70% of industry's assets. One key catalyst supporting the growth of passive investing is low cost for investors—management fees for index (or other passive) funds are often a fraction of those for active strategies. Another catalyst is the challenge that many active asset managers face in generating ex ante alpha, especially in somewhat more-efficiently priced markets, such as large-cap US equities.

Exhibit 20 Top Five ETP Managers Globally (as of 30 July 2017)

ETP Provider	Assets (US$ billions)	Market Share (%)
iShares	1,583	37
Vanguard	803	19
State Street Global Advisors	596	14
PowerShares	132	3
Nomura	100	2

Source: ETFGI.

6.4.2 *Use of "Big Data" in the Investment Process*

The prevalence of new data is extraordinary: In 2013, IBM estimated that 90% of the world's entire universe of data was created in the previous two years. The digitization of data and an exponential increase in computing power and data storage capacity have expanded additional information sources for asset managers. Massive amounts of data containing information of potential value to investors are created and captured daily. These data include both structured data—such as order book data and security returns—and data lacking recognizable structure, which is generated by a vast number of activities on the internet and elsewhere (e.g., compiled search information). The term "big data" is used to refer to these massively large datasets and their analysis.

Asset managers are using advanced statistical and machine-learning techniques to help process and analyze these new sources of data. Such techniques are used in both fundamentally driven and quantitatively driven investment processes. For example, computers are used to "read" earnings and economic data releases much faster than humans can and react with short-term trading strategies.

Third-party research vendors are supplying a vast range of relevant new data for asset managers, such as data used for time-series and predictive models. Among the most popular new sources of data are social media data and imagery and sensor data.

- **Social media data**. Real-time media and content outlets, such as Twitter and Facebook, provide meaningful market and company-specific announcements for investors and asset managers. In addition, the aggregation and analysis of social media users can aid key market sentiment indicators (e.g., short-term directional market movements) and indicate potential specific user trends related to products and services.

- **Imagery and sensor data**. Satellite imagery and geolocation devices provide vast real-time data to investment professionals. As the cost of launching and maintaining satellites has decreased, more satellites have been launched to track sensors and imagery that are relevant to economic considerations (e.g., weather conditions, cargo ship traffic patterns) and company-specific considerations (e.g., retailer parking capacity/usage, tracking of retail customers).

The challenge for asset managers is to discover data with predictive potential and to do so faster than fellow market participants. Many market participants are participating in an "information arms race" that has required substantial investments in specialized human capital (e.g., programmers, data scientists), technology, and information technology infrastructure to effectively convert various forms of structured and unstructured data into alpha-generating portfolio and security-level decisions.

6.4.3 *Robo-Advisers: An Expanding Wealth Management Channel*

Robo-advisers represent technology solutions that use automation and investment algorithms to provide several wealth management services—notably, investment planning, asset allocation, tax loss harvesting, and investment strategy selection. Investment and advice services provided by robo-advisers typically reflect an investor's general investment goals and risk tolerance preferences (often obtained from an investor questionnaire). Robo-adviser platforms range from exclusively digital investment advice platforms to hybrid offerings that offer both digital investment advice and the services of a human financial adviser.

At the end of 2017, robo-advisers managed an estimated US$180 billion in assets,[13] and market participants expect that number to grow considerably over time. This expected rapid growth in robo-advisory assets is based on several industry trends:

- **Growing demand from "mass affluent" and younger investors**: Traditional investment advice has often underserved younger and "mass affluent" investors with lower relative levels of investable assets. Given the efficiencies of robo-advisers and the scalability of technology, customized but standardized investment advice now can be offered to a wide range and size of investors.
- **Lower fees**: The cost of digital investment advice provided by robo-advisers is often a fraction of traditional investment advice channels because of scalability. For example, in the United States, a typical financial adviser may charge a 1% annual advisory fee[14] based on a client's assets, while robo-adviser fees typically average 0.20% annually.[15] Additionally, robo-advisers often rely on lower fee underlying portfolio investment options, such as index funds or ETFs, when constructing portfolios for clients.
- **New entrants**: Reflecting low barriers to entry, large wealth management firms have introduced robo-adviser solutions to service certain customer segments and appeal to a new generation of investors. In addition to these large wealth managers, other less-traditional entrants, such as insurance companies and asset managers, are developing solutions to cross-sell into their existing clients.

---

[13] S&P Global Market Intelligence.

[14] http://www.riainabox.com/blog/2016-ria-industry-study-average-investment-advisory-fee-is-0-99-percent.

[15] Deloitte, "Robo-Advisors Capitalizing on a Growing Opportunity" (https://www2.deloitte.com/content/dam/Deloitte/us/Documents/strategy/us-cons-robo-advisors.pdf).

Many market observers expect that non-financial firms (large technology leaders) will also become key players in the robo-adviser industry as they look to monetize their access to user data.

7 POOLED INTEREST - MUTUAL FUNDS

f describe mutual funds and compare them with other pooled investment products.

In the asset management industry, a challenge faced by all investors is to and find the right set of investment products to meet their needs. There is a diverse set of investment products available to investors, ranging from a simple brokerage account in which the individual creates her own portfolio by assembling individual securities, to large institutions that employ individual portfolio managers to meet clients' investment management needs. Among the major investment products offered by asset managers are mutual funds and other pooled investment products, such as separately managed accounts, exchange-traded funds, hedge funds, and private equity/venture capital funds.

7.1 Mutual Funds

Rather than assemble a portfolio on their own, individual investors and institutions can turn over the selection and management of their investment portfolio to a third party. One way of doing this is through a **mutual fund**. This type of fund is a commingled investment pool in which investors in the fund each have a pro-rata claim on the income and value of the fund. The value of a mutual fund is referred to as the "net asset value." It is computed daily based on the closing price of the securities in the portfolio.

Mutual funds represent a primary investment product of individual investors globally. According to the International Investment Funds Association, worldwide regulated open-end fund assets totaled US$50 trillion as of the first quarter of 2018. Exhibit 21 shows the growth of global open-end funds over the past five years by region. Mutual funds provide several advantages, including low investment minimums, diversified portfolios, daily liquidity, and standardized performance and tax reporting.

Exhibit 21 Worldwide Regulated Open-End Funds: Total Net Assets (as of year-end, in US$ trillions)

	2011	2012	2013	2014	2015	2016	Q1 2018
World	**27.9**	**31.9**	**36.3**	**38.0**	**38.2**	**40.4**	**50.0**
Americas	14.6	16.5	18.9	20.0	19.6	21.1	24.9
Europe	10.3	11.9	13.6	13.8	13.7	14.1	18.1
Asia and Pacific	2.9	3.3	3.7	4.1	4.7	5.0	6.8
Africa	0.1	0.1	0.1	0.1	0.1	0.1	0.2

Notes: Components may not add to the total because of rounding. Regulated open-end funds include mutual funds, exchange-traded funds (ETFs), and institutional funds.
Source: International Investment Funds Association (IIFA).

Mutual funds are one of the most important investment vehicles for individuals and institutions. The best way to understand how a mutual fund works is to consider a simple example. Suppose that an investment firm wishes to start a mutual fund with a target amount of US$10 million. It is able to reach this goal through investments from five individuals and two institutions. The investment of each is as follows:

Investor	Amount Invested (US$)	Percent of Total	Number of Shares
Individuals			
A	$1.0 million	10%	10,000
B	1.0	10	10,000
C	0.5	5	5,000
D	2.0	20	20,000
E	0.5	5	5,000
Institutions			
X	2.0	20	20,000
Y	3.0	30	30,000
Totals	**$10.0 million**	**100%**	**100,000**

Based on the US$10 million value (net asset value), the investment firm sets a total of 100,000 shares at an initial value of US$100 per share (US$10 million/100,000 = US$100). The investment firm will appoint a portfolio manager to be responsible for the investment of the US$10 million. Going forward, the total value of the fund or net asset value will depend on the value of the assets in the portfolio.

The fund can be set up as an open-end fund or a closed-end fund. If it is an **open-end fund**, it will accept new investment money and issue additional shares at a value equal to the net asset value of the fund at the time of investment. For example, assume that at a later date the net asset value of the fund increases to US$12 million and the new net asset value per share is US$120. A new investor, F, wishes to invest US$0.96 million in the fund. If the total value of the assets in the fund is now US$12 million or US$120 per share, in order to accommodate the new investment the fund would create 8,000 (US$0.96 million/US$120) new shares. After this investment, the net asset value of the fund would be US$12.96 million and there would be a total of 108,000 shares.

Funds can also be withdrawn at the net asset value per share. Suppose on the same day Investor E wishes to withdraw all her shares in the mutual fund. To accommodate this withdrawal, the fund will have to liquidate US$0.6 million in assets to retire 5,000 shares at a net asset value of US$120 per share (US$0.6 million/US$120). The combination of the inflow and outflow on the same day would be as follows:

Type	Investment (US$)	Shares
Inflow (Investor F buys)	$960,000	8,000
Outflow (Investor E sells)	–$600,000	–5,000
Net	$360,000	3,000

The net of the inflows and outflows on that day would be US$360,000 of new funds to be invested and 3,000 new shares created. However, the number of shares held and the value of the shares of all remaining investors, except Investor E, would remain the same.

An alternative to setting the fund up as an open-end fund would be to create a **closed-end fund** in which no new investment money is accepted into the fund. New investors invest by buying existing shares, and investors in the fund liquidate by selling their shares to other investors. Hence, the number of outstanding shares does not change. One consequence of this fixed share base is that, unlike open-end funds in which new shares are created and sold at the current net asset value per share, closed-end funds can sell for a premium or discount to net asset value depending on the demand for the shares.

There are advantages and disadvantages to each type of fund. The open-end fund structure makes it easy to grow in size but creates pressure on the portfolio manager to manage the cash inflows and outflows. One consequence of this structure is the need to liquidate assets that the portfolio manager might not want to sell at the time to meet redemptions. Conversely, the inflows require finding new assets in which to invest. As such, open-end funds tend not to be fully invested but rather keep some cash for redemptions not covered by new investments. Closed-end funds do not have these problems, but they do have a limited ability to grow. Of the total net asset value of all US mutual funds at the end of 2017 (US$19 trillion), only approximately 1 percent were in the form of closed-end funds.

In addition to open-end or closed-end funds, mutual funds can be classified as load or no-load funds. The primary difference between the two is whether the investor pays a sales charge (a "load") to purchase, hold, or redeem shares in the fund. In the case of the **no-load fund**, there is no fee for investing in the fund or for redemption but there is an annual fee based on a percentage of the fund's net asset value. **Load funds** are funds in which, in addition to the annual fee, a percentage fee is charged to invest in the fund and/or for redemptions from the fund. In addition, load funds are usually sold through retail brokers who receive part of the upfront fee. Overall, the number and importance of load funds has declined over time.

8

POOLED INTEREST - TYPE OF MUTUAL FUNDS

f describe mutual funds and compare them with other pooled investment products.

The following section introduces the major types of mutual funds differentiated by the asset type that they invest in: money market funds, bond mutual funds, stock mutual funds, and hybrid or balanced funds.

8.1 Money Market Funds

Money market funds are mutual funds that invest in short-term money market instruments such as treasury bills, certificates of deposit, and commercial paper. They aim to provide security of principal, high levels of liquidity, and returns in line with money market rates. Many funds operate on a constant net asset value (CNAV) basis where the share price is maintained at $1 (or local currency equivalent). Others operate on a variable net asset value (VNAV) basis where the unit price can vary. In the United States, there are two basic types of money market funds: taxable and tax-free. Taxable money market funds invest in high-quality, short-term corporate debt and federal government debt. Tax-free money market funds invest in short-term state and local government debt. Although money market funds have been a substitute for bank savings accounts since the early 1980s, they are not insured in the same way as bank deposits.

8.2 Bond Mutual Funds

A bond mutual fund is an investment fund consisting of a portfolio of individual bonds and, occasionally, preferred shares. The net asset value of the fund is the sum of the value of each bond in the portfolio divided by the number of shares. Investors in the mutual fund hold shares, which account for their pro-rata share or interest in the portfolio. The major difference between a bond mutual fund and a money market fund is the maturity of the underlying assets. In a money market fund the maturity is as short as overnight and rarely longer than 90 days. A bond mutual fund, however, holds bonds with maturities as short as one year and as long as 30 years (or more). Exhibit 22 illustrates the general categories of bond mutual funds.[16]

Exhibit 22 Bond Mutual Funds

Type of Bond Mutual Fund	Securities Held
Global	Domestic and non-domestic government, corporate, and securitized debt
Government	Government bonds and other government-affiliated bonds
Corporate	Corporate debt
High yield	Below investment-grade corporate debt
Inflation protected	Inflation-protected government debt
National tax-free bonds	National tax-free bonds (e.g., US municipal bonds)

8.3 Stock Mutual Funds

Historically, the largest types of mutual funds based on market value of assets under management are stock (equity) funds.

There are two types of stock mutual funds. The first is an actively managed fund in which the portfolio manager seeks outstanding performance through the selection of the appropriate stocks to be included in the portfolio. Passive management is followed by index funds that are very different from actively managed funds. Their goal is to match or track the performance of different indexes. The first index fund was introduced in 1976 by the Vanguard Group.

There are several major differences between actively managed funds and index funds. First, management fees for actively managed funds are higher than for index funds. The higher fees for actively managed funds reflect its goal to outperform an index, whereas the index fund simply aims to match the return on the index. Higher fees are required to pay for the research conducted to actively select securities. A second difference is that the level of trading in an actively managed fund is much higher than in an index fund, which depending on the jurisdiction, has tax implications. Mutual funds are often required to distribute all income and capital gains realized in the portfolio, so the actively managed fund tends to have more opportunity to realize capital gains. This results in higher taxes relative to an index fund, which uses a buy-and-hold strategy. Consequently, there is less buying and selling in an index fund and less likelihood of realizing capital gains distributions.

---

[16] In the United States, judicial rulings on federal powers of taxation have created a distinction between (federally) taxable and (federally) tax-exempt bonds and a parallel distinction for US bond mutual funds.

8.4 Hybrid/Balanced Funds

Hybrid or balanced funds are mutual funds that invest in both bonds and stocks. These types of funds represent a small fraction of the total investment in US mutual funds but are more common in Europe. These types of funds, however, have gained popularity with the growth of lifecycle funds. Lifecycle or Target Date funds manage the asset mix based on a desired retirement date. For example, if an investor is 40 years old in 2019 and planned to retire at the age of 67, he could invest in a mutual fund with a target date of 2046 and the fund would manage the appropriate asset mix over the next 27 years. In 2019 it might be 90 percent invested in shares and 10 percent in bonds. As time passes, however, the fund would gradually change the mix of shares and bonds to reflect the appropriate mix given the time to retirement.

9

POOLED INTEREST - OTHER INVESTMENT PRODUCTS

f describe mutual funds and compare them with other pooled investment products.

A fund management service for institutions or individual investors with substantial assets is the **separately managed account** (SMA), which is also commonly referred to as a "managed account," "wrap account," or "individually managed account."

SMAs are managed exclusively for the benefit of a single individual or institution. Unlike a mutual fund, the assets of an SMA are owned directly by the individual or institution. The main disadvantage of an SMA is that the required minimum investment is usually much higher than with a mutual fund.

Large institutional investors are generally the dominant users of SMAs. SMAs enable asset managers to implement an investment strategy that matches an investor's specific objectives, portfolio constraints, and tax considerations, where applicable. For example, a public pension plan investing in an asset manager's large value equity strategy might have a socially responsible investment preference. In this case, the plan sponsor may wish to exclude certain industries, such as tobacco and defense, while also including additional companies that are deemed favorable according to other environmental, social, and governance (ESG) considerations.

9.1 Exchange-Traded Funds

Exchange-traded funds (ETFs) are investment funds that trade on exchanges (similar to individual stocks) and are generally structured as open-end funds. ETFs represent one of the fastest-growing investment products in the asset management industry. According to BlackRock, global ETF assets increased from US$428 billion in 2005 to US$4.9 trillion as of June 2018. Long-term investors—both institutional and retail—use ETFs in building a diversified asset allocation. While ETFs are structured similarly to open-end mutual funds, some key differences exist between the two products. One difference relates to transaction price. Because they are traded on exchanges, ETFs can be transacted (and are priced) intraday. That is, ETF investors buy the shares from other investors just as if they were buying or selling shares of stock. ETF investors can also short shares or purchase the shares on margin. In contrast, mutual funds typically can be purchased or sold only once a day, and short sales or purchasing shares on margin is not allowed. Mutual fund investors buy the fund shares directly from the fund, and all investments are settled at the net asset value. In practice, the market price of the ETF is likely to be close to the net asset value of the underlying investments.

Other key differences between ETFs and mutual funds relate to transaction costs and treatment of dividends and the minimum investment amount. Dividends on ETFs are paid out to the shareholders whereas mutual funds usually reinvest the dividends. Finally, the minimum required investment in ETFs is usually smaller than that of mutual funds.

9.2 Hedge Funds

Hedge funds are private investment vehicles that typically use leverage, derivatives, and long and short investment strategies. The origin of hedge funds can be traced back as far as 1949 to a fund managed by A.W. Jones & Co. It offered a strategy of a non-correlated offset to the "long-only" position typical of most portfolios. Since then, the hedge fund industry has grown considerably, with global hedge fund assets totaling US$3.3 trillion as of May 2017.

Hedge fund investment strategies are diverse and can range from specific niche strategies (e.g., long–short financial services) to global multi-strategy approaches. Consequently, hedge funds are often used by investors for portfolio diversification purposes. In general, hedge funds share a few distinguishing characteristics:

- *Short selling*: Many hedge funds implement short positions directly or synthetically using such derivatives as options, futures, and credit default swaps.
- *Absolute return seeking:* Hedge funds often seek positive returns in all market environments.
- *Leverage:* Many hedge funds use financial leverage (bank borrowing) or implicit leverage (using derivatives). The use and amount of leverage are dependent on the investment strategy being implemented.
- *Low correlation:* Some hedge funds have historically exhibited low return correlations with traditional equity and/or fixed-income asset classes.
- *Fee structures:* Hedge funds typically charge two distinct fees: a traditional asset-based management fee (AUM fee) and an incentive (or performance) fee in which the hedge fund earns a portion of the fund's realized capital gains.[17] Hedge funds have traditionally charged management fees of 2% and incentive fees of up to 20%, although there has been downward pressure on those fees amid increased competition and the availability of competing products.

Hedge funds are not readily available to all investors. They typically require a high minimum investment and often have restricted liquidity by allowing only periodic (e.g., quarterly) withdrawals or having a long fixed-term commitment.

9.3 Private Equity and Venture Capital Funds

Private equity funds and **venture capital funds** are alternative funds that seek to buy, optimize, and ultimately sell portfolio companies to generate profits. As of December 2017, assets under management in the private equity industry totaled US$3.1 trillion, a historical high point.[18] Most private equity and venture capital funds have a lifespan of approximately 7–10 years (but usually subject to contractual extensions). Unlike most traditional asset managers that trade in public securities, private equity and venture firms often take a "hands-on" approach to their portfolio companies through a combination of financial engineering (e.g., realizing expense synergies, changing capital structures), installment of executive management and board members, and significant contributions to the development of a target company's business strategy. The final investment stage, often referred to as the "exit" or "harvesting" stage, occurs when a private equity or venture capital fund divests its portfolio companies through a merger with another company, the acquisition by another company, or an initial public offering (IPO).

---

[17] Performance fees are often subject to high-water mark provisions, which preclude a manager from earning a performance fee unless the value of a fund at the end of a predefined measurement period is higher than the value of the fund at the beginning of the measurement period. The unpredictability of future performance leads to uncertainty in performance fee revenue, which is regarded as less reliable than revenue derived from management fees.

[18] https://www.pionline.com/article/20180724/ONLINE/180729930/preqin-private-equity-aum-grows-20-in-2017-to-record-306-trillion# (accessed 13 November 2018)

As with most alternative funds, the majority of private equity and venture capital funds are structured as limited partnerships. These limited partnership agreements exist between the fund manager, called the general partner (GP), and the fund's investors, called limited partners (LPs). The funds generate revenue through several types of fees:

- *Management fees:* Fees are based on committed capital (or sometimes net asset value or invested capital) and typically range from 1–3% annually. Sometimes these fees step down several years into the investment period of a fund.
- *Transaction fees:* Fees are paid by portfolio companies to the fund for various corporate and structuring services. Typically, a percentage of the transaction fee is shared with the LPs by offsetting the management fee.
- *Carried interest:* Carried interest is the GP's share of profits (typically 20%) on sales of portfolio companies. Most GPs do not earn the incentive fee until LPs have recovered their initial investment.
- *Investment income.* Investment income includes profits generated on capital contributed to the fund by the GP.

SUMMARY

- A portfolio approach to investing could be preferable to simply investing in individual securities.
- The problem with focusing on individual securities is that this approach may lead to the investor "putting all her eggs in one basket."
- Portfolios provide important diversification benefits, allowing risk to be reduced without necessarily affecting or compromising return.
- Understanding the needs of your client and preparing an investment policy statement represent the first steps of the portfolio management process. Those steps are followed by asset allocation, security analysis, portfolio construction, portfolio monitoring and rebalancing, and performance measurement and reporting.
- Types of investors include individual and institutional investors. Institutional investors include defined benefit pension plans, endowments and foundations, banks, insurance companies, and sovereign wealth funds.
- The asset management industry is an integral component of the global financial services sector. Asset managers offer either active management, passive management, or both. Asset managers are typically categorized as traditional or alternative, although the line between traditional and alternative has blurred.

- Three key trends in the asset management industry include the growth of passive investing, “big data” in the investment process, and robo-advisers in the wealth management industry.
- Investors use different types of investment products in their portfolios. These include mutual funds, separately managed accounts, exchange-traded funds, hedge funds, and private equity and venture capital funds.

REFERENCES

Lintner, John. 1965. “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets.” *Review of Economics and Statistics*, vol. 47, no. 1 (February):13–37.

Markowitz, Harry M. 1952. “Portfolio Selection.” *Journal of Finance*, vol. 7, no. 1 (March):77–91.

Sharpe, William F. 1964. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk.” *Journal of Finance*, vol. 19, no. 3 (September):425–442.

Singletary, Michelle. 2001. “Cautionary Tale of an Enron Employee Who Went for Broke.” Seattlepi.com (10 December): http://www.seattlepi.com/money/49894_singletary10.shtml.

Treynor, J. L. 1961. “Toward a Theory of Market Value of Risky Assets.” Unpublished manuscript.

PRACTICE PROBLEMS

1 Investors should use a portfolio approach to:

A reduce risk.

B monitor risk.

C eliminate risk.

2 Which of the following is the *best* reason for an investor to be concerned with the composition of a portfolio?

A Risk reduction.

B Downside risk protection.

C Avoidance of investment disasters.

3 With respect to the formation of portfolios, which of the following statements is *most accurate*?

A Portfolios affect risk less than returns.

B Portfolios affect risk more than returns.

C Portfolios affect risk and returns equally.

4 Which of the following institutions will *on average* have the greatest need for liquidity?

A Banks.

B Investment companies.

C Non-life insurance companies.

5 Which of the following institutional investors will *most likely* have the longest time horizon?

A Defined benefit plan.

B University endowment.

C Life insurance company.

6 A defined benefit plan with a large number of retirees is *likely* to have a high need for:

A income.

B liquidity.

C insurance.

7 Which of the following institutional investors is *most likely* to manage investments in mutual funds?

A Insurance companies.

B Investment companies.

C University endowments.

8 With respect to the portfolio management process, the asset allocation is determined in the:

A planning step.

B feedback step.

C execution step.

9 The planning step of the portfolio management process is *least likely* to include an assessment of the client's

© 2011 CFA Institute. All rights reserved.

A securities.

B constraints.

C risk tolerance.

10 With respect to the portfolio management process, the rebalancing of a portfolio's composition is *most likely* to occur in the:

A planning step.

B feedback step.

C execution step.

11 An analyst gathers the following information for the asset allocations of three portfolios:

Portfolio	Fixed Income (%)	Equity (%)	Alternative Assets (%)
1	25	60	15
2	60	25	15
3	15	60	25

Which of the portfolios is *most likely* appropriate for a client who has a high degree of risk tolerance?

A Portfolio 1.

B Portfolio 2.

C Portfolio 3.

12 Which of the following investment products is *most likely* to trade at their net asset value per share?

A Exchange traded funds.

B Open-end mutual funds.

C Closed-end mutual funds.

13 Which of the following financial products is *least likely* to have a capital gain distribution?

A Exchange traded funds.

B Open-end mutual funds.

C Closed-end mutual funds.

14 Which of the following forms of pooled investments is subject to the *least* amount of regulation?

A Hedge funds.

B Exchange traded funds.

C Closed-end mutual funds.

15 Which of the following pooled investments is *most likely* characterized by a few large investments?

A Hedge funds.

B Buyout funds.

C Venture capital funds.

SOLUTIONS

1 A is correct. Combining assets into a portfolio should reduce the portfolio's volatility. Specifically, "individuals and institutions should hold portfolios to reduce risk." As illustrated in the reading, however, risk reduction may not be as great during a period of dramatic economic change.

2 A is correct. Combining assets into a portfolio should reduce the portfolio's volatility. The portfolio approach does not necessarily provide downside protection or guarantee that the portfolio always will avoid losses.

3 B is correct. As illustrated in the reading, portfolios reduce risk more than they increase returns.

4 A is correct. The excess reserves invested by banks need to be relatively liquid. Although investment companies and non-life insurance companies have high liquidity needs, the liquidity need for banks is on average the greatest.

5 B is correct. Most foundations and endowments are established with the intent of having perpetual lives. Although defined benefit plans and life insurance companies have portfolios with a long time horizon, they are not perpetual.

6 A is correct. Income is necessary to meet the cash flow obligation to retirees. Although defined benefit plans have a need for income, the need for liquidity typically is quite low. A retiree may need life insurance; however, a defined benefit plan does not need insurance.

7 B is correct. Investment companies manage investments in mutual funds. Although endowments and insurance companies may own mutual funds, they do not issue or redeem shares of mutual funds.

8 C is correct. The client's objectives and constraints are established in the investment policy statement and are used to determine the client's target asset allocation, which occurs in the execution step of the portfolio management process.

9 A is correct. Securities are analyzed in the execution step. In the planning step, a client's objectives and constraints are used to develop the investment policy statement.

10 B is correct. Portfolio monitoring and rebalancing occurs in the feedback step of the portfolio management process.

11 C is correct. Portfolio 3 has the same equity exposure as Portfolio 1 and has a higher exposure to alternative assets, which have greater volatility (as discussed in the section of the reading comparing the endowments from Yale University and the University of Virginia).

12 B is correct. Open-end funds trade at their net asset value per share, whereas closed-end funds and exchange traded funds can trade at a premium or a discount.

13 A is correct. Exchange traded funds do not have capital gain distributions. If an investor sells shares of an ETF (or open-end mutual fund or closed-end mutual fund), the investor may have a capital gain or loss on the shares sold; however, the gain (or loss) from the sale is not a distribution.

14 A is correct. Hedge funds are currently exempt from the reporting requirements of a typical public investment company.

15 B is correct. Buyout funds or private equity firms make only a few large investments in private companies with the intent of selling the restructured companies in three to five years. Venture capital funds also have a short time horizon; however, these funds consist of many small investments in companies with the expectation that only a few will have a large payoff (and that most will fail).

READING

49

Portfolio Risk and Return: Part I

by Vijay Singal, PhD, CFA

Vijay Singal, PhD, CFA, is at Virginia Tech (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** calculate and interpret major return measures and describe their appropriate uses;
☐	**b.** compare the money-weighted and time-weighted rates of return and evaluate the performance of portfolios based on these measures;
☐	**c.** describe characteristics of the major asset classes that investors consider in forming portfolios;
☐	**d.** calculate and interpret the mean, variance, and covariance (or correlation) of asset returns based on historical data;
☐	**e.** explain risk aversion and its implications for portfolio selection;
☐	**f.** calculate and interpret portfolio standard deviation;
☐	**g.** describe the effect on a portfolio's risk of investing in assets that are less than perfectly correlated;
☐	**h.** describe and interpret the minimum-variance and efficient frontiers of risky assets and the global minimum-variance portfolio;
☐	**i.** explain the selection of an optimal portfolio, given an investor's utility (or risk aversion) and the capital allocation line.

INTRODUCTION

1

Construction of an optimal portfolio is an important objective for an investor. In this reading, we will explore the process of examining the risk and return characteristics of individual assets, creating all possible portfolios, selecting the most efficient portfolios, and ultimately choosing the optimal portfolio tailored to the individual in question.

© 2019 CFA Institute. All rights reserved.

During the process of constructing the optimal portfolio, several factors and investment characteristics are considered. The most important of those factors are risk and return of the individual assets under consideration. Correlations among individual assets along with risk and return are important determinants of portfolio risk. Creating a portfolio for an investor requires an understanding of the risk profile of the investor. Although we will not discuss the process of determining risk aversion for individuals or institutional investors, it is necessary to obtain such information for making an informed decision. In this reading, we will explain the broad types of investors and how their risk–return preferences can be formalized to select the optimal portfolio from among the infinite portfolios contained in the investment opportunity set.

The reading is organized as follows: Sections 2–8 discuss the investment characteristics of assets. In particular, we show the various types of returns and risks, their computation and their applicability to the selection of appropriate assets for inclusion in a portfolio. Sections 9–11 discuss risk aversion and how indifference curves, which incorporate individual preferences, can be constructed. The indifference curves are then applied to the selection of an optimal portfolio using two risky assets. Sections 12–14 provides an understanding and computation of portfolio risk. The role of correlation and diversification of portfolio risk are examined in detail. Sections 15–17 begins with the risky assets available to investors and constructs a large number of risky portfolios. It illustrates the process of narrowing the choices to an efficient set of risky portfolios before identifying the optimal risky portfolio. The risky portfolio is combined with investor risk preferences to generate the investor's optimal portfolio. A summary concludes this reading.

2 INVESTMENT CHARACTERISTICS OF ASSETS: RETURN

a calculate and interpret major return measures and describe their appropriate uses;

Financial assets are frequently defined in terms of their risk and return characteristics. Comparison along these two dimensions simplifies the process of building a portfolio from among the multitude of available assets. In this section, we will compute, evaluate, and compare various measures of return and risk.

2.1 Return

Financial assets normally generate two types of return for investors. First, they may provide periodic income through cash dividends or interest payments. Second, the price of a financial asset can increase or decrease, leading to a capital gain or loss.

Some financial assets provide return through only one of these mechanisms. For example, investors in non-dividend-paying stocks obtain their return from price movement only. Similarly, you could also own or have a claim to assets that only generate periodic income. For example, defined benefit pension plans and retirement annuities make income payments as long as you live.

In the following section, we consider the computation and application of various types of returns.

2.1.1 *Holding Period Return*

Returns can be measured over a single period or over multiple periods. Single period returns are straightforward because there is only one way to calculate them. Multiple period returns, however, can be calculated in various ways and it is important to be aware of these differences to avoid confusion.

A **holding period return** is the return earned from holding an asset for a single specified period of time. The period may be 1 day, 1 week, 1 month, 5 years, or any specified period. If the asset (bond, stock, etc.) is bought now, time ($t = 0$), at a price of 100 and sold later, say at time ($t = 1$), at a price of 105 with no dividends or other income, then the holding period return is 5 percent [(105 − 100)/100)]. If the asset also pays an income of 2 units at time ($t = 1$), then the total return is 7 percent. This return can be generalized and shown as a mathematical expression in which P is the price and I is the income:

$$R = \frac{\left(P_1 - P_0\right) + I_1}{P_0}$$

The subscript indicates the time of the price or income, ($t = 0$), is the beginning of the period and ($t = 1$) is the end of the period. The following two observations are important.

- We computed a capital gain of 5 percent and a dividend yield of 2 percent in the above example. For ease of illustration, we assumed that the dividend is paid at time $t = 1$. If the dividend was received any time before $t = 1$, our holding period return would have been higher because we would have earned a return by reinvesting the dividend for the remainder of the period.
- Return can be expressed in decimals (0.07), fractions (7/100), or as a percent (7%). They are all equivalent.

The holding period return can be computed for a period longer than one year. For example, you may need to compute a three-year holding period return from three annual returns. In that case, the holding period return is computed by compounding the three annual returns: $R = [(1 + R_1) \times (1 + R_2) \times (1 + R_3)] - 1$, where R_1, R_2, and R_3 are the three annual returns.

2.1.2 *Arithmetic or Mean Return*

When assets have returns for multiple holding periods, it is necessary to aggregate those returns into one overall return for ease of comparison and understanding. Most holding period returns are reported as daily, monthly, or annual returns. When comparing returns across assets, it is important that the returns are computed using a common time period.

There are different methods for aggregating returns across several holding periods. The remainder of this section presents various ways of computing average returns and discusses their applicability.

The simplest way to compute the return is to take a simple arithmetic average of all holding period returns. Thus, three annual returns of −50 percent, 35 percent, and 27 percent will give us an average of 4 percent per year $= \left(\frac{-50\% + 35\% + 27\%}{3}\right)$. The arithmetic average return is easy to compute and has known statistical properties, such as standard deviation. We can calculate its standard deviation to determine how dispersed the observations are around the mean or if the mean return is statistically different from zero.

In general, the arithmetic or mean return is denoted by $\overline{R}_i$ and given by the following equation for asset *i*, where R_{it} is the return in period *t* and *T* is the total number of periods:

$$\overline{R}_i = \frac{R_{i1} + R_{i2} + \ldots + R_{i,T-1} + R_{iT}}{T} = \frac{1}{T}\sum_{t=1}^{T} R_{it}$$

2.1.3 *Geometric Mean Return*

The arithmetic mean return assumes that the amount invested at the beginning of each period is the same. In an investment portfolio, however, even if there are no cash flows into or out of the portfolio, the base amount changes each year. (The previous year's earnings must be added to the beginning value of the subsequent year's investment—these earnings will be "compounded" by the returns earned in that subsequent year.) We can use the geometric mean return to account for the compounding of returns.

A geometric mean return provides a more accurate representation of the growth in portfolio value over a given time period than does an arithmetic mean return. In general, the geometric mean return is denoted by $\overline{R}_{Gi}$ and given by the following equation for asset *i*:

$$\overline{R}_{Gi} = \sqrt[T]{(1 + R_{i1}) \times (1 + R_{i2}) \times \ldots \times (1 + R_{i,T-1}) \times (1 + R_{iT})} - 1$$

where R_{it} is the return in period *t* and *T* is the total number of periods.

In the example in Section 2, we calculated the arithmetic mean to be 4 percent. Exhibit 1 shows the actual return for each year and the balance at the end of each year using actual returns. Beginning with an initial investment of €1.0000, we will have a balance of €0.8573 at the end of the three-year period as shown in the third column. Note that we compounded the returns because, unless otherwise stated, we earn a return on the balance as of the end of the prior year. That is, we will receive a return of 35 percent in the second year on the balance at the end of the first year, which is only €0.5000, not the initial balance of €1.0000. Let us compare the balance at the end of the three-year period computed using geometric returns with the balance we would calculate using the 4 percent annual arithmetic mean return from our earlier example. The ending value using the arithmetic mean return is €1.1249 (=1.0000 × 1.04^3). This is much larger than the actual balance of €0.8573. In general, the arithmetic return is biased upward unless each of the underlying holding period returns are equal. The bias in arithmetic mean returns is particularly severe if holding period returns are a mix of both positive and negative returns, as in the example.

Exhibit 1

	Actual Return for the Year (%)	Year-End Amount	Year-End Amount Using Arithmetic Return of 4%	Year-End Amount Using Geometric Return of −5%
Year 0		€1.0000	€1.0000	€1.0000
Year 1	−50	0.5000	1.0400	0.9500
Year 2	35	0.6750	1.0816	0.9025
Year 3	27	0.8573	1.1249	0.8574

MONEY-WEIGHTED RETURN OR INTERNAL RATE OF RETURN

3

b compare the money-weighted and time-weighted rates of return and evaluate the performance of portfolios based on these measures;

The arithmetic and geometric return computations do not account for the cash flows into and out of a portfolio If the investor had invested €10,000 in the first year, €1,000 in the second year, and €1,000 in the third year, then the return of −50 percent in the first year significantly hurts her. On the other hand, if she had invested only €100 in the first year, the effect of the −50 percent return is drastically reduced.

The **money-weighted return** accounts for the money invested and provides the investor with information on the return she earns on her actual investment. The money-weighted return and its calculation are similar to the **internal rate of return** and the yield to maturity. Just like the internal rate of return, amounts invested are cash outflows from the investor's perspective and amounts returned or withdrawn by the investor, or the money that remains at the end of an investment cycle, is a cash inflow for the investor.

The money-weighted return can be illustrated most effectively with an example. In this example, we use the returns from the previous example. Assume that the investor invests €100 in a mutual fund at the beginning of the first year, adds another €950 at the beginning of the second year, and withdraws €350 at the end of the second year. The cash flows are shown in Exhibit 2.

Exhibit 2

Year	1	2	3
Balance from previous year	€0	€50	€1,000
New investment by the investor (cash inflow for the mutual fund) at the start of the year	100	950	0
Net balance at the beginning of year	100	1,000	1,000
Investment return for the year	−50%	35%	27%
Investment gain (loss)	−50	350	270
Withdrawal by the investor (cash outflow for the mutual fund) at the end of the year	0	−350	0
Balance at the end of year	€50	€1,000	€1,270

The internal rate of return is the discount rate at which the sum of present values of these cash flows will equal zero. In general, the equation may be expressed as follows, where T is the number of periods, CF_t is the cash flow at time t, and IRR is the internal rate of return or the money-weighted rate of return:

$$\sum_{t=0}^{T} \frac{CF_t}{(1 + \text{IRR})^t} = 0$$

A cash flow can be positive or negative; a positive cash flow is an inflow where money flows to the investor, whereas a negative cash flow is an outflow where money flows away from the investor. We can compute the internal rate of return by using the above equation. The flows are expressed as follows, where each cash inflow or outflow occurs at the end of each year. Thus, CF_0 refers to the cash flow at the end of Year 0 or beginning of Year 1, and CF_3 refers to the cash flow at end of Year 3 or beginning of Year 4. Because cash flows are being discounted to the present—that is, end of Year 0 or beginning of Year 1—the period of discounting CF_0 is zero.

$$CF_0 = -100$$
$$CF_1 = -950$$
$$CF_2 = +350$$
$$CF_3 = +1{,}270$$
$$\frac{CF_0}{(1+IRR)^0} + \frac{CF_1}{(1+IRR)^1} + \frac{CF_2}{(1+IRR)^2} + \frac{CF_3}{(1+IRR)^3}$$
$$= \frac{-100}{1} + \frac{-950}{(1+IRR)^1} + \frac{+350}{(1+IRR)^2} + \frac{+1270}{(1+IRR)^3} = 0$$
$$IRR = 26.11\%$$

IRR = 26.11% is the internal rate of return, or the money-weighted rate of return, which tells the investor what she earned on the actual euros invested for the entire period. This return is much greater than the arithmetic and geometric mean returns because only a small amount was invested when the mutual fund's return was −50 percent.

Next, we'll illustrate calculating the money-weighted return for a dividend paying stock. Consider an investment that covers a two-year horizon. At time $t = 0$, an investor buys one share at \$200. At time $t = 1$, he purchases an additional share at \$225. At the end of Year 2, $t = 2$, he sells both shares for \$235 each. During both years, the stock pays a per-share dividend of \$5. The $t = 1$ dividend is not reinvested. Exhibit 3 shows the total cash inflows and outflows.

Exhibit 3 Cash Flows

Time	Outflows
0	\$200 to purchase the first share
1	\$225 to purchase the second share

Time	Inflows
1	\$5 dividend received from first share (and not reinvested)
2	\$10 dividend (\$5 per share × 2 shares) received
2	\$470 received from selling two shares at \$235 per share

To solve for the money-weighted return, we use either a financial calculator that allows us to enter cash flows or a spreadsheet with an IRR function.[1] The first step is to group net cash flows by time. For this example, we have −\$200 for the $t = 0$ net cash flow, −\$220 = −\$225 + \$5 for the $t = 1$ net cash flow, and \$480 for the $t = 2$ net cash flow. After entering these cash flows, we use the spreadsheet's or calculator's IRR function to find that the money-weighted rate of return is 9.39 percent.[2]

---

[1] In this particular case we could solve for r by solving the quadratic equation $480x^2 - 220x - 200 = 0$ with $x = 1/(1 + r)$, using standard results from algebra. In general, however, we rely on a calculator or spreadsheet software to compute a money-weighted rate of return.

$$CF_0 = -200$$
$$CF_1 = -220$$
$$CF_2 = +480$$
$$\frac{CF_0}{(1+IRR)^0} + \frac{CF_1}{(1+IRR)^1} + \frac{CF_2}{(1+IRR)^2}$$
$$= \frac{-200}{1} + \frac{-220}{(1+IRR)^1} + \frac{480}{(1+IRR)^2} = 0$$
$$IRR = 9.39\%$$

Now we take a closer look at what has happened to the portfolio during each of the two years. In the first year, the portfolio generated a one-period holding period return of (\$5 + \$225 − \$200)/\$200 = 15 percent. At the beginning of the second year, the amount invested is \$450, calculated as \$225 (per share price of stock) × 2 shares, because the \$5 dividend was spent rather than reinvested. At the end of the second year, the proceeds from the liquidation of the portfolio are \$470 (as detailed in Exhibit 3) plus \$10 in dividends (as also detailed in Exhibit 3). So in the second year the portfolio produced a holding period return of (\$10 + \$470 − \$450)/\$450 = 6.67 percent. The mean holding period return was (15% + 6.67%)/2 = 10.84 percent. The money-weighted rate of return, which we calculated as 9.39 percent, puts a greater weight on the second year's relatively poor performance (6.67 percent) than the first year's relatively good performance (15 percent), as more money was invested in the second year than in the first. That is the sense in which returns in this method of calculating performance are "money weighted." Although the money-weighted return is an accurate measure of what the investor actually earned on the money invested, it is limited in its applicability to other situations. For example, it does not allow for return comparison between different individuals or different investment opportunities. Two investors in the *same* mutual fund or with the same portfolio of underlying investments may have different money-weighted returns because they invested different amounts in different years.

EXAMPLE 1

Computation of Returns

Ulli Lohrmann and his wife, Suzanne Lohrmann, are planning for retirement and want to compare the past performance of a few mutual funds they are considering for investment. They believe that a comparison over a five-year period would be appropriate. They are given the following information about the Rhein Valley Superior Fund that they are considering.

Year	Assets Under Management at the Beginning of Year (€)	Net Return (%)
1	30 million	15
2	45 million	−5
3	20 million	10

(continued)

---

[2] Note that the calculator or spreadsheet will give the IRR as a periodic rate. If the periods are not annual, we annualize the periodic rate.

Year	Assets Under Management at the Beginning of Year (€)	Net Return (%)
4	25 million	15
5	35 million	3

The Lohrmanns are interested in aggregating this information for ease of comparison with other funds.

1 Compute the holding period return for the five-year period.

2 Compute the arithmetic mean annual return.

3 Compute the geometric mean annual return. How does it compare with the arithmetic mean annual return?

4 The Lohrmanns want to earn a minimum annual return of 5 percent. Is the money-weighted annual return greater than 5 percent?

Solution to 1:

The holding period return is $R = (1 + R_1)(1 + R_2)(1 + R_3)(1 + R_4)(1 + R_5) - 1 = (1.15)(0.95)(1.10)(1.15)(1.03) - 1 = 0.4235 = 42.35\%$ for the five-year period.

Solution to 2:

The arithmetic mean annual return can be computed as an arithmetic mean of the returns given by this equation:

$$\bar{R}_i = \frac{15\% - 5\% + 10\% + 15\% + 3\%}{5} = 7.60\%$$

Solution to 3:

The geometric mean annual return can be computed using this equation:

$$\begin{aligned}\bar{R}_{Gi} &= \sqrt[T]{(1 + R_{i1}) \times (1 + R_{i2}) \times \ldots \times (1 + R_{i,T-1}) \times (1 + R_{iT})} - 1 \\ &= \sqrt[5]{1.15 \times 0.95 \times 1.10 \times 1.15 \times 1.03} - 1 \\ &= \sqrt[5]{1.4235} - 1 = 0.0732 = 7.32\%\end{aligned}$$

Thus, the geometric mean annual return is 7.32 percent, slightly less than the arithmetic mean return.

Solution to 4:

To calculate the money-weighted rate of return, tabulate the annual returns and investment amounts to determine the cash flows, as shown in Exhibit 4. All amounts are in millions of euros.

Exhibit 4

Year	1	2	3	4	5
Balance from previous year	0	34.50	42.75	22.00	28.75
New investment by the investor (cash inflow for the Rhein fund)	30.00	10.50	0	3.00	6.25
Withdrawal by the investor (cash outflow for the Rhein fund)	0	0	−22.75	0	0

Exhibit 4 (Continued)

Year	1	2	3	4	5
Net balance at the beginning of year	30.00	45.00	20.00	25.00	35.00
Investment return for the year	15%	–5%	10%	15%	3%
Investment gain (loss)	4.50	–2.25	2.00	3.75	1.05
Balance at the end of year	34.50	42.75	22.00	28.75	36.05

$CF_0 = -30.00$, $CF_1 = -10.50$, $CF_2 = +22.75$, $CF_3 = -3.00$, $CF_4 = -6.25$, $CF_5 = +36.05$.

For clarification, it may be appropriate to explain the notation for cash flows. Each cash inflow or outflow occurs at the end of each year. Thus, CF_0 refers to the cash flow at the end of Year 0 or beginning of Year 1, and CF_5 refers to the cash flow at end of Year 5 or beginning of Year 6. Because cash flows are being discounted to the present—that is, end of Year 0 or beginning of Year 1—the period of discounting CF_0 is zero whereas the period of discounting for CF_5 is 5 years.

To get the exact money-weighted rate of return (IRR), the following equation would be equal to zero. Instead of calculating, however, use the 5 percent return to see whether the value of the expression is positive or not. If it is positive, then the money-weighted rate of return is greater than 5 percent, because a 5 percent discount rate could not reduce the value to zero.

$$\frac{-30.00}{(1.05)^0} + \frac{-10.50}{(1.05)^1} + \frac{22.75}{(1.05)^2} + \frac{-3.00}{(1.05)^3} + \frac{-6.25}{(1.05)^4} + \frac{36.05}{(1.05)^5} = 1.1471$$

Because the value is positive, the money-weighted rate of return is greater than 5 percent. Using a financial calculator, the exact money-weighted rate of return is 5.86 percent.

TIME-WEIGHTED RATE OF RETURN

4

b compare the money-weighted and time-weighted rates of return and evaluate the performance of portfolios based on these measures;

An investment measure that is not sensitive to the additions and withdrawals of funds is the time-weighted rate of return. The **time-weighted rate of return** measures the compound rate of growth of $1 initially invested in the portfolio over a stated measurement period. For the evaluation of portfolios of publicly traded securities, the time-weighted rate of return is the preferred performance measure as it neutralizes the effect of cash withdrawals or additions to the portfolio, which are generally outside of the control of the portfolio manager. To compute an exact time-weighted rate of return on a portfolio, take the following three steps:

1 Price the portfolio immediately prior to any significant addition or withdrawal of funds. Break the overall evaluation period into subperiods based on the dates of cash inflows and outflows.

2 Calculate the holding period return on the portfolio for each subperiod.

3 Link or compound holding period returns to obtain an annual rate of return for the year (the time-weighted rate of return for the year). If the investment is for more than one year, take the geometric mean of the annual returns to obtain the time-weighted rate of return over that measurement period.

Let us return to our dividend stock money-weighted example and calculate the time-weighted rate of return for that investor's portfolio. In that example, we computed the holding period returns on the portfolio, Step 2 in the procedure for finding the time-weighted rate of return. Given that the portfolio earned returns of 15 percent during the first year and 6.67 percent during the second year, what is the portfolio's time-weighted rate of return over an evaluation period of two years?

We find this time-weighted return by taking the geometric mean of the two holding period returns, Step 3 in the procedure above. The calculation of the geometric mean exactly mirrors the calculation of a compound growth rate. Here, we take the product of 1 plus the holding period return for each period to find the terminal value at t = 2 of \$1 invested at t = 0. We then take the square root of this product and subtract 1 to get the geometric mean. We interpret the result as the annual compound growth rate of \$1 invested in the portfolio at t = 0. Thus we have

$$(1 + \text{Time-weighted return})^2 = (1.15)(1.0667)$$

$$\text{Time-weighted return} = \sqrt{(1.15)(1.0667)} - 1 = 10.76\%$$

The time-weighted return on the portfolio was 10.76 percent, compared with the money-weighted return of 9.39 percent, which gave larger weight to the second year's return. We can see why investment managers find time-weighted returns more meaningful. If a client gives an investment manager more funds to invest at an unfavorable time, the manager's money-weighted rate of return will tend to be depressed. If a client adds funds at a favorable time, the money-weighted return will tend to be elevated. The time-weighted rate of return removes these effects.

In defining the steps to calculate an exact time-weighted rate of return, we said that the portfolio should be valued immediately prior to any significant addition or withdrawal of funds. With the amount of cash flow activity in many portfolios, this task can be costly. We can often obtain a reasonable approximation of the time-weighted rate of return by valuing the portfolio at frequent, regular intervals, particularly if additions and withdrawals are unrelated to market movements. The more frequent the valuation, the more accurate the approximation. Daily valuation is commonplace. Suppose that a portfolio is valued daily over the course of a year. To compute the time-weighted return for the year, we first compute each day's holding period return. We compute 365 such daily returns, denoted r_1, r_2, …, r_{365}. We obtain the annual return for the year by linking the daily holding period returns in the following way: $(1 + r_1) \times (1 + r_2) \times \ldots \times (1 + r_{365}) - 1$. If withdrawals and additions to the portfolio happen only at day's end, this annual return is a precise time-weighted rate of return for the year. Otherwise, it is an approximate time-weighted return for the year.

If we have a number of years of data, we can calculate a time-weighted return for each year individually, as above. If r_i is the time-weighted return for year i, we calculate an annualized time-weighted return as the geometric mean of N annual returns, as follows:

$$r_{\text{TW}} = \left[(1 + r_1) \times (1 + r_2) \times \ldots \times (1 + r_N)\right]^{1/N} - 1$$

Example 2 illustrates the calculation of the time-weighted rate of return.

EXAMPLE 2

Time-Weighted Rate of Return

Strubeck Corporation sponsors a pension plan for its employees. It manages part of the equity portfolio in-house and delegates management of the balance to Super Trust Company. As chief investment officer of Strubeck, you want to review the performance of the in-house and Super Trust portfolios over the last four quarters. You have arranged for outflows and inflows to the portfolios to be made at the very beginning of the quarter. Exhibit 5 summarizes the inflows and outflows as well as the two portfolios' valuations. In the table, the ending value is the portfolio's value just prior to the cash inflow or outflow at the beginning of the quarter. The amount invested is the amount each portfolio manager is responsible for investing.

Exhibit 5 Cash Flows for the In-House Strubeck Account and the Super Trust Account

	Quarter			
	1 ($)	2 ($)	3 ($)	4 ($)
In-House Account				
Beginning value	4,000,000	6,000,000	5,775,000	6,720,000
Beginning of period inflow (outflow)	1,000,000	(500,000)	225,000	(600,000)
Amount invested	5,000,000	5,500,000	6,000,000	6,120,000
Ending value	6,000,000	5,775,000	6,720,000	5,508,000
Super Trust Account				
Beginning value	10,000,000	13,200,000	12,240,000	5,659,200
Beginning of period inflow (outflow)	2,000,000	(1,200,000)	(7,000,000)	(400,000)
Amount invested	12,000,000	12,000,000	5,240,000	5,259,200
Ending value	13,200,000	12,240,000	5,659,200	5,469,568

Based on the information given, address the following.

1 Calculate the time-weighted rate of return for the in-house account.

2 Calculate the time-weighted rate of return for the Super Trust account.

Solution to 1:

To calculate the time-weighted rate of return for the in-house account, we compute the quarterly holding period returns for the account and link them into an annual return. The in-house account's time-weighted rate of return is 27 percent, calculated as follows:

$$\text{1Q HPR: } r_1 = (\$6{,}000{,}000 - \$5{,}000{,}000)/\$5{,}000{,}000 = 0.20$$

$$\text{2Q HPR: } r_2 = (\$5{,}775{,}000 - \$5{,}500{,}000)/\$5{,}500{,}000 = 0.05$$

$$\text{3Q HPR: } r_3 = (\$6{,}720{,}000 - \$6{,}000{,}000)/\$6{,}000{,}000 = 0.12$$

$$\text{4Q HPR: } r_4 = (\$5{,}508{,}000 - \$6{,}120{,}000)/\$6{,}120{,}000 = -0.10$$

$$(1 + r_1)(1 + r_2)(1 + r_3)(1 + r_4) - 1 = (1.20)(1.05)(1.12)(0.90) - 1 = 0.27 \text{ or } 27\%$$

Solution to 2:

The account managed by Super Trust has a time-weighted rate of return of 26 percent, calculated as follows:

$$\text{1Q HPR: } r_1 = (\$13{,}200{,}000 - \$12{,}000{,}000)/\$12{,}000{,}000 = 0.10$$

$$\text{2Q HPR: } r_2 = (\$12{,}240{,}000 - \$12{,}000{,}000)/\$12{,}000{,}000 = 0.02$$

$$\text{3Q HPR: } r_3 = (\$5{,}659{,}200 - \$5{,}240{,}000)/\$5{,}240{,}000 = 0.08$$

$$\text{4Q HPR: } r_4 = (\$5{,}469{,}568 - \$5{,}259{,}200)/\$5{,}259{,}200 = 0.04$$

$$(1 + r_1)(1 + r_2)(1 + r_3)(1 + r_4) - 1 = (1.10)(1.02)(1.08)(1.04) - 1 = 0.26 \text{ or } 26\%$$

The in-house portfolio's time-weighted rate of return is higher than the Super Trust portfolio's by 100 basis points.

Having worked through this exercise, we are ready to look at a more detailed case.

EXAMPLE 3

Time-Weighted and Money-Weighted Rates of Return Side by Side

Your task is to compute the investment performance of the Walbright Fund during 2014. The facts are as follows:

- On 1 January 2014, the Walbright Fund had a market value of $100 million.
- During the period 1 January 2014 to 30 April 2014, the stocks in the fund showed a capital gain of $10 million.
- On 1 May 2014, the stocks in the fund paid a total dividend of $2 million. All dividends were reinvested in additional shares.
- Because the fund's performance had been exceptional, institutions invested an additional $20 million in Walbright on 1 May 2014, raising assets under management to $132 million ($100 + $10 + $2 + $20).
- On 31 December 2014, Walbright received total dividends of $2.64 million. The fund's market value on 31 December 2014, not including the $2.64 million in dividends, was $140 million.
- The fund made no other interim cash payments during 2014.

Based on the information given, address the following.

1 Compute the Walbright Fund's time-weighted rate of return.

2 Compute the Walbright Fund's money-weighted rate of return.

3 Interpret the differences between the time-weighted and money-weighted rates of return.

Solution to 1:

Because interim cash flows were made on 1 May 2014, we must compute two interim total returns and then link them to obtain an annual return. Exhibit 6 lists the relevant market values on 1 January, 1 May, and 31 December as well as the associated interim four-month (1 January to 1 May) and eight-month (1 May to 31 December) holding period returns.

Exhibit 6 Cash Flows for the Walbright Fund

1 January 2014	Beginning portfolio value = $100 million
1 May 2014	Dividends received before additional investment = $2 million
	Ending portfolio value = $110 million
	$\text{Holding period return} = \frac{\$2 + \$10}{\$100} = 12\%$
	New investment = $20 million
	Beginning market value for last 2/3 of year = $132 million
31 December 2014	Dividends received = $2.64 million
	Ending portfolio value = $140 million
	$\text{Holding period return} = \frac{\$2.64 + \$140 - \$132}{\$132} = 8.06\%$

Now we must geometrically link the four- and eight-month returns to compute an annual return. We compute the time-weighted return as follows:

$$\text{Time-weighted return} = 1.12 \times 1.0806 - 1 = 0.2103$$

In this instance, we compute a time-weighted rate of return of 21.03 percent for one year. The four-month and eight-month intervals combine to equal one year. (Taking the square root of the product 1.12 × 1.0806 would be appropriate only if 1.12 and 1.0806 each applied to one full year.)

Solution to 2:

To calculate the money-weighted return, we find the discount rate that sets the sum of the present value of cash inflows and outflows equal to zero. The initial market value of the fund and all additions to it are treated as cash outflows. (Think of them as expenditures.) Withdrawals, receipts, and the ending market value of the fund are counted as inflows. (The ending market value is the amount investors receive on liquidating the fund.) Because interim cash flows have occurred at four-month intervals, we must solve for the four-month internal rate of return. Exhibit 6 details the cash flows and their timing.

$CF_0 = -100$

$CF_1 = -20$

$CF_2 = 0$

$CF_3 = 142.64$

CF_0 refers to the initial investment of $100 million made at the beginning of the first four-month interval on 1 January 2014. CF_1 refers to the cash flows made at end of the first four-month interval or the beginning of the second four-month interval on 1 May 2014. Those cash flows include a cash inflow of $2 million for the dividend received and cash outflows of $22 million for the dividend reinvested and additional investment respectively. The second four-month interval had no cash flow so CF_2 is equal to zero. CF_3 refers to the cash inflows at the end of the third four-month interval. Those cash inflows include a $2.64 million dividend received and the fund's terminal market value of $140 million

Using a spreadsheet or IRR-enabled calculator, we use −100, −20, 0, and \$142.64 for the $t = 0$, $t = 1$, $t = 2$, and $t = 3$ net cash flows, respectively.[3] Using either tool, we get a four-month IRR of 6.28 percent. The quick way to annualize this is to multiply by 3. A more accurate way is $(1.0628)^3 - 1 = 0.20$ or 20 percent.

$$\frac{CF_0}{(1+\text{IRR})^0}+\frac{CF_1}{(1+\text{IRR})^1}+\frac{CF_2}{(1+\text{IRR})^2}+\frac{CF_3}{(1+\text{IRR})^3}$$

$$=\frac{-100}{1}+\frac{-20}{(1+\text{IRR})^1}+\frac{0}{(1+\text{IRR})^2}+\frac{142.64}{(1+\text{IRR})^3}$$

$$\text{IRR} = 6.28\%$$

Solution to 3:

In this example, the time-weighted return (21.03 percent) is greater than the money-weighted return (20 percent). The Walbright Fund's performance was relatively poorer during the eight-month period, when the fund owned more shares, than it was overall. This fact is reflected in a lower money-weighted rate of return compared with time-weighted rate of return, as the money-weighted return is sensitive to the timing and amount of withdrawals and additions to the portfolio.

The accurate measurement of portfolio returns is important to the process of evaluating portfolio managers. In addition to considering returns, however, analysts must also weigh risk. When we worked through Example 2, we stopped short of suggesting that in-house management was superior to Super Trust because it earned a higher time-weighted rate of return. A judgement as to whether performance was "better" or "worse" must include the risk dimension, which will be covered later in your study materials.

5 ANNUALIZED RETURN

a calculate and interpret major return measures and describe their appropriate uses;

The period during which a return is earned or computed can vary and often we have to annualize a return that was calculated for a period that is shorter (or longer) than one year. You might buy a short-term treasury bill with a maturity of 3 months, or you might take a position in a futures contract that expires at the end of the next quarter. How can we compare these returns? In, many cases, it is most convenient to annualize all available returns. Thus, daily weekly, monthly, and quarterly returns are converted to annualize all available returns. Many formulas used for calculating certain values or prices also require all returns and periods to be expressed as annualized rates of return. For example, the most common version of the Black–Scholes option-pricing model requires annualized returns and periods to be in years.

To annualize any return for a period shorter than one year, the return for the period must be compounded by the number of periods in a year. A monthly return is compounded 12 times, a weekly return is compounded 52 times, and a quarterly return is compounded 4 times. Daily returns are normally compounded 365 times. For an uncommon number of days, we compound by the ratio of 365 to the number of days.

---

[3] By convention, we denote outflow with a negative sign, and we need 0 as a placeholder for the $t = 2$.

If the weekly return is 0.2 percent, then the compound annual return is computed as shown because there are 52 weeks in a year:

$$\begin{aligned} r_{annual} &= \left(1 + r_{weekly}\right)^{52} - 1 = \left(1 + 0.2\%\right)^{52} - 1 \\ &= \left(1.002\right)^{52} - 1 = 0.1095 = 10.95\% \end{aligned}$$

If the return for 15 days is 0.4 percent, the annualized return is computed assuming 365 days in a year. Thus,

$$\begin{aligned} r_{annual} &= \left(1 + r_{15}\right)^{365/15} - 1 = \left(1 + 0.4\%\right)^{365/15} - 1 \\ &= \left(1.004\right)^{365/15} - 1 = 0.1020 = 10.20\% \end{aligned}$$

A general equation to annualize returns is given, where c is the number of periods in a year. For a quarter, $c = 4$ and for a month, $c = 12$:

$$r_{annual} = \left(1 + r_{period}\right)^{c} - 1$$

How can we annualize a return when the holding period return is more than one year? For example, how do we annualize an 18-month holding period return? Because one year contains two-thirds of 18-month periods, $c = 2/3$ in the above equation. An 18-month return of 20 percent can be annualized, as shown:

$$r_{annual} = \left(1 + r_{18month}\right)^{2/3} - 1 = \left(1 + 0.20\right)^{2/3} - 1 = 0.1292 = 12.92\%$$

Similar expressions can be constructed when quarterly or weekly returns are needed for comparison instead of annual returns. In such cases, c is equal to the number of holding periods in a quarter or in a week. For example, assume that you want to convert daily returns to weekly returns or annual returns to weekly returns for comparison between weekly returns. For converting daily returns to weekly returns, $c = 5$, assuming that there are five trading days in a week. For converting annual returns to weekly returns, $c = 1/52$. The expressions for annual returns can then be rewritten as expressions for weekly returns, as shown:

$$r_{weekly} = \left(1 + r_{daily}\right)^{5} - 1; \; r_{weekly} = \left(1 + r_{annual}\right)^{1/52} - 1$$

One major limitation of annualizing returns is the implicit assumption that returns can be repeated precisely, that is, money can be reinvested repeatedly while earning a similar return. This type of return is not always possible. An investor may earn a return of 5 percent during a week because the market went up that week or he got lucky with his stock, but it is highly unlikely that he will earn a return of 5 percent every week for the next 51 weeks, resulting in an annualized return of 1,164.3 percent ($= 1.05^{52} - 1$). Therefore, it is important to annualize short-term returns with this limitation in mind.

EXAMPLE 4

Annualized Returns

An analyst is trying to evaluate three securities that have been in her portfolio for different periods of time.

- In the last 100 days, Security A has earned a return of 6.2 percent.
- Security B has earned 2 percent over the last 4 weeks.
- Security C has earned a return of 5 percent over the last 3 months

The analyst is trying to assess the relative performance of the 3 securities.

Solution:

Annualized return for Security A: $R_{SA} = (1 + 0.062)^{365/100} - 1 = 0.2455 = 24.55\%$

Annualized return for Security B: $R_{SB} = (1 + 0.02)^{52/4} - 1 = 0.2936 = 29.36\%$

Annualized return for Security C: $R_{SC} = (1 + 0.05)^{4} - 1 = 0.2155 = 21.55\%$

Security B has the highest annualized return.

6

OTHER MAJOR RETURN MEASURES AND THEIR APPLICATIONS

a calculate and interpret major return measures and describe their appropriate uses;

The statistical measures of return discussed in the previous section are generally applicable across a wide range of assets and time periods. Special assets, however, such as mutual funds, and other considerations, such as taxes or inflation, may require return measures that are specific to a particular application.

Although it is not possible to consider all types of special applications, we will discuss the effect of fees (gross versus net returns), taxes (pre-tax and after-tax returns), inflation (nominal and real returns), and **leverage**. Many investors use mutual funds or other external entities (i.e., investment vehicles) for investment. In those cases, funds charge management fees and expenses to the investors. Consequently, gross and net-of-fund-expense returns should also be considered. Of course, an investor may be interested in the net-of-expenses after-tax real return, which is in fact what an investor truly receives. We consider these additional return measures in the following sections.

6.1 Gross and Net Return

A gross return is the return earned by an asset manager prior to deductions for management expenses, custodial fees, taxes, or any other expenses that are not directly related to the generation of returns but rather related to the management and administration of an investment. These expenses are not deducted from the gross return because they may vary with the amount of assets under management or may vary because of the tax status of the investor. Trading expenses, however, such as commissions, *are* accounted for in (i.e., deducted from) the computation of gross return because trading expenses contribute directly to the return earned by the manager. Thus, gross return is an appropriate measure for evaluating and comparing the investment skill of asset managers because it does not include any fees related to the management and administration of an investment.

Net return is a measure of what the investment vehicle (mutual fund, etc.) has earned for the investor. Net return accounts for (i.e., deducts) all managerial and administrative expenses that reduce an investor's return. Because individual investors are most concerned about the net return (i.e., what they actually receive), small mutual funds with a limited amount of assets under management are at a disadvantage compared with the larger funds that can spread their largely fixed administrative expenses over a larger asset base. As a result, many small mutual funds waive part of the expenses to keep the funds competitive.

6.2 Pre-tax and After-tax Nominal Return

All return measures discussed previously are pre-tax nominal returns—that is, no adjustment has been made for taxes or inflation. In general, all returns are pre-tax nominal returns unless they are otherwise designated.

Many investors are concerned about the possible tax liability associated with their returns because taxes reduce the net return that they receive. Capital gains and income may be taxed differently, depending on the jurisdiction. Capital gains come in two forms: short-term capital gains and long-term capital gains. Long-term capital gains receive preferential tax treatment in a number of countries. Interest income is taxed as ordinary income in most countries. Dividend income may be taxed as ordinary income, may have a lower tax rate, or may be exempt from taxes depending on the country and the type of investor. The after-tax nominal return is computed as the total return minus any allowance for taxes on dividends, interest and realized gains.[4]

Because taxes are paid on realized capital gains and income, the investment manager can minimize the tax liability by selecting appropriate securities (e.g., those subject to more favorable taxation, all other investment considerations equal) and reducing trading turnover. Therefore, taxable investors evaluate investment managers based on the after-tax nominal return.

6.3 Real Returns

A nominal return (r) consists of three components: a real risk-free return as compensation for postponing consumption (r_{rF}), inflation as compensation for loss of purchasing power (π), and a **risk premium** for assuming risk (RP). Thus, nominal return and real return can be expressed as:

$$(1 + r) = (1 + r_{rF}) \times (1 + \pi) \times (1 + RP)$$

$$(1 + r_{real}) = (1 + r_{rF}) \times (1 + RP) \text{ or}$$

$$(1 + r_{real}) = (1 + r) \div (1 + \pi)$$

Often the real risk-free return and the risk premium are combined to arrive at the real “risky” rate as given in the second equation above, simply referred to as the real return. Real returns are particularly useful in comparing returns across time periods because inflation rates may vary over time. Real returns are also useful in comparing returns among countries when returns are expressed in local currencies instead of a constant investor currency and when inflation rates vary between countries (which are usually the case). Finally, the after-tax real return is what the investor receives as compensation for postponing consumption and assuming risk after paying taxes on investment returns. As a result, the after-tax real return becomes a reliable benchmark for making investment decisions. Although it is a measure of an investor’s benchmark return, it is not commonly calculated by asset managers because it is difficult to estimate a general tax component applicable to all investors. For example, the tax component depends on an investor’s specific taxation rate (marginal tax rate), how long the investor holds an investment (long-term versus short-term), and the type of account the asset is held in (tax-exempt, tax-deferred, or normal).

---

[4] Bonds issued at a discount to the par value may be taxed based on accrued gains instead of realized gains.

EXAMPLE 5

Computation of Special Returns

Let's return to Example 1. After reading this section, Mr. Lohrmann decided that he was not being fair to the fund manager by including the asset management fee and other expenses because the small size of the fund would put it at a competitive disadvantage. He learns that the fund spends a fixed amount of €500,000 every year on expenses that are unrelated to the manager's performance.

Mr. Lohrmann has become concerned that both taxes and inflation may reduce his return. Based on the current tax code, he expects to pay 20 percent tax on the return he earns from his investment. Historically, inflation has been around 2 percent and he expects the same rate of inflation to be maintained.

1 Estimate the annual gross return for the first year by adding back the fixed expenses.
2 What is the net return that investors in the Rhein Valley Superior Fund earned during the five-year period?
3 What is the after-tax net return for the first year that investors earned from the Rhein Valley Superior Fund? Assume that all gains are realized at the end of the year and the taxes are paid immediately at that time.
4 What is the anticipated after-tax real return that investors would have earned in the fifth year?

Solution to 1:

The gross return for the first year is higher by 1.67 percent (= €500,000/€30,000,000) than the investor return reported by the fund. Thus, the gross return is 16.67 percent (= 15% + 1.67%).

Solution to 2:

The investor return reported by the mutual fund is the net return of the fund after accounting for all direct and indirect expenses. The net return is also the pre-tax nominal return because it has not been adjusted for taxes or inflation. The net return for the five-year holding period was 42.35 percent.

Solution to 3:

The net return earned by investors during the first year was 15 percent. Applying a 20 percent tax rate, the after-tax return that accrues to the investors is 12 percent [= 15% − (0.20 × 15%)].

Solution to 4:

As in Part 3, the after-tax return earned by investors in the fifth year is 2.4 percent [= 3% − (0.20 × 3%)]. Inflation reduces the return by 2 percent so the after-tax real return earned by investors in the fifth year is 0.39 percent, as shown:

$$\frac{(1 + 2.40\%)}{(1 + 2.00\%)} - 1 = \frac{(1 + 0.0240)}{(1 + 0.0200)} - 1 = 1.0039 - 1 = 0.0039 = 0.39\%$$

Note that taxes are paid before adjusting for inflation.

6.4 Leveraged Return

In the previous calculations, we have assumed that the investor's position in an asset is equal to the total investment made by an investor using his or her own money. This section differs in that the investor creates a leveraged position. There are two ways of creating a claim on asset returns that are greater than the investment of one's own money. First, an investor may trade futures contracts in which the money required to take a position may be as little as 10 percent of the notional value of the asset. In this case, the leveraged return, the return on the investor's own money, is 10 times the actual return of the underlying security. Both the gains and losses are amplified by a factor of 10.

Investors can also invest more than their own money by borrowing money to purchase the asset. This approach is easily done in stocks and bonds, and very common when investing in real estate. If half (50 percent) of the money invested is borrowed, then the gross return to the investor is doubled but the interest to be paid on borrowed money must be deducted in order to calculate the net return.

HISTORICAL RETURN AND RISK

7

c describe characteristics of the major asset classes that investors consider in forming portfolios;

At this time, it is helpful to look at historical risk and returns for the three main asset categories: stocks, bonds, and Treasury bills. Stocks refer to corporate ownership, bonds refer to long-term fixed-income securities, and Treasury bills refer to short-term government debt securities. Although there is generally no expectation of default on government securities, long-term government bond prices are volatile (risky) because of possible future changes in interest rates. In addition, bondholders also face the risk that inflation will reduce the purchasing power of their cash flows.

7.1 Historical Mean Return and Expected Return

Before examining historical data, it is useful to distinguish between the historical mean return and expected return, which are very different concepts but easy to confuse. Historical return is what was actually earned in the *past*, whereas expected return is what an investor anticipates to earn in the *future*.

Expected return is the nominal return that would cause the marginal investor to invest in an asset based on the real risk-free interest rate (r_{rF}), expected inflation [$E(\pi)$], and expected risk premium for the risk of the asset [$E(RP)$]. The real risk-free interest rate is expected to be positive as compensation for postponing consumption. Similarly, the risk premium is expected to be positive in most cases.[5] The expected inflation rate is generally positive, except when the economy is in a deflationary state and prices are falling. Thus, expected return is generally positive. The relationship between the expected return and the real risk-free interest rate, inflation rate, and risk premium can be expressed by the following equation:

$$1 + E(R) = (1 + r_{rF}) \times [1 + E(\pi)] \times [1 + E(RP)]$$

The historical mean return for investment in a particular asset, however, is obtained from the actual return that was earned by an investor. Because the investment is risky, there is no guarantee that the actual return will be equal to the expected return. In fact, it is very unlikely that the two returns are equal for a specific time period being considered. Given a long enough period of time, we can *expect* that the future (expected) return will equal the average historical return. Unfortunately, we do not know how long that period is—10 years, 50 years, or 100 years. As a practical matter, we often assume that the historical mean return is an adequate representation of the expected return, although this assumption may not be accurate. For example, Exhibit 7 shows that the historical equity returns in the last eight years (2010–2017) for large US company stocks were positive whereas the actual return was negative the prior decade, but nearly always positive historically. Nonetheless, longer-term returns (1926–2017) were positive and could be consistent with expected return. Though it is unknown if the historical mean returns accurately represent expected returns, it is an assumption that is commonly made.

---

[5] There are exceptions when an asset reduces overall risk of a portfolio. We will consider those exceptions in Section 14.

Exhibit 7 Risk and Return for US Asset Classes by Decade (%)

		1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s*	1926–2017
Large company stocks	Return	−0.1	9.2	19.4	7.8	5.9	17.6	18.2	−1.0	13.9	10.2
	Risk	41.6	17.5	14.1	13.1	17.2	19.4	15.9	16.3	13.6	19.8
Small company stocks	Return	1.4	20.7	16.9	15.5	11.5	15.8	15.1	6.3	14.8	12.1
	Risk	78.6	34.5	14.4	21.5	30.8	22.5	20.2	26.1	19.4	31.7
Long-term corporate bonds	Return	6.9	2.7	1	1.7	6.2	13	8.4	7.7	8.3	6.1
	Risk	5.3	1.8	4.4	4.9	8.7	14.1	6.9	11.7	8.8	8.3
Long-term government bonds	Return	4.9	3.2	−0.1	1.4	5.5	12.6	8.8	7.7	6.8	5.5
	Risk	5.3	2.8	4.6	6	8.7	16	8.9	12.4	10.8	9.9
Treasury bills	Return	0.6	0.4	1.9	3.9	6.3	8.9	4.9	2.8	0.2	3.4
	Risk	0.2	0.1	0.2	0.4	0.6	0.9	0.4	0.6	0.1	3.1
Inflation	Return	−2.0	5.4	2.2	2.5	7.4	5.1	2.9	2.5	1.7	2.9
	Risk	2.5	3.1	1.2	0.7	1.2	1.3	0.7	1.6	1.1	4.0

* Through 31 December 2017

Note: Returns are measured as annualized geometric mean returns.

Risk is measured by annualizing monthly standard deviations.

Source: 2018 SBBI Yearbook (Exhibits 1.2,1.3, 2.3 and 6.2).

Going forward, be sure to distinguish between expected return and historical mean return. We will alert the reader whenever historical returns are used to estimate expected returns.

7.2 Nominal Returns of Major US Asset Classes

We focus on three major asset categories in Exhibit 7: stocks, bonds, and T-bills. The mean nominal returns for US asset classes are reported decade by decade since the 1930s. The total for the 1926–2017 period is in the last column. All returns are annual geometric mean returns. Large company stocks had an overall annual return of 10.2 percent during the 92-year period. The return was negative in the 1930s and 2000s, and positive in all remaining decades. The 1950s and 1990s were the best decades for large company stocks. Small company stocks fared even better. The nominal return was never negative for any decade, and had double-digit growth in all decades except two, leading to an overall 92-year annual return of 12.1 percent.

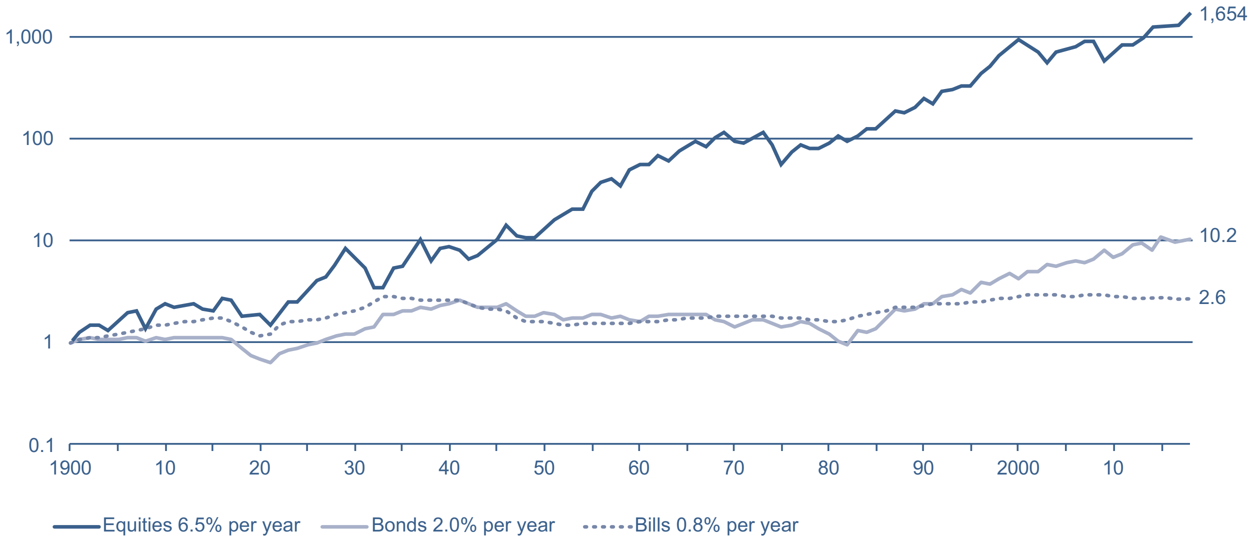
Long-term corporate bonds and long-term government bonds earned overall returns of 6.1 percent and 5.5 percent, respectively. The corporate bonds did not have a single negative decade, although government bonds recorded a negative return in the 1950s when stocks were doing extremely well. Bonds also had some excellent decades, earning double-digit returns in the 1980s and 2000s.

Treasury bills (short-term government securities) did not earn a negative return in any decade. In fact, Treasury bills earned a negative return only in 1938 (–0.02 percent) when the inflation rate was –2.78 percent. Consistently positive returns for Treasury bills are not surprising because nominal interest rates are almost never negative and the Treasury bills suffer from little interest rate or inflation risk. Since the Great Depression, there has been no deflation in any decade, although inflation rates were highly negative in 1930 (–6.03 percent), 1931 (–9.52 percent), and 1932 (–10.30 percent). Conversely, inflation rates were very high in the late 1970s and early 1980s, reaching 13.31 percent in 1979. Inflation rates have been largely range bound between 1 and 3 percent from 1991 to 2017. Overall, the inflation rate was 2.9 percent for the 92-year period.

7.3 Real Returns of Major US Asset Classes

Because annual inflation rates can vary greatly, from –10.30 percent to +13.31 percent in the last 92 years, comparisons across various time periods are difficult and misleading using nominal returns. Therefore, it is more effective to rely on real returns. Real returns on stocks, bonds, and T-bills are reported from 1900 in Exhibits 8 and 9.

Exhibit 8 Cumulative Returns on US Asset Classes in Real Terms, 1900–2017



Source: E. Dimson, P. Marsh, and M. Staunton, *Credit Suisse Global Investment Returns Yearbook 2018*, Credit Suisse Research Institute (February 2018). This chart is updated annually and can be found at https://www.credit-suisse.com/media/assets/corporate/docs/about-us/media/media-release/2018/02/giry-summary-2018.pdf.

Exhibit 8 shows that $1 would have grown to $1,654 if invested in stocks, to only $10.20 if invested in bonds, and to $2.60 if invested in T-bills. The difference in growth among the three asset categories is huge, although the difference in real returns does not seem that large: 6.5 percent per year for equities compared with 2.0 percent per year for bonds. This difference represents the effect of compounding over a 118-year period.

Exhibit 9 reports real rates of return. As we discussed earlier and as shown in the table, geometric mean is never greater than the arithmetic mean. Our analysis of returns focuses on the geometric mean because it is a more accurate representation of returns for multiple holding periods than the arithmetic mean. We observe that the real returns for stocks are higher than the real returns for bonds.

Exhibit 9 Real Returns and Risk Premiums for Asset Classes (1900–2017)

		United States			World			World excluding United States		
	Asset	*GM* (%)	*AM* (%)	*SD* (%)	*GM* (%)	*AM* (%)	*SD* (%)	*GM* (%)	*AM* (%)	*SD* (%)
Real Returns	Equities	6.5	8.4	20.0	5.2	6.6	17.4	4.5	6.2	18.9
	Bonds	2.0	2.5	10.4	2.0	2.5	11.0	1.7	2.7	14.4
Premiums	Equities vs. bonds	4.4	6.5	20.7	3.2	4.4	15.3	2.8	3.8	14.4

Note: All returns are in percent per annum measured in US$. GM = geometric mean, AM = arithmetic mean, SD = standard deviation. "World" consists of 21 developed countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, South Africa, Spain, Sweden, Switzerland, United Kingdom, and the United States. Weighting is by each country's relative market capitalization size. See source for details of calculations.
Source: Credit Suisse Global Investment Returns Sourcebook, 2018.

7.4 Nominal and Real Returns of Asset Classes in Major Countries

Along with US returns, real returns of major asset classes for a 21-country world and the world excluding the United States are also presented in Exhibit 9. Equity returns are weighted by each country's GDP before 1968 because of a lack of reliable market capitalization data. Returns are weighted by a country's market capitalization beginning with 1968. Similarly, bond returns are defined by a 21-country bond index, except GDP is used to create the weights because equity market capitalization weighting is inappropriate for a bond index and bond market capitalizations were not readily available.

The real geometric mean return for the world stock index over the last 117 years was 5.2 percent, and bonds had a real geometric mean return of 2.0 percent. The real geometric mean return for the world excluding the United States were 4.5 percent for stocks and 1.7 percent for bonds. For both stocks and bonds, the United States earned higher returns than the world excluding the United States. Similarly, real returns for stocks and bonds in the United States were higher than the real returns for rest of the world.

7.5 Risk of Major Asset Classes

Risk for major asset classes in the United States is reported for 1926–2017 in Exhibit 7, and the risk for major asset classes for the United States, the world, and the world excluding the United States are reported for 1900–2017 in Exhibit 9. Exhibit 7 shows that US small company stocks had the highest risk, 31.7 percent, followed by US large company stocks, 19.8 percent. Long-term government bonds and long-term corporate bonds had lower risk at 9.9 percent and 8.3 percent, with Treasury bills having the lowest risk at about 3.1 percent.

Exhibit 9 shows that the risk for world stocks is 17.4 percent and for world bonds is 11.0 percent. The world excluding the United States has risks of 18.9 percent for stocks and 14.4 percent for bonds. The effect of diversification is apparent when world risk is compared with US risk and world excluding US risk. Although the risk of US stocks is 20.0 percent and the risk of world excluding US stocks is 18.9 percent, the combination gives a risk of only 17.4 percent for world stocks.

7.6 Risk–Return Trade-off

The expression "risk–return trade-off" refers to the positive relationship between expected risk and return. In other words, a higher return is not possible to attain in **efficient markets** and over long periods of time without accepting higher risk. Expected returns should be greater for assets with greater risk.

The historical data presented above show the risk–return trade-off. Exhibit 7 shows for the United States that small company stocks had higher risk and higher return than large company stocks. Large company stocks had higher returns and higher risk than both long-term corporate bonds and government bonds. Bonds had higher returns and higher risk than Treasury bills. Uncharacteristically, however, long-term government bonds had higher total risk than long-term corporate bonds, although the returns of corporate bonds were slightly higher. These factors do not mean that long-term government bonds had greater default risk, just that they were more variable than corporate bonds during this historic period.

Exhibit 9 reveals that the risk and return for stocks were the highest of the asset classes, and the risk and return for bonds were lower than stocks for the United States, the world, and the world excluding the United States.

Another way of looking at the risk–return trade-off is to focus on the **risk premium**, which is the extra return investors can expect for assuming additional risk, after accounting for the risk-free interest rate. The nominal risk premium is the nominal risky return minus the nominal risk-free rate (which includes both compensation for expected inflation and the real risk-free interest rate). The real risk premium is the real risky return minus the real risk-free rate. Worldwide equity risk premiums reported at the bottom of Exhibit 9 show that equities outperformed bonds. Investors in equities earned a higher return than investors in bonds because of the higher risk in equities.

A more dramatic representation of the risk–return trade-off is shown in Exhibit 8, which shows the cumulative returns of US asset classes in real terms. The line representing T-bills is much less volatile than the other lines. Adjusted for inflation, the average real return on T-bills was 0.8 percent per year. The line representing bonds is more volatile than the line for T-bills but less volatile than the line representing stocks. The total return for equities including dividends and capital gains shows how $1 invested at the beginning of 1900 grows to $1,654, generating an annualized return of 6.5 percent in real terms.

Over long periods of time, we observe that higher risk does result in higher mean returns. Thus, it is reasonable to claim that, over the long term, market prices reward higher risk with higher returns, which is a characteristic of a risk-averse investor, a topic that we discuss in Section 9.

OTHER INVESTMENT CHARACTERISTICS

8

c describe characteristics of the major asset classes that investors consider in forming portfolios;

In evaluating investments using only the mean (expected return) and variance (risk), we are implicitly making two important assumptions: 1) that the returns are normally distributed and can be fully characterized by their means and variances and 2) that markets are not only informationally efficient but that they are also operationally efficient. To the extent that these assumptions are violated, we need to consider additional investment characteristics. These are discussed below.

8.1 Distributional Characteristics

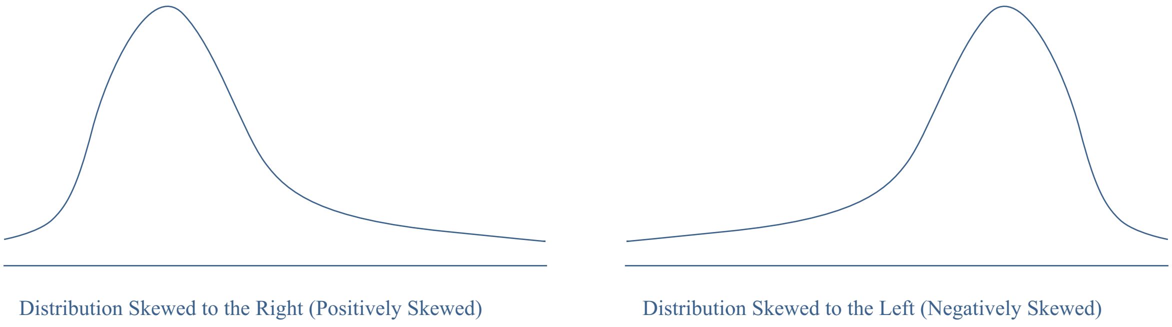
As explained in an earlier reading, a **normal distribution** has three main characteristics: its mean and median are equal; it is completely defined by two parameters, mean and variance; and it is symmetric around its mean with:

- 68 percent of the observations within ±1σ of the mean,
- 95 percent of the observations within ±2σ of the mean, and
- 99 percent of the observations within ±3σ of the mean.

Using only mean and variance would be appropriate to evaluate investments if returns were distributed normally. Returns, however, are not normally distributed; deviations from normality occur both because the returns are skewed, which means they are not symmetric around the mean, and because the probability of extreme events is significantly greater than what a normal distribution would suggest. The latter deviation is referred to as kurtosis or fat tails in a return distribution. The next sections discuss these deviations more in-depth.

Skewness **Skewness** refers to asymmetry of the return distribution, that is, returns are not symmetric around the mean. A distribution is said to be left skewed or negatively skewed if most of the distribution is concentrated to the right, and right skewed or positively skewed if most is concentrated to the left. Exhibit 10 shows a typical representation of negative and positive skewness, whereas Exhibit 11 demonstrates the negative skewness of stock returns by plotting a histogram of US large company stock returns for 1926–2017.

Exhibit 10 Skewness

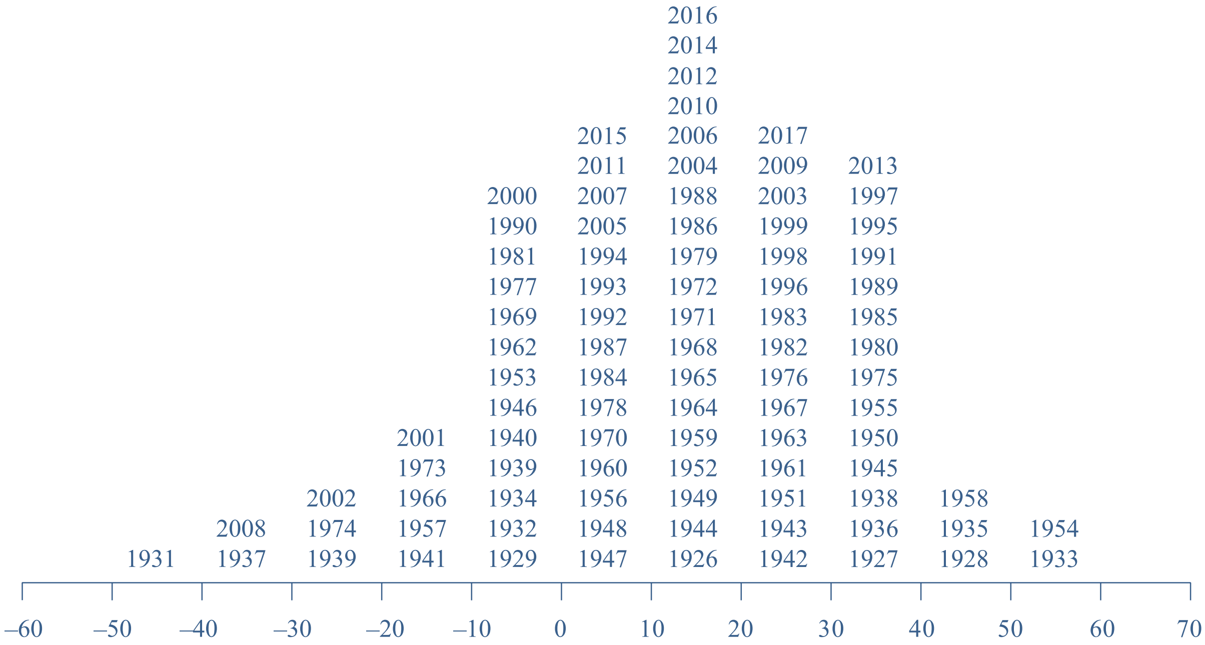


Distribution Skewed to the Right (Positively Skewed)

Distribution Skewed to the Left (Negatively Skewed)

Source: Reprinted from *Fixed Income Readings for the Chartered Financial Analyst® Program.* Copyright CFA Institute.

Exhibit 11 Histogram of US Large Company Stock Returns, 1926–2017 (Percent)



Source: 2018 SBBI Yearbook (Appendix A1)

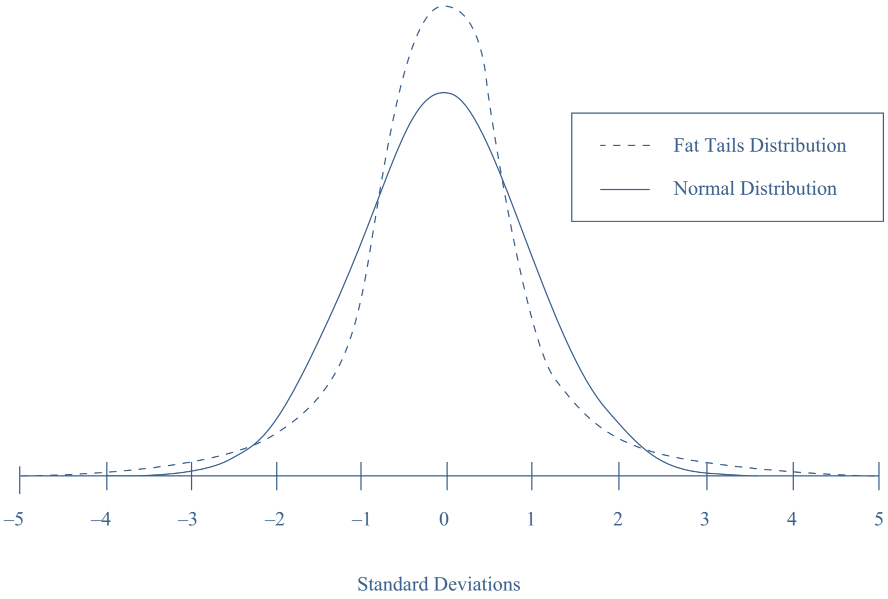
Kurtosis **Kurtosis** refers to fat tails or higher than normal probabilities for extreme returns and has the effect of increasing an asset's risk that is not captured in a mean–variance framework, as illustrated in Exhibit 12. Investors try to evaluate the effect of kurtosis by using such statistical techniques as value at risk (VaR) and conditional tail expectations.[6] Several market participants note that the probability and the magnitude of extreme events is underappreciated and was a primary contributing factor to the financial crisis of 2008.[7] The higher probability of extreme negative outcomes among stock returns can also be observed in Exhibit 11.

---

[6] Value at risk is a money measure of the minimum losses expected on a portfolio during a specified time period at a given level of probability. It is commonly used to measure the losses a portfolio can suffer under normal market conditions. For example, if a portfolio's one-day 10 percent VaR is £200,000, it implies that there is a 10 percent probability that the value of the portfolio will decrease by more than £200,000 over a single one-day period (under normal market conditions). This probability implies that the portfolio will experience a loss of at least £200,000 on one out of every ten days.

[7] For example, see Bogle (2008) and Taleb (2007).

Exhibit 12 Kurtosis



Source: Reprinted from *Fixed Income Readings for the Chartered Financial Analyst® Program.* Copyright CFA Institute.

8.2 Market Characteristics

In the previous analysis, we implicitly assumed that markets are both informationally and operationally efficient. Although informational efficiency of markets is a topic beyond the purview of this reading, we should highlight certain operational limitations of the market that affect the choice of investments. One such limitation is **liquidity**.

The cost of trading has three main components—brokerage commission, bid–ask spread, and price impact. Liquidity affects the latter two. Stocks with low liquidity can have wide bid–ask spreads. The bid–ask spread, which is the difference between the buying price and the selling price, is incurred as a cost of trading a security. The larger the bid–ask spread, the higher the cost of trading. If a $100 stock has a spread of 10 cents, the bid–ask spread is only 0.1 percent ($0.10/$100). On the other hand, if a $10 stock has a spread of 10 cents, the bid–ask spread is 1 percent. Clearly, the $10 stock is more expensive to trade and an investor will need to earn 0.9 percent extra to make up the higher cost of trading relative to the $100 stock.

Liquidity also has implications for the price impact of trade. Price impact refers to how the price moves in response to an order in the market. Small orders usually have little impact, especially for liquid stocks. For example, an order to buy 100 shares of a $100 stock with a spread of 1 cent may have no effect on the price. On the other hand, an order to buy 100,000 shares may have a significant impact on the price as the buyer has to induce more and more stockholders to tender their shares. The extent of the price impact depends on the liquidity of the stock. A stock that trades millions of shares a day may be less affected than a stock that trades only a few hundred thousand shares a day. Investors, especially institutional investors managing large sums of money, must keep the liquidity of a stock in mind when making investment decisions.

Liquidity is a bigger concern in emerging markets than in developed markets because of the smaller volume of trading in those markets. Similarly, liquidity is a more important concern in corporate bond markets and especially for bonds of lower credit quality than in equity markets because an individual corporate bond issue may not trade for several days or weeks. This certainly became apparent during the global financial crisis.

There are other market-related characteristics that affect investment decisions because they might instill greater confidence in the security or might affect the costs of doing business. These include analyst coverage, availability of information, firm size, etc. These characteristics about companies and financial markets are essential components of investment decision making.

RISK AVERSION AND PORTFOLIO SELECTION & THE CONCEPT OF RISK AVERSION

9

e explain risk aversion and its implications for portfolio selection;

As we have seen, stocks, bonds, and T-bills provide different levels of returns and have different levels of risk. Although investment in equities may be appropriate for one investor, another investor may not be inclined to accept the risk that accompanies a share of stock and may prefer to hold more cash. In the last section, we considered investment characteristics of assets in understanding their risk and return. In this section, we consider the characteristics of investors, both individual and institutional, in an attempt to pair the right kind of investors with the right kind of investments.

First, we discuss risk aversion and utility theory. Later we discuss their implications for portfolio selection.

9.1 The Concept of Risk Aversion

The concept of **risk aversion** is related to the behavior of individuals under uncertainty. Assume that an individual is offered two alternatives: one where he will get £50 for sure and the other is a gamble with a 50 percent chance that he gets £100 and 50 percent chance that he gets nothing. The expected value in both cases is £50, one with certainty and the other with uncertainty. What will an investor choose? There are three possibilities: an investor chooses the gamble, the investor chooses £50 with certainty, or the investor is indifferent. Let us consider each in turn. However, please understand that this is only a representative example, and a single choice does not determine the risk aversion of an investor.

Risk Seeking

If an investor chooses the gamble, then the investor is said to be risk loving or risk seeking. The gamble has an uncertain outcome, but with the same expected value as the guaranteed outcome. Thus, an investor choosing the gamble means that the investor gets extra "utility" from the uncertainty associated with the gamble. How much is that extra utility worth? Would the investor be willing to accept a smaller expected value because he gets extra utility from risk? Indeed, risk seekers will accept less return because of the risk that accompanies the gamble. For example, a risk seeker may choose a gamble with an expected value of £45 in preference to a guaranteed outcome of £50.

There is a little bit of gambling instinct in many of us. People buy lottery tickets although the expected value is less than the money they pay to buy it. Or people gamble at casinos with the full knowledge that the expected return is negative, a characteristic of risk seekers. These or any other isolated actions, however, cannot be taken at face value except for compulsive gamblers.

Risk Neutral

If an investor is indifferent about the gamble or the guaranteed outcome, then the investor may be risk neutral. Risk neutrality means that the investor cares only about return and not about risk, so higher return investments are more desirable even if they come with higher risk. Many investors may exhibit characteristics of risk neutrality when the investment at stake is an insignificant part of their wealth. For example, a billionaire may be indifferent about choosing the gamble or a £50 guaranteed outcome.

Risk Averse

If an investor chooses the guaranteed outcome, he/she is said to be **risk averse** because the investor does not want to take the chance of not getting anything at all. Depending on the level of aversion to risk, an investor may be willing to accept a guaranteed outcome of £45 instead of a gamble with an expected value of £50.

In general, investors are likely to shy away from risky investments for a lower, but guaranteed return. That is why they want to minimize their risk for the same amount of return, and maximize their return for the same amount of risk. The risk–return trade-off discussed earlier is an indicator of risk aversion. A risk-neutral investor would maximize return irrespective of risk and a risk-seeking investor would maximize both risk and return.

Data presented in the last section illustrate the historically positive relationship between risk and return, which demonstrates that market prices were based on transactions and investments by risk-averse investors and reflect risk aversion. Therefore, for all practical purposes and for our future discussion, we will assume that the representative investor is a risk-averse investor. This assumption is the standard approach taken in the investment industry globally.

Risk Tolerance

Risk tolerance refers to the amount of risk an investor can tolerate to achieve an investment goal. The higher the risk tolerance, the greater is the willingness to take risk. Thus, risk tolerance is negatively related to risk aversion.

10

UTILITY THEORY AND INDIFFERENCE CURVES

e explain risk aversion and its implications for portfolio selection;

Continuing with our previous example, a risk-averse investor would rank the guaranteed outcome of £50 higher than the uncertain outcome with an expected value of £50. We can say that the utility that an investor or an individual derives from the guaranteed outcome of £50 is greater than the utility or satisfaction or happiness he/she derives from the alternative. In general terms, utility is a measure of relative satisfaction from consumption of various goods and services or in the case of investments, the satisfaction that an investor derives from a portfolio.

Because individuals are different in their preferences, all risk-averse individuals may not rank investment alternatives in the same manner. Consider the £50 gamble again. All risk-averse individuals will rank the guaranteed outcome of £50 higher than the gamble. What if the guaranteed outcome is only £40? Some risk-averse investors might consider £40 inadequate, others might accept it, and still others may now be indifferent about the uncertain £50 and the certain £40.

A simple implementation of utility theory allows us to quantify the rankings of investment choices using risk and return. There are several assumptions about individual behavior that we make in the definition of utility given in the equation below.

We assume that investors are risk averse. They always prefer more to less (greater return to lesser return). They are able to rank different portfolios in the order of their preference and that the rankings are internally consistent. If an individual prefers X to Y and Y to Z, then he/she must prefer X to Z. This property implies that the indifference curves (see Exhibit 13) for the same individual can never touch or intersect. An example of a utility function is given below

$$U = E(r) - \frac{1}{2}A\sigma^2$$

where, U is the utility of an investment, $E(r)$ is the expected return, and σ^2 is the variance of the investment.

In the above equation, A is a measure of risk aversion, which is measured as the marginal reward that an investor requires to accept additional risk. More risk-averse investors require greater compensation for accepting additional risk. Thus, A is higher for more risk-averse individuals. As was mentioned previously, a risk-neutral investor would maximize return irrespective of risk and a risk-seeking investor would maximize both risk and return.

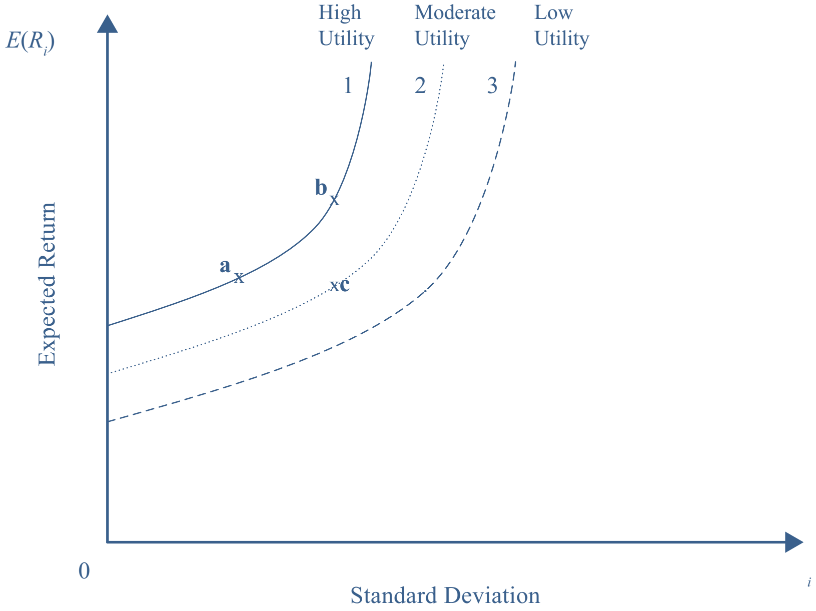
We can draw several conclusions from the utility function. First, utility is unbounded on both sides. It can be highly positive or highly negative. Second, higher return contributes to higher utility. Third, higher variance reduces the utility but the reduction in utility gets amplified by the risk aversion coefficient, A. Utility can always be increased, albeit marginally, by getting higher return or lower risk. Fourth, utility does not indicate or measure satisfaction itself—it can be useful only in ranking various investments. For example, a portfolio with a utility of 4 is not necessarily two times better than a portfolio with a utility of 2. The portfolio with a utility of 4 could increase our happiness 10 times or just marginally. But we do prefer a portfolio with a utility of 4 to a portfolio with a utility of 2. Utility cannot be compared among individuals or investors because it is a very personal concept. From a societal point of view, by the same argument, utility cannot be summed among individuals.

Let us explore the utility function further. The risk aversion coefficient, A, is greater than zero for a risk-averse investor. So any increase in risk reduces his/her utility. The risk aversion coefficient for a risk-neutral investor is 0, and changes in risk do not affect his/her utility. For a risk lover, the risk aversion coefficient is negative, creating an inverse situation so that additional risk contributes to an increase in his/her utility. Note that a risk-free asset ($\sigma^2 = 0$) generates the same utility for all individuals.

10.1 Indifference Curves

An **indifference curve** plots the combinations of risk–return pairs that an investor would accept to maintain a given level of utility (i.e., the investor is indifferent about the combinations on any one curve because they would provide the same level of overall utility). Indifference curves are thus defined in terms of a trade-off between expected rate of return and variance of the rate of return. Because an infinite number of combinations of risk and return can generate the same utility for the same investor, indifference curves are continuous at all points.

Exhibit 13 Indifference Curves for Risk-Averse Investors



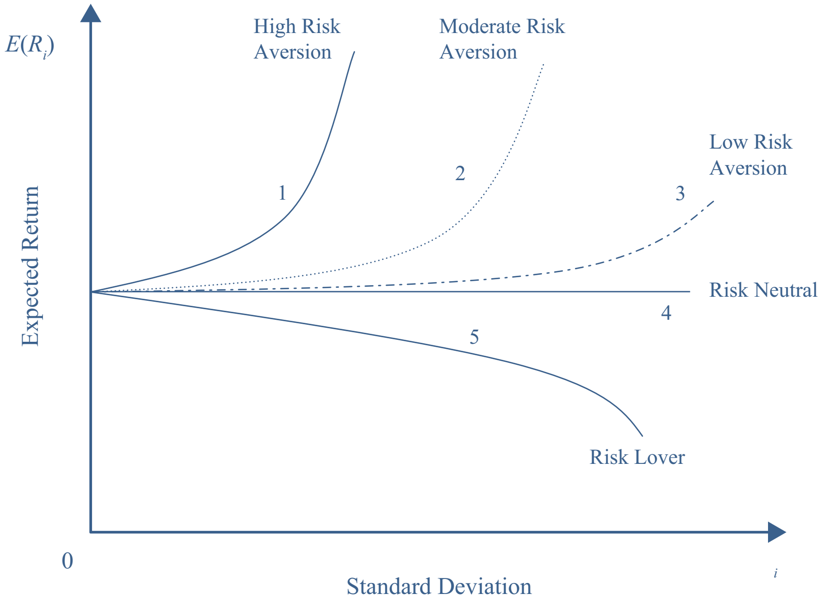
A set of indifference curves is plotted in Exhibit 13. By definition, all points on any one of the three curves have the same utility. An investor does not care whether he/she is at Point **a** or Point **b** on indifference Curve 1. Point **a** has lower risk and lower return than Point **b**, but the utility of both points is the same because the higher return at Point **b** is offset by the higher risk.

Like Curve 1, all points on Curve 2 have the same utility and an investor is indifferent about where he/she is on Curve 2. Now compare Point **c** with Point **b**. Point **c** has the same risk but significantly lower return than Point **b**, which means that the utility at Point **c** is less than the utility at Point **b**. Given that all points on Curve 1 have the same utility and all points on Curve 2 have the same utility and Point **b** has higher utility than Point **c**, Curve 1 has higher utility than Curve 2. Therefore, a risk-averse investor with indifference Curves 1 and 2 will prefer Curve 1 to Curve 2. The utility of a risk-averse investor always increases as you move northwest—higher return with lower risk. Because all investors prefer more utility to less, investors want to move northwest to the indifference curve with the highest utility.

The indifference curve for risk-averse investors runs from the southwest to the northeast because of the risk–return trade-off. If risk increases (going east) then it must be compensated by higher return (going north) to generate the same utility. The indifference curves are convex because of diminishing marginal utility of return (or wealth). As risk increases, an investor needs greater return to compensate for higher risk at an increasing rate (i.e., the curve gets steeper). The upward-sloping convex indifference curve has a slope coefficient closely related to the risk aversion coefficient. The greater the slope, the higher is the risk aversion of the investor as a greater increment in return is required to accept a given increase in risk.

Indifference curves for investors with different levels of risk aversion are plotted in Exhibit 14. The most risk-averse investor has an indifference curve with the greatest slope. As volatility increases, this investor demands increasingly higher returns to compensate for risk. The least risk-averse investor has an indifference curve with the least slope and so the demand for higher return as risk increases is not as acute as for the more risk-averse investor. The risk-loving investor's indifference curve, however, exhibits a negative slope, implying that the risk-lover is happy to substitute risk for return. For a risk lover, the utility increases both with higher risk and higher return. Finally, the indifference curves of risk-neutral investors are horizontal because the utility is invariant with risk.

Exhibit 14 Indifference Curves for Various Types of Investors



In the remaining parts of this reading, all investors are assumed to be risk averse unless stated otherwise.

EXAMPLE 6

Comparing a Gamble with a Guaranteed Outcome

Assume that you are given an investment with an expected return of 10 percent and a risk (standard deviation) of 20 percent, and your risk aversion coefficient is 3.

1 What is your utility of this investment?

2 What must be the minimum risk-free return you should earn to get the same utility?

Solution to 1:

$$U = 0.10 - 0.5 \times 3 \times 0.20^2 = 0.04.$$

Solution to 2:

A risk-free return's σ is zero, so the second term disappears. To get the same utility (0.04), the risk-free return must be at least 4 percent. Thus, in your mind, a risky return of 10 percent is equivalent to a risk-free return or a guaranteed outcome of 4 percent.

EXAMPLE 7

Computation of Utility

Based on investment information given below and the utility formula $U = E(r) - 0.5A\sigma^2$, answer the following questions. Returns and standard deviations are both expressed as percent per year. When using the utility formula, however, returns and standard deviations must be expressed in decimals.

Investment	Expected Return $E(r)$	Standard Deviation σ
1	12%	30%
2	15	35
3	21	40
4	24	45

1. Which investment will a risk-averse investor with a risk aversion coefficient of 4 choose?
2. Which investment will a risk-averse investor with a risk aversion coefficient of 2 choose?
3. Which investment will a risk-neutral investor choose?
4. Which investment will a risk-loving investor choose?

Solutions to 1 and 2:

The utility for risk-averse investors with $A = 4$ and $A = 2$ for each of the four investments are shown in the following table. Complete calculations for Investment 1 with $A = 4$ are as follows: $U = 0.12 - 0.5 \times 4 \times 0.30^2 = -0.06$.

Investment	Expected Return $E(r)$	Standard Deviation σ	Utility $A = 4$	Utility $A = 2$
1	12%	30%	−0.0600	0.0300
2	15	35	−0.0950	0.0275
3	21	40	−0.1100	0.0500
4	24	45	−0.1650	0.0375

The risk-averse investor with a risk aversion coefficient of 4 should choose Investment 1. The risk-averse investor with a risk aversion coefficient of 2 should choose Investment 3.

Solution to 3:

A risk-neutral investor cares only about return. In other words, his risk aversion coefficient is 0. Therefore, a risk-neutral investor will choose Investment 4 because it has the highest return.

Solution to 4:

A risk-loving investor likes both higher risk and higher return. In other words, his risk aversion coefficient is negative. Therefore, a risk-loving investor will choose Investment 4 because it has the highest return and highest risk among the four investments.

APPLICATION OF UTILITY THEORY TO PORTFOLIO SELECTION

e explain risk aversion and its implications for portfolio selection;

i explain the selection of an optimal portfolio, given an investor's utility (or risk aversion) and the capital allocation line.

The simplest application of utility theory and risk aversion is to a portfolio of two assets, a risk-free asset and a risky asset. The risk-free asset has zero risk and a return of R_f. The risky asset has a risk of σ_i (> 0) and an expected return of $E(R_i)$. Because the risky asset has risk that is greater than that of the risk-free asset, the expected return from the risky asset will be greater than the return from the risk-free asset, that is, $E(R_i) > R_f$.

We can construct a portfolio of these two assets with a portfolio expected return, $E(R_p)$, and portfolio risk, σ_p, based on the formulas provided below. In the equations given below, w_1 is the weight in the risk-free asset and $(1 - w_1)$ is the weight in the risky asset. Because $\sigma_f = 0$ for the risk-free asset, the first and third terms in the formula for variance are zero leaving only the second term. We arrive at the last equation by taking the square root of both sides, which shows the expression for standard deviation for a portfolio of two assets when one asset is the risk-free asset:

$$E\left(R_p\right) = w_1 R_f + \left(1 - w_1\right) E\left(R_i\right)$$

$$\sigma_P^2 = w_1^2 \sigma_f^2 + \left(1 - w_1\right)^2 \sigma_i^2 + 2 w_1 \left(1 - w_1\right) \rho_{12} \sigma_f \sigma_i = \left(1 - w_1\right)^2 \sigma_i^2$$

$$\sigma_p = \left(1 - w_1\right) \sigma_i$$

The two-asset portfolio is drawn in Exhibit 15 by varying w_1 from 0 percent to 100 percent. The portfolio standard deviation is on the horizontal axis and the portfolio return is on the vertical axis. If only these two assets are available in the economy and the risky asset represents the market, the line in Exhibit 15 is called the **capital allocation line**. The capital allocation line represents the portfolios available to an investor. The equation for this line can be derived from the above two equations by rewriting the second equation as $w_1 = 1 - \frac{\sigma_p}{\sigma_i}$. Substituting the value of w_1 in the equation for expected return, we get the following equation for the capital allocation line:

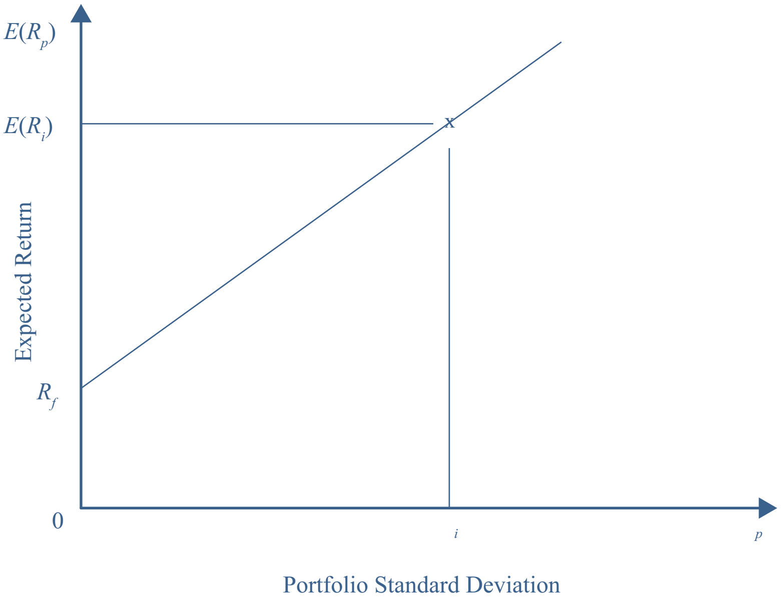
$$E\left(R_p\right) = \left(1 - \frac{\sigma_p}{\sigma_i}\right) R_f + \frac{\sigma_p}{\sigma_i} E\left(R_i\right)$$

This equation can be rewritten in a more usable form:

$$E\left(R_p\right) = R_f + \frac{\left(E\left(R_i\right) - R_f\right)}{\sigma_i} \sigma_p$$

The capital allocation line has an intercept of R_f, and a slope of $\frac{\left(E\left(R_i\right) - R_f\right)}{\sigma_i}$, which is the additional required return for every increment in risk, and is sometimes referred to as the market price of risk.

Exhibit 15 Capital Allocation Line with Two Assets

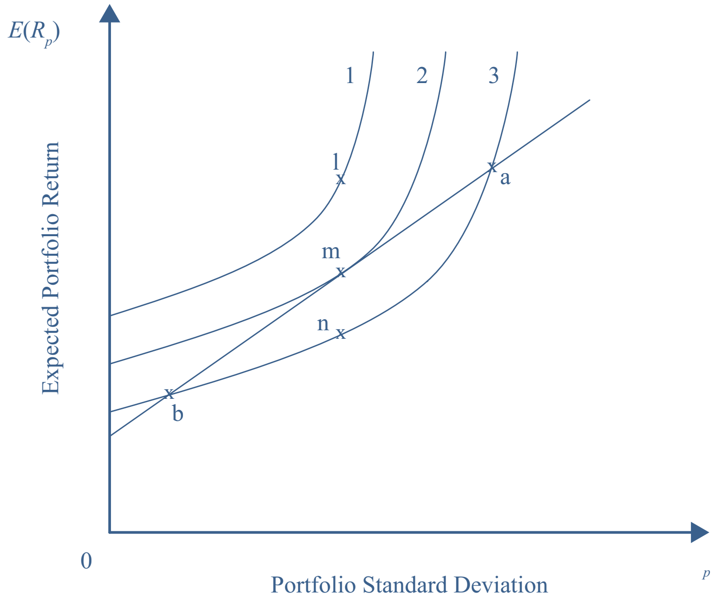


Because the equation is linear, the plot of the capital allocation line is a straight line. The line begins with the risk-free asset as the leftmost point with zero risk and a risk-free return, R_f. At that point, the portfolio consists of only the risk-free asset. If 100 percent is invested in the portfolio of all risky assets, however, we have a return of $E(R_i)$ with a risk of σ_i.

We can move further along the line in pursuit of higher returns by borrowing at the risk-free rate and investing the borrowed money in the portfolio of all risky assets. If 50 percent is borrowed at the risk-free rate, then $w_1 = -0.50$ and 150 percent is placed in the risky asset, giving a return $= 1.50E(R_i) - 0.50R_f$, which is $> E(R_i)$ because $E(R_i) > R_f$.

The line plotted in Exhibit 15 is comprised of an unlimited number of risk–return pairs or portfolios. Which *one* of these portfolios should be chosen by an investor? The answer lies in combining indifference curves from utility theory with the capital allocation line from portfolio theory. Utility theory gives us the utility function or the indifference curves for an individual, as in Exhibit 13, and the capital allocation line gives us the set of feasible investments. Overlaying each individual's indifference curves on the capital allocation line will provide us with the optimal portfolio for that investor. Exhibit 16 illustrates this process of portfolio selection.

Exhibit 16 Portfolio Selection



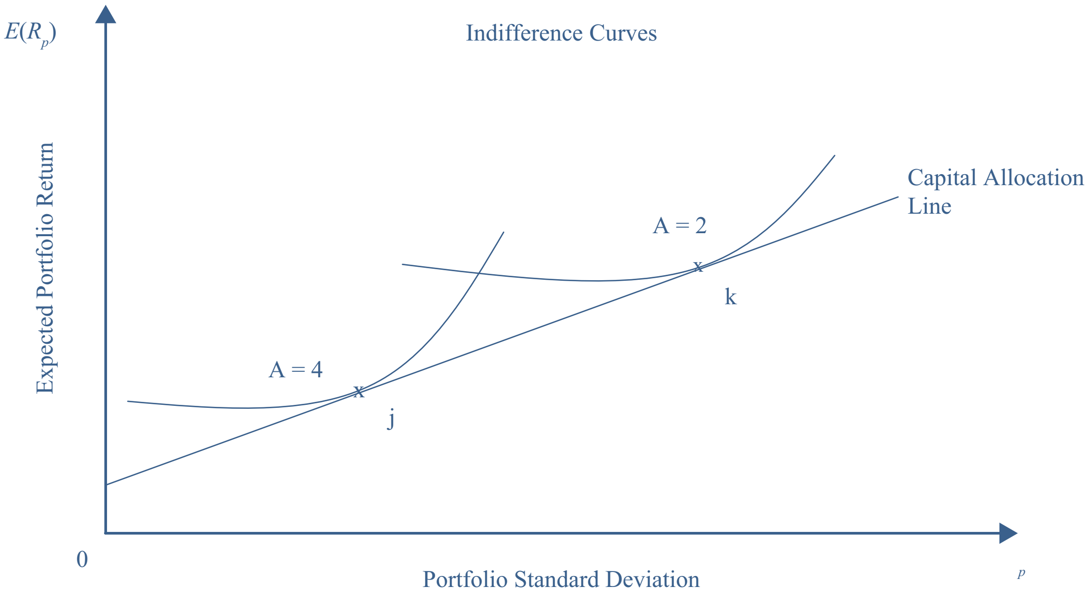
The capital allocation line consists of the set of feasible portfolios. Points under the capital allocation line may be attainable but are not preferred by any investor because the investor can get a higher return for the same risk by moving up to the capital allocation line. Points above the capital allocation line are desirable but not achievable with available assets.

Three indifference curves for the same individual are also shown in Exhibit 16. Curve 1 is above the capital allocation line, Curve 2 is tangential to the line, and Curve 3 intersects the line at two points. Curve 1 has the highest utility and Curve 3 has the lowest utility. Because Curve 1 lies completely above the capital allocation line, points on Curve 1 are not achievable with the available assets on the capital allocation line. Curve 3 intersects the capital allocation line at two Points, **a** and **b**. The investor is able to invest at either Point **a** or **b** to derive the risk–return trade-off and utility associated with Curve 3. Comparing points with the same risk, observe that Point **n** on Curve 3 has the same risk as Point **m** on Curve 2, yet Point **m** has the higher expected return. Therefore, all investors will choose Curve 2 instead of Curve 3. Curve 2 is tangential to the capital allocation line at Point **m**. Point **m** is on the capital allocation line and investable. Point **m** and the utility associated with Curve 2 is the best that the investor can do because he/she cannot move to a higher utility indifference curve. Thus, we have been able to select the optimal portfolio for the investor with indifference Curves 1, 2, and 3. Point **m**, the optimal portfolio for one investor, may not be optimal for another investor. We can follow the same process, however, for finding the optimal portfolio for other investors: the optimal portfolio is the point of tangency between the capital allocation line and the indifference curve for that investor. In other words, the optimal portfolio maximizes the return per unit of risk (as it is on the capital allocation line), and it simultaneously supplies the investor with the most satisfaction (utility).

As an illustration, Exhibit 17 shows two indifference curves for two different investors: Kelly with a risk aversion coefficient of 2 and Jane with a risk aversion coefficient of 4. The indifference curve for Kelly is to the right of the indifference curve for Jane because Kelly is less risk averse than Jane and can accept a higher amount of risk, i.e. has a higher tolerance for risk. Accordingly, their optimal portfolios are different: Point **k** is the optimal portfolio for Kelly and Point **j** is the optimal portfolio for Jane.

In addition, for the same return, the slope of Jane's curve is higher than Kelly's suggesting that Jane needs greater incremental return as compensation for accepting an additional amount of risk compared with Kelly.

Exhibit 17 Portfolio Selection for Two Investors with Various Levels of Risk Aversion



12 PORTFOLIO RISK & PORTFOLIO OF TWO RISKY ASSETS

d calculate and interpret the mean, variance, and covariance (or correlation) of asset returns based on historical data;

f calculate and interpret portfolio standard deviation;

g describe the effect on a portfolio's risk of investing in assets that are less than perfectly correlated;

We have seen before that investors are risk averse and demand a higher return for a riskier investment. Therefore, ways of controlling portfolio risk without affecting return are valuable. As a precursor to managing risk, this section explains and analyzes the components of portfolio risk. In particular, it examines and describes how a portfolio consisting of assets with low correlations have the potential of reducing risk without necessarily reducing return.

12.1 Portfolio of Two Risky Assets

The return and risk of a portfolio of two assets was introduced in Sections 2–8 of this reading. In this section, we briefly review the computation of return and extend the concept of portfolio risk and its components.

12.1.1 *Portfolio Return*

When several individual assets are combined into a portfolio, we can compute the portfolio return as a weighted average of the returns in the portfolio. The portfolio return is simply a weighted average of the returns of the individual investments, or assets. If Asset 1 has a return of 20 percent and constitutes 25 percent of the portfolio's investment, then the contribution to the portfolio return is 5 percent (= 25% of 20%). In general, if Asset i has a return of R_i and has a weight of w_i in the portfolio, then the portfolio return, R_P, is given as:

$$R_P = \sum_{i=1}^{N} w_i R_i, \quad \sum_{i=1}^{N} w_i = 1$$

Note that the weights must add up to 1 because the assets in a portfolio, including cash, must account for 100 percent of the investment. Also, note that these are single period returns, so there are no cash flows during the period and the weights remain constant.

When two individual assets are combined in a portfolio, we can compute the portfolio return as a weighted average of the returns of the two assets. Consider Assets 1 and 2 with weights of 25 percent and 75 percent in a portfolio. If their returns are 20 percent and 5 percent, the weighted average return = (0.25 × 20%) + (0.75 × 5%) = 8.75%. More generally, the portfolio return can be written as below, where R_p is return of the portfolio, w_1 and w_2 are the weights of the two assets, and R_1, R_2 are returns on the two assets:

$$R_p = w_1 R_1 + (1 - w_1) R_2$$

12.1.2 *Portfolio Risk*

Like a portfolio's return, we can calculate a portfolio's variance. Although the return of a portfolio is simply a weighted average of the returns of each security, this is not the case with the standard deviation of a portfolio (unless all securities are perfectly correlated—that is, correlation equals one). Variance can be expressed more generally for N securities in a portfolio using the notation from the portfolio return calculation above:

$$\sum_{i=1}^{N} w_i = 1$$

$$\sigma_P^2 = \text{Var}(R_P) = \text{Var}\left(\sum_{i=1}^{N} w_i R_i\right)$$

Note that the weights must add up to 1. The right side of the equation is the variance of the weighted average returns of individual securities. Weight is a constant, but the returns are variables whose variance is shown by Var(R_i). We can rewrite the equation as shown next. Because the covariance of an asset with itself is the variance of the asset, we can separate the variances from the covariances in the second equation:

$$\sigma_P^2 = \sum_{i,j=1}^{N} w_i w_j \text{Cov}(R_i, R_j)$$

$$\sigma_P^2 = \sum_{i=1}^{N} w_i^2 \text{Var}(R_i) + \sum_{i,j=1, i \neq j}^{N} w_i w_j \text{Cov}(R_i, R_j)$$

Cov(R_i,R_j) is the covariance of returns, R_i and R_j, and can be expressed as the product of the correlation between the two returns ($\rho_{1,2}$) and the standard deviations of the two assets. Thus, Cov(R_i,R_j) = $\rho_{ij}\sigma_i\sigma_j$.

For a two asset portfolio, the expression for portfolio variance simplifies to the following using covariance and then using correlation:

$$\sigma_P^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\text{Cov}(R_1, R_2)$$

$$\sigma_P^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{12}\sigma_1\sigma_2$$

The standard deviation of a two asset portfolio is given by the square root of the portfolio's variance:

$$\sigma_P = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\text{Cov}(R_1, R_2)}$$

or,

$$\sigma_P = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{12}\sigma_1\sigma_2}$$

EXAMPLE 8

Return and Risk of a Two-Asset Portfolio

Assume that as a US investor, you decide to hold a portfolio with 80 percent invested in the S&P 500 US stock index and the remaining 20 percent in the MSCI Emerging Markets index. The expected return is 9.93 percent for the S&P 500 and 18.20 percent for the Emerging Markets index. The risk (standard deviation) is 16.21 percent for the S&P 500 and 33.11 percent for the Emerging Markets index. What will be the portfolio's expected return and risk given that the covariance between the S&P 500 and the Emerging Markets index is 0.5 percent or 0.0050? Note that units for covariance and variance are written as $\%^2$ when not expressed as a fraction. These are units of measure like squared feet and the numbers themselves are not actually squared.

Solution:

Portfolio return, $R_P = w_1R_1 + (1 - w_1), R_2 = (0.80 \times 0.0993) + (0.20 \times 0.1820) = 0.1158 = 11.58\%$.

$$\text{Portfolio risk} = \sigma_P = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2Cov(R_1, R_2)}$$

$$\sigma_p^2 = w_{US}^2\sigma_{US}^2 + w_{EM}^2\sigma_{EM}^2 + 2w_{US}w_{EM}Cov_{US,EM}$$

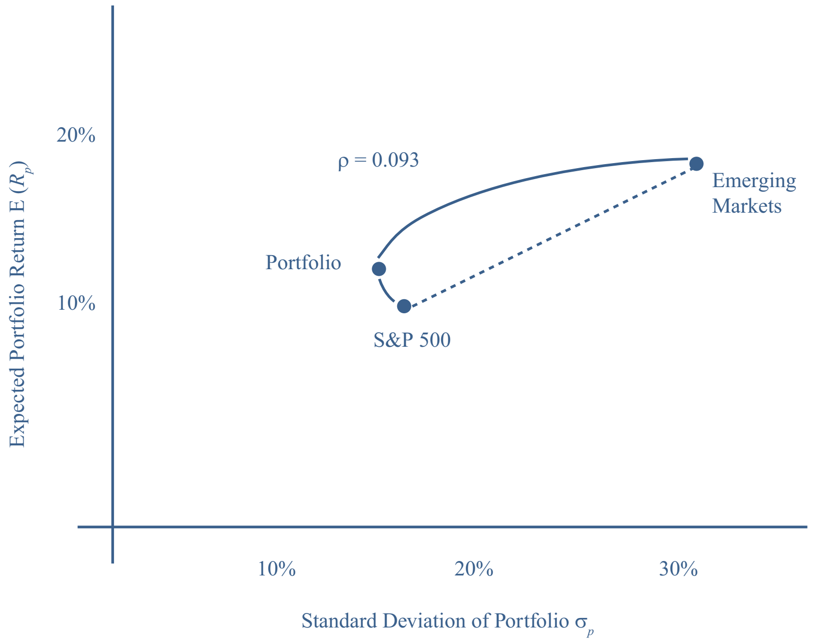
$$\sigma_p^2 = \left(0.80^2 \times 0.1621^2\right) + \left(0.20^2 \times 0.3311^2\right) + \left(2 \times 0.80 \times 0.20 \times 0.0050\right)$$

$$\sigma_p^2 = 0.01682 + 0.00439 + 0.00160 = 0.02281$$

$$\sigma_p = 0.15103 = 15.10\%$$

The portfolio's expected return is 11.58 percent and the portfolio's risk is 15.10 percent. Look at this example closely. It shows that we can take the portfolio of a US investor invested only in the S&P 500, combine it with a *riskier* portfolio consisting of emerging markets securities, and the return of the US investor increases from 9.93 percent to 11.58 percent while the risk of the portfolio actually falls from 16.21 percent to 15.10 percent. Exhibit 18 depicts how the combination of the two assets results in a superior risk–return trade-off. Not only does the investor get a higher return, but he also gets it at a lower risk. That is the power of diversification as you will see later in this reading.

Exhibit 18 Combination of Two Assets



12.1.3 *Covariance and Correlation*

The **covariance** in the formula for portfolio standard deviation can be expanded as $\text{Cov}(R_1,R_2) = \rho_{12}\sigma_1\sigma_2$ where ρ_{12} is the correlation between returns, R_1, R_2. Although covariance is important, it is difficult to interpret because it is unbounded on both sides. It is easier to understand the **correlation coefficient** (ρ_{12}), which is bounded but provides similar information.

Correlation is a measure of the consistency or tendency for two investments to act in a similar way. The correlation coefficient, ρ_{12}, can be positive or negative and ranges from −1 to +1. Consider three different values of the correlation coefficient:

- $\rho_{12} = +1$: Returns of the two assets are perfectly *positively* correlated. Assets 1 and 2 move together 100 percent of the time.
- $\rho_{12} = -1$: Returns of the two assets are perfectly *negatively* correlated. Assets 1 and 2 move in opposite directions 100 percent of the time.
- $\rho_{12} = 0$: Returns of the two assets are *uncorrelated*. Movement of Asset 1 provides no prediction regarding the movement of Asset 2.

The correlation coefficient between two assets determines the effect on portfolio risk when the two assets are combined. To see how this works, consider two different values of ρ_{12}. You will find that portfolio risk is unaffected when the two assets are perfectly correlated ($\rho_{12} = +1$). In other words, the portfolio's standard deviation is simply a weighted average of the standard deviations of the two assets and as such a portfolio's risk is unchanged with the addition of assets with the same risk parameters. Portfolio risk falls, however, when the two assets are not perfectly correlated ($\rho_{12} <$ +1). Sufficiently low values of the correlation coefficient can make the portfolio riskless under certain conditions.

First, let $\rho_{12} = +1$

$$\begin{aligned}\sigma_p^2 &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{12}\sigma_1\sigma_2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2\\ &= \left(w_1\sigma_1 + w_2\sigma_2\right)^2\\ \sigma_p &= w_1\sigma_1 + w_2\sigma_2\end{aligned}$$

The first set of terms on the right side of the first equation contain the usual terms for portfolio variance. Because the correlation coefficient is equal to +1, the right side can be rewritten as a perfect square. The third row shows that portfolio risk is a weighted average of the risks of the individual assets' risks. We showed earlier that the portfolio return is a weighted average of the assets' returns. Because both risk and return are just weighted averages of the two assets in the portfolio there is no reduction in risk when $\rho_{12} = +1$.

Now let $\rho_{12} < +1$

The above analysis showed that portfolio risk is a weighted average of asset risks when $\rho_{12} = +1$. When $\rho_{12} < +1$, the portfolio risk is less than the weighted average of the individual assets' risks.

To show this, we begin by reproducing the general formula for portfolio risk, which is expressed by the terms to the left of the "<" sign below. The term to the right of "<" shows the portfolio risk when $\rho_{12} = +1$:

$$\begin{aligned}\sigma_p &= \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{12}\sigma_1\sigma_2} < \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2}\\ &= \left(w_1\sigma_1 + w_2\sigma_2\right)\\ \sigma_p &< \left(w_1\sigma_1 + w_2\sigma_2\right)\end{aligned}$$

The left side is smaller than the right side because the correlation coefficient on the left side for the new portfolio is <1. Thus, the portfolio risk is less than the weighted average of risks while the portfolio return is still a weighted average of returns.

As you can see, we have achieved diversification by combining two assets that are not perfectly correlated. For an extreme case in which $\rho_{12} = -1$ (that is, the two asset returns move in opposite directions), the portfolio can be made risk free.

EXAMPLE 9

Effect of Correlation on Portfolio Risk

Two stocks have the same return and risk (standard deviation): 10 percent return with 20 percent risk. You form a portfolio with 50 percent each of Stock 1 and Stock 2 to examine the effect of correlation on risk.

1 Calculate the portfolio return and risk if the correlation is 1.0.
2 Calculate the portfolio return and risk if the correlation is 0.0.
3 Calculate the portfolio return and risk if the correlation is –1.0.
4 Compare the return and risk of portfolios with different correlations.

Solution to 1:

$$R_1 = R_2 = 10\% = 0.10;\ \sigma_1 = \sigma_2 = 20\% = 0.20;\ w_1 = w_2 = 50\%$$
$$= 0.50.\ \text{Case 1: } \rho_{12} = +1$$
$$R_p = w_1 R_1 + w_2 R_2$$
$$R_p = (0.5 \times 0.1) + (0.5 \times 0.1) = 0.10 = 10\%$$
$$\sigma_p^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2\rho_{12}$$
$$\sigma_p^2 = \left(0.5^2 \times 0.2^2\right) + \left(0.5^2 \times 0.2^2\right) + (2 \times 0.5 \times 0.5 \times 0.2 \times 0.2 \times 1) = 0.04$$
$$\sigma_p = \sqrt{0.04} = 0.20 = 20\%$$

This equation demonstrates the earlier point that with a correlation of 1.0 the risk of the portfolio is the same as the risk of the individual assets.

Solution to 2:

$$\rho_{12} = 0$$
$$R_p = w_1 R_1 + w_2 R_2 = 0.10 = 10\%$$
$$\sigma_p^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2\rho_{12}$$
$$\sigma_p^2 = \left(0.5^2 \times 0.2^2\right) + \left(0.5^2 \times 0.2^2\right)$$
$$+ (2 \times 0.5 \times 0.5 \times 0.2 \times 0.2 \times 0) = 0.02$$
$$\sigma_p = \sqrt{0.02} = 0.14 = 14\%$$

This equation demonstrates the earlier point that, when assets have correlations of less than 1.0, they can be combined in a portfolio that has less risk than either of the assets individually.

Solution to 3:

$$\rho_{12} = -1$$
$$R_p = w_1 R_1 + w_2 R_2 = 0.10 = 10\%$$
$$\sigma_p^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2\rho_{12}$$
$$\sigma_p^2 = \left(0.5^2 \times 0.2^2\right) + \left(0.5^2 \times 0.2^2\right)$$
$$+ (2 \times 0.5 \times 0.5 \times 0.2 \times 0.2 \times -1) = 0$$
$$\sigma_p = 0\%$$

This equation demonstrates the earlier point that, if the correlation of assets is low enough, in this case 100 percent negative correlation or –1.00 (exactly inversely related), a portfolio can be designed that eliminates risk. The individual assets retain their risk characteristics, but the portfolio is risk free.

Solution to 4:

The expected return is 10 percent in all three cases; however, the returns will be more volatile in Case 1 and least volatile in Case 3. In the first case, there is no diversification of risk (same risk as before of 20 percent) and the return remains the same. In the second case, with a correlation coefficient of 0, we have achieved diversification of risk (risk is now 14 percent instead of 20 percent), again with the same return. In the third case with a correlation coefficient of

–1, the portfolio is risk free, although we continue to get the same return of 10 percent. This example shows the power of diversification that we expand on further in Section 14.

12.1.4 *Relationship between Portfolio Risk and Return*

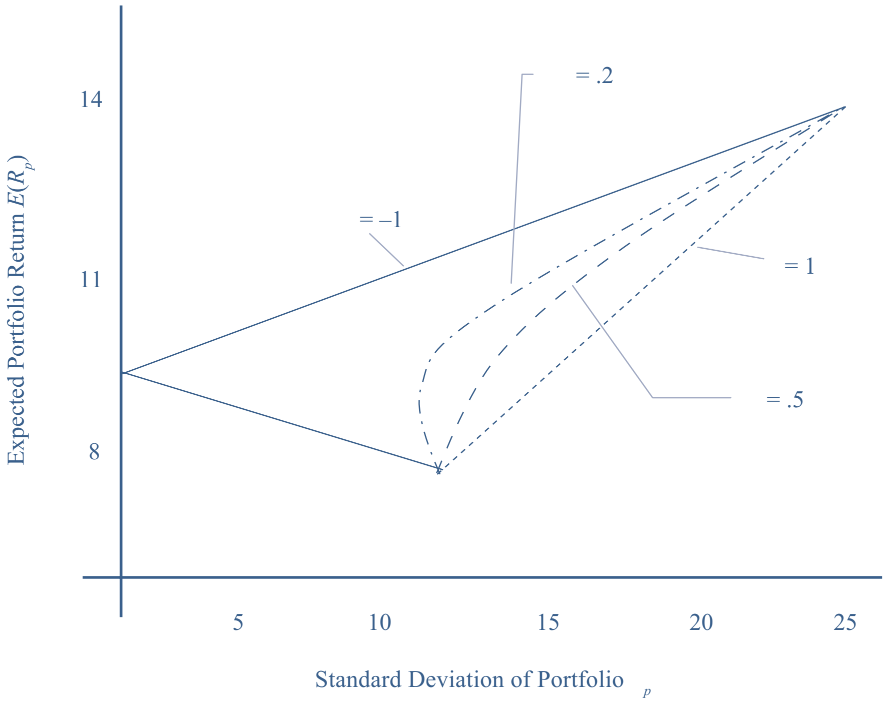
The previous example illustrated the effect of correlation on portfolio risk while keeping the weights in the two assets equal and unchanged. In this section, we consider how portfolio risk and return vary with different portfolio weights and different correlations.

Asset 1 has an annual return of 7 percent and annualized risk of 12 percent, whereas Asset 2 has an annual return of 15 percent and annualized risk of 25 percent. The relationship is tabulated in Exhibit 19 for the two assets and graphically represented in Exhibit 20.

Exhibit 19 Relationship between Risk and Return

Weight in Asset 1 (%)	Portfolio Return	Portfolio Risk with Correlation of			
		1.0	0.5	0.2	−1.0
0	15.0	25.0	25.0	25.0	25.0
10	14.2	23.7	23.1	22.8	21.3
20	13.4	22.4	21.3	20.6	17.6
30	12.6	21.1	19.6	18.6	13.9
40	11.8	19.8	17.9	16.6	10.2
50	11.0	18.5	16.3	14.9	6.5
60	10.2	17.2	15.0	13.4	2.8
70	9.4	15.9	13.8	12.3	0.9
80	8.6	14.6	12.9	11.7	4.6
90	7.8	13.3	12.2	11.6	8.3
100	7.0	12.0	12.0	12.0	12.0

Exhibit 20 Relationship between Risk and Return



The table shows the portfolio return and risk for four correlation coefficients ranging from +1.0 to –1.0 and 11 weights ranging from 0 percent to 100 percent. The portfolio return and risk are 15 percent and 25 percent, respectively, when 0 percent is invested in Asset 1, versus 7 percent and 12 percent when 100 percent is invested in Asset 1. The portfolio return varies with weights but is unaffected by the correlation coefficient.

Portfolio risk becomes smaller with each successive decrease in the correlation coefficient, with the smallest risk when $\rho_{12} = -1$. The graph in Exhibit 20 shows that the risk–return relationship is a straight line when $\rho_{12} = +1$. As the correlation falls, the risk becomes smaller and smaller as in the table. The curvilinear nature of a portfolio of assets is recognizable in all investment opportunity sets (except at the extremes where $\rho_{12} = -1$ or $+1$).

EXAMPLE 10

Portfolio of Two Assets

An investor is considering investing in a small-cap stock fund and a general bond fund. Their returns and standard deviations are given below and the correlation between the two fund returns is 0.10.

	Expected Annual Return (%)	Standard Deviation of Returns (%)
Small-cap fund, S	19	33
Bond fund, B	8	13

1 If the investor requires a portfolio return of 12 percent, what should the proportions in each fund be?

2 What is the standard deviation of the portfolio constructed in Part 1?

Solution to 1:

We can calculate the weights by setting the portfolio return equal to 12 percent. $12\% = w_1 \times 19\% + (1 - w_1) \times 8\%$; $w_1 = 36.4\%$, $(1 - w_1) = 63.6\%$. Thus, 36.4 percent should be invested in the small-cap fund and 63.6 percent should be invested in the bond fund.

Solution to 2:

$$\begin{aligned}\sigma_p &= \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{12}\sigma_1\sigma_2}\\ &= \sqrt{\left(0.364^2 \times 0.33^2\right) + \left(0.636^2 \times 0.13^2\right) + \left(2 \times 0.364 \times 0.636 \times 0.10 \times 0.33 \times 0.13\right)}\\ &= 15.23\%\end{aligned}$$

The portfolio risk is 15.23 percent, which is much less than a weighted average of risks of 20.28% (63.6% × 13% + 36.4% × 33%).

13 PORTFOLIO OF MANY RISKY ASSETS

f calculate and interpret portfolio standard deviation;

g describe the effect on a portfolio's risk of investing in assets that are less than perfectly correlated;

In the previous section, we discussed how the correlation between two assets can affect the risk of a portfolio and the smaller the correlation the lower is the risk. The above analysis can be extended to a portfolio with many risky assets (N). Recall the previous equations for portfolio return and variance:

$$E\left(R_p\right) = \sum_{i=1}^{N} w_i E\left(R_i\right),\ \sigma_P^2 = \left(\sum_{i=1}^{N} w_i^2\sigma_i^2 + \sum_{i,j=1,i\neq j}^{N} w_i w_j \text{Cov}(i,j)\right),\ \sum_{i=1}^{N} w_i = 1$$

To examine how a portfolio with many risky assets works and the ways in which we can reduce the risk of a portfolio, assume that the portfolio has equal weights ($1/N$) for all N assets. In addition, assume that $\bar{\sigma}^2$ and $\overline{\text{Cov}}$ are the average variance and average covariance. Given equal weights and average variance/covariance, we can rewrite the portfolio variance as below (intermediate steps are omitted to focus on the main result):

$$\sigma_P^2 = \left(\sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i,j=1, i \neq j}^{N} w_i w_j \text{Cov}(i,j)\right)$$

$$\sigma_P^2 = \frac{\bar{\sigma}^2}{N} + \frac{(N-1)}{N}\overline{\text{Cov}}$$

The equation in the second line shows that as N becomes large, the first term on the right side with the denominator of N becomes smaller and smaller, implying that the contribution of one asset's variance to portfolio variance gradually becomes negligible. The second term, however, approaches the average covariance as N increases. It is reasonable to say that for portfolios with a large number of assets, covariance among the assets accounts for almost all of the portfolio's risk.

13.1 Importance of Correlation in a Portfolio of Many Assets

The analysis becomes more instructive and interesting if we assume that all assets in the portfolio have the same variance and the same correlation among assets. In that case, the portfolio risk can then be rewritten as:

$$\sigma_p = \sqrt{\frac{\sigma^2}{N} + \frac{(N-1)}{N}\rho\sigma^2}$$

The first term under the root sign becomes negligible as the number of assets in the portfolio increases leaving the second term (correlation) as the main determining factor for portfolio risk. If the assets are unrelated to one another, the portfolio can have close to zero risk. In the next section, we review these concepts to learn how portfolios can be diversified.

THE POWER OF DIVERSIFICATION

14

c describe characteristics of the major asset classes that investors consider in forming portfolios;

g describe the effect on a portfolio's risk of investing in assets that are less than perfectly correlated;

Diversification is one of the most important and powerful concepts in investments. Because investors are risk averse, they are interested in reducing risk preferably without reducing return. In other cases, investors may accept a lower return if it will reduce the chance of catastrophic losses. In previous sections of this reading, you learned the importance of correlation and covariance in managing risk. This section applies those concepts to explore ways for risk diversification. We begin with a simple but intuitive example.

EXAMPLE 11

Diversification with Rain and Shine

Assume a company Beachwear rents beach equipment. The annual return from the company's operations is 20 percent in years with many sunny days but falls to 0 percent in rainy years with few sunny days. The probabilities of a sunny year and a rainy year are equal at 50 percent. Thus, the average return is 10 percent, with a 50 percent chance of 20 percent return and a 50 percent chance of 0 percent return. Because Beachwear can earn a return of 20 percent or 0 percent, its average return of 10 percent is risky.

You are excited about investing in Beachwear but do not like the risk. Having heard about diversification, you decide to add another business to the portfolio to reduce your investment risk.

- There is a snack shop on the beach that sells all the healthy food you like. You estimate that the annual return from the Snackshop is also 20 percent in years with many sunny days and 0 percent in other years. As with the Beachwear shop, the average return is 10 percent.

You decide to invest 50 percent each in Snackshop and Beachwear. The average return is still 10 percent, with 50 percent of 10 percent from Snackshop and 50 percent of 10 percent from Beachwear. In a sunny year, you would earn 20 percent (= 50% of 20% from Beachwear + 50% of 20% from Snackshop). In a rainy year, you would earn 0 percent (=50% of 0% from Beachwear + 50% of 0% from Snackshop). The results are tabulated in Exhibit 21.

Exhibit 21

Type	Company	Percent Invested	Return in Sunny Year (%)	Return in Rainy Year (%)	Average Return (%)
Single stock	Beachwear	100	20	0	10
Single stock	Snackshop	100	20	0	10
Portfolio of two stocks	Beachwear	50	20	0	10
	Snackshop	50	20	0	10
	Total	100	20	0	10

These results seem counterintuitive. You thought that by adding another business you would be able to diversify and reduce your risk, but the risk is exactly the same as before. What went wrong? Note that both businesses do well when it is sunny and both businesses do poorly when it rains. The correlation between the two businesses is +1.0. No reduction in risk occurs when the correlation is +1.0.

- To reduce risk, you must consider a business that does well in a rainy year. You find a company that rents DVDs. DVDrental company is similar to the Beachwear company, except that its annual return is 20 percent in a rainy year and 0 percent in a sunny year, with an average return of 10 percent. DVDrental's 10 percent return is also risky just like Beachwear's return.

If you invest 50 percent each in DVDrental and Beachwear, then the average return is still 10 percent, with 50 percent of 10 percent from DVDrental and 50 percent of 10 percent from Beachwear. In a sunny year, you would earn 10 percent (= 50% of 20% from Beachwear + 50% of 0% from DVDrental). In a rainy year also, you would earn 10 percent (=50% of 0% from Beachwear + 50% of 20% from DVDrental). You have no risk because you earn 10 percent in both sunny and rainy years. Thus, by adding DVDrental to Beachwear, you have reduced (eliminated) your risk without affecting your return. The results are tabulated in Exhibit 22.

Exhibit 22

Type	Company	Percent Invested	Return in Sunny Year (%)	Return in Rainy Year (%)	Average Return (%)
Single stock	Beachwear	100	20	0	10
Single stock	DVDrental	100	0	20	10
Portfolio of two stocks	Beachwear	50	20	0	10
	DVDrental	50	0	20	10
	Total	100	10	10	10

In this case, the two businesses have a correlation of –1.0. When two businesses with a correlation of –1.0 are combined, risk can always be reduced to zero.

14.1 Correlation and Risk Diversification

Correlation is the key in diversification of risk. Notice that the returns from Beachwear and DVDRental always go in the opposite direction. If one of them does well, the other does not. Therefore, adding assets that do not behave like other assets in your portfolio is good and can reduce risk. The two companies in the above example have a correlation of –1.0.

Even when we expand the portfolio to many assets, correlation among assets remains the primary determinant of portfolio risk. Lower correlations are associated with lower risk. Unfortunately, most assets have high positive correlations. The challenge in diversifying risk is to find assets that have a correlation that is much lower than +1.0.

14.2 Historical Risk and Correlation

When we previously discussed asset returns, we were careful to distinguish between historical or past returns and expected or future returns because historical returns may not be a good indicator of future returns. Returns may be highly positive in one period and highly negative in another period depending on the risk of that asset. Exhibit 7 showed that returns for large US company stocks were high in the 1990s but were very low in the 2000s.

Risk for an asset class, however, does not usually change dramatically from one period to the next. Stocks have been risky even in periods of low returns. T-bills are always less risky even when they earn high returns. From Exhibit 7, we can see that risk has typically not varied much from one decade to the next, except that risk for bonds has been much higher in recent decades when compared with earlier decades. Therefore, it is not unreasonable to assume that historical risk can work as a good proxy for future risk.

As with risk, correlations are quite stable among assets of the same country. Intercountry correlations, however, have been on the rise in the last few decades as a result of globalization and the liberalization of many economies. A correlation above 0.90 is considered high because the assets do not provide much opportunity for diversification of risk Low correlations—generally less than 0.50—are desirable for portfolio diversification.

14.3 Historical Correlation among Asset Classes

Correlations among major US asset classes and international stocks are reported in Exhibit 23 for 1970–2017. The highest correlation is between US large company stocks and US small company stocks at about 70 percent, whereas the correlation between US large company stocks and international stocks is approximately 66 percent. Although these are the highest correlations, they still provide diversification benefits because the correlations are less than 100 percent. The correlation between international stocks and US small company stocks is lower, at 50 percent. The lowest correlations are between stocks and bonds, with some correlations being negative, such as that between US small company stocks and US long-term government bonds. Similarly, the correlation between T-bills and stocks is close to zero and is negative for international stocks.[8]

Exhibit 23 Correlation Among US Assets and International Stocks (1970–2017)

Series	International Stocks	US Large Company Stocks	US Small Company Stocks	US Long-Term Corporate Bonds	US Long-Term Treasury Bonds	US T-Bills	US Inflation
International stocks	1.00						
US large company stocks	0.66	1.00					
US small company stocks	0.50	0.72	1.00				
US long-term corporate bonds	0.02	0.23	0.06	1.00			
US long-term Treasury bonds	–0.13	0.01	–0.15	0.89	1.00		
US T-bills	0.01	0.04	0.02	0.05	0.09	1.00	
US inflation	–0.06	–0.11	0.04	–0.32	–0.26	0.69	1.00

Source: 2018 SBBI Yearbook (Exhibit 12.13).

---

[8] In any short period, T-bills are riskless and uncorrelated with other asset classes. For example, a 3-month US Treasury bill is redeemable at its face value upon maturity irrespective of what happens to other assets. When we consider multiple periods, however, returns on T-bills may be related to other asset classes because short-term interest rates vary depending on the strength of the economy and outlook for inflation.

The low correlations between stocks and bonds are attractive for portfolio diversification. Similarly, including international securities in a portfolio can also control portfolio risk. It is not surprising that most diversified portfolios of investors contain domestic stocks, domestic bonds, foreign stocks, foreign bonds, real estate, cash, and other asset classes.

14.4 Avenues for Diversification

The reason for diversification is simple. By constructing a portfolio with assets that do not move together, you create a portfolio that reduces the ups and downs in the short term but continues to grow steadily in the long term. Diversification thus makes a portfolio more resilient to gyrations in financial markets.

We describe a number of approaches for diversification, some of which have been discussed previously and some of which might seem too obvious. Diversification, however, is such an important part of investing that it cannot be emphasized enough, especially when we continue to meet and see many investors who are not properly diversified.

- *Diversify with asset classes.* Correlations among major asset classes[9] are not usually high, as can be observed from the few US asset classes listed in Exhibit 23. Correlations for other asset classes and other countries are also typically low, which provides investors the opportunity to benefit from diversifying among many asset classes to achieve the biggest benefit from diversification. A partial list of asset classes includes domestic large caps, domestic small caps, growth stocks, value stocks, domestic corporate bonds, long-term domestic government bonds, domestic Treasury bills (cash), emerging market stocks, emerging market bonds, developed market stocks (i.e., developed markets excluding domestic market), developed market bonds, real estate, and gold and other commodities. In addition, industries and sectors are used to diversify portfolios. For example, energy stocks may not be well correlated with health care stocks. The exact proportions in which these assets should be included in a portfolio depend on the risk, return, and correlation characteristics of each and the home country of the investor.
- *Diversify with index funds.* Diversifying among asset classes can become costly for small portfolios because of the number of securities required. For example, creating diversified exposure to a single category, such as a domestic large company asset class, may require a group of at least 30 stocks. Exposure to 10 asset classes may require 300 securities, which can be expensive to trade and track. Instead, it may be effective to use exchange-traded funds or mutual funds that track the respective indexes, which could bring down the costs associated with building a well-diversified portfolio. Therefore, many investors should consider index mutual funds as an investment vehicle as opposed to individual securities.
- *Diversification among countries.* Countries are different because of industry focus, economic policy, and political climate. The US economy produces many financial and technical services and invests a significant amount in innovative research. The Chinese and Indian economies, however, are focused on manufacturing. Countries in the European Union are vibrant democracies whereas East Asian countries are experimenting with democracy. Thus, financial returns in one country over time are not likely to be highly correlated with returns in another country. Country returns may also be different because of different currencies. In other words, the return on a foreign investment may be different

---

[9] Major asset classes are distinguished from sub-classes, such as US value stocks and US growth stocks.

when translated to the home country's currency. Because currency returns are uncorrelated with stock returns, they may help reduce the risk of investing in a foreign country even when that country, in isolation, is a very risky emerging market from an equity investment point of view. Investment in foreign countries is an essential part of a well-diversified portfolio.

- *Diversify by* not *owning your employer's stock.* Companies encourage their employees to invest in company stock through employee stock plans and retirement plans. You should evaluate investing in your company, however, just as you would evaluate any other investment. In addition, you should consider the nonfinancial investments that you have made, especially the human capital you have invested in your company. Because you work for your employer, you are already heavily invested in it—your earnings depend on your employer. The level of your earnings, whether your compensation improves or whether you get a promotion, depends on how well your employer performs. If a competitor drives your employer out of the market, you will be out of a job. Additional investments in your employer will concentrate your wealth in one asset even more so and make you less diversified.
- *Evaluate each asset before adding to a portfolio.* Every time you add a security or an asset class to the portfolio, recognize that there is a cost associated with diversification. There is a cost of trading an asset as well as the cost of tracking a larger portfolio. In some cases, the securities or assets may have different names but belong to an asset class in which you already have sufficient exposure. A general rule to evaluate whether a new asset should be included to an existing portfolio is based on the following risk–return trade-off relationship:

$$E\left(R_{new}\right) = R_f + \frac{\sigma_{new}\rho_{new,p}}{\sigma_p} \times \left[E\left(R_p\right) - R_f\right]$$

  where $E(R)$ is the return from the asset, R_f is the return on the risk-free asset, σ is the standard deviation, ρ is the correlation coefficient, and the subscripts *new* and *p* refer to the new stock and existing portfolio. If the new asset's risk-adjusted return benefits the portfolio, then the asset should be included. The condition can be rewritten using the Sharpe ratio on both sides of the equation as:

$$\frac{E\left(R_{new}\right) - R_f}{\sigma_{new}} > \frac{E\left(R_p\right) - R_f}{\sigma_p} \times \rho_{new,p}$$

  If the Sharpe ratio of the new asset is greater than the Sharpe ratio of the current portfolio times the correlation coefficient, it is beneficial to add the new asset.
- *Buy insurance for risky portfolios.* It may come as a surprise, but insurance is an investment asset—just a different kind of asset. Insurance has a negative correlation with your assets and is thus very valuable. Insurance gives you a positive return when your assets lose value, but pays nothing if your assets maintain their value. Over time, insurance generates a negative average return. Many individuals, however, are willing to accept a small negative return because insurance reduces their exposure to an extreme loss. In general, it is reasonable to add an investment with a negative return if that investment significantly reduces risk (an example of a classic case of the risk–return trade-off).

Alternatively, investments with negative correlations also exist. Historically, gold has a negative correlation with stocks; however, the expected return is usually small and sometimes even negative. Investors often include gold and other commodities in their portfolios as a way of reducing their overall portfolio risk, including currency risk and inflation risk.

Buying put options is another way of reducing risk. Because put options pay when the underlying asset falls in value (negative correlation), they can protect an investor's portfolio against catastrophic losses. Of course, put options cost money, and the expected return is zero or marginally negative.

EFFICIENT FRONTIER: INVESTMENT OPPORTUNITY SET & MINIMUM VARIANCE PORTFOLIOS

15

g describe the effect on a portfolio's risk of investing in assets that are less than perfectly correlated;

h describe and interpret the minimum-variance and efficient frontiers of risky assets and the global minimum-variance portfolio;

In this section, we formalize the effect of diversification and expand the set of investments to include all available risky assets in a mean–variance framework. The addition of a risk-free asset generates an optimal risky portfolio and the capital allocation line. We can then derive an investor's optimal portfolio by overlaying the capital allocation line with the indifference curves of investors.

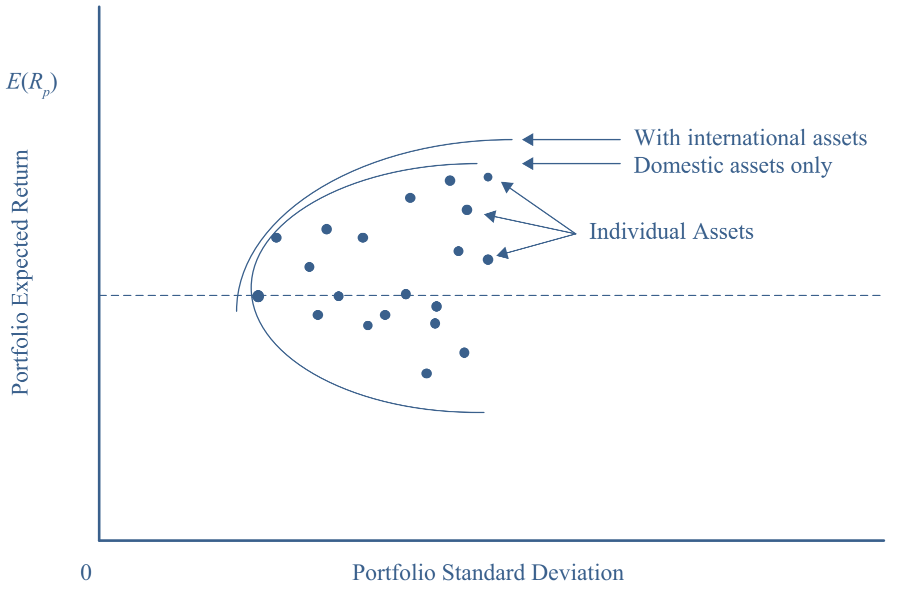
15.1 Investment Opportunity Set

If two assets are perfectly correlated, the risk–return opportunity set is represented by a straight line connecting those two assets. The line contains portfolios formed by changing the weight of each asset invested in the portfolio. This correlation was depicted by the straight line (with $\rho = 1$) in Exhibit 20. If the two assets are not perfectly correlated, the portfolio's risk is less than the weighted average risk of the components, and the portfolio formed from the two assets bulges on the left as shown by curves with the correlation coefficient (ρ) less than 1.0 in Exhibit 20. All of the points connecting the two assets are achievable (or feasible). The addition of new assets to this portfolio creates more and more portfolios that are either a linear combination of the existing portfolio and the new asset or a curvilinear combination, depending on the correlation between the existing portfolio and the new asset.

As the number of available assets increases, the number of possible combinations increases rapidly. When all investable assets are considered, and there are hundreds and thousands of them, we can construct an opportunity set of investments. The opportunity set will ordinarily span all points within a frontier because it is also possible to reach every possible point within that curve by judiciously creating a portfolio from the investable assets.

We begin with individual investable assets and gradually form portfolios that can be plotted to form a curve as shown in Exhibit 24. All points on the curve and points to the right of the curve are attainable by a combination of one or more of the investable assets. This set of points is called the investment opportunity set. Initially, the opportunity set consists of domestic assets only and is labeled as such in Exhibit 24.

Exhibit 24 Investment Opportunity Set



15.1.1 *Addition of Asset Classes*

Exhibit 24 shows the effect of adding a new asset class, such as international assets. As long as the new asset class is not perfectly correlated with the existing asset class, the investment opportunity set will expand out further to the northwest, providing a superior risk–return trade-off.

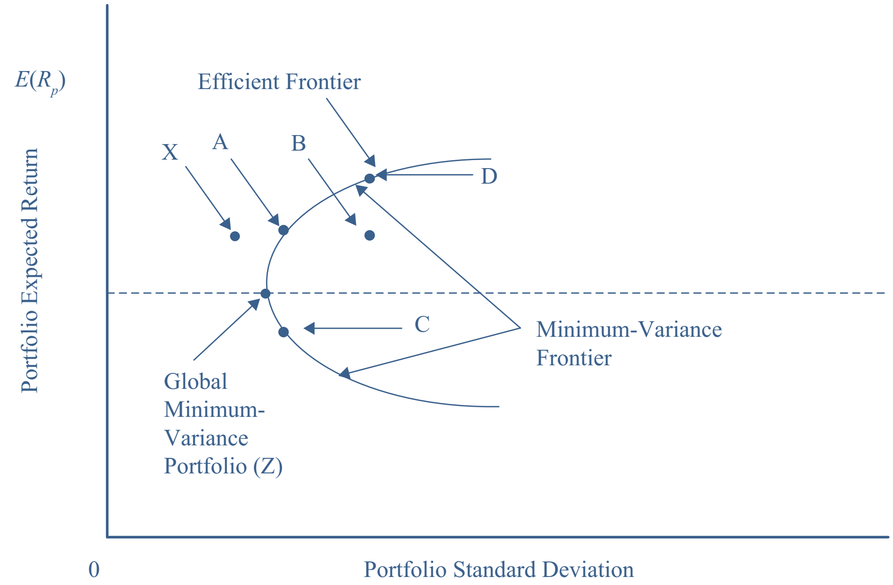
The investment opportunity set with international assets dominates the opportunity set that includes only domestic assets. Adding other asset classes will have the same impact on the opportunity set. Thus, we should continue to add asset classes until they do not further improve the risk–return trade-off. The benefits of diversification can be fully captured in this way in the construction of the investment opportunity set, and eventually in the selection of the optimal portfolio.

In the discussion that follows in this section, we will assume that *all* investable assets available to an investor are included in the investment opportunity set and no special attention needs to be paid to new asset classes or new investment opportunities.

15.2 Minimum-Variance Portfolios

The investment opportunity set consisting of all available investable sets is shown in Exhibit 25. There are a large number of portfolios available for investment, but we must choose a single optimal portfolio. In this subsection, we begin the selection process by narrowing the choice to fewer portfolios.

Exhibit 25 Minimum-Variance Frontier


$E(R_p)$
Portfolio Expected Return
Efficient Frontier
X
A
B
D
C
Minimum-Variance Frontier
Global Minimum-Variance Portfolio (Z)
0
Portfolio Standard Deviation

15.2.1 *Minimum-Variance Frontier*

Risk-averse investors seek to minimize risk for a given return. Consider Points A, B, and X in Exhibit 25 and assume that they are on the same horizontal line by construction. Thus, the three points have the same expected return, $E(R_1)$, as do all other points on the imaginary line connecting A, B, and X. Given a choice, an investor will choose the point with the minimum risk, which is Point X. Point X, however, is unattainable because it does not lie within the investment opportunity set. Thus, the minimum risk that we can attain for $E(R_1)$ is at Point A. Point B and all points to the right of Point A are feasible but they have higher risk. Therefore, a risk-averse investor will choose only Point A in preference to any other portfolio with the same return.

Similarly, Point C is the minimum variance point for the return earned at C. Points to the right of C have higher risk. We can extend the above analysis to all possible returns. In all cases, we find that the **minimum-variance portfolio** is the one that lies on the solid curve drawn in Exhibit 25. The entire collection of these minimum-variance portfolios is referred to as the minimum-variance frontier. The minimum-variance frontier defines the smaller set of portfolios in which investors would want to invest. Note that no risk-averse investor will choose to invest in a portfolio to the right of the minimum-variance frontier because a portfolio on the minimum-variance frontier can give the same return but at a lower risk.

15.2.2 *Global Minimum-Variance Portfolio*

The left-most point on the minimum-variance frontier is the portfolio with the minimum variance among all portfolios of risky assets, and is referred to as the **global minimum-variance portfolio**. An investor cannot hold a portfolio consisting of *risky* assets that has less risk than that of the global minimum-variance portfolio. Note the emphasis on “risky” assets. Later, the introduction of a risk-free asset will allow us to relax this constraint.

15.2.3 *Efficient Frontier of Risky Assets*

The minimum-variance frontier gives us portfolios with the minimum variance for a given return. However, investors also want to maximize return for a given risk. Observe Points A and C on the minimum-variance frontier shown in Exhibit 25. Both of them have the same risk. Given a choice, an investor will choose Portfolio A because it has a higher return. No one will choose Portfolio C. The same analysis applies to all points on the minimum-variance frontier that lie below the global minimum-variance portfolio. Thus, portfolios on the curve below the global minimum-variance portfolio and to the right of the global minimum-variance portfolio are not beneficial and are inefficient portfolios for an investor.

The curve that lies above and to the right of the global minimum-variance portfolio is referred to as the **Markowitz efficient frontier** because it contains all portfolios of risky assets that rational, risk-averse investors will choose.

An important observation that is often ignored is the slope at various points on the efficient frontier. As we move right from the global minimum-variance portfolio (Point Z) in Exhibit 25, there is an increase in risk with a concurrent increase in return. The increase in return with every unit increase in risk, however, keeps decreasing as we move from left to the right because the slope continues to decrease. The slope at Point D is less than the slope at Point A, which is less than the slope at Point Z. The increase in return by moving from Point Z to Point A is the same as the increase in return by moving from Point A to Point D. It can be seen that the additional risk in moving from Point A to Point D is 3 to 4 times more than the additional risk in moving from Point Z to Point A. Thus, investors obtain decreasing increases in returns as they assume more risk.

16

EFFICIENT FRONTIER: A RISK-FREE ASSET AND MANY RISKY ASSETS

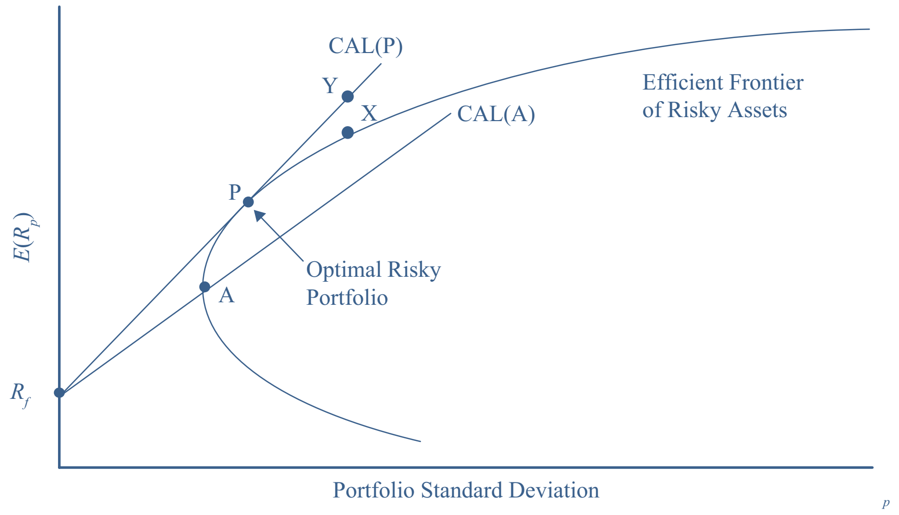
i. explain the selection of an optimal portfolio, given an investor's utility (or risk aversion) and the capital allocation line.

Until now, we have only considered risky assets in which the return is risky or uncertain. Most investors, however, have access to a risk-free asset, most notably from securities issued by the government. The addition of a risk-free asset makes the investment opportunity set much richer than the investment opportunity set consisting only of risky assets.

16.1 Capital Allocation Line and Optimal Risky Portfolio

By definition, a risk-free asset has zero risk so it must lie on the y-axis in a mean-variance graph. A risk-free asset with a return of R_f is plotted in Exhibit 26. This asset can now be combined with a portfolio of risky assets. The combination of a risk-free asset with a portfolio of risky assets is a straight line, such as in Section 11 (see Exhibit 15). Unlike in Section 11, however, we have many risky portfolios to choose from instead of a single risky portfolio.

Exhibit 26 Optimal Risky Portfolio



All portfolios on the efficient frontier are candidates for being combined with the risk-free asset. Two combinations are shown in Exhibit 26: one between the risk-free asset and efficient Portfolio A and the other between the risk-free asset and efficient Portfolio P. Comparing capital allocation line A and capital allocation line P reveals that there is a point on CAL(P) with a higher return and same risk for each point on CAL(A). In other words, the portfolios on CAL(P) dominate the portfolios on CAL(A). Therefore, an investor will choose CAL(P) over CAL(A). We would like to move further northwest to achieve even better portfolios. None of those portfolios, however, is attainable because they are above the efficient frontier.

What about other points on the efficient frontier? For example, Point X is on the efficient frontier and has the highest return of all risky portfolios for its risk. However, Point Y on CAL(P), achievable by leveraging Portfolio P as seen in Section 11, lies above Point X and has the same risk but higher return. In the same way, we can observe that not only does CAL(P) dominate CAL(A) but it also dominates the Markowitz efficient frontier of risky assets.

CAL(P) is the optimal capital allocation line and Portfolio P is the optimal risky portfolio. Thus, with the addition of the risk-free asset, we are able to narrow our selection of risky portfolios to a single optimal risky portfolio, P, which is at the tangent of CAL(P) and the efficient frontier of risky assets.

16.2 The Two-Fund Separation Theorem

The **two-fund separation theorem** states that all investors regardless of taste, risk preferences, and initial wealth will hold a combination of two portfolios or funds: a risk-free asset and an optimal portfolio of risky assets.[10]

The separation theorem allows us to divide an investor's investment problem into two distinct steps: the investment decision and the financing decision. In the first step, as in the previous analysis, the investor identifies the optimal risky portfolio. The optimal risky portfolio is selected from numerous risky portfolios without considering the investor's preferences. The investment decision at this step is based on the optimal risky portfolio's (a single portfolio) return, risk, and correlations.

---

[10] In the next reading, you will learn that the optimal portfolio of risky assets is the market portfolio.

The capital allocation line connects the optimal risky portfolio and the risk-free asset. All optimal investor portfolios must be on this line. Each investor's optimal portfolio on the CAL(P) is determined in the second step. Considering each individual investor's risk preference, using indifference curves, determines the investor's allocation to the risk-free asset (lending) and to the optimal risky portfolio. Portfolios beyond the optimal risky portfolio are obtained by borrowing at the risk-free rate (i.e., buying on margin). Therefore, the individual investor's risk preference determines the amount of financing (i.e., lending to the government instead of investing in the optimal risky portfolio or borrowing to purchase additional amounts of the optimal risky portfolio).

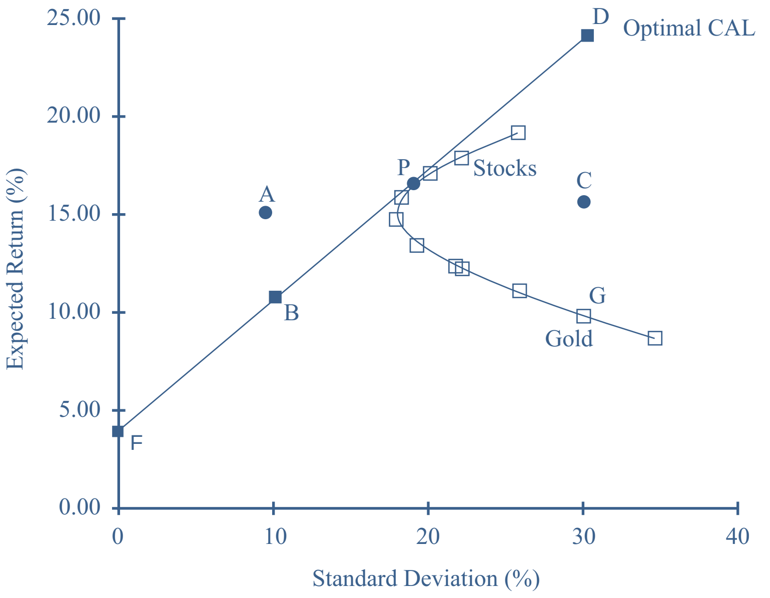
EXAMPLE 12

Choosing the Right Portfolio

In Exhibit 27, the risk and return of the points marked are as follows:

Point	Return (%)	Risk (%)	Point (%)	Return (%)	Risk (%)
A	15	10	B	11	10
C	15	30	D	25	30
F	4	0	G (gold)	10	30
P	16	17			

Exhibit 27



Answer the following questions with reference to the points plotted on Exhibit 27 and explain your answers. The investor is choosing one portfolio based on the graph.

1. Which of the above points is not achievable?
2. Which of these portfolios will not be chosen by a rational, risk-averse investor?
3. Which of these portfolios is most suitable for a risk-neutral investor?

4 Gold is on the inefficient part of the feasible set. Nonetheless, gold is owned by many rational investors as part of a larger portfolio. Why?

5 What is the utility of an investor at point P with a risk aversion coefficient of 3?

Solution to 1:

Portfolio A is not attainable because it lies outside the feasible set and not on the capital allocation line.

Solution to 2:

Portfolios G and C will not be chosen because D provides higher return for the same risk. G and C are the only investable points that do not lie on the capital allocation line.

Solution to 3:

Portfolio D is most suitable because a risk-neutral investor cares only about return and portfolio D provides the highest return. A = 0 in the utility formula.

Solution to 4:

Gold may be owned as part of a portfolio (not as *the* portfolio) because gold has low or negative correlation with many risky assets, such as stocks. Being part of a portfolio can thus reduce overall risk even though its standalone risk is high and return is low. Note that gold's price is not stable—its return is very risky (30 percent). Even risk seekers will choose D over G, which has the same risk but higher return.

Solution to 5:

$$U = E(r) - 0.5A\sigma^2 = 0.16 - 0.5 \times 3 \times 0.0289 = 0.1167 = 11.67\%.$$

EFFICIENT FRONTIER: OPTIMAL INVESTOR PORTFOLIO

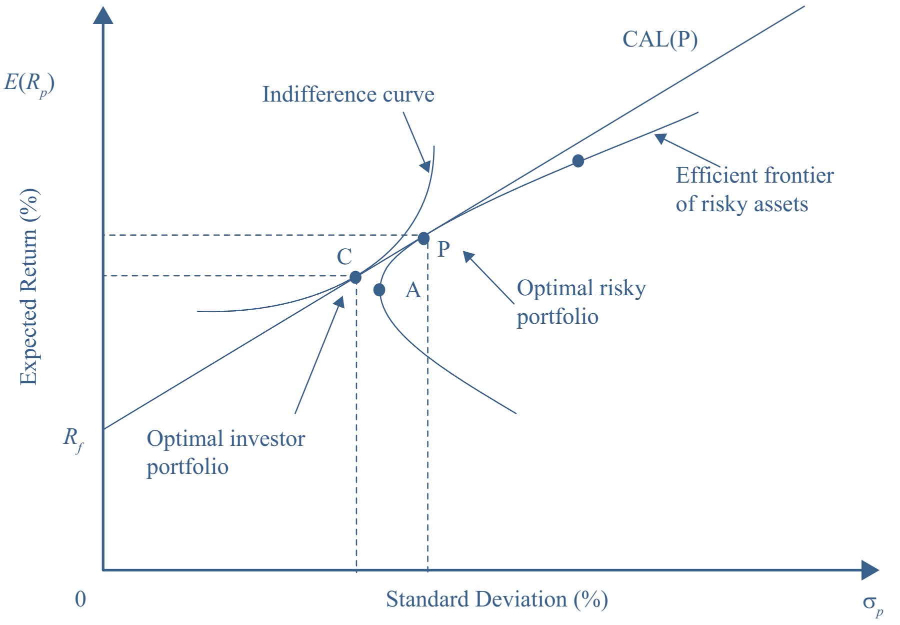
17

i. explain the selection of an optimal portfolio, given an investor's utility (or risk aversion) and the capital allocation line.

The CAL(P) in Exhibits 26 and 28 contains the best possible portfolios available to investors. Each of those portfolios is a linear combination of the risk-free asset and the optimal risky portfolio. Among the available portfolios, the selection of each investor's optimal portfolio depends on the risk preferences of an investor. In Sections 9–11, we discussed that the individual investor's risk preferences are incorporated into their indifference curves. These can be used to select the optimal portfolio.

Exhibit 28 shows an indifference curve that is tangent to the capital allocation line, CAL(P). Indifference curves with higher utility than this one lie above the capital allocation line, so their portfolios are not achievable. Indifference curves that lie below this one are not preferred because they have lower utility. Thus, the optimal portfolio for the investor with this indifference curve is portfolio C on CAL(P), which is tangent to the indifference curve.

Exhibit 28 Optimal Investor Portfolio


E(R_p)
Expected Return (%)
CAL(P)
Indifference curve
Efficient frontier
of risky assets
C
P
A
Optimal risky
portfolio
R_f
Optimal investor
portfolio
0
Standard Deviation (%)
σ_p

EXAMPLE 13

Comprehensive Example on Portfolio Selection

This comprehensive example reviews many concepts learned in this reading. The example begins with simple information about available assets and builds an optimal investor portfolio for the Lohrmanns.

Suppose the Lohrmanns can invest in only two risky assets, A and B. The expected return and standard deviation for asset A are 20 percent and 50 percent, and the expected return and standard deviation for asset B are 15 percent and 33 percent. The two assets have zero correlation with one another.

1 Calculate portfolio expected return and portfolio risk (standard deviation) if an investor invests 10 percent in A and the remaining 90 percent in B.

Solution to 1:

The subscript "*rp*" means risky portfolio.

$$R_{rp} = [0.10 \times 20\%] + [(1 - 0.10) \times 15\%] = 0.155 = 15.50\%$$

$$\sigma_{rp} = \sqrt{w_A^2\sigma_A^2 + w_B^2\sigma_B^2 + 2w_A w_B \rho_{AB}\sigma_A\sigma_B}$$

$$= \sqrt{(0.10^2 \times 0.50^2) + (0.90^2 \times 0.33^2) + (2 \times 0.10 \times 0.90 \times 0.0 \times 0.50 \times 0.33)}$$

$$= 0.3012 = 30.12\%$$

Note that the correlation coefficient is 0, so the last term for standard deviation is zero.

2 Generalize the above calculations for portfolio return and risk by assuming an investment of w_A in Asset A and an investment of $(1 - w_A)$ in Asset B.

Solution to 2:

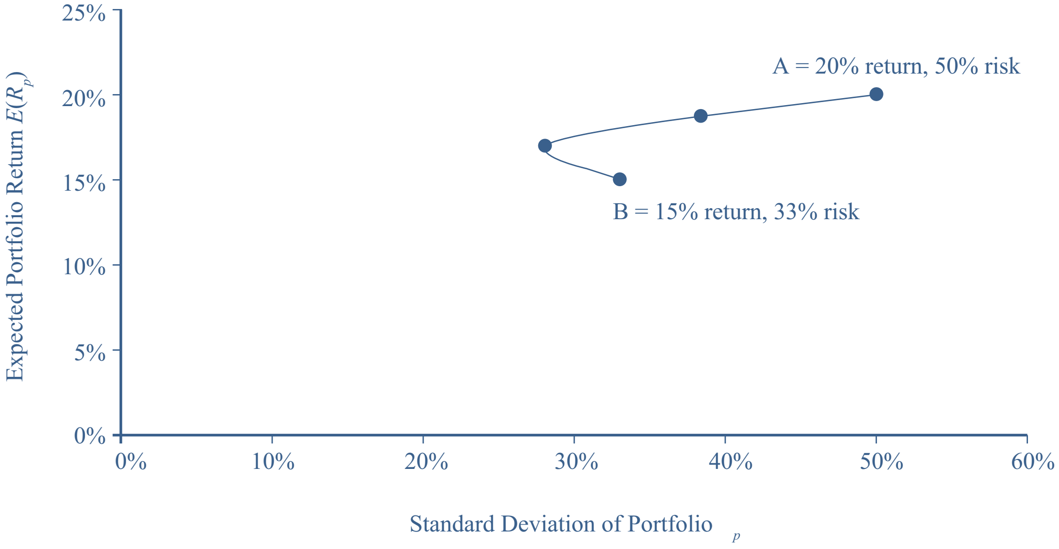
$$R_{rp} = w_A \times 20\% + \left(1 - w_A\right) \times 15\% = 0.05w_A + 0.15$$

$$\sigma_{rp} = \sqrt{w_A^2 \times 0.5^2 + \left(1 - w_A\right)^2 \times 0.33^2} = \sqrt{0.25w_A^2 + 0.1089\left(1 - 2w_A + w_A^2\right)}$$

$$= \sqrt{0.3589w_A^2 - 0.2178w_A + 0.1089}$$

The investment opportunity set can be constructed by using different weights in the expressions for $E(R_{rp})$ and σ_{rp} in Part 1 of this example. Exhibit 29 shows the combination of Assets A and B.

Exhibit 29



3 Now introduce a risk-free asset with a return of 3 percent. Write an equation for the capital allocation line in terms of w_A that will connect the risk-free asset to the portfolio of risky assets. (Hint: use the equation in Section 11 and substitute the expressions for a risky portfolio's risk and return from Part 2 above).

Solution to 3:

The equation of the line connecting the risk-free asset to the portfolio of risky assets is given below (see Section 11), where the subscript "*rp*" refers to the risky portfolio instead of "*i*," and the subscript "*p*" refers to the new portfolio of two risky assets and one risk-free asset.

$$E(R_p) = R_f + \frac{E(R_i) - R_f}{\sigma_i}\sigma_p,$$

Rewritten as

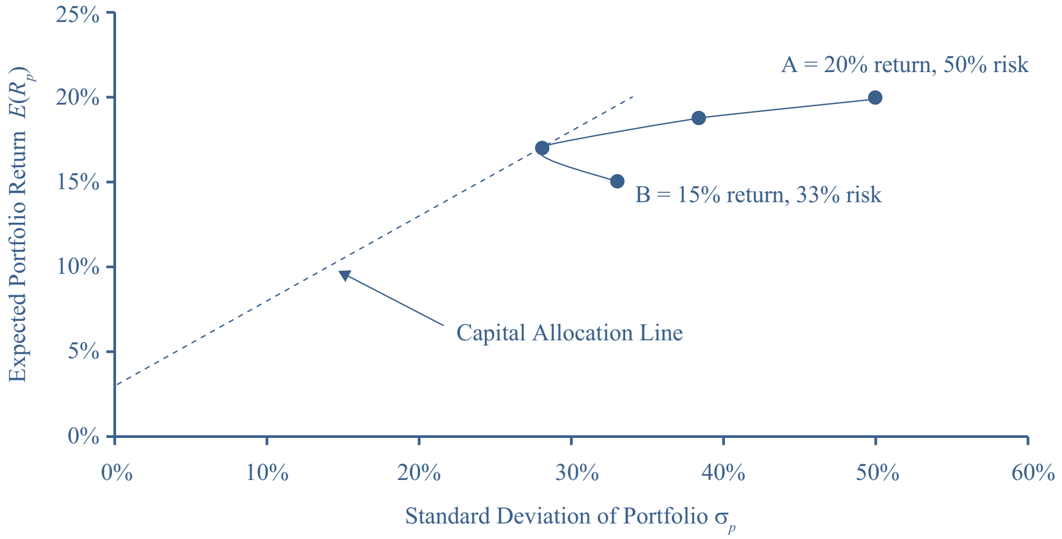
$$E(R_p) = R_f + \frac{E(R_{rp}) - R_f}{\sigma_{rp}}\sigma_p$$

$$= 0.03 + \frac{0.05w_A + 0.15 - 0.03}{\sqrt{0.3589w_A^2 - 0.2178w_A + 0.1089}}\sigma_p$$

$$= 0.03 + \frac{0.05w_A + 0.12}{\sqrt{0.3589w_A^2 - 0.2178w_A + 0.1089}}\sigma_p$$

The capital allocation line is the line that has the maximum slope because it is tangent to the curve formed by portfolios of the two risky assets. Exhibit 30 shows the capital allocation line based on a risk-free asset added to the group of assets.

Exhibit 30



4 The slope of the capital allocation line is maximized when the weight in Asset A is 38.20 percent.[11] What is the equation for the capital allocation line using w_A of 38.20 percent?

---

[11] You can maximize $\frac{0.05w_A + 0.12}{\sqrt{0.3589w_A^2 - 0.2178w_A + 0.1089}}$ by taking the first derivative of the slope with respect to w_A and setting it to 0.

Solution to 4:

By substituting 38.20 percent for w_A in the equation in Part 3, we get $E(R_p) = 0.03 + 0.4978\sigma_p$ as the capital allocation line.

5 Having created the capital allocation line, we turn to the Lohrmanns. What is the standard deviation of a portfolio that gives a 20 percent return and is on the capital allocation line? How does this portfolio compare with asset A?

Solution to 5:

Solve the equation for the capital allocation line to get the standard deviation: $0.20 = 0.03 + 0.4978\sigma_p$. $\sigma_p = 34.2\%$. The portfolio with a 20 percent return has the same return as Asset A but a lower standard deviation, 34.2 percent instead of 50.0 percent.

6 What is the risk of portfolios with returns of 3 percent, 9 percent, 15 percent, and 20 percent?

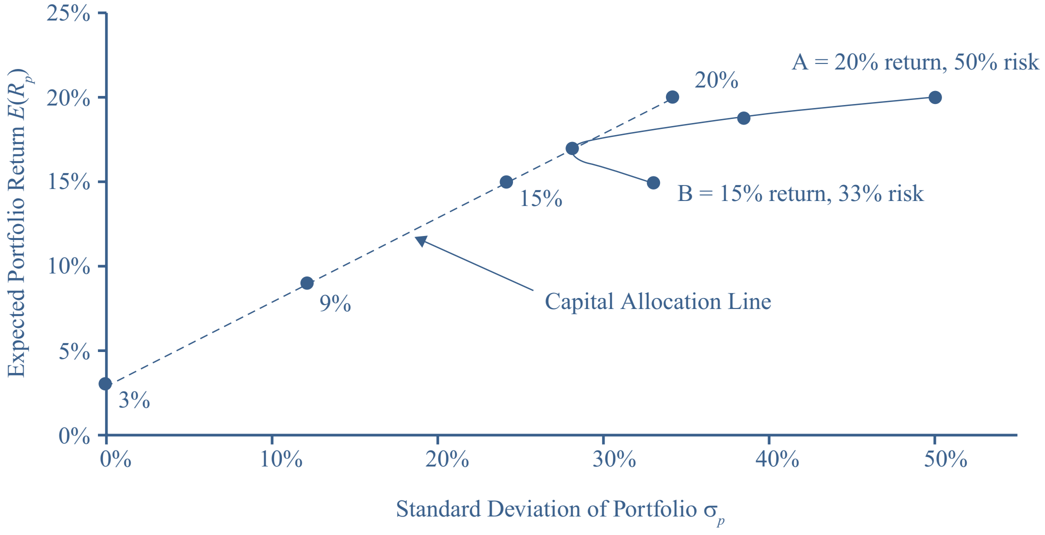
Solution to 6:

You can find the risk of the portfolio using the equation for the capital allocation line: $E(R_p) = 0.03 + 0.4978\sigma_p$.

For a portfolio with a return of 15 percent, write $0.15 = 0.03 + 0.4978\sigma_p$. Solving for σ_p gives 24.1 percent. You can similarly calculate risks of other portfolios with the given returns.

The risk of the portfolio for a return of 3 percent is 0.0 percent, for a return of 9 percent is 12.1 percent, for a return of 15 percent is 24.1 percent, and for a return of 20 percent is 34.2 percent. The points are plotted in Exhibit 31.

Exhibit 31



7 What is the utility that the Lohrmanns derive from a portfolio with a return of 3 percent, 9 percent, 15 percent, and 20 percent? The risk aversion coefficient for the Lohrmanns is 2.5.

Solution to 7:

To find the utility, use the utility formula with a risk aversion coefficient of 2.5:

$$\text{Utility} = E\left(R_p\right) - 0.5 \times 2.5\sigma_p^2$$

$$\text{Utility}\left(3\%\right) = 0.0300$$

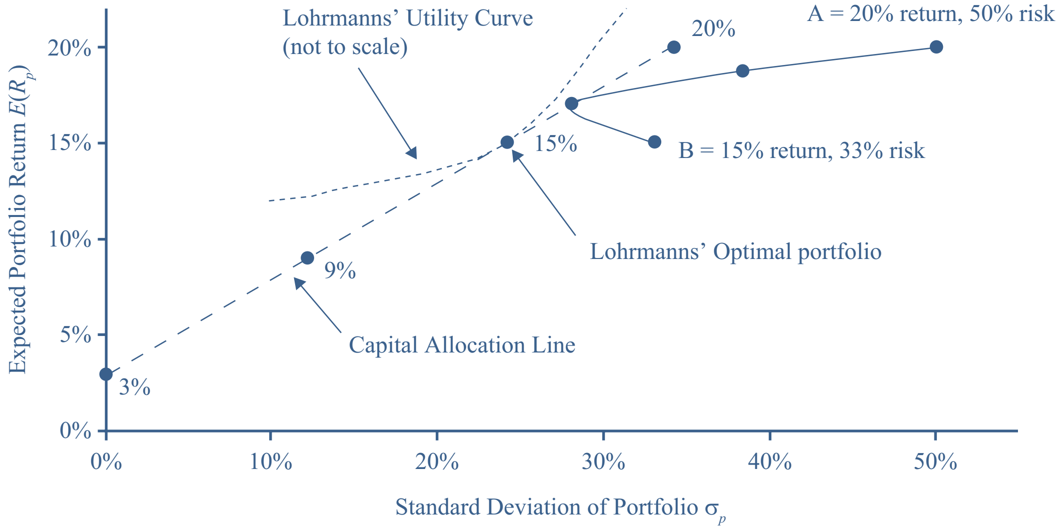
$$\text{Utility}\left(9\%\right) = 0.09 - 0.5 \times 2.5 \times 0.121^2 = +0.0717$$

$$\text{Utility}\left(15\%\right) = 0.15 - 0.5 \times 2.5 \times 0.241^2 = +0.0774$$

$$\text{Utility}\left(20\%\right) = 0.20 - 0.5 \times 2.5 \times 0.341^2 = +0.0546$$

Based on the above information, the Lohrmanns choose a portfolio with a return of 15 percent and a standard deviation of 24.1 percent because it has the highest utility: 0.0774. Finally, Exhibit 32 shows the indifference curve that is tangent to the capital allocation line to generate Lohrmanns' optimal investor portfolio.

Exhibit 32



17.1 Investor Preferences and Optimal Portfolios

The location of an optimal investor portfolio depends on the investor's risk preferences. A highly risk-averse investor may invest a large proportion, even 100 percent, of his/her assets in the risk-free asset. The optimal portfolio in this investor's case will be located close to the *y*-axis. A less risk-averse investor, however, may invest a large portion of his/her wealth in the optimal risky asset. The optimal portfolio in this investor's case will lie closer to Point P in Exhibit 28.

Some less risk-averse investors (i.e., with a high risk tolerance) may wish to accept even more risk because of the chance of higher return. Such an investor may borrow money to invest more in the risky portfolio. If the investor borrows 25 percent of his wealth, he/she can invest 125 percent in the optimal risky portfolio. The optimal investor portfolio for such an investor will lie to the right of Point P on the capital allocation line.

Thus, moving from the risk-free asset along the capital allocation line, we encounter investors who are willing to accept more risk. At Point P, the investor is 100 percent invested in the optimal risky portfolio. Beyond Point P, the investor accepts even more risk by borrowing money and investing in the optimal risky portfolio.

Note that we are able to accommodate all types of investors with just two portfolios: the risk-free asset and the optimal risky portfolio. Exhibit 28 is also an illustration of the two-fund separation theorem. Portfolio P is the optimal risky portfolio that is selected without regard to investor preferences. The optimal investor portfolio is selected on the capital allocation line by overlaying the indifference curves that incorporate investor preferences.

SUMMARY

This reading provides a description and computation of investment characteristics, such as risk and return, that investors use in evaluating assets for investment. This was followed by sections about portfolio construction, selection of an optimal risky portfolio, and an understanding of risk aversion and indifference curves. Finally, the tangency point of the indifference curves with the capital allocation line allows identification of the optimal investor portfolio. Key concepts covered in the reading include the following:

- Holding period return is most appropriate for a single, predefined holding period.
- Multiperiod returns can be aggregated in many ways. Each return computation has special applications for evaluating investments.
- Risk-averse investors make investment decisions based on the risk–return trade-off, maximizing return for the same risk, and minimizing risk for the same return. They may be concerned, however, by deviations from a normal return distribution and from assumptions of financial markets' operational efficiency.
- Investors are risk averse, and historical data confirm that financial markets price assets for risk-averse investors.
- The risk of a two-asset portfolio is dependent on the proportions of each asset, their standard deviations and the correlation (or covariance) between the assets' returns. As the number of assets in a portfolio increases, the correlation among asset risks becomes a more important determinate of portfolio risk.
- Combining assets with low correlations reduces portfolio risk.
- The two-fund separation theorem allows us to separate decision making into two steps. In the first step, the optimal risky portfolio and the capital allocation line are identified, which are the same for all investors. In the second step, investor risk preferences enable us to find a unique optimal investor portfolio for each investor.
- The addition of a risk-free asset creates portfolios that are dominant to portfolios of risky assets in all cases except for the optimal risky portfolio.

By successfully understanding the content of this reading, you should be comfortable calculating an investor's optimal portfolio given the investor's risk preferences and universe of investable assets available.

REFERENCES

2009 Ibbotson Stocks, Bonds, Bills, and Inflation (SBBI) Classic Yearbook. 2009. Chicago, IL: Morningstar.

Bogle, John C. 2008. "Black Monday and Black Swans." *Financial Analysts Journal*, vol. 64, no. 2: 30–40.

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2009. *Credit Suisse Global Investment Returns Sourcebook*. 2009. Zurich, Switzerland: Credit Suisse Research Institute.

Taleb, Nassim N. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House Inc.

PRACTICE PROBLEMS

1. An investor purchased 100 shares of a stock for $34.50 per share at the beginning of the quarter. If the investor sold all of the shares for $30.50 per share after receiving a $51.55 dividend payment at the end of the quarter, the holding period return is *closest* to:

   A –13.0%.

   B –11.6%.

   C –10.1%.

2. An analyst obtains the following annual rates of return for a mutual fund:

Year	Return (%)
2008	14
2009	–10
2010	–2

   The fund's holding period return over the three-year period is *closest* to:

   A 0.18%.

   B 0.55%.

   C 0.67%.

3. An analyst observes the following annual rates of return for a hedge fund:

Year	Return (%)
2008	22
2009	–25
2010	11

   The hedge fund's annual geometric mean return is *closest* to:

   A 0.52%.

   B 1.02%.

   C 2.67%.

4. Which of the following return calculating methods is *best* for evaluating the annualized returns of a buy-and-hold strategy of an investor who has made annual deposits to an account for each of the last five years?

   A Geometric mean return.

   B Arithmetic mean return.

   C Money-weighted return.

5. An investor performs the following transactions on the shares of a firm.

   - At t = 0, she purchases a share for $1,000.
   - At t = 1, she receives a dividend of $25 and then purchases three additional shares for $1,055 each.
   - At t = 2, she receives a total dividend of $100 and then sells the four shares for $1,100 each.

   The money-weighted rate of return is *closest* to:

   A 4.5%.

© 2019 CFA Institute. All rights reserved.

B 6.9%.

C 7.3%.

6 A fund receives investments at the beginning of each year and generates returns as shown in the table.

Year of Investment	Assets Under Management at the beginning of each year	Return during Year of Investment
1	\$1,000	15%
2	\$4,000	14%
3	\$45,000	−4%

Which return measure over the three-year period is negative?

A Geometric mean return

B Time-weighted rate of return

C Money-weighted rate of return

7 At the beginning of Year 1, a fund has \$10 million under management; it earns a return of 14% for the year. The fund attracts another \$100 million at the start of Year 2 and earns a return of 8% for that year. The money-weighted rate of return is *most likely*:

A less than the time-weighted rate of return.

B the same as the time-weighted rate of return.

C greater than the time-weighted rate of return.

8 An investor evaluating the returns of three recently formed exchange-traded funds gathers the following information:

ETF	Time Since Inception	Return Since Inception (%)
1	146 days	4.61
2	5 weeks	1.10
3	15 months	14.35

The ETF with the highest annualized rate of return is:

A ETF 1.

B ETF 2.

C ETF 3.

9 With respect to capital market theory, which of the following asset characteristics is *least likely* to impact the variance of an investor's equally weighted portfolio?

A Return on the asset.

B Standard deviation of the asset.

C Covariances of the asset with the other assets in the portfolio.

10 A portfolio manager creates the following portfolio:

Security	Security Weight (%)	Expected Standard Deviation (%)
1	30	20
2	70	12

If the correlation of returns between the two securities is 0.40, the expected standard deviation of the portfolio is *closest* to:

A 10.7%.

B 11.3%.

C 12.1%.

11 A portfolio manager creates the following portfolio:

Security	Security Weight (%)	Expected Standard Deviation (%)
1	30	20
2	70	12

If the covariance of returns between the two securities is −0.0240, the expected standard deviation of the portfolio is *closest* to:

A 2.4%.

B 7.5%.

C 9.2%.

The following information relates to Questions 12–13

A portfolio manager creates the following portfolio:

Security	Security Weight (%)	Expected Standard Deviation (%)
1	30	20
2	70	12

12 If the standard deviation of the portfolio is 14.40%, the correlation between the two securities is equal to:

A −1.0.

B 0.0.

C 1.0.

13 If the standard deviation of the portfolio is 14.40%, the covariance between the two securities is equal to:

A 0.0006.

B 0.0240.

C 1.0000.

The following information relates to Questions 14–17

An analyst observes the following historic geometric returns:

Asset Class	Geometric Return (%)
Equities	8.0
Corporate Bonds	6.5
Treasury bills	2.5
Inflation	2.1

14 The real rate of return for equities is *closest* to:

- **A** 5.4%.
- **B** 5.8%.
- **C** 5.9%.

15 The real rate of return for corporate bonds is *closest* to:

- **A** 4.3%.
- **B** 4.4%.
- **C** 4.5%.

16 The risk premium for equities is *closest* to:

- **A** 5.4%.
- **B** 5.5%.
- **C** 5.6%.

17 The risk premium for corporate bonds is *closest* to:

- **A** 3.5%.
- **B** 3.9%.
- **C** 4.0%.

---

18 With respect to trading costs, liquidity is *least likely* to impact the:

- **A** stock price.
- **B** bid–ask spreads.
- **C** brokerage commissions.

19 Evidence of risk aversion is *best* illustrated by a risk–return relationship that is:

- **A** negative.
- **B** neutral.
- **C** positive.

20 With respect to risk-averse investors, a risk-free asset will generate a numerical utility that is:

- **A** the same for all individuals.
- **B** positive for risk-averse investors.
- **C** equal to zero for risk seeking investors.

21 With respect to utility theory, the most risk-averse investor will have an indifference curve with the:

- **A** most convexity.
- **B** smallest intercept value.
- **C** greatest slope coefficient.

22 With respect to an investor's utility function expressed as: $U = E(r) - \frac{1}{2}A\sigma^2$, which of the following values for the measure for risk aversion has the *least* amount of risk aversion?

A −4.

B 0.

C 4.

The following information relates to Questions 23–26

A financial planner has created the following data to illustrate the application of utility theory to portfolio selection:

Investment	Expected Return (%)	Expected Standard Deviation (%)
1	18	2
2	19	8
3	20	15
4	18	30

23 A risk-neutral investor is *most likely* to choose:

A Investment 1.

B Investment 2.

C Investment 3.

24 If an investor's utility function is expressed as $U = E(r) - \frac{1}{2}A\sigma^2$ and the measure for risk aversion has a value of −2, the risk-seeking investor is *most likely* to choose:

A Investment 2.

B Investment 3.

C Investment 4.

25 If an investor's utility function is expressed as $U = E(r) - \frac{1}{2}A\sigma^2$ and the measure for risk aversion has a value of 2, the risk-averse investor is *most likely* to choose:

A Investment 1.

B Investment 2.

C Investment 3.

26 If an investor's utility function is expressed as $U = E(r) - \frac{1}{2}A\sigma^2$ and the measure for risk aversion has a value of 4, the risk-averse investor is *most likely* to choose:

A Investment 1.

B Investment 2.

C Investment 3.

---

27 With respect to the mean–variance portfolio theory, the capital allocation line, CAL, is the combination of the risk-free asset and a portfolio of all:

A risky assets.

B equity securities.

C feasible investments.

28 Two individual investors with different levels of risk aversion will have optimal portfolios that are:

A below the capital allocation line.

B on the capital allocation line.

C above the capital allocation line.

The following information relates to Questions 29–31

A portfolio manager creates the following portfolio:

Security	Expected Annual Return (%)	Expected Standard Deviation (%)
1	16	20
2	12	20

29 If the portfolio of the two securities has an expected return of 15%, the proportion invested in Security 1 is:

A 25%.

B 50%.

C 75%.

30 If the correlation of returns between the two securities is −0.15, the expected standard deviation of an equal-weighted portfolio is *closest* to:

A 13.04%.

B 13.60%.

C 13.87%.

31 If the two securities are uncorrelated, the expected standard deviation of an equal-weighted portfolio is *closest* to:

A 14.00%.

B 14.14%.

C 20.00%.

---

32 As the number of assets in an equally-weighted portfolio increases, the contribution of each individual asset's variance to the volatility of the portfolio:

A increases.

B decreases.

C remains the same.

33 With respect to an equally-weighted portfolio made up of a large number of assets, which of the following contributes the *most* to the volatility of the portfolio?

A Average variance of the individual assets.

B Standard deviation of the individual assets.

C Average covariance between all pairs of assets.

34 The correlation between assets in a two-asset portfolio increases during a market decline. If there is no change in the proportion of each asset held in the portfolio or the expected standard deviation of the individual assets, the volatility of the portfolio is *most likely* to:

A increase.

B decrease.

C remain the same.

The following information relates to Questions 35–37

An analyst has made the following return projections for each of three possible outcomes with an equal likelihood of occurrence:

Asset	Outcome 1 (%)	Outcome 2 (%)	Outcome 3 (%)	Expected Return (%)
1	12	0	6	6
2	12	6	0	6
3	0	6	12	6

35 Which pair of assets is perfectly negatively correlated?

A Asset 1 and Asset 2.

B Asset 1 and Asset 3.

C Asset 2 and Asset 3.

36 If the analyst constructs two-asset portfolios that are equally-weighted, which pair of assets has the *lowest* expected standard deviation?

A Asset 1 and Asset 2.

B Asset 1 and Asset 3.

C Asset 2 and Asset 3.

37 If the analyst constructs two-asset portfolios that are equally weighted, which pair of assets provides the *least* amount of risk reduction?

A Asset 1 and Asset 2.

B Asset 1 and Asset 3.

C Asset 2 and Asset 3.

---

38 Which of the following statements is *least* accurate? The efficient frontier is the set of all attainable risky assets with the:

A highest expected return for a given level of risk.

B lowest amount of risk for a given level of return.

C highest expected return relative to the risk-free rate.

39 The portfolio on the minimum-variance frontier with the lowest standard deviation is:

A unattainable.

B the optimal risky portfolio.

C the global minimum-variance portfolio.

40 The set of portfolios on the minimum-variance frontier that dominates all sets of portfolios below the global minimum-variance portfolio is the:

A capital allocation line.

B Markowitz efficient frontier.

C set of optimal risky portfolios.

41 The dominant capital allocation line is the combination of the risk-free asset and the:

A optimal risky portfolio.

B levered portfolio of risky assets.

C global minimum-variance portfolio.

42 Compared to the efficient frontier of risky assets, the dominant capital allocation line has higher rates of return for levels of risk greater than the optimal risky portfolio because of the investor's ability to:

A lend at the risk-free rate.

B borrow at the risk-free rate.

C purchase the risk-free asset.

43 With respect to the mean–variance theory, the optimal portfolio is determined by each individual investor's:

A risk-free rate.

B borrowing rate.

C risk preference.

SOLUTIONS

1 C is correct. −10.1% is the holding period return, which is calculated as: (3,050 − 3,450 + 51.55)/3,450, which is comprised of a dividend yield of 1.49% = 51.55/(3,450) and a capital loss yield of −11.59% = −400/(3,450).

2 B is correct. $[(1 + 0.14)(1 - 0.10)(1 - 0.02)] - 1 = 0.0055 = 0.55\%$.

3 A is correct. $[(1 + 0.22)(1 - 0.25)(1 + 0.11)]^{(1/3)} - 1 = 1.0157^{(1/3)} - 1 = 0.0052 = 0.52\%$

4 A is correct. The geometric mean return compounds the returns instead of the amount invested.

5 B is correct. Computation of the money-weighted return, r, requires finding the discount rate that sums the present value of cash flows to zero.

The first step is to group net cash flows by time. For this example, we have −\$1,000 for the $t = 0$ net cash flow, −\$3,140 = −\$3,165 + \$25 for the $t = 1$ net cash flow, and \$4,500 = \$4,400 + \$100 for the $t = 2$ net cash flow

Solving for r,

$$\text{CF}_0 = -1{,}000$$

$$\text{CF}_1 = -3{,}140$$

$$\text{CF}_2 = +4{,}500$$

$$\frac{\text{CF}_0}{(1+\text{IRR})^0} + \frac{\text{CF}_1}{(1+\text{IRR})^1} + \frac{\text{CF}_2}{(1+\text{IRR})^2}$$

$$= \frac{-1{,}000}{1} + \frac{-3{,}140}{(1+\text{IRR})^1} + \frac{4{,}500}{(1+\text{IRR})^2} = 0$$

results in a value of $r = 6.91\%$

6 C is correct. The money-weighted rate of return considers both the timing and amounts of investments into the fund. To calculate the money-weighted rate of return, tabulate the annual returns and investment amounts to determine the cash flows

Year	1	2	3
Balance from previous year	0	\$1,150	\$4,560
New investment	\$1,000	\$2,850	\$40,440
Net balance at the beginning of year	\$1,000	\$4,000	\$45,000
Investment return for the year	15%	14%	−4%
Investment gain (loss)	\$150	\$560	−\$1,800
Balance at the end of year	\$1,150	\$4,560	\$43,200

$$\text{CF}_0 = -\$1{,}000,\ \text{CF}_1 = -\$2{,}850,\ \text{CF}_2 = -\$40{,}440,\ \text{CF}_3 = +\$43{,}200$$

Each cash inflow or outflow occurs at the end of each year. Thus, CF_0 refers to the cash flow at the end of Year 0 or beginning of Year 1, and CF_3 refers to the cash flow at end of Year 3 or beginning of Year 4. Because cash flows are being discounted to the present—that is, end of Year 0 or beginning of Year 1—the period of discounting CF_0 is zero whereas the period of discounting for CF_3 is 3 years.

Solving for *r*,

$$CF_0 = -1{,}000$$

$$CF_1 = -2{,}850$$

$$CF_2 = -40{,}440$$

$$CF_3 = +43{,}200$$

$$\frac{CF_0}{(1+IRR)^0}+\frac{CF_1}{(1+IRR)^1}+\frac{CF_2}{(1+IRR)^2}+\frac{CF_3}{(1+IRR)^3}$$

$$=\frac{-1{,}000}{1}+\frac{-2{,}850}{(1+IRR)^1}+\frac{-40{,}440}{(1+IRR)^2}+\frac{43{,}200}{(1+IRR)^3}=0$$

results in a value of $r = -2.22\%$

Note that B is incorrect because the time-weighted rate of return (TWR) of the fund is the same as the geometric mean return of the fund and is thus positive:

$$TWR = \sqrt[3]{(1.15)(1.14)(0.96)} - 1 = 7.97\%$$

7 A is correct. Computation of the money-weighted return, *r*, requires finding the discount rate that sums the present value of cash flows to zero. Because most of the investment came during Year 2, the measure will be biased toward the performance of Year 2. The cash flows are as follows:

$$CF_0 = -10$$

$$CF_1 = -100$$

$$CF_2 = +120.31$$

The terminal value is determined by summing the investment returns for each period [(10 × 1.14 × 1.08) + (100 × 1.08)]

$$\frac{CF_0}{(1+IRR)^0}+\frac{CF_1}{(1+IRR)^1}+\frac{CF_2}{(1+IRR)^2}$$

$$=\frac{-10}{1}+\frac{-100}{(1+IRR)^1}+\frac{120.31}{(1+IRR)^2}$$

results in a value of $r = 8.53\%$

The time-weighted return of the fund is = $\sqrt[2]{(1.14)(1.08)} - 1 = 10.96\%$.

8 B is correct. The annualized rate of return for ETF 2 is 12.05% = $(1.0110^{52/5})$ – 1, which is greater than the annualized rate of ETF 1, 11.93% = $(1.0461^{365/146})$ – 1, and ETF 3, 11.32% = $(1.1435^{12/15})$ – 1. Despite having the lowest value for the periodic rate, ETF 2 has the highest annualized rate of return because of the reinvestment rate assumption and the compounding of the periodic rate.

9 A is correct. The asset's returns are not used to calculate the portfolio's variance [only the assets' weights, standard deviations (or variances), and covariances (or correlations) are used].

10 C is correct.

$$\sigma_{port} = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{1,2}\sigma_1\sigma_2}$$

$$= \sqrt{(0.3)^2(20\%)^2 + (0.7)^2(12\%)^2 + 2(0.3)(0.7)(0.40)(20\%)(12\%)}$$

$$= (0.3600\% + 0.7056\% + 0.4032\%)^{0.5} = (1.4688\%)^{0.5} = 12.11\%$$

11 A is correct.

$$\begin{aligned}\sigma_{port} &= \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2Cov\left(R_1R_2\right)}\\ &= \sqrt{(0.3)^2(20\%)^2 + (0.7)^2(12\%)^2 + 2(0.3)(0.7)(-0.0240)}\\ &= (0.3600\% + 0.7056\% - 1.008\%)^{0.5} = (0.0576\%)^{0.5} = 2.40\%\end{aligned}$$

12 C is correct. A portfolio standard deviation of 14.40% is the weighted average, which is possible only if the correlation between the securities is equal to 1.0.

13 B is correct. A portfolio standard deviation of 14.40% is the weighted average, which is possible only if the correlation between the securities is equal to 1.0. If the correlation coefficient is equal to 1.0, then the covariance must equal 0.0240, calculated as: $Cov(R_1,R_2) = \rho_{12}\sigma_1\sigma_2 = (1.0)(20\%)(12\%) = 2.40\% = 0.0240$.

14 B is correct. (1 + 0.080)/(1 + 0.0210) − 1 = 5.8%

15 A is correct. (1 + 0.065)/(1 + 0.0210) − 1 = 4.3%

16 A is correct. (1 + 0.080)/(1 + 0.0250) − 1 = 5.4%

17 B is correct. (1 + 0.0650)/(1 + 0.0250) − 1 = 3.9%

18 C is correct. Brokerage commissions are negotiated with the brokerage firm. A security's liquidity impacts the operational efficiency of trading costs. Specifically, liquidity impacts the bid–ask spread and can impact the stock price (if the ability to sell the stock is impaired by the uncertainty associated with being able to sell the stock).

19 C is correct. Historical data over long periods of time indicate that there exists a positive risk–return relationship, which is a reflection of an investor's risk aversion.

20 A is correct. A risk-free asset has a variance of zero and is not dependent on whether the investor is risk neutral, risk seeking or risk averse. That is, given that the utility function of an investment is expressed as $U = E(r) - \frac{1}{2}A\sigma^2$, where A is the measure of risk aversion, then the sign of A is irrelevant if the variance is zero (like that of a risk-free asset).

21 C is correct. The most risk-averse investor has the indifference curve with the greatest slope.

22 A is correct. A negative value in the given utility function indicates that the investor is a risk seeker.

23 C is correct. Investment 3 has the highest rate of return. Risk is irrelevant to a risk-neutral investor, who would have a measure of risk aversion equal to 0. Given the utility function, the risk-neutral investor would obtain the greatest amount of utility from Investment 3.

Investment	Expected Return (%)	Expected Standard Deviation (%)	Utility *A* = 0
1	18	2	0.1800
2	19	8	0.1900
3	20	15	0.2000
4	18	30	0.1800

24 C is correct. Investment 4 provides the highest utility value (0.2700) for a risk-seeking investor, who has a measure of risk aversion equal to −2.

Investment	Expected Return (%)	Expected Standard Deviation (%)	Utility *A* = –2
1	18	2	0.1804
2	19	8	0.1964
3	20	15	0.2225
4	18	30	0.2700

25 B is correct. Investment 2 provides the highest utility value (0.1836) for a risk-averse investor who has a measure of risk aversion equal to 2.

Investment	Expected Return (%)	Expected Standard Deviation (%)	Utility *A* = 2
1	18	2	0.1796
2	19	8	0.1836
3	20	15	0.1775
4	18	30	0.0900

26 A is correct. Investment 1 provides the highest utility value (0.1792) for a risk-averse investor who has a measure of risk aversion equal to 4.

Investment	Expected Return (%)	Expected Standard Deviation (%)	Utility *A* = 4
1	18	2	0.1792
2	19	8	0.1772
3	20	15	0.1550
4	18	30	0.0000

27 A is correct. The CAL is the combination of the risk-free asset with zero risk and the portfolio of all risky assets that provides for the set of feasible investments. Allowing for borrowing at the risk-free rate and investing in the portfolio of all risky assets provides for attainable portfolios that dominate risky assets below the CAL.

28 B is correct. The CAL represents the set of all feasible investments. Each investor's indifference curve determines the optimal combination of the risk-free asset and the portfolio of all risky assets, which must lie on the CAL.

29 C is correct.

$$R_p = w_1 \times R_1 + (1 - w_1) \times R_2$$

$$R_p = w_1 \times 16\% + (1 - w_1) \times 12\%$$

$$15\% = 0.75(16\%) + 0.25(12\%)$$

30 A is correct.

$$\sigma_{port} = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{1,2}\sigma_1\sigma_2}$$

$$= \sqrt{(0.5)^2(20\%)^2 + (0.5)^2(20\%)^2 + 2(0.5)(0.5)(-0.15)(20\%)(20\%)}$$

$$= (1.0000\% + 1.0000\% - 0.3000\%)^{0.5} = (1.7000\%)^{0.5} = 13.04\%$$

31 B is correct.

$$\begin{aligned}\sigma_{port} &= \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1 w_2 \rho_{1,2}\sigma_1\sigma_2} \\ &= \sqrt{(0.5)^2(20\%)^2 + (0.5)^2\left(20\%^2\right) + 2(0.5)(0.5)(0.00)(20\%)(20\%)} \\ &= (1.0000\% + 1.0000\% + 0.0000\%)^{0.5} = (2.0000\%)^{0.5} = 14.14\%\end{aligned}$$

32 B is correct. The contribution of each individual asset's variance (or standard deviation) to the portfolio's volatility decreases as the number of assets in the equally weighted portfolio increases. The contribution of the co-movement measures between the assets increases (i.e., covariance and correlation) as the number of assets in the equally weighted portfolio increases. The following equation for the variance of an equally weighted portfolio illustrates these points: $\sigma_p^2 = \frac{\bar{\sigma}^2}{N} + \frac{N-1}{N}\overline{COV} = \frac{\bar{\sigma}^2}{N} + \frac{N-1}{N}\bar{\rho}\bar{\sigma}^2$.

33 C is correct. The co-movement measures between the assets increases (i.e., covariance and correlation) as the number of assets in the equally weighted portfolio increases. The contribution of each individual asset's variance (or standard deviation) to the portfolio's volatility decreases as the number of assets in the equally weighted portfolio increases. The following equation for the variance of an equally weighted portfolio illustrates these points:

$$\sigma_p^2 = \frac{\bar{\sigma}^2}{N} + \frac{N-1}{N}\overline{COV} = \frac{\bar{\sigma}^2}{N} + \frac{N-1}{N}\bar{\rho}\bar{\sigma}^2$$

34 A is correct. Higher correlations will produce less diversification benefits provided that the other components of the portfolio standard deviation do not change (i.e., the weights and standard deviations of the individual assets).

35 C is correct. Asset 2 and Asset 3 have returns that are the same for Outcome 2, but the exact opposite returns for Outcome 1 and Outcome 3; therefore, because they move in opposite directions at the same magnitude, they are perfectly negatively correlated.

36 C is correct. An equally weighted portfolio of Asset 2 and Asset 3 will have the lowest portfolio standard deviation, because for each outcome, the portfolio has the same expected return (they are perfectly negatively correlated).

37 A is correct. An equally weighted portfolio of Asset 1 and Asset 2 has the highest level of volatility of the three pairs. All three pairs have the same expected return; however, the portfolio of Asset 1 and Asset 2 provides the least amount of risk reduction.

38 C is correct. The efficient frontier does not account for the risk-free rate. The efficient frontier is the set of all attainable risky assets with the highest expected return for a given level of risk or the lowest amount of risk for a given level of return.

39 C is correct. The global minimum-variance portfolio is the portfolio on the minimum-variance frontier with the lowest standard deviation. Although the portfolio is attainable, when the risk-free asset is considered, the global minimum-variance portfolio is not the optimal risky portfolio.

40 B is correct. The Markowitz efficient frontier has higher rates of return for a given level of risk. With respect to the minimum-variance portfolio, the Markowitz efficient frontier is the set of portfolios above the global minimum-variance portfolio that dominates the portfolios below the global minimum-variance portfolio.

41 A is correct. The use of leverage and the combination of a risk-free asset and the optimal risky asset will dominate the efficient frontier of risky assets (the Markowitz efficient frontier).

42 B is correct. The CAL dominates the efficient frontier at all points except for the optimal risky portfolio. The ability of the investor to purchase additional amounts of the optimal risky portfolio by borrowing (i.e., buying on margin) at the risk-free rate makes higher rates of return for levels of risk greater than the optimal risky asset possible.

43 C is correct. Each individual investor's optimal mix of the risk-free asset and the optimal risky asset is determined by the investor's risk preference.

READING

50

Portfolio Risk and Return: Part II

by Vijay Singal, PhD, CFA

Vijay Singal, PhD, CFA, is at Virginia Tech (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the implications of combining a risk-free asset with a portfolio of risky assets;
☐	**b.** explain the capital allocation line (CAL) and the capital market line (CML);
☐	**c.** explain systematic and nonsystematic risk, including why an investor should not expect to receive additional return for bearing nonsystematic risk;
☐	**d.** explain return generating models (including the market model) and their uses;
☐	**e.** calculate and interpret beta;
☐	**f.** explain the capital asset pricing model (CAPM), including its assumptions, and the security market line (SML);
☐	**g.** calculate and interpret the expected return of an asset using the CAPM;
☐	**h.** describe and demonstrate applications of the CAPM and the SML;
☐	**i.** calculate and interpret the Sharpe ratio, Treynor ratio, M^2, and Jensen's alpha.

INTRODUCTION

1

Our objective in this reading is to identify the optimal risky portfolio for all investors by using the capital asset pricing model (CAPM). The foundation of this reading is the computation of risk and return of a portfolio and the role that correlation plays in diversifying portfolio risk and arriving at the efficient frontier. The efficient frontier and the capital allocation line consist of portfolios that are generally acceptable to all

© 2019 CFA Institute. All rights reserved.

investors. By combining an investor's individual indifference curves with the market-determined capital allocation line, we are able to illustrate that the only optimal risky portfolio for an investor is the portfolio of all risky assets (i.e., the market).

Additionally, we discuss the capital market line, a special case of the capital allocation line that is used for passive investor portfolios. We also differentiate between systematic and nonsystematic risk, and explain why investors are compensated for bearing systematic risk but receive no compensation for bearing nonsystematic risk. We discuss in detail the CAPM, which is a simple model for estimating asset returns based only on the asset's systematic risk. Finally, we illustrate how the CAPM allows security selection to build an optimal portfolio for an investor by changing the asset mix beyond a passive market portfolio.

The reading is organized as follows. In Sections 2–4, we discuss the consequences of combining a risk-free asset with the market portfolio and provide an interpretation of the capital market line. Sections 5–7 decompose total risk into systematic and nonsystematic risk and discuss the characteristics of and differences between the two kinds of risk. We also introduce return-generating models, including the single-index model, and illustrate the calculation of beta. In Sections 8 and 9, we introduce the capital asset pricing model and the security market line. Our focus on the CAPM does not suggest that the CAPM is the only viable asset pricing model. Although the CAPM is an excellent starting point, more advanced readings expand on these discussions and extend the analysis to other models that account for multiple explanatory factors. Section 10 covers several post-CAPM developments in theory. Section 11 covers measures for evaluating the performance of a portfolio which take account of risk. Section 12 covers some applications of the CAPM in portfolio construction. A summary and practice problems conclude the reading.

2 CAPITAL MARKET THEORY: RISK-FREE AND RISKY ASSETS

a describe the implications of combining a risk-free asset with a portfolio of risky assets;

b explain the capital allocation line (CAL) and the capital market line (CML);

You have learned how to combine a risk-free asset with one risky asset and with many risky assets to create a capital allocation line. In this section, we will expand our discussion of multiple risky assets and consider a special case of the capital allocation line, called the capital market line. While discussing the capital market line, we will define the market and its role in passive portfolio management. Using these concepts, we will illustrate how leveraged portfolios can enhance both risk and return.

2.1 Portfolio of Risk-Free and Risky Assets

Although investors desire an asset that produces the highest return and carries the lowest risk, such an asset does not exist. As the risk–return capital market theory illustrates, one must assume higher risk in order to earn a higher return. We can improve an investor's portfolio, however, by expanding the opportunity set of risky assets because this allows the investor to choose a superior mix of assets.

Similarly, an investor's portfolio improves if a risk-free asset is added to the mix. In other words, a combination of the risk-free asset and a risky asset can result in a better risk–return trade-off than an investment in only one type of asset because the risk-free asset has zero correlation with the risky asset. The combination is called the **capital allocation line** (and is depicted in Exhibit 2). Superimposing an investor's indifference curves on the capital allocation line will lead to the optimal investor portfolio.

Investors with different levels of risk aversion will choose different portfolios. Highly risk-averse investors choose to invest most of their wealth in the risk-free asset and earn low returns because they are not willing to assume higher levels of risk. Less risk-averse investors, in contrast, invest more of their wealth in the risky asset, which is expected to yield a higher return. Obviously, the higher return cannot come without higher risk, but the less risk-averse investor is willing to accept the additional risk.

2.1.1 *Combining a Risk-Free Asset with a Portfolio of Risky Assets*

We can extend the analysis of one risky asset to a portfolio of risky assets. For convenience, assume that the portfolio contains all available risky assets (*N*), although an investor may not wish to include all of these assets in the portfolio because of the investor's specific preferences. If an asset is not included in the portfolio, its weight will be zero. The risk–return characteristics of a portfolio of *N* risky assets are given by the following equations:

$$E(R_p) = \sum_{i=1}^{N} w_i E(R_i)$$

$$\sigma_p^2 = \left(\sum_{i=1,j=1}^{N} w_i w_j \text{Cov}(i,j) \right), \text{ and } \sum_{i=1}^{N} w_i = 1$$

The expected return on the portfolio, $E(R_p)$, is the weighted average of the expected returns of individual assets, where w_i is the fractional weight in asset i and R_i is the expected return of asset i. The risk of the portfolio (σ_p), however, depends on the weights of the individual assets, the risk of the individual assets, and their interrelationships. The **covariance** between assets i and j, Cov(i,j), is a statistical measure of the interrelationship between each pair of assets in the portfolio and can be expressed as follows, where ρ_{ij} is the **correlation** between assets i and j and σ_i is the risk of asset i:

$$\text{Cov}(i,j) = \rho_{ij}\sigma_i\sigma_j$$

Note from the equation below that the correlation of an asset with itself is 1; therefore:

$$\text{Cov}(i,i) = \rho_{ii}\sigma_i\sigma_i = \sigma_i^2$$

By substituting the above expressions for covariance, we can rewrite the portfolio variance equation as

$$\sigma_p^2 = \left(\sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i,j=1,i \neq j}^{N} w_i w_j \rho_{ij} \sigma_i \sigma_j \right)$$

The suggestion that portfolios have lower risk than the assets they contain may seem counterintuitive. These portfolios can be constructed, however, as long as the assets in the portfolio are not perfectly correlated. As an illustration of the effect of asset weights on portfolio characteristics, consider a simple two-asset portfolio with zero weights in all other assets. Assume that Asset 1 has a return of 10 percent and a standard deviation (risk) of 20 percent. Asset 2 has a return of 5 percent and a standard deviation (risk) of 10 percent. Furthermore, the correlation between the two assets is zero. Exhibit 1 shows risks and returns for Portfolio X with a weight of 25 percent in Asset 1 and 75 percent in Asset 2, Portfolio Y with a weight of 50 percent in each of the two assets, and Portfolio Z with a weight of 75 percent in Asset 1 and 25 percent in Asset 2.

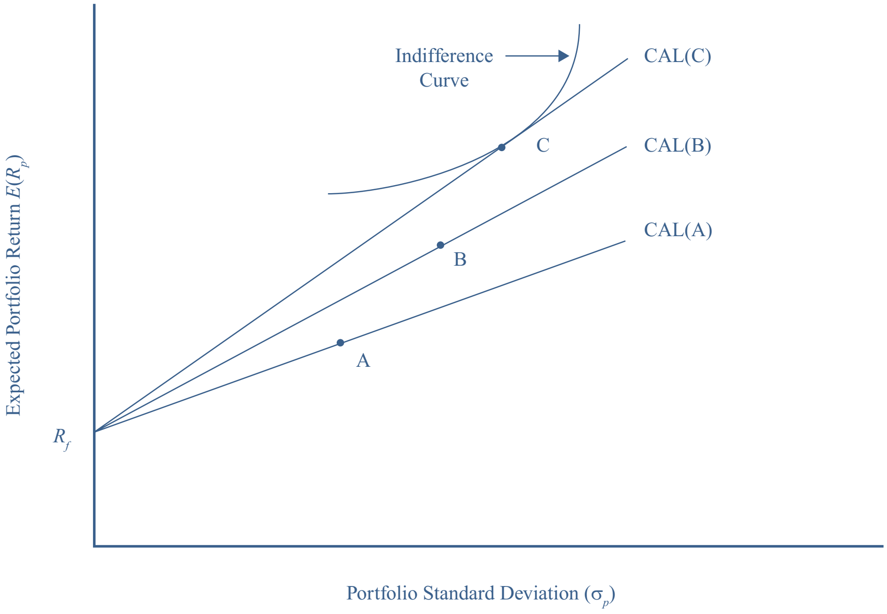
Exhibit 1 Portfolio Risk and Return

Portfolio	Weight in Asset 1 (%)	Weight in Asset 2 (%)	Portfolio Return (%)	Portfolio Standard Deviation (%)
X	25.0	75.0	6.25	9.01
Y	50.0	50.0	7.50	11.18
Z	75.0	25.0	8.75	15.21
Return	10.0	5.0		
Standard deviation	20.0	10.0		
Correlation between Assets 1 and 2		0.0		

From this example we observe that the three portfolios are quite different in terms of their risk and return. Portfolio X has a 6.25 percent return and only 9.01 percent standard deviation, whereas the standard deviation of Portfolio Z is more than two-thirds higher (15.21 percent), although the return is only slightly more than one-third higher (8.75 percent). These portfolios may become even more dissimilar as other assets are added to the mix.

Consider three portfolios of risky assets, A, B, and C, as in Exhibit 2, that may have been presented to a representative investor by three different investment advisers. Each portfolio is combined with the risk-free asset to create three capital allocation lines, CAL(A), CAL(B), and CAL(C). The exhibit shows that Portfolio C is superior to the other two portfolios because it has a greater expected return for any given level of risk. As a result, an investor will choose the portfolio that lies on the capital allocation line for Portfolio C. The combination of the risk-free asset and the risky Portfolio C that is selected for an investor depends on the investor's degree of risk aversion.

Exhibit 2 Risk-Free Asset and Portfolio of Risky Assets



2.1.2 *Does a Unique Optimal Risky Portfolio Exist?*

We assume that all investors have the same economic expectation and thus have the same expectations of prices, cash flows, and other investment characteristics. This assumption is referred to as **homogeneity of expectations**. Given these investment characteristics, everyone goes through the same calculations and should arrive at the same optimal risky portfolio. Therefore, assuming homogeneous expectations, only one optimal portfolio exists. If investors have different expectations, however, they might arrive at different optimal risky portfolios. To illustrate, we begin with an expression for the price of an asset:

$$P = \sum_{t=0}^{T} \frac{\mathrm{CF}_t}{\left(1 + r_t\right)^t}$$

where CF_t is the cash flow at the end of period t and r_t is the discount rate or the required rate of return for that asset for period t. Period t refers to all periods beginning from now until the asset ceases to exist at the end of time T. Because the current time is the end of period 0, which is the same as the beginning of period 1, there are $(T + 1)$ cash flows and $(T + 1)$ required rates of return. These conditions are based on the assumption that a cash flow, such as an initial investment, can occur now (t = 0). Ordinarily, however, CF_0 is zero.

We use the formula for the price of an asset to estimate the intrinsic value of an asset. Assume that the asset we are valuing is a share of Siemens AG which trades on Xetra. In the case of corporate stock, there is no expiration date, so T could be extremely large, meaning we will need to estimate a large number of cash flows and rates of return. Fortunately, the denominator reduces the importance of distant cash flows, so it may be sufficient to estimate, say, 20 annual cash flows and 20 rates of returns. How much will Siemens earn next year and the year after next? What will the product markets Siemens operates in look like in five years' time? Different analysts and investors will have their own estimates that may be quite different from one another. Also, as we delve further into the future, more serious issues in estimating future revenue, expenses, and growth rates arise. Therefore, to assume that cash flow estimates for Siemens will vary among these investors is reasonable. In addition to the numerator (cash flows), it is also necessary to estimate the denominator, the required rates of return. We know that riskier companies will require higher returns because risk and return are positively correlated. Siemens stock is riskier than a risk-free asset, but by how much? And what should the compensation for that additional risk be? Again, it is evident that different analysts will view the riskiness of Siemens differently and, therefore, arrive at different required rates of return.

Siemens closed at €111.84 on Xetra on 31 August 2018. The traded price represents the value that a marginal investor attaches to a share of Siemens, say, corresponding to Analyst A's expectation. Analyst B may think that the price should be €95, however, and Analyst C may think that the price should be €125. Given a price of €111.84, the expected returns of Siemens are quite different for the three analysts. Analyst B, who believes the price should be €95, concludes that Siemens is overvalued and may assign a weight of zero to Siemens in the recommended portfolio even though the market capitalization of Siemens was in excess of €100 billion as of the date of the quotation. In contrast, Analyst C, with a valuation of €125, thinks Siemens is undervalued and may significantly overweight Siemens in a portfolio.

Our discussion illustrates that analysts can arrive at different valuations that necessitate the assignment of different asset weights in a portfolio. Given the existence of many asset classes and numerous assets in each asset class, one can visualize that each investor will have his or her own optimal risky portfolio depending on his or her assumptions underlying the valuation computations. Therefore, market participants will have their own and possibly different optimal risky portfolios.

If investors have different valuations of assets, then the construction of a unique optimal risky portfolio is not possible. If we make a simplifying assumption of homogeneity in investor expectations, we will have a single optimal risky portfolio as previously mentioned. Even if investors have different expectations, market prices are a proxy of what the marginal, informed investor expects, and the market portfolio becomes the base case, the benchmark, or the reference portfolio that other portfolios can be judged against. For Siemens, the market price was €111.84 per share and the market capitalization was about €108 billion. In constructing the market portfolio, Siemens's weight in the market portfolio will be equal to its market value divided by the value of all other assets included in the market portfolio.

3

CAPITAL MARKET THEORY: THE CAPITAL MARKET LINE

b explain the capital allocation line (CAL) and the capital market line (CML);

In the previous section, we discussed how the risk-free asset could be combined with a risky portfolio to create a capital allocation line (CAL). In this section, we discuss a specific CAL that uses the market portfolio as the optimal risky portfolio and is known as the capital market line. We also discuss the significance of the market portfolio and applications of the capital market line (CML).

3.1 Passive and Active Portfolios

In the above subsection, we hypothesized three possible valuations for each share of Siemens: €95, €111.84, and €125. Which one is correct?

If the market is an **informationally efficient market**, the price in the market, €111.84, is an unbiased estimate of all future discounted cash flows (recall the formula for the price of an asset). In other words, the price aggregates and reflects all information that is publicly available, and investors cannot expect to earn a return that is greater than the required rate of return for that asset. If, however, the price reflects all publicly available information and there is no way to outperform the market, then there is little point in investing time and money in evaluating Siemens to arrive at your price using your own estimates of cash flows and rates of return.

In that case, a simple and convenient approach to investing is to rely on the prices set by the market. Portfolios that are based on the assumption of unbiased market prices are referred to as passive portfolios. Passive portfolios most commonly replicate and track market indexes, which are passively constructed on the basis of market prices and market capitalizations. Examples of market indexes are the S&P 500 Index, the Nikkei 300, and the CAC 40. Passive portfolios based on market indexes are called index funds and generally have low costs because no significant effort is expended in valuing securities that are included in an index.

In contrast to passive investors' reliance on market prices and index funds, active investors may not rely on market valuations. They have more confidence in their own ability to estimate cash flows, growth rates, and discount rates. Based on these estimates, they value assets and determine whether an asset is fairly valued. In an actively managed portfolio, assets that are undervalued, or have a chance of offering above-normal returns, will have a positive weight (i.e., overweight compared to the market weight in the benchmark index), whereas other assets will have a zero weight, or even a negative weight if short selling is permitted (i.e., some assets will be underweighted compared with the market weight in the benchmark index). (**Short selling** is a transaction in which borrowed securities are sold with the intention to repurchase them at a lower price at a later date and return them to the lender.) This style of investing is called active investment management, and the portfolios are referred to as active portfolios. Most open-end mutual funds and hedge funds practice active investment management, and most analysts believe that active investing adds value. Whether these analysts are right or wrong is the subject of continuing debate.

3.2 What Is the "Market"?

In the previous discussion, we referred to the "market" on numerous occasions without actually defining the market. The optimal risky portfolio and the capital market line depend on the definition of the market. So what is the market?

Theoretically, the market includes all risky assets or anything that has value, which includes stocks, bonds, real estate, and even human capital. Not all assets are tradable, however, and not all tradable assets are investable. For example, the Taj Mahal in India is an asset but is not a tradable asset. Similarly, human capital is an asset that is not tradable. Moreover, assets may be tradable but not investable because of restrictions placed on certain kinds of investors. For example, all stocks listed on the Shanghai Stock Exchange are tradable. However, whereas Class A shares are listed in RMB and open to domestic investors and qualified foreign investors, Class B shares are listed in USD and open to foreign investors and domestic investors holding foreign currency dealing accounts.

If we consider all stocks, bonds, real estate assets, commodities, etc., probably hundreds of thousands of assets are tradable and investable. The "market" should contain as many assets as possible; we emphasize the word "possible" because it is not practical to include all assets in a single risky portfolio. Even though advancements in technology and interconnected markets have made it much easier to span the major equity markets, we are still not able to easily invest in other kinds of assets like bonds and real estate except in the most developed countries.

For the rest of this reading, we will define the "market" quite narrowly because it is practical and convenient to do so. Typically, a local or regional stock market index is used as a proxy for the market because of active trading in stocks and because a local or regional market is most visible to the local investors. For our purposes, we will use the S&P 500 Index as the market's proxy. The S&P 500 is commonly used by analysts as a benchmark for market performance throughout the United States. It contains 500 of the largest stocks that are domiciled in the United States, and these stocks are weighted by their market capitalization (price times the number of outstanding shares).

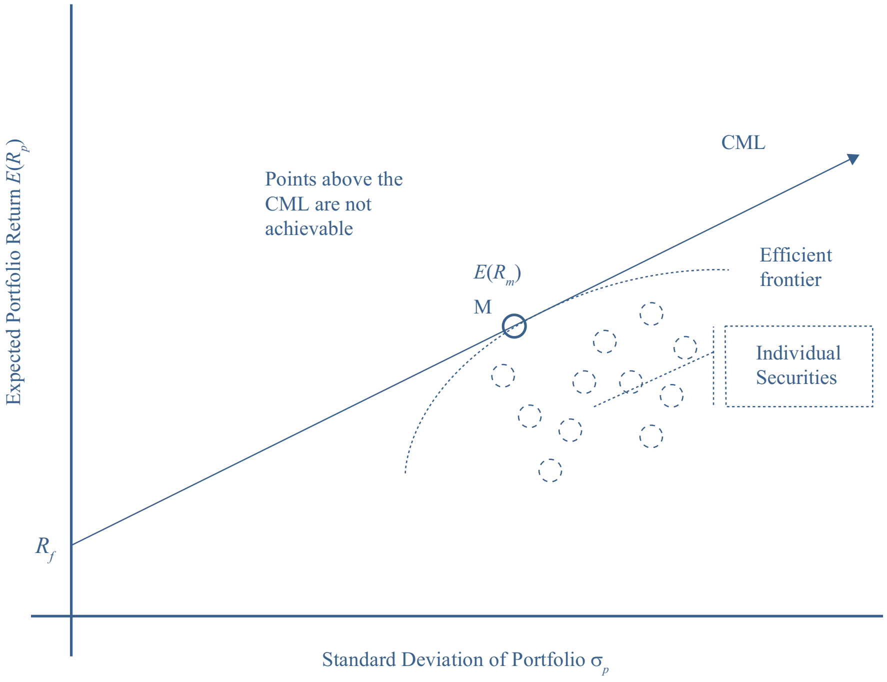
As of mid-2018, the stocks in the S&P 500 account for approximately 80 percent of the total equity market capitalization in the United States, and because the US stock markets represent about 40 percent of the world markets, the S&P 500 represents roughly 32 percent of worldwide publicly traded equity. Our definition of the market does not include non-US stock markets, bond markets, real estate, and many other asset classes, and therefore, "market" return and the "market" risk premium refer to US equity return and the US equity risk premium, respectively. The use of this proxy, however, is sufficient for our discussion, and is relatively easy to expand to include other tradable assets.

3.3 The Capital Market Line (CML)

A capital allocation line includes all possible combinations of the risk-free asset and an investor's optimal risky portfolio. The **capital market line** is a special case of the capital allocation line, where the risky portfolio is the market portfolio. The risk-free asset is a debt security with no default risk, no inflation risk, no liquidity risk, no interest rate risk, and no risk of any other kind. US Treasury bills are usually used as a proxy of the risk-free return, R_f.

The S&P 500 is a proxy of the market portfolio, which is the optimal risky portfolio. Therefore, the expected return on the risky portfolio is the expected market return, expressed as $E(R_m)$. The capital market line is shown in Exhibit 3, where the standard deviation (σ_p), or total risk, is on the x-axis and expected portfolio return, $E(R_p)$, is on the y-axis. Graphically, the market portfolio is the point on the Markowitz efficient frontier where a line from the risk-free asset is tangent to the Markowitz efficient frontier. All points on the interior of the Markowitz efficient frontier are inefficient portfolios in that they provide the same level of return with a higher level of risk or a lower level of return with the same amount of risk. When plotted together, the point at which the CML is tangent to the Markowitz efficient frontier is the optimal combination of risky assets, on the basis of market prices and market capitalizations. The optimal risky portfolio is the market portfolio.

Exhibit 3 Capital Market Line


Expected Portfolio Return $E(R_p)$
CML
Points above the CML are not achievable
$E(R_m)$
M
Efficient frontier
Individual Securities
R_f
Standard Deviation of Portfolio σ_p

The CML's intercept on the y-axis is the risk-free return (R_f) because that is the return associated with zero risk. The CML passes through the point represented by the market return, $E(R_m)$. With respect to capital market theory, any point above the CML is not achievable and any point below the CML is dominated by and inferior to any point on the CML.

Note that we identify the CML and CAL as lines even though they are a combination of two assets. Unlike a combination of two risky assets, which is usually not a straight line, a combination of the risk-free asset and a risky portfolio is a straight line, as illustrated below by computing the combination's risk and return.

Risk and return characteristics of the portfolio represented by the CML can be computed by using the return and risk expressions for a two-asset portfolio:

$$E(R_p) = w_1 R_f + (1 - w_1)E(R_m),$$

and

$$\sigma_p = \sqrt{w_1^2\sigma_f^2 + (1 - w_1)^2\sigma_m^2 + 2w_1(1 - w_1)\text{Cov}(R_f, R_m)}$$

The proportion invested in the risk-free asset is given by w_1, and the balance is invested in the market portfolio, $(1 - w_1)$. The risk of the risk-free asset is given by σ_f, the risk of the market is given by σ_m, the risk of the portfolio is given by σ_p, and the covariance between the risk-free asset and the market portfolio is represented by $\text{Cov}(R_f, R_m)$.

By definition, the standard deviation of the risk-free asset is zero. Because its risk is zero, the risk-free asset does not co-vary or move with any other asset. Therefore, its covariance with all other assets, including the market portfolio, is zero, making the first and third terms under the square root sign zero. As a result, the portfolio return and portfolio standard deviation can be simplified and rewritten as:

$$E(R_p) = w_1 R_f + (1 - w_1)E(R_m),$$

and

$$\sigma_p = (1 - w_1)\sigma_m$$

By substitution, we can express $E(R_p)$ in terms of σ_p. Substituting for w_1, we get:

$$E(R_p) = R_f + \left(\frac{E(R_m) - R_f}{\sigma_m}\right) \times \sigma_p$$

Note that the expression is in the form of a line, $y = a + bx$. The y-intercept is the risk-free rate, and the slope of the line referred to as the market price of risk is $[E(R_m) - R_f]/\sigma_m$. The CML has a positive slope because the market's risky return is larger than the risk-free return. As the amount of the total investment devoted to the market increases—that is, as we move up the line—both standard deviation (risk) and expected return increase.

EXAMPLE 1

Risk and Return on the CML

Mr. Miles is a first time investor and wants to build a portfolio using only US T-bills and an index fund that closely tracks the S&P 500 Index. The T-bills have a return of 5 percent. The S&P 500 has a standard deviation of 20 percent and an expected return of 15 percent.

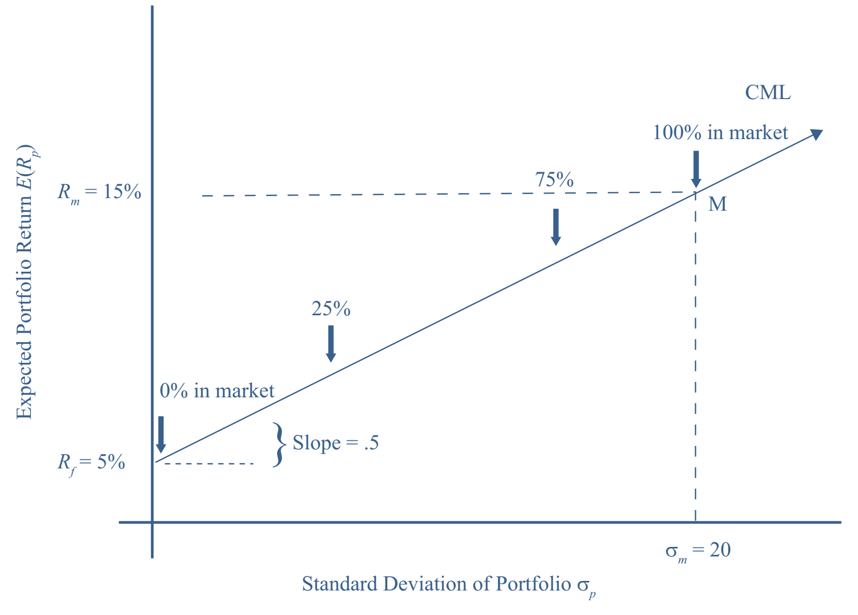
1. Draw the CML and mark the points where the investment in the market is 0 percent, 25 percent, 75 percent, and 100 percent.
2. Mr. Miles is also interested in determining the exact risk and return at each point.

Solution to 1:

We calculate the equation for the CML as $E(R_p) = 5\% + 0.50 \times \sigma_p$ by substituting the given information into the general CML equation. The intercept of the line is 5 percent, and its slope is 0.50. We can draw the CML by arbitrarily taking any two points on the line that satisfy the above equation.

Alternatively, the CML can be drawn by connecting the risk-free return of 5 percent on the y-axis with the market portfolio at (20 percent, 15 percent). The CML is shown in Exhibit 4.

Exhibit 4 Risk and Return on the CML



Solution to 2:

Return with 0 percent invested in the market = 5 percent, which is the risk-free return.

Standard deviation with 0 percent invested in the market = 0 percent because T-bills are not risky.

Return with 25 percent invested in the market = (0.75 × 5%) + (0.25 × 15%) = 7.5%.

Standard deviation with 25 percent invested in the market = 0.25 × 20% = 5%.

Return with 75 percent invested in the market = (0.25 × 5%) + (0.75 × 15%) = 12.50%.

Standard deviation with 75 percent invested in the market = 0.75 × 20% = 15%.

Return with 100 percent invested in the market = 15 percent, which is the return on the S&P 500.

Standard deviation with 100 percent invested in the market = 20 percent, which is the risk of the S&P 500.

4

CAPITAL MARKET THEORY: CML - LEVERAGED PORTFOLIOS

b explain the capital allocation line (CAL) and the capital market line (CML);

In the previous example, Mr. Miles evaluated an investment of between 0 percent and 100 percent in the market and the balance in T-bills. The line connecting R_f and M (market portfolio) in Exhibit 4 illustrates these portfolios with their respective levels of investment. At R_f, an investor is investing all of his or her wealth into risk-free securities, which is equivalent to lending 100 percent at the risk-free rate. At Point *M* he or she is holding the market portfolio and not lending any money at the risk-free rate. The combinations of the risk-free asset and the market portfolio, which may be achieved by the points between these two limits, are termed "lending" portfolios. In effect, the investor is lending part of his or her wealth at the risk-free rate.

If Mr. Miles is willing to take more risk, he may be able to move to the right of the market portfolio (Point *M* in Exhibit 4) by borrowing money and purchasing more of Portfolio M. Assume that he is able to borrow money at the same risk-free rate of interest, R_f, at which he can invest. He can then supplement his available wealth with borrowed money and construct a borrowing portfolio. If the straight line joining R_f and *M* is extended to the right of Point *M*, this extended section of the line represents borrowing portfolios. As one moves further to the right of Point *M*, an increasing amount of borrowed money is being invested in the market. This means that there is *negative* investment in the risk-free asset, which is referred to as a *leveraged position* in the risky portfolio. The particular point chosen on the CML will depend on the individual's utility function, which, in turn, will be determined by his risk and return preferences.

EXAMPLE 2

Risk and Return of a Leveraged Portfolio with Equal Lending and Borrowing Rates

Mr. Miles decides to set aside a small part of his wealth for investment in a portfolio that has greater risk than his previous investments because he anticipates that the overall market will generate attractive returns in the future. He assumes that he can borrow money at 5 percent and achieve the same return on the S&P 500 as before: an expected return of 15 percent with a standard deviation of 20 percent.

Calculate his expected risk and return if he borrows 25 percent, 50 percent, and 100 percent of his initial investment amount.

Solution:

The leveraged portfolio's standard deviation and return can be calculated in the same manner as before with the following equations:

$$E(R_p) = w_1 R_f + (1 - w_1)E(R_m)$$

and

$$\sigma_p = (1 - w_1)\sigma_m$$

The proportion invested in T-bills becomes negative instead of positive because Mr. Miles is borrowing money. If 25 percent of the initial investment is borrowed, $w_1 = -0.25$, and $(1 - w_1) = 1.25$, etc.

Return with $w_1 = -0.25 = (-0.25 \times 5\%) + (1.25 \times 15\%) = 17.5\%$.

Standard deviation with $w_1 = -0.25 = 1.25 \times 20\% = 25\%$.

Return with $w_1 = -0.50 = (-0.50 \times 5\%) + (1.50 \times 15\%) = 20.0\%$.

Standard deviation with $w_1 = -0.50 = 1.50 \times 20\% = 30\%$.

Return with $w_1 = -1.00 = (-1.00 \times 5\%) + (2.00 \times 15\%) = 25.0\%$.

Standard deviation with $w_1 = -1.00 = 2.00 \times 20\% = 40\%$.

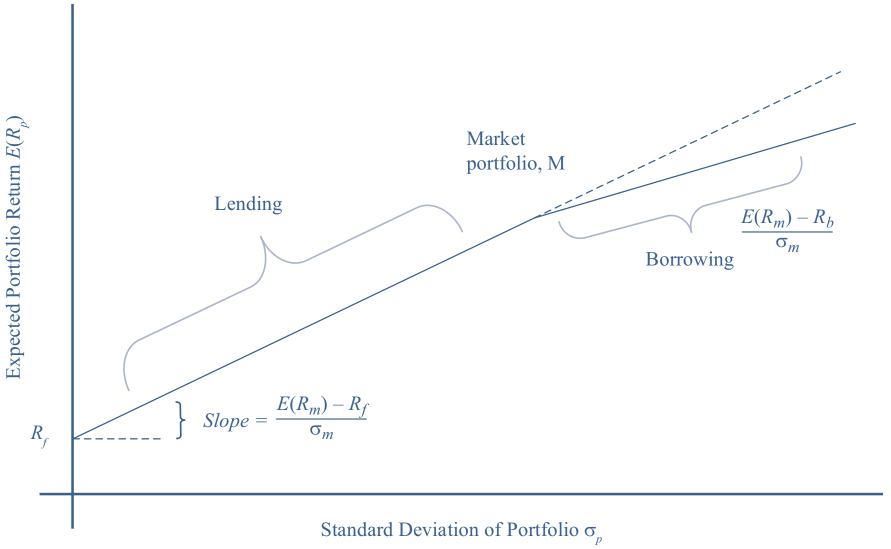
Note that negative investment (borrowing) in the risk-free asset provides a higher expected return for the portfolio but that higher return is also associated with higher risk.

4.1 Leveraged Portfolios with Different Lending and Borrowing Rates

Although we assumed that Mr. Miles can borrow at the same rate as the US government, it is more likely that he will have to pay a higher interest rate than the government because his ability to repay is not as certain as that of the government. Now consider that although Mr. Miles can invest (lend) at R_f, he can borrow at only R_b, a rate that is higher than the risk-free rate.

With different lending and borrowing rates, the CML will no longer be a single straight line. The line will have a slope of $[E(R_m) - R_f]/\sigma_m$ between Points R_f and M, where the lending rate is R_f, but will have a smaller slope of $[E(R_m) - R_b]/\sigma_m$ at points to the right of M, where the borrowing rate is R_b. Exhibit 5 illustrates the CML with different lending and borrowing rates.

Exhibit 5 CML with Different Lending and Borrowing Rates



The equations for the two lines are given below.

$$w_1 \geq 0\colon E\left(R_p\right) = R_f + \left(\frac{E\left(R_m\right) - R_f}{\sigma_m}\right) \times \sigma_p$$

and

$$w_1 < 0\colon E\left(R_p\right) = R_b + \left(\frac{E\left(R_m\right) - R_b}{\sigma_m}\right) \times \sigma_p$$

The first equation is for the line where the investment in the risk-free asset is zero or positive—that is, at M or to the left of M in Exhibit 5. The second equation is for the line where borrowing, or negative investment in the risk-free asset, occurs. Note that the only difference between the two equations is in the interest rates used for borrowing and lending.

All passive portfolios will lie on the kinked CML, although the investment in the risk-free asset may be positive (lending), zero (no lending or borrowing), or negative (borrowing). Leverage allows less risk-averse investors to increase the amount of risk they take by borrowing money and investing more than 100 percent in the passive portfolio.

EXAMPLE 3

Leveraged Portfolio with Different Lending and Borrowing Rates

Mr. Miles approaches his broker to borrow money against securities held in his portfolio. Even though Mr. Miles' loan will be secured by the securities in his portfolio, the broker's rate for lending to customers is 7 percent. Assuming a risk-free rate of 5 percent and a market return of 15 percent with a standard deviation of 20 percent, estimate Mr. Miles' expected return and risk if he invests 25 percent and 75 percent in the market and if he decides to borrow 25 percent and 75 percent of his initial investment and invest the money in the market.

Solution:

The unleveraged portfolio's standard deviation and return are calculated using the same equations as before:

$$E(R_p) = w_1 R_f + (1 - w_1)E(R_m),$$

and

$$\sigma_p = (1 - w_1)\sigma_m$$

The results are unchanged. The slope of the line for the unleveraged portfolio is 0.50, just as before:

Return with 25 percent invested in the market = (0.75 × 5%) + (0.25 × 15%) = 7.5%.

Standard deviation with 25 percent invested in the market = 0.25 × 20% = 5%.

Return with 75 percent invested in the market = (0.25 × 5%) + (0.75 × 15%) = 12.5%.

Standard deviation with 75 percent invested in the market = 0.75 × 20% = 15%.

For the leveraged portfolio, everything remains the same except that R_f is replaced with R_b.

$$E(R_p) = w_1 R_b + (1 - w_1)E(R_m),$$

and

$$\sigma_p = (1 - w_1)\sigma_{m.}$$

Return with $w_1 = -0.25 = (-0.25 \times 7\%) + (1.25 \times 15\%) = 17.0\%$.

Standard deviation with $w_1 = -0.25 = 1.25 \times 20\% = 25\%$.

Return with $w_1 = -0.75 = (-0.75 \times 7\%) + (1.75 \times 15\%) = 21.0\%$.

Standard deviation with $w_1 = -0.75 = 1.75 \times 20\% = 35\%$.

The risk and return of the leveraged portfolio is higher than that of the unleveraged portfolio. As Mr. Miles borrows more money to invest in the market, the expected return increases but so does the standard deviation of the portfolio.

The slope of the line for the leveraged portfolio is 0.40, compared with 0.50 for the unleveraged portfolio, which means that for every 1 percent increase in risk, the investor gets a 0.40 percent increase in expected return in the leveraged part of the portfolio, compared with a 0.50 percent increase in expected return in the unleveraged part of the portfolio. Only investors who are less risk averse will choose leveraged portfolios.

5 SYSTEMATIC AND NONSYSTEMATIC RISK

c explain systematic and nonsystematic risk, including why an investor should not expect to receive additional return for bearing nonsystematic risk;

In constructing a portfolio, it is important to understand the concept of correlation and how less than perfect correlation can diversify the risk of a portfolio. As a consequence, the risk of an asset held alone may be greater than the risk of that same asset when it is part of a portfolio. Because the risk of an asset varies from one environment to another, which kind of risk should an investor consider and how should that risk be priced? This section addresses the question of pricing of risk by decomposing the total risk of a security or a portfolio into systematic and nonsystematic risk. The meaning of these risks, how they are computed, and their relevance to the pricing of assets are also discussed.

5.1 Systematic Risk and Nonsystematic Risk

Systematic risk, also known as non-diversifiable or market risk, is the risk that affects the entire market or economy. In contrast, nonsystematic risk is the risk that pertains to a single company or industry and is also known as company-specific, industry-specific, diversifiable, or idiosyncratic risk.

Systematic risk is risk that cannot be avoided and is inherent in the overall market. It is non-diversifiable because it includes risk factors that are innate within the market and affect the market as a whole. Examples of factors that constitute systematic risk include interest rates, inflation, economic cycles, political uncertainty, and widespread natural disasters. These events affect the entire market, and there is no way to avoid their effect. Systematic risk can be magnified through selection or by using leverage, or diminished by including securities that have a low correlation with the portfolio, assuming they are not already part of the portfolio.

Nonsystematic risk is risk that is local or limited to a particular asset or industry that need not affect assets outside of that asset class. Examples of nonsystematic risk could include the failure of a drug trial or an airliner crash. All these events will directly affect their respective companies and possibly industries, but have no effect on assets that are far removed from these industries. Investors can avoid nonsystematic risk through diversification by forming a portfolio of assets that are not highly correlated with one another.

We will derive expressions for each kind of risk later in this reading. You will see that the sum of systematic variance and nonsystematic variance equals the total variance of the security or portfolio:

Total variance = Systematic variance + Nonsystematic variance

Although the equality relationship is between variances, you will find frequent references to total risk as the sum of systematic risk and nonsystematic risk. In those cases, the statements refer to variance, not standard deviation.

5.1.1 *Pricing of Risk*

Pricing or valuing an asset is equivalent to estimating its expected rate of return. If an asset has a known terminal value, such as the face value of a bond, then a lower current price implies a higher future return and a higher current price implies a lower future return. The relationship between price and return can also be observed in the valuation expression shown in Section 2. Therefore, we will occasionally use price and return interchangeably when discussing the price of risk.

Consider an asset with both systematic and nonsystematic risk. Assume that both kinds of risk are priced—that is, you receive a return for both systematic risk and nonsystematic risk. What will you do? Realizing that nonsystematic risk can be diversified away, you would buy assets that have a large amount of nonsystematic risk. Once you have bought those assets with nonsystematic risk, you would diversify, or reduce that risk, by including other assets that are not highly correlated. In the process, you will minimize nonsystematic risk and eventually eliminate it altogether from your portfolio. You would now have a diversified portfolio with only systematic risk, yet you would be compensated for nonsystematic risk that you no longer have. Just like everyone else, you would have an incentive to take on more and more diversifiable risk because you are compensated for it even though you can get rid of it. The demand for diversifiable risk would keep increasing until its price becomes infinite and its expected return falls to zero. This means that our initial assumption of a non-zero return for diversifiable risk was incorrect and that the correct assumption is zero return for diversifiable risk. Therefore, according to theory, in an efficient market no incremental reward is earned for taking on diversifiable risk.

We have argued that investors should not be compensated for taking on nonsystematic risk. Therefore, investors who have nonsystematic risk must diversify it away by investing in many industries, many countries, and many asset classes. Because future returns are unknown and it is not possible to pick only winners, diversification helps in offsetting poor returns in one asset class by garnering good returns in another asset class, thereby reducing the overall risk of the portfolio. In contrast, investors must be compensated for accepting systematic risk because that risk cannot be diversified away. If investors do not receive a return commensurate with the amount of systematic risk they are taking, they will refuse to accept systematic risk.

In summary, according to theory, systematic or non-diversifiable risk is priced and investors are compensated for holding assets or portfolios based only on that investment's systematic risk. Investors do not receive any return for accepting non-systematic or diversifiable risk. Therefore, it is in the interest of risk-averse investors to hold only well-diversified portfolios.

EXAMPLE 4

Systematic and Nonsystematic Risk

1. Describe the systematic and nonsystematic risk components of the following assets:
   - **A** A risk-free asset, such as a three-month Treasury bill
   - **B** The market portfolio, such as the S&P 500.
2. Consider two assets, A and B. Asset A has twice the amount of total risk as Asset B. For Asset A, systematic risk comprises two-thirds of total risk. For Asset B, all of total risk is systematic risk. Which asset should have a higher expected rate of return?

Solution to 1A:

By definition, a risk-free asset has no risk. Therefore, a risk-free asset has zero systematic risk and zero nonsystematic risk.

Solution to 1B:

As we mentioned earlier, a market portfolio is a diversified portfolio, one in which no more risk can be diversified away. We have also described it as an efficient portfolio. Therefore, a market portfolio does not contain any nonsystematic risk.

Solution to 2:

Based on the facts given, Asset A's systematic risk is one-third greater than Asset B's systematic risk. Because only systematic risk is priced or receives a return, the expected rate of return must be higher for Asset A.

6

RETURN GENERATING MODELS

d explain return generating models (including the market model) and their uses;

As previously mentioned, in order to form the market portfolio, you should combine all available risky assets. Knowledge of the correlations among those assets allows us to estimate portfolio risk. You also learned that a fully diversified portfolio will include all asset classes and essentially all assets in those asset classes. The work required for construction of the market portfolio is formidable. For example, for a portfolio of 1,000 assets, we will need 1,000 return estimates, 1,000 standard deviation estimates, and 499,500 (1,000 × 999 ÷ 2) correlations. Other related questions that arise with this analysis are whether we really need all 1,000 assets and what happens if there are errors in these estimates.

An alternate method of constructing an optimal portfolio is simpler and easier to implement. An investor begins with a known portfolio, such as the S&P 500, and then adds other assets one at a time on the basis of the asset's standard deviation, expected return, and impact on the portfolio's risk and return. This process continues until the addition of another asset does not have a significant impact on the performance of the portfolio. The process requires only estimates of systematic risk for each asset because investors will not be compensated for nonsystematic risk. Expected returns can be calculated by using return-generating models, as we will discuss in this section. In addition to using return-generating models, we will also decompose total variance into systematic variance and nonsystematic variance and establish a formal relationship between systematic risk and return. In the next section, we will expand on this discussion and introduce the CAPM as the preferred return-generating model.

6.1 Return-Generating Models

A **return-generating model** is a model that can provide an estimate of the expected return of a security given certain parameters. If systematic risk is the only relevant parameter for return, then the return-generating model will estimate the expected return for any asset given the level of systematic risk.

As with any model, the quality of estimates of expected return will depend on the quality of input estimates and the accuracy of the model. Because it is difficult to decide which factors are appropriate for generating returns, the most general form of a return-generating model is a multi-factor model. A **multi-factor model** allows more than one variable to be considered in estimating returns and can be built using different kinds of factors, such as macroeconomic, fundamental, and statistical factors.

Macroeconomic factor models use economic factors that are correlated with security returns. These factors may include economic growth, the interest rate, the inflation rate, productivity, employment, and consumer confidence. Past relationships with returns are estimated to obtain parameter estimates, which are, in turn, used for computing expected returns. Fundamental factor models analyze and use relationships between security returns and the company's underlying fundamentals, such as, for example, earnings, earnings growth, cash flow generation, investment in research, advertising, and number of patents. Finally, in a statistical factor model, historical and cross-sectional return data are analyzed to identify factors that explain variance or covariance in observed returns. These statistical factors, however, may or may not have an economic or fundamental connection to returns. For example, the conference to which the American football Super Bowl winner belongs, whether the American Football Conference or the National Football Conference, may be a factor in US stock returns, but no obvious economic connection seems to exist between the winner's conference and US stock returns. Moreover, data mining may generate many spurious factors that are devoid of any economic meaning. Because of this limitation, analysts prefer the macroeconomic and fundamental factor models for specifying and estimating return-generating models.

A general return-generating model is expressed in the following manner:

$$E(R_i) - R_f = \sum_{j=1}^{k} \beta_{ij} E(F_j) = \beta_{i1}\left[E(R_m) - R_f\right] + \sum_{j=2}^{k} \beta_{ij} E(F_j)$$

The model has k factors, $E(F_1)$, $E(F_2)$, ... $E(F_k)$. The coefficients, β_{ij}, are the factor weights (sometimes called factor loadings) associated with each factor. The left-hand side of the model has the expected excess return (i.e., the expected return over the risk-free rate). The right-hand side provides the risk factors that would generate the return or premium required to assume that risk. We have separated out one factor, $E(R_m)$, which represents the market return. All models contain return on the market portfolio as a key factor.

Three-Factor and Four-Factor Models

Eugene Fama and Kenneth French[1] suggested that a return-generating model for stock returns should include relative market capitalization of the company ("size") relative book-to-market value of the company in addition to beta. Fama and French found that past returns could be explained better with their model than with other models available at that time, most notably, the capital asset pricing model. Mark Carhart (1997) extended the Fama and French model by adding another factor: momentum, defined as relative past stock returns.

The Single-Index Model

The simplest form of a return-generating model is a single-factor linear model, in which only one factor is considered. The most common implementation is a single-index model, which uses the market factor in the following form: $E(R_i) - R_f = \beta_i[E(R_m) - R_f]$.

---

[1] Fama and French (1992).

Although the single-index model is simple, it fits nicely with the capital market line. Recall that the CML is linear, with an intercept of R_f and a slope of $[E(R_m) - R_f]/\sigma_m$. We can rewrite the CML by moving the intercept to the left-hand side of the equation, rearranging the terms, and generalizing the subscript from p to i, for any security:

$$E(R_i) - R_f = \left(\frac{\sigma_i}{\sigma_m}\right)\left[E(R_m) - R_f\right]$$

The factor loading or factor weight, σ_i/σ_m, refers to the ratio of total security risk to total market risk. To obtain a better understanding of factor loading and to illustrate that the CML reduces to a single-index model, we decompose total risk into its components.

6.2 Decomposition of Total Risk for a Single-Index Model

With the introduction of return-generating models, particularly the single-index model, we are able to decompose total variance into systematic and nonsystematic variances. Instead of using expected returns in the single index, let us use realized returns. The difference between expected returns and realized returns is attributable to non-market changes, as an error term, e_i, in the second equation below:

$$E(R_i) - R_f = \beta_i[E(R_m) - R_f]$$

and

$$R_i - R_f = \beta_i(R_m - R_f) + e_i$$

The variance of realized returns can be expressed in the equation below (note that R_f is a constant). We can further drop the covariance term in this equation because, by definition, any non-market return is uncorrelated with the market. Thus, we are able to decompose total variance into systematic and nonsystematic variances in the second equation below:

$$\sigma_i^2 = \beta_i^2\sigma_m^2 + \sigma_e^2 + 2\mathrm{Cov}(R_m, e_i)$$

Total variance = Systematic variance + Nonsystematic variance, which can be written as

$$\sigma_i^2 = \beta_i^2\sigma_m^2 + \sigma_e^2$$

Total risk can be expressed as

$$\sigma_i = \sqrt{\beta_i^2\sigma_m^2 + \sigma_e^2}$$

Because nonsystematic risk is zero for well-diversified portfolios, such as the market portfolio, the total risk of a market portfolio and other similar portfolios is only systematic risk, which is $\beta_i\sigma_m$. We can now return to the CML discussed in the previous subsection and replace σ_i with $\beta_i\sigma_m$ because the CML assumes that the market is a diversified portfolio. By making this substitution for the above equation, we get the following single-index model:

$$E(R_i) - R_f = \left(\frac{\sigma_i}{\sigma_m}\right) \times \left[E(R_m) - R_f\right] = \left(\frac{\beta_i\sigma_m}{\sigma_m}\right) \times \left[E(R_m) - R_f\right],$$

$$E(R_i) - R_f = \beta_i[E(R_m) - R_f]$$

Thus, the CML, which holds only for well-diversified portfolios, is fully consistent with a single-index model.

In summary, total variance may be decomposed into systematic and nonsystematic variances and the CML is the same as a single-index model for diversified portfolios.

6.3 Return-Generating Models: The Market Model

The most common implementation of a single-index model is the **market model**, in which the market return is the single factor or single index. In principle, the market model and the single-index model are similar. The difference is that the market model is easier to work with and is normally used for estimating beta risk and computing abnormal returns. The market model is

$$R_i = \alpha_i + \beta_i R_m + e_i$$

To be consistent with the previous section, $\alpha_i = R_f(1 - \beta)$. The intercept, α_i, and slope coefficient, β_i, can be estimated by using historical security and market returns. These parameter estimates are then used to predict company-specific returns that a security may earn in a future period. Assume that a regression of Wal-Mart's historical daily returns on S&P 500 daily returns gives an α_i of 0.0001 and a β_i of 0.9. Thus, Wal-Mart's expected daily return = 0.0001 + 0.90 × R_m. If, on a given day the market rises by 1 percent and Wal-Mart's stock rises by 2 percent, then Wal-Mart's company-specific return (e_i) for that day = $R_i - E(R_i) = R_i - (\alpha_i + \beta_i R_m) = 0.02 - (0.0001 + 0.90 \times 0.01) = 0.0109$, or 1.09%. In other words, Wal-Mart earned an abnormal return of 1.09 percent on that day.

CALCULATION AND INTERPRETATION OF BETA

7

e calculate and interpret beta;

We begin with the single-index model introduced earlier using realized returns and rewrite it as

$$R_i = (1 - \beta_i)R_f + \beta_i \times R_m + e_i$$

Because systematic risk depends on the correlation between the asset and the market, we can arrive at a measure of systematic risk from the covariance between R_i and R_m, where R_i is defined using the above equation. Note that the risk-free rate is a constant, so the first term in R_i drops out.

$$\begin{aligned}\mathrm{Cov}(R_i, R_m) &= \mathrm{Cov}(\beta_i \times R_m + e_i, R_m) \\ &= \beta_i \mathrm{Cov}(R_m, R_m) + \mathrm{Cov}(e_i, R_m) \\ &= \beta_i \sigma_m^2 + 0\end{aligned}$$

The first term is beta multiplied by the variance of R_m. Because the error term is uncorrelated with the market, the second term drops out. Then, we can rewrite the equation in terms of beta as follows:

$$\beta_i = \frac{\mathrm{Cov}(R_i, R_m)}{\sigma_m^2} = \frac{\rho_{i,m}\sigma_i\sigma_m}{\sigma_m^2} = \frac{\rho_{i,m}\sigma_i}{\sigma_m}$$

The above formula shows the expression for beta, β_i, which is similar to the factor loading in the single-index model presented earlier. For example, if the correlation between an asset and the market is 0.70 and the asset and market have standard deviations of return of 0.25 and 0.15, respectively, the asset's beta would be (0.70)(0.25)/0.15 = 1.17. If the asset's covariance with the market and market variance were given as 0.026250 and 0.02250, respectively, the calculation would be 0.026250/0.02250 = 1.17. The beta in the market model includes an adjustment for the correlation between asset i and the market because the market model covers all assets whereas the CML works only for fully diversified portfolios.

As shown in the above equation, **beta** is a measure of how sensitive an asset's return is to the market as a whole and is calculated as the covariance of the return on *i* and the return on the market divided by the variance of the market return; that expression is equivalent to the product of the asset's correlation with the market with a ratio of standard deviations of return (i.e., the ratio of the asset's standard deviation to the market's). As we have shown, beta captures an asset's systematic risk, or the portion of an asset's risk that cannot be eliminated by diversification. The variances and correlations required for the calculation of beta are usually based on historical returns.

A positive beta indicates that the return of an asset follows the general market trend, whereas a negative beta shows that the return of an asset generally follows a trend that is opposite to that of the market. In other words, a positive beta indicates that the return of an asset moves in the same direction of the market, whereas a negative beta indicates that the return of an asset moves in the opposite direction of the market. A risk-free asset's beta is zero because its covariance with other assets is zero. In other words, a beta of zero indicates that the asset's return has no correlation with movements in the market. The market's beta can be calculated by substituting σ_m for σ_i in the numerator. Also, any asset's correlation with itself is 1, so the beta of the market is 1:

$$\beta_i = \frac{\rho_{i,m}\sigma_i}{\sigma_m} = \frac{\rho_{m,m}\sigma_m}{\sigma_m} = 1$$

Because the market's beta is 1, the average beta of stocks in the market, by definition, is 1. In terms of correlation, most stocks, especially in developed markets, tend to be highly correlated with the market, with correlations in excess of 0.70. Some US broad market indexes, such as the S&P 500, the Dow Jones 30, and the NASDAQ 100, have even higher correlations that are in excess of 0.90. The correlations among different sectors are also high, which shows that companies have similar reactions to the same economic and market changes. As a consequence and as a practical matter, finding assets that have a consistently negative beta is unusual because of the market's broad effects on all assets.

EXAMPLE 5

Calculation of Beta

Assuming that the risk (standard deviation) of the market is 25 percent, calculate the beta for the following assets:

1. A short-term US Treasury bill.
2. Gold, which has a standard deviation equal to the standard deviation of the market but a zero correlation with the market.
3. A new emerging market that is not currently included in the definition of "market"—the emerging market's standard deviation is 60 percent, and the correlation with the market is –0.1.
4. An initial public offering or new issue of stock with a standard deviation of 40 percent and a correlation with the market of 0.7 (IPOs are usually very risky but have a relatively low correlation with the market).

We use the formula for beta in answering the above questions: $\beta_i = \dfrac{\rho_{i,m}\sigma_i}{\sigma_m}$

Solution to 1:

By definition, a short-term US Treasury bill has zero risk. Therefore, its beta is zero.

Solution to 2:

Because the correlation of gold with the market is zero, its beta is zero.

Solution to 3:

Beta of the emerging market is –0.1 × 0.60 ÷ 0.25 = –0.24.

Solution to 4:

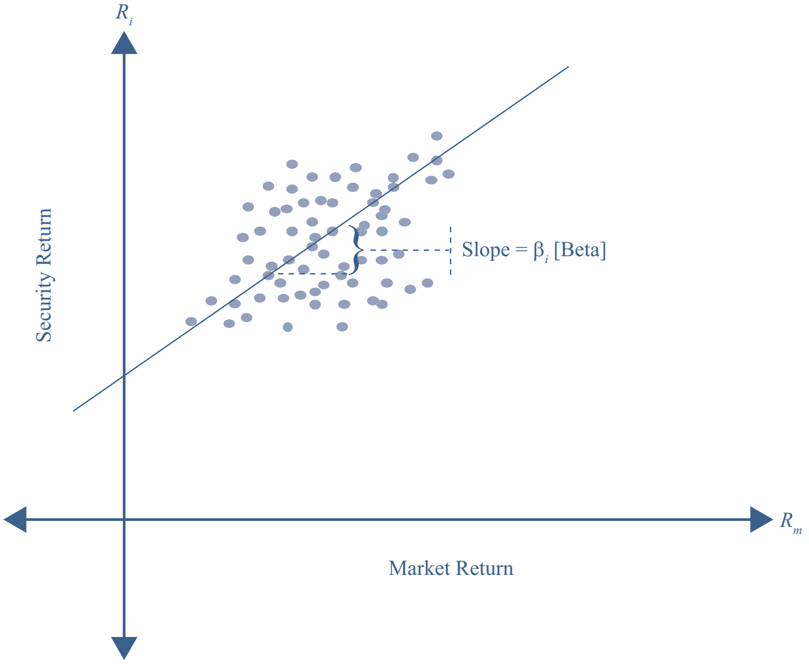
Beta of the initial public offering is 0.7 × 0.40 ÷ 0.25 = 1.12.

7.1 Estimation of Beta

An alternative and more practical approach is to estimate beta directly by using the market model described above. The market model, $R_i = \alpha_i + \beta_i R_m + e_i$, is estimated by using regression analysis, which is a statistical process that evaluates the relationship between a given variable (the dependent variable) and one or more other (independent) variables. Historical security returns (R_i) and historical market returns (R_m) are inputs used for estimating the two parameters α_i and β_i.

Regression analysis is similar to plotting all combinations of the asset's return and the market return (R_i, R_m) and then drawing a line through all points such that it minimizes the sum of squared linear deviations from the line. Exhibit 6 illustrates the market model and the estimated parameters. The intercept, α_i (sometimes referred to as the constant), and the slope term, β_i, are all that is needed to define the security characteristic line and obtain beta estimates.

Exhibit 6 Beta Estimation Using a Plot of Security and Market Returns



Although beta estimates are important for forecasting future levels of risk, there is much concern about their accuracy. In general, shorter periods of estimation (e.g., 12 months) represent betas that are closer to the asset's current level of systematic risk. Shorter period beta estimates, however, are also less accurate than beta estimates measured over three to five years because they may be affected by special events in that short period. Although longer period beta estimates are more accurate, they may be a poor representation of future expectations, especially if major changes in the asset have occurred. Therefore, it is necessary to recognize that estimates of beta, whether obtained through calculation or regression analysis, may or may not represent current or future levels of an asset's systematic risk.

7.2 Beta and Expected Return

Although the single-index model, also called the **capital asset pricing model** (CAPM), will be discussed in greater detail in the next section, we will use the CAPM in this section to estimate returns, given asset betas. The CAPM is usually written with the risk-free rate on the right-hand side:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f]$$

The model shows that the primary determinant of expected return for a security is its beta, or how well the security correlates with the market. The higher the beta of an asset, the higher its expected return will be. Assets with a beta greater than 1 have an expected return that is higher than the market return, whereas assets with a beta of less than 1 have an expected return that is less than the market return.

In certain cases, assets may require a return less than the risk-free return. For example, if an asset's beta is negative, the required return will be less than the risk-free rate. When combined with the market, the asset reduces the risk of the overall portfolio, which makes the asset very valuable. Insurance is one such asset. Insurance gives a positive return when the insured's wealth is reduced because of a catastrophic loss. In the absence of such a loss or when the insured's wealth is growing, the insured is required to pay an insurance premium. Thus, insurance has a negative beta and a negative expected return, but helps in reducing overall risk.

EXAMPLE 6

Calculation of Expected Return

1 Alpha Natural Resources (ANR), a coal producer, buys a large but privately held coal producer in China. As a result of the cross-border acquisition of a private company, ANR's standard deviation of returns is reduced from 50 percent to 30 percent and its correlation with the market falls from 0.95 to 0.75. Assume that the standard deviation and return of the market remain unchanged at 25 percent and 10 percent, respectively, and that the risk-free rate is 3 percent.

   A Calculate the beta of ANR stock and its expected return before the acquisition.

   B Calculate the expected return after the acquisition.

Solution to 1A:

Using the formula for β_i, we can calculate β_i and then the return.

$$\beta_i = \frac{\rho_{i,m}\sigma_i}{\sigma_m} = \frac{0.95 \times 0.50}{0.25} = 1.90$$

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] = 0.03 + 1.90 \times (0.10 - 0.03) = 0.163 = 16.3\%$$

Solution to 1B:

We follow the same procedure but with the after-acquisition correlation and risk.

$$\beta_i = \frac{\rho_{i,m}\sigma_i}{\sigma_m} = \frac{0.75 \times 0.30}{0.25} = 0.90$$

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] = 0.03 + 0.90 \times (0.10 - 0.03) = 0.093 = 9.3\%$$

The market risk premium is 7 percent (10% – 3%). As the beta changes, the change in the security's expected return is the market risk premium multiplied by the change in beta. In this scenario, ANR's beta decreased by 1.0, so the new expected return for ANR is 7 percentage points lower.

2 Mr. Miles observes the strong demand for iPods and iPhones and wants to invest in Apple stock. Unfortunately, Mr. Miles doesn't know the return he should expect from his investment. He has been given a risk-free rate of 3 percent, a market return of 10 percent, and Apple's beta of 1.5.

- **A** Calculate Apple's expected return.
- **B** An analyst looking at the same information decides that the past performance of Apple is not representative of its future performance. He decides that, given the increase in Apple's market capitalization, Apple acts much more like the market than before and thinks Apple's beta should be closer to 1.1. What is the analyst's expected return for Apple stock?

Solution to 2A:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] = 0.03 + 1.5 \times (0.10 - 0.03) = 0.135 = 13.5\%$$

Solution to 2B:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] = 0.03 + 1.1 \times (0.10 - 0.03) = 0.107 = 10.7\%$$

This example illustrates the lack of connection between estimation of past returns and projection into the future. Investors should be aware of the limitations of using past returns for estimating future returns.

CAPITAL ASSET PRICING MODEL: ASSUMPTIONS AND THE SECURITY MARKET LINE

8

f explain the capital asset pricing model (CAPM), including its assumptions, and the security market line (SML);

g calculate and interpret the expected return of an asset using the CAPM;

The capital asset pricing model is one of the most significant innovations in portfolio theory. The model is simple, yet powerful; is intuitive, yet profound. The CAPM was introduced independently by William Sharpe, John Lintner, Jack Treynor, and Jan Mossin and builds on Harry Markowitz's earlier work on diversification and modern portfolio theory.[2] The model provides a linear expected return–beta relationship that precisely determines the expected return given the beta of an asset. In doing so, it makes the transition from total risk to systematic risk, the primary determinant of expected return. Recall the following equation:

---

[2] See, for example, Markowitz (1952), Sharpe (1964), Lintner (1965a, 1965b), Treynor (1961, 1962), and Mossin (1966).

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f]$$

The CAPM asserts that the expected returns of assets vary only by their systematic risk as measured by beta. Two assets with the same beta will have the same expected return irrespective of the nature of those assets. Given the relationship between risk and return, all assets are defined only by their beta risk, which we will explain as the assumptions are described.

In the remainder of this section, we will examine the assumptions made in arriving at the CAPM and the limitations those assumptions entail. Second, we will implement the CAPM through the security market line to price any portfolio or asset, both efficient and inefficient. Finally, we will discuss ways in which the CAPM can be applied to investments, valuation, and capital budgeting.

8.1 Assumptions of the CAPM

Similar to all other models, the CAPM ignores many of the complexities of financial markets by making simplifying assumptions. These assumptions allow us to gain important insights into how assets are priced without complicating the analysis. Once the basic relationships are established, we can relax the assumptions and examine how our insights need to be altered. Some of these assumptions are constraining, whereas others are benign. And other assumptions affect only a particular set of assets or only marginally affect the hypothesized relationships.

1 ***Investors are risk-averse, utility-maximizing, rational individuals.***

Risk aversion means that investors expect to be compensated for accepting risk. Note that the assumption does not require investors to have the same degree of risk aversion; it only requires that they are averse to risk. Utility maximization implies that investors want higher returns, not lower returns, and that investors always want more wealth (i.e., investors are never satisfied). Investors are understood to be rational in that they correctly evaluate and analyze available information to arrive at rational decisions. Although rational investors may use the same information to arrive at different estimates of expected risk and expected returns, homogeneity among investors (see Assumption 4) requires that investors be rational individuals.

Risk aversion and utility maximization are generally accepted as reflecting a realistic view of the world. Yet, rationality among investors has been questioned because investors may allow their personal biases and experiences to disrupt their decision making, resulting in suboptimal investments. Nonetheless, the model's results are unaffected by such irrational behavior as long as it does not affect prices in a significant manner (i.e., the trades of irrational investors cancel each other or are dominated by the trades of rational investors).

2 ***Markets are frictionless, including no transaction costs and no taxes.***

Frictionless markets allow us to abstract the analysis from the operational characteristics of markets. In doing so, we do not allow the risk–return relationship to be affected by, for example, the trading volume on the New York Stock Exchange or the difference between buying and selling prices. Specifically, frictionless markets do not have transaction costs, taxes, or any costs or restrictions on short selling. We also assume that borrowing and lending at the risk-free rate is possible.

The transaction costs of many large institutions are negligible, and many institutions do not pay taxes. Even the presence of non-zero transaction costs, taxes, or the inability to borrow at the risk-free rate does not materially affect the general conclusions of the CAPM. Costs of short selling or restrictions on short selling, however, can introduce an upward bias in asset prices, potentially jeopardizing important conclusions of the CAPM.

3 ***Investors plan for the same single holding period.***

The CAPM is a single-period model, and all investor decisions are made on the basis of that one period. The assumption of a single period is applied for convenience because working with multi-period models is more difficult. A single-period model, however, does not allow learning to occur, and bad decisions can persist. In addition, maximizing utility at the end of a multi-period horizon may require decisions in certain periods that may seem suboptimal when examined from a single-period perspective. Nonetheless, the single holding period does not severely limit the applicability of the CAPM to multi-period settings.

4 ***Investors have homogeneous expectations or beliefs.***

This assumption means that all investors analyze securities in the same way using the same probability distributions and the same inputs for future cash flows. In addition, given that they are rational individuals, the investors will arrive at the same valuations. Because their valuations of all assets are identical, they will generate the same optimal risky portfolio, which we call the market portfolio.

The assumption of homogeneous beliefs can be relaxed as long as the differences in expectations do not generate significantly different optimal risky portfolios.

5 ***All investments are infinitely divisible.***

This assumption implies that an individual can invest as little or as much as he or she wishes in an asset. This supposition allows the model to rely on continuous functions rather than on discrete jump functions. The assumption is made for convenience only and has an inconsequential impact on the conclusions of the model.

6 ***Investors are price takers.***

The CAPM assumes that there are many investors and that no investor is large enough to influence prices. Thus, investors are price takers, and we assume that security prices are unaffected by investor trades. This assumption is generally true because even though investors may be able to affect prices of small stocks, those stocks are not large enough to affect the primary results of the CAPM.

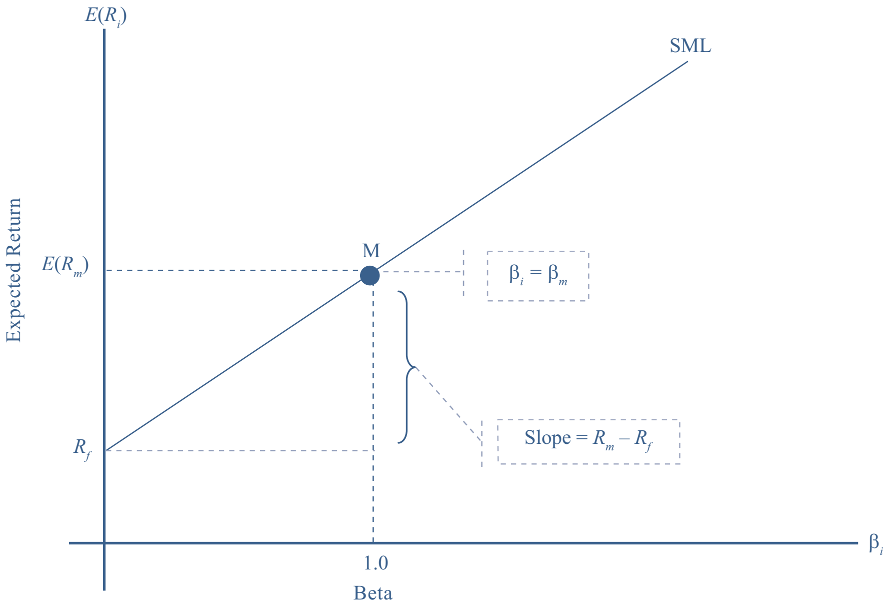
The main objective of these assumptions is to create a marginal investor who rationally chooses a mean–variance-efficient portfolio in a predictable fashion. We assume away any inefficiency in the market from both operational and informational perspectives. Although some of these assumptions may seem unrealistic, relaxing most of them will have only a minor influence on the model and its results. Moreover, the CAPM, with all its limitations and weaknesses, provides a benchmark for comparison and for generating initial return estimates.

8.2 The Security Market Line

In this subsection, we apply the CAPM to the pricing of securities. The **security market line** (SML) is a graphical representation of the capital asset pricing model with beta, reflecting systematic risk, on the *x*-axis and expected return on the *y*-axis. Using the same concept as the capital market line, the SML intersects the *y*-axis at the risk-free rate of return, and the slope of this line is the market risk premium, $R_m - R_f$. Recall that the capital market line (CML) does not apply to all securities or assets but only to portfolios on the efficient frontier. The efficient frontier gives optimal combinations of expected return and total risk. In contrast, the security market line applies to any security, efficient or not. Total risk and systematic risk are equal only for efficient portfolios because those portfolios have no diversifiable risk remaining.

Exhibit 7 is a graphical representation of the CAPM, the security market line. As shown earlier in this reading, the beta of the market is 1 (*x*-axis) and the market earns an expected return of R_m (*y*-axis). Using this line, it is possible to calculate the expected return of an asset. The next example illustrates the beta and return calculations.

Exhibit 7 The Security Market Line



EXAMPLE 7

Security Market Line and Expected Return

1. Suppose the risk-free rate is 3 percent, the expected return on the market portfolio is 13 percent, and its standard deviation is 23 percent. An Indian company, Bajaj Auto, has a standard deviation of 50 percent but is uncorrelated with the market. Calculate Bajaj Auto's beta and expected return.
2. Suppose the risk-free rate is 3 percent, the expected return on the market portfolio is 13 percent, and its standard deviation is 23 percent. A German company, Mueller Metals, has a standard deviation of 50 percent and a correlation of 0.65 with the market. Calculate Mueller Metal's beta and expected return.

Solution to 1:

Using the formula for β_i, we can calculate β_i and then the return.

$$\beta_i = \frac{\rho_{i,m}\sigma_i}{\sigma_m} = \frac{0.0 \times 0.50}{0.23} = 0$$

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] = 0.03 + 0 \times (0.13 - 0.03) = 0.03 = 3.0\%$$

Because of its zero correlation with the market portfolio, Bajaj Auto's beta is zero. Because the beta is zero, the expected return for Bajaj Auto is the risk-free rate, which is 3 percent.

Solution to 2:

Using the formula for β_i, we can calculate β_i and then the return.

$$\beta_i = \frac{\rho_{i,m}\sigma_i}{\sigma_m} = \frac{0.65 \times 0.50}{0.23} = 1.41$$

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] = 0.03 + 1.41\times (0.13 - 0.03) = 0.171 = 17.1\%$$

Because of the high degree of correlation with the market, the beta for Mueller Metals is 1.41 and the expected return is 17.1 percent. Because Mueller Metals has systematic risk that is greater than that of the market, it has an expected return that exceeds the expected return of the market.

8.2.1 *Portfolio Beta*

As we stated above, the security market line applies to all securities. But what about a combination of securities, such as a portfolio? Consider two securities, 1 and 2, with a weight of w_i in Security 1 and the balance in Security 2. The return for the two securities and return of the portfolio can be written as:

$$\begin{aligned}
E(R_1) &= R_f + \beta_1\left[E(R_m) - R_f\right] \\
E(R_2) &= R_f + \beta_2\left[E(R_m) - R_f\right] \\
E(R_p) &= w_1 E(R_1) + w_2 E(R_2) \\
&= w_1 R_f + w_1\beta_1\left[E(R_m) - R_f\right] + w_2 R_f + w_2\beta_2\left[E(R_m) - R_f\right] \\
&= R_f + (w_1\beta_1 + w_2\beta_2)\left[E(R_m) - R_f\right]
\end{aligned}$$

The last equation gives the expression for the portfolio's expected return. From this equation, we can conclude that the portfolio's beta = $w_1\beta_1 + w_2\beta_2$. In general, the portfolio beta is a weighted sum of the betas of the component securities and is given by:

$$\beta_p = \sum_{i=1}^{n} w_i\beta_i; \sum_{i=1}^{n} w_i = 1$$

The portfolio's return given by the CAPM is

$$E(R_p) = R_f + \beta_p[E(R_m) - R_f]$$

This equation shows that a linear relationship exists between the expected return of a portfolio and the systematic risk of the portfolio as measured by β_p.

EXAMPLE 8

Portfolio Beta and Return

You invest 20 percent of your money in the risk-free asset, 30 percent in the market portfolio, and 50 percent in RedHat, a US stock that has a beta of 2.0. Given that the risk-free rate is 4 percent and the market return is 16 percent, what are the portfolio's beta and expected return?

Solution:

The beta of the risk-free asset = 0, the beta of the market = 1, and the beta of RedHat is 2.0. The portfolio beta is

$$\beta_p = w_1\beta_1 + w_2\beta_2 + w_3\beta_3 = (0.20 \times 0.0) + (0.30 \times 1.0) + (0.50 \times 2.0) = 1.30$$

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] = 0.04 + 1.30 \times (0.16 - 0.04) = 0.196 = 19.6\%$$

The portfolio beta is 1.30, and its expected return is 19.6 percent.

Alternate Method:

Another method for calculating the portfolio's return is to calculate individual security returns and then use the portfolio return formula (i.e., weighted average of security returns) to calculate the overall portfolio return.

Return of the risk-free asset = 4 percent; return of the market = 16 percent

RedHat's return based on its beta = 0.04 + 2.0 × (0.16 – 0.04) = 0.28

Portfolio return = (0.20 × 0.04) + (0.30 × 0.16) + (0.50 × 0.28) = 0.196 = 19.6%

Not surprisingly, the portfolio return is 19.6 percent, as calculated in the first method.

9 CAPITAL ASSET PRICING MODEL: APPLICATIONS

g calculate and interpret the expected return of an asset using the CAPM;

h describe and demonstrate applications of the CAPM and the SML;

The CAPM offers powerful and intuitively appealing predictions about risk and the relationship between risk and return. The CAPM is not only important from a theoretical perspective but is also used extensively in practice. In this section, we will discuss some common applications of the model. When applying these tools to different scenarios, it is important to understand that the CAPM and the SML are functions that give an indication of what the return in the market *should* be, given a certain level of risk. The actual return may be quite different from the expected return.

Applications of the CAPM include estimates of the expected return for capital budgeting, comparison of the actual return of a portfolio or portfolio manager with the CAPM return for performance appraisal, and the analysis of alternate return estimates and the CAPM returns as the basis for security selection. The applications are discussed in more detail in this section.

9.1 Estimate of Expected Return

Given an asset's systematic risk, the expected return can be calculated using the CAPM. Recall that the price of an asset is the sum of all future cash flows discounted at the required rate of return, where the discount rate or the required rate of return is commensurate with the asset's risk. The expected rate of return obtained from the CAPM is normally the first estimate that investors use for valuing assets, such as stocks, bonds, real estate, and other similar assets. The required rate of return from the CAPM is also used for capital budgeting and determining the economic feasibility of projects. Again, recall that when computing the net present value of a project, investments and net revenues are considered cash flows and are discounted at the required rate of return. The required rate of return, based on the project's risk, is calculated using the CAPM.

Because risk and return underlie almost all aspects of investment decision making, it is not surprising that the CAPM is used for estimating expected return in many scenarios. Other examples include calculating the cost of capital for regulated companies by regulatory commissions and setting fair insurance premiums. The next example shows an application of the CAPM to capital budgeting.

EXAMPLE 9

Application of the CAPM to Capital Budgeting

GlaxoSmithKline Plc is examining the economic feasibility of developing a new medicine. The initial investment in Year 1 is \$500 million. The investment in Year 2 is \$200 million. There is a 50 percent chance that the medicine will be developed and will be successful. If that happens, GlaxoSmithKline must spend another \$100 million in Year 3, but its income from the project in Year 3 will be \$500 million, not including the third-year investment. In Years 4, 5, and 6, it will earn \$400 million a year if the medicine is successful. At the end of Year 6, it intends to sell all rights to the medicine for \$600 million. If the medicine is unsuccessful, none of GlaxoSmithKline's investments can be salvaged. Assume that the market return is 12 percent, the risk-free rate is 2 percent, and the beta risk of the project is 2.3. All cash flows occur at the end of each year.

1 Calculate the expected annual cash flows using the probability of success.
2 Calculate the expected return.
3 Calculate the net present value.

Solution to 1:

There is a 50 percent chance that the cash flows in Years 3–6 will occur. Taking that into account, the expected annual cash flows are:

Year 1: –\$500 million (outflow)

Year 2: –\$200 million (outflow)

Year 3: 50% of –\$100 million (outflow) + 50% of \$500 million = \$200 million

Year 4: 50% of \$400 million = \$200 million

Year 5: 50% of \$400 million = \$200 million

Year 6: 50% of \$400 million + 50% of \$600 million = \$500 million

Solution to 2:

The expected or required return for the project can be calculated using the CAPM, which is $= 0.02 + 2.3 \times (0.12 - 0.02) = 0.25$.

Solution to 3:

The net present value is the discounted value of all cash flows:

$$
\begin{aligned}
NPV &= \sum_{t=0}^{T} \frac{CF_t}{(1+r_t)^t} \\
&= \frac{-500}{(1+0.25)} + \frac{-200}{(1+0.25)^2} + \frac{200}{(1+0.25)^3} + \frac{200}{(1+0.25)^4} \\
&\quad + \frac{200}{(1+0.25)^5} + \frac{500}{(1+0.25)^6} \\
&= -400 - 128 + 102.40 + 81.92 + 65.54 + 131.07 = -147.07.
\end{aligned}
$$

Because the net present value is negative (–$147.07 million), the project should not be accepted by GlaxoSmithKline.

10 BEYOND CAPM: LIMITATIONS AND EXTENSIONS OF CAPM

h describe and demonstrate applications of the CAPM and the SML;

In general, return-generating models allow us to estimate an asset's return given its characteristics, where the asset characteristics required for estimating the return are specified in the model. Estimating an asset's return is important for investment decision making. These models are also important as a benchmark for evaluating portfolio, security, or manager performance. The return-generating models were briefly introduced in Section 6, and one of those models, the capital asset pricing model, was discussed in detail in Sections 8 and 9.

The purpose of this section is to make readers aware that, although the CAPM is an important concept and model, the CAPM is not the only return-generating model. In this section, we revisit and highlight the limitations of the CAPM and preview return-generating models that address some of those limitations.

10.1 Limitations of the CAPM

The CAPM is subject to theoretical and practical limitations. Theoretical limitations are inherent in the structure of the model, whereas practical limitations are those that arise in implementing the model.

10.1.1 *Theoretical Limitations of the CAPM*

- Single-factor model: Only systematic risk or beta risk is priced in the CAPM. Thus, the CAPM states that no other investment characteristics should be considered in estimating returns. As a consequence, it is prescriptive and easy to understand and apply, although it is very restrictive and inflexible.
- Single-period model: The CAPM is a single-period model that does not consider multi-period implications or investment objectives of future periods, which can lead to myopic and suboptimal investment decisions. For example, it may be optimal to default on interest payments in the current period to

maximize current returns, but the consequences may be negative in the next period. A single-period model like the CAPM is unable to capture factors that vary over time and span several periods.

10.1.2 *Practical Limitations of the CAPM*

In addition to the theoretical limitations, implementation of the CAPM raises several practical concerns, some of which are listed below.

- Market portfolio: The true market portfolio according to the CAPM includes all assets, financial and nonfinancial, which means that it also includes many assets that are not investable, such as human capital and assets in closed economies. Richard Roll[3] noted that one reason the CAPM is not testable is that the true market portfolio is unobservable.
- Proxy for a market portfolio: In the absence of a true market portfolio, market participants generally use proxies. These proxies, however, vary among analysts, the country of the investor, etc. and generate different return estimates for the same asset, which is impermissible in the CAPM.
- Estimation of beta risk: A long history of returns (three to five years) is required to estimate beta risk. The historical state of the company, however, may not be an accurate representation of the current or future state of the company. More generally, the CAPM is an *ex ante* model, yet it is usually applied using *ex post* data. In addition, using different periods for estimation results in different estimates of beta. For example, a three-year beta is unlikely to be the same as a five-year beta, and a beta estimated with daily returns is unlikely to be the same as the beta estimated with monthly returns. Thus, we are likely to estimate different returns for the same asset depending on the estimate of beta risk used in the model.
- The CAPM is a poor predictor of returns: If the CAPM is a good model, its estimate of asset returns should be closely associated with realized returns. However, empirical support for the CAPM is weak.[4] In other words, tests of the CAPM show that asset returns are not determined only by systematic risk. Poor predictability of returns when using the CAPM is a serious limitation because return-generating models are used to estimate future returns.
- Homogeneity in investor expectations: The CAPM assumes that homogeneity exists in investor expectations for the model to generate a single optimal risky portfolio (the market) and a single security market line. Without this assumption, there will be numerous optimal risky portfolios and numerous security market lines. Clearly, investors can process the same information in a rational manner and arrive at different optimal risky portfolios.

10.2 Extensions to the CAPM

Given the limitations of the CAPM, it is not surprising that other models have been proposed to address some of these limitations. These new models are not without limitations of their own, which we will mention while discussing the models. We divide the models into two categories—theoretical models and practical models—and provide one example of each type.

---

[3] Roll (1977).

[4] See, for example, Fama and French (1992).

10.2.1 *Theoretical Models*

Theoretical models are based on the same principle as the CAPM but expand the number of risk factors. The best example of a theoretical model is the arbitrage pricing theory (APT), which was developed by Stephen Ross.[5] Like the CAPM, APT proposes a linear relationship between expected return and risk:

$$E(R_p) = R_F + \lambda_1\beta_{p,1} + \ldots + \lambda_K\beta_{p,K}$$

where

$E(R_p)$ = the expected return of portfolio p
R_F = the risk-free rate
λ_j = the risk premium (expected return in excess of the risk-free rate) for factor j
$\beta_{p,j}$ = the sensitivity of the portfolio to factor j
K = the number of risk factors

Unlike the CAPM, however, APT allows numerous risk factors—as many as are relevant to a particular asset. Moreover, other than the risk-free rate, the risk factors need not be common and may vary from one asset to another. A no-arbitrage condition in asset markets is used to determine the risk factors and estimate betas for the risk factors.

Although it is theoretically elegant, flexible, and superior to the CAPM, APT is not commonly used in practice because it does not specify any of the risk factors and it becomes difficult to identify risk factors and estimate betas for each asset in a portfolio. So from a practical standpoint, the CAPM is preferred to APT.

10.2.2 *Practical Models*

If beta risk in the CAPM does not explain returns, which factors do? Practical models seek to answer this question through extensive research. As mentioned in Section 6, the best example of such a model is the four-factor model proposed by Fama and French (1992) and Carhart (1997).

Based on an analysis of the relationship between past returns and a variety of different factors, Fama and French (1992) proposed that three factors seem to explain asset returns better than just systematic risk. Those three factors are relative size, relative book-to-market value, and beta of the asset. With Carhart's (1997) addition of relative past stock returns, the model can be written as follows:

$$E\left(R_{it}\right) = \alpha_i + \beta_{i,MKT}MKT_t + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,UMD}UMD_t$$

where

$E(R_i)$ = the return on an asset in excess of the one-month T-bill return
MKT = the excess return on the market portfolio
SMB = the difference in returns between small-capitalization stocks and large-capitalization stocks (size)
HML = the difference in returns between high-book-to-market stocks and low-book-to-market stocks (value versus growth)
UMD = the difference in returns of the prior year's winners and losers (momentum)

---

[5] Ross (1976).

Historical analysis shows that the coefficient on *MKT* is not significantly different from zero, which implies that stock return is unrelated to the market. The factors that explain stock returns are size (smaller companies outperform larger companies), book-to-market ratio (value companies outperform glamour companies), and momentum (past winners outperform past losers).

The four-factor model has been found to predict asset returns much better than the CAPM and is extensively used in estimating returns for US stocks.

Two observations are in order. First, the model is not underpinned by a theory of market equilibrium, as is the case for the CAPM. Second, there is no assurance that the model will continue to work well in the future.

PORTFOLIO PERFORMANCE APPRAISAL MEASURES

11

i. calculate and interpret the Sharpe ratio, Treynor ratio, M^2, and Jensen's alpha.

In the investment industry, **performance evaluation** refers to the measurement, attribution, and appraisal of investment results. In particular, performance evaluation provides information about the return and risk of investment portfolios over specified investment period(s). By providing accurate data and analysis on investment decisions and their consequences, performance evaluation allows portfolio managers to take corrective measures to improve investment decision-making and management processes. Performance evaluation information helps in understanding and controlling investment risk and should, therefore, lead to improved risk management. Performance evaluation seeks to answer the following questions:

- What was the investment portfolio's past performance, and what may be expected in the future?

Answering this question is the subject of performance measurement. *Performance measurement is concerned with the measurement of return and risk.*

- How did the investment portfolio produce its observed performance, and what are the expected sources of expected future performance?

Answering this question is the subject of performance attribution. *Performance attribution is concerned with identifying and quantifying the sources of performance of a portfolio.*

- Was the observed investment portfolio's performance the result of investment skill or luck?

Answering this question is the subject of performance appraisal. *Performance appraisal is concerned with identifying and measuring investment skill.*

The information provided by performance evaluation is of great interest to all stakeholders in the investment management process because of its value in evaluating the overall quality of the investment management process as well as individual investment decisions.

In this reading, performance appraisal is based only on the CAPM. However, it is easy to extend this analysis to multi-factor models that may include industry or other special factors. Four ratios are commonly used in performance appraisal.

11.1 The Sharpe Ratio

Performance has two components, risk and return. Although return maximization is a laudable objective, comparing just the return of a portfolio with that of the market is not sufficient. Because investors are risk averse, they will require compensation for higher risk in the form of higher returns. A commonly used measure of performance is the **Sharpe ratio**, which is defined as the portfolio's risk premium divided by its risk. An appealing feature of the Sharpe ratio is that its use can be justified on a theoretical *ex ante* (before the fact) basis and *ex post* (after the fact) values can easily be determined by using readily available market data. The Sharpe ratio is also easy to interpret, essentially being an efficiency ratio relating reward to risks taken. It is the most widely recognized and used appraisal measure.

The equation below defines the *ex ante* Sharpe ratio in terms of three inputs: (1) the portfolio's expected return, $E(R_p)$; (2) the risk-free rate of interest, R_f; and (3) the portfolio's *ex ante* standard deviation of returns (return volatility), σ_p, a quantitative measure of total risk.

$$\text{SR} = \frac{E\left(R_p\right) - R_F}{\sigma_p}$$

The Sharpe ratio can also be used on an *ex post* basis to evaluate historical risk-adjusted returns. Assume we have a sample of historical data that can be used to determine the sample mean portfolio return, $\overline{R}_p$; the standard deviation of the sample returns, here denoted by $\hat{\sigma}_p$ (s_p is a familiar notation in other contexts); and the sample mean risk-free rate, $\overline{R}_f$. The *ex post* (or realized or historical) Sharpe ratio can then be determined by using the following:

$$\widehat{\text{SR}} = \frac{\overline{R}_p - \overline{R}_F}{\hat{\sigma}_p}$$

Recalling the CAL from earlier in the reading, one can see that the Sharpe ratio, also called the reward-to-variability ratio, is simply the slope of the capital allocation line. Note, however, that the ratio uses the *total risk* of the portfolio, not its systematic risk. The use of total risk is appropriate if the portfolio is an investor's total portfolio—that is, the investor does not own any other assets. Sharpe ratios of the market and other portfolios can also be calculated in a similar manner. The portfolio with the highest Sharpe ratio has the best risk-adjusted performance, and the one with the lowest Sharpe ratio has the worst risk-adjusted performance, provided that the numerator is positive for all comparison portfolios. If the numerator is negative, the ratio will be less negative for riskier portfolios, resulting in incorrect rankings.

The Sharpe ratio, however, suffers from two limitations. First, it uses total risk as a measure of risk when only systematic risk is priced. Second, the ratio itself (e.g., 0.2 or 0.3) is not informative. To rank portfolios, the Sharpe ratio of one portfolio must be compared with the Sharpe ratio of another portfolio. Nonetheless, the ease of computation makes the Sharpe ratio a popular tool.

11.2 The Treynor Ratio

The **Treynor ratio** is a simple extension of the Sharpe ratio and resolves the Sharpe ratio's first limitation by substituting beta (systematic risk) for total risk. The *ex ante* and *ex post* Treynor ratios are provided below.

$$\text{TR} = \frac{E\left(R_p\right) - R_f}{\beta_p}$$

$$\widehat{\mathrm{TR}} = \frac{\bar{R}_p - \bar{R}_f}{\hat{\beta}_p}$$

Just like the Sharpe ratio, the numerators must be positive for the Treynor ratio to give meaningful results. In addition, the Treynor ratio does not work for negative-beta assets—that is, the denominator must also be positive for obtaining correct estimates and rankings. Although both the Sharpe and Treynor ratios allow for ranking of portfolios, neither ratio gives any information about the economic significance of differences in performance. For example, assume the Sharpe ratio of one portfolio is 0.75 and the Sharpe ratio for another portfolio is 0.80. The second portfolio is superior, but is that difference meaningful? In addition, we do not know whether either of the portfolios is better than the passive market portfolio. The remaining two measures, M^2 and Jensen's alpha, attempt to address that problem by comparing portfolios while also providing information about the extent of the overperformance or underperformance.

11.3 M^2: Risk-Adjusted Performance (RAP)

$\mathbf{M^2}$ provides a measure of portfolio return that is adjusted for the total risk of the portfolio relative to that of some benchmark. In 1997, Nobel Prize winner Franco Modigliani and his granddaughter, Leah Modigliani, developed what they called a risk-adjusted performance measure, or RAP. The RAP measure has since become more commonly known as M^2 reflecting the Modigliani names. It is related to the Sharpe ratio and ranks portfolios identically, but it has the useful advantage of being denominated in familiar terms of percentage return advantage assuming the same level of total risk as the market

M^2 borrows from capital market theory by assuming a portfolio is leveraged or de-leveraged until its volatility (as measured by standard deviation) matches that of the market. This adjustment produces a portfolio-specific leverage ratio that equates the portfolio's risk to that of the market. The portfolio's excess return times the leverage ratio plus the risk-free rate is then compared with the markets actual return to determine whether the portfolio has outperformed or underperformed the market on a risk-adjusted basis.

The equations below provide the *ex ante* and *ex post* formulas for M^2, where σ_m is the standard deviation of the market portfolio and σ_m/σ_p is the portfolio-specific leverage ratio. Because the Sharpe ratio is defined as $\frac{E(R_p) - R_f}{\sigma_p}$, the equation shows that M^2 can be thought of as a rescaling of the Sharpe ratio that allows for easier comparisons among different portfolios. The reason that M^2 and Sharpe ratios rank portfolios identically is because, in a given time period—and for any given comparison of the market portfolio—both the risk-free rate and the market volatility are constant across all comparisons. Only the Sharpe ratio differs, so it determines all rankings.

$$\mathrm{M}^2 = \left[E(R_p) - R_f\right]\frac{\sigma_m}{\sigma_p} + R_f = \mathrm{SR} \times \sigma_m + R_f \text{ (ex ante)}$$

$$\widehat{\mathrm{M}^2} = \left(\bar{R}_p - \bar{R}_f\right)\frac{\hat{\sigma}_m}{\hat{\sigma}_p} + R_f = \widehat{\mathrm{SR}} \times \hat{\sigma}_m + R_f \text{ (ex post)}$$

For example, assume that $\bar{R}_f = 4.0\%$, $\bar{R}_p = 14.0\%$, $\hat{\sigma}_p = 25.0\%$ and $\hat{\sigma}_m = 20.0\%$. The Sharpe ratio is 0.4, $\widehat{\mathrm{SR}} = \frac{0.14 - 0.04}{0.25} = 0.4$, and $\widehat{M^2}$ is 12.0%, $\widehat{M^2} = 0.4(0.2) + 0.04 = 0.12 = 12.0\%$. If the market return was 10%, then the portfolio outperformed the market on a risk-adjusted basis by 12.0% – 10.0% = 2.0%. This difference between the risk-adjusted performance of the portfolio and the performance of the market is frequently referred to as **M^2 alpha**.

The Sharpe ratio of the market portfolio is $\widehat{\mathrm{SR}} = \frac{0.10 - 0.04}{0.20} = 0.3$. Comparing the Sharpe ratio of the portfolio with the Sharpe ratio of the market portfolio shows that the fund outperformed the market. But the 2.0% difference between M^2 and the market's return tells us the risk-adjusted outperformance as a percentage return.

11.4 Jensen's Alpha

Like the Treynor ratio, Jensen's alpha is based on systematic risk. We can measure a portfolio's systematic risk by estimating the market model, which is done by regressing the portfolio's daily return on the market's daily return. The coefficient on the market return is an estimate of the beta risk of the portfolio. We can calculate the risk-adjusted return of the portfolio using the beta of the portfolio and the CAPM. The difference between the actual portfolio return and the calculated risk-adjusted return is a measure of the portfolio's performance relative to the market portfolio and is called Jensen's alpha. By definition, α_m of the market is zero. Jensen's alpha is also the vertical distance from the SML measuring the excess return for the same risk as that of the market and is given by

$$\alpha_p = R_p - \{R_f + \beta_p[\mathrm{E}(R_m) - R_f]\}$$

If the period is long, it may contain different risk-free rates, in which case R_f represents the average risk-free rate. Furthermore, the returns in the equation are all realized, actual returns. The sign of α_p indicates whether the portfolio has outperformed the market. If α_p is positive, then the portfolio has outperformed the market; if α_p is negative, the portfolio has underperformed the market. Jensen's alpha is commonly used for evaluating most institutional managers, pension funds, and mutual funds. Values of alpha can be used to rank different managers and the performance of their portfolios, as well as the magnitude of underperformance or overperformance. For example, if a portfolio's alpha is 2 percent and another portfolio's alpha is 5 percent, the second portfolio has outperformed the first portfolio by 3 percentage points and the market by 5 percentage points. Jensen's alpha is the maximum amount that you should be willing to pay the manager to manage your money. As with other performance appraisal measures, Jensen's alpha has *ex ante* and *ex post* forms. The use context usually clarifies which one is being referred to. Where we want to underscore a reference to *ex post* Jensen's alpha based on an estimated beta, $\hat{\beta}_p$, and an average market return, the notation $\hat{\alpha}_p$ is used.

EXAMPLE 10

Portfolio Performance Evaluation

A British pension fund has employed three investment managers, each of whom is responsible for investing in one-third of all asset classes so that the pension fund has a well-diversified portfolio. Information about the managers is given below.

Manager	**Average Return**	$\hat{\sigma}$	$\hat{\beta}$
X	10%	20%	1.1
Y	11	10	0.7

Manager	Average Return	$\hat{\sigma}$	$\hat{\beta}$
Z	12	25	0.6
Market (M)	9	19	
Risk-free rate (R_f)	3		

Calculate the expected return for each manager, based on using the average market return and the CAPM. Then also calculate for the managers (ex post) Sharpe ratio, Treynor ratio, M^2 alpha, and Jensen's alpha. Analyze your results and plot the returns and betas of these portfolios.

Solution:

In each case, the calculations are shown only for Manager X. All answers are tabulated below. Note that the β of the market is 1 and the σ and β of the risk-free rate are both zero.

$$\text{Expected return: } E(R_X) = R_f + \beta_X\left[E(R_m) - R_f\right] = 0.03 + 1.10 \times (0.09 - 0.03) = 0.096 = 9.6\%$$

$$\widehat{\text{SR}} = \frac{\bar{R}_x - \bar{R}_f}{\hat{\sigma}_x} = \frac{0.10 - 0.03}{0.20} = 0.35$$

$$\widehat{\text{TR}} = \frac{\bar{R}_x - \bar{R}_f}{\hat{\beta}_x} = \frac{0.10 - 0.03}{1.1} = 0.064$$

$$\widehat{M^2} = \left(\bar{R}_x - \bar{R}_f\right)\frac{\hat{\sigma}_m}{\hat{\sigma}_x} + \bar{R}_f = \widehat{SR} \times \hat{\sigma}_m + \bar{R}_f$$

$$= 0.35 \times 0.19 + 0.03 = 0.0965 = 9.65\%$$

Since the market return is 9%, M^2 alpha is 0.65% (9.65% – 9%).

$$\hat{\alpha}_X = R_X - \left[\bar{R}_f + \hat{\beta}_X\left(\bar{R}_m - \bar{R}_f\right)\right] = 0.10 - (0.03 + 1.1 \times 0.06)$$

$$= 0.004 = 0.40\%$$

Exhibit 8 Measures of Portfolio Performance Evaluation

Manager	$\bar{R}_i$	$\hat{\sigma}_i$	$\hat{\beta}_i$	$E(R_i)$	Sharpe Ratio	Treynor Ratio	M^2 alpha	$\hat{\alpha}_i$
X	10.0%	20.0%	1.10	9.6%	0.35	0.064	0.65%	0.40%
Y	11.0	10.0	0.70	7.2	0.80	0.114	9.20	3.80
Z	12.0	25.0	0.60	6.6	0.36	0.150	0.84	5.40
M	9.0	19.0	1.00	9.0	0.32	0.060	0.00	0.00
R_f	3.0	0.0	0.00	3.0	–	–	–	0.00

Let us begin with an analysis of the risk-free asset. Because the risk-free asset has zero risk and a beta of zero, calculating the Sharpe ratio, Treynor ratio, or M^2 is not possible because they all require the portfolio risk in the denominator. The risk-free asset's alpha, however, is zero. Turning to the market portfolio, we see that the absolute measures of performance, the Sharpe ratio and the Treynor ratio, are positive for the market portfolio. These ratios are positive as long as the portfolio earns a return that is in excess of that of the risk-free asset. $\widehat{M}^2$ and $\hat{\alpha}_i$ are performance measures relative to the market, so they are both equal to zero for the market portfolio.

All three managers have Sharpe and Treynor ratios greater than those of the market, and all three managers' M^2 alpha and α_i are positive; therefore, the pension fund should be satisfied with their performance. Among the three managers, Manager X has the worst performance, irrespective of whether total risk or systematic risk is considered for measuring performance. The relative rankings are depicted in Exhibit 9.

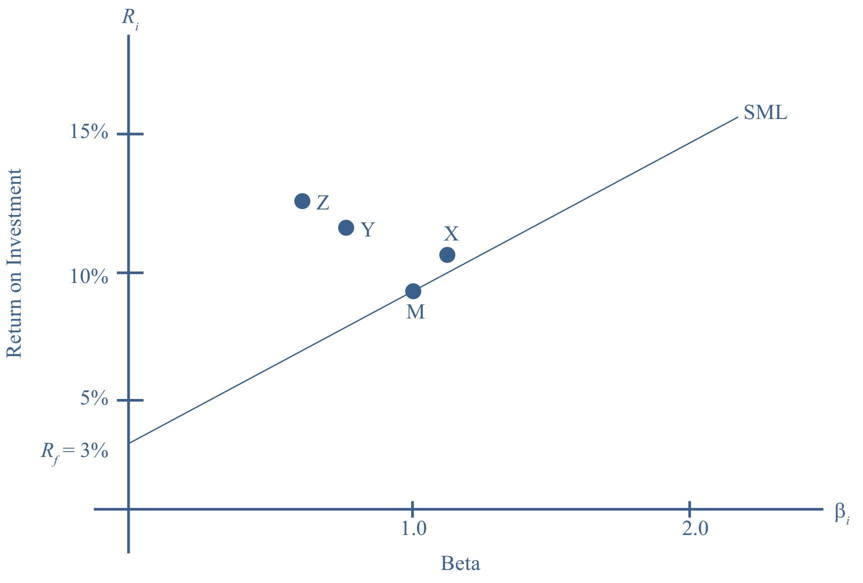
Exhibit 9 Ranking of Portfolios by Performance Measure

Rank	Sharpe Ratio	Treynor Ratio	M^2 alpha	α_i
1	Y	Z	Y	Z
2	Z	Y	Z	Y
3	X	X	X	X
4	M	M	M	M
5	–	–	–	R_f

Comparing Y and Z, we can observe that Y performs much better than Z when total risk is considered. Y has a Sharpe ratio of 0.80, compared with a Sharpe ratio of 0.36 for Z. Similarly, M^2 alpha is higher for Y (9.20 percent) than for Z (0.84 percent). In contrast, when systematic risk is used, Z outperforms Y. The Treynor ratio is higher for Z (0.150) than for Y (0.114), and Jensen's alpha is also higher for Z (5.40 percent) than for Y (3.80 percent), which indicates that Z has done a better job of generating excess return relative to systematic risk than Y.

Exhibit 10 confirms these observations in that all three managers outperform the benchmark because all three points lie above the SML. Among the three portfolios, Z performs the best when we consider risk-adjusted returns because it is the point in Exhibit 10 that is located northwest relative to the portfolios X and Y.

Exhibit 10 Portfolios Along the SML



When do we use total risk performance measures like the Sharpe ratio and M^2, and when do we use beta risk performance measures like the Treynor ratio and Jensen's alpha? Total risk is relevant for an investor when he or she holds a portfolio that is not fully diversified, which is not a desirable portfolio. In such cases, the Sharpe ratio and M^2 are appropriate performance measures. Thus, if the pension fund were to choose only one fund manager to manage all its assets, it should choose Manager Y. Performance measures relative to beta risk—Treynor ratio and Jensen's alpha—are relevant when the investor holds a well-diversified portfolio with negligible diversifiable risk. In other words, if the pension fund is well diversified and only the systematic risk of the portfolio matters, the fund should choose Manager Z.

The measures of performance evaluation assume that the market portfolio is the correct benchmark. As a result, an error in the benchmark may cause the results to be misleading. For example, evaluating a real estate fund against the S&P 500 is incorrect because real estate has different characteristics than equity. In addition to errors in benchmarking, errors could occur in the measurement of risk and return of the market portfolio and the portfolios being evaluated. Finally, many estimates are based on historical data. Any projections based on such estimates assume that this level of performance will continue in the future.

APPLICATIONS OF THE CAPM IN PORTFOLIO CONSTRUCTION

12

g calculate and interpret the expected return of an asset using the CAPM;

h describe and demonstrate applications of the CAPM and the SML;

This section introduces applications of the CAPM in portfolio construction. First, the security characteristic line, which graphically indicates *ex post* Jensen's alpha, is described. If we relax the assumption that investors have the same expectations about risk and return, a positive Jensen's alpha can be interpreted as an indication of superior information or investment ability. The section on security selection covers that possibility. The last section summarizes how the CAPM and related concepts can be applied to portfolio construction.

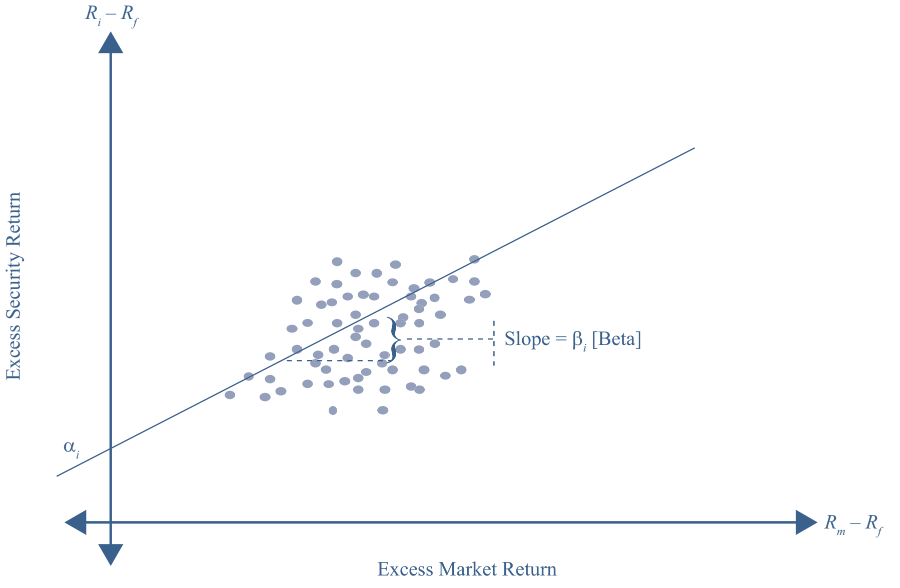
12.1 Security Characteristic Line

Similar to the SML, we can draw a **security characteristic line** (SCL) for a security. The SCL is a plot of the excess return of the security on the excess return of the market. In Exhibit 8, Jensen's alpha is the intercept and the beta is the slope. The equation of the line can be obtained by rearranging the terms in the expression for Jensen's alpha and replacing the subscript p with i:

$$R_i - R_f = \alpha_i + \beta_i(R_m - R_f)$$

As an example, the SCL is drawn in Exhibit 11 using Manager X's portfolio from Exhibit 8. The security characteristic line can also be estimated by regressing the excess security return, $R_i - R_f$, on the excess market return, $R_m - R_f$.

Exhibit 11 The Security Characteristic Line



12.2 Security Selection

When discussing the CAPM, we assumed that investors have homogeneous expectations and are rational, risk-averse, utility-maximizing investors. With these assumptions, we were able to state that all investors assign the same value to all assets and, therefore, have the same optimal risky portfolio, which is the market portfolio. In other words, we assumed that there is commonality among beliefs about an asset's future cash flows and the required rate of return. Given the required rate of return, we can discount the future cash flows of the asset to arrive at its current value, or price, which is agreed upon by all or most investors.

In this section, we introduce heterogeneity in beliefs of investors. Because investors are price takers, it is assumed that such heterogeneity does not significantly affect the market price of an asset. The difference in beliefs can relate to future cash flows, the systematic risk of the asset, or both. Because the current price of an asset is the discounted value of the future cash flows, the difference in beliefs could result in an investor-estimated price that is different from the CAPM-calculated price. The CAPM-calculated price is the current market price because it reflects the beliefs of all other investors in the market. If the investor-estimated current price is higher (lower) than the market price, the asset is considered undervalued (overvalued). Therefore, the CAPM is an effective tool for determining whether an asset is undervalued or overvalued and whether an investor should buy or sell the asset.

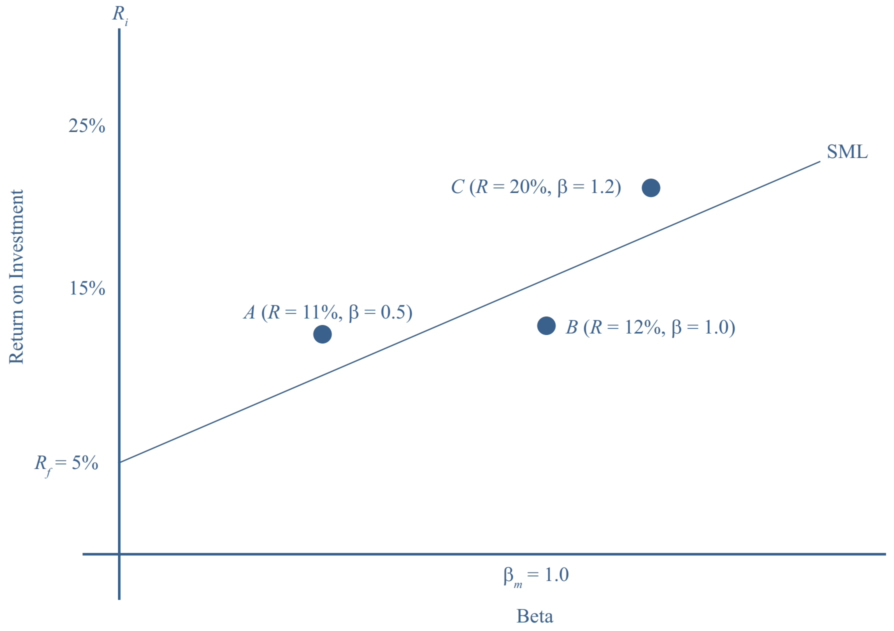
Although portfolio performance evaluation is backward looking and security selection is forward looking, we can apply the concepts of portfolio evaluation to security selection. The best measure to apply is Jensen's alpha because it uses systematic risk and is meaningful even on an absolute basis. A positive Jensen's alpha indicates a superior security, whereas a negative Jensen's alpha indicates a security that is likely to underperform the market when adjusted for risk.

Another way of presenting the same information is with the security market line. Potential investors can plot a security's expected return and beta against the SML and use this relationship to decide whether the security is overvalued or undervalued in the market.[6] Exhibit 12 shows a number of securities along with the SML. All securities that reflect the consensus market view are points directly on the SML (i.e., properly valued). If a point representing the estimated return of an asset is above the SML (Points A and C), the asset has a low level of risk relative to the amount of expected return and would be a good choice for investment. In contrast, if the point representing a particular asset is below the SML (Point B), the stock is considered overvalued. Its return does not compensate for the level of risk and should not be considered for investment. Of course, a short position in Asset B can be taken if short selling is permitted.

---

[6] In this reading, we do not consider transaction costs, which are important whenever deviations from a passive portfolio are considered. Thus, the magnitude of undervaluation or overvaluation should be considered in relation to transaction costs prior to making an investment decision.

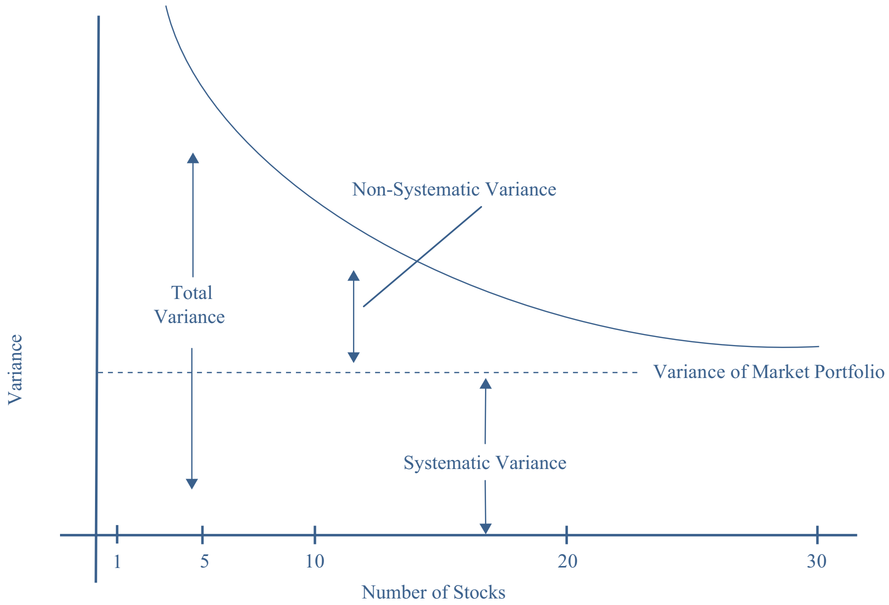
Exhibit 12 Security Selection Using SML



12.3 Implications of the CAPM for Portfolio Construction

Based on the CAPM, investors should hold a combination of the risk-free asset and the market portfolio. The true market portfolio consists of a large number of securities, and an investor would have to own all of them in order to be completely diversified. Because owning all existing securities is not practical, in this section, we will consider an alternate method of constructing a portfolio that may not require a large number of securities and will still be sufficiently diversified. Exhibit 13 shows the reduction in risk as we add more and more securities to a portfolio. As can be seen from the exhibit, much of the nonsystematic risk can be diversified away in as few as 30 securities. These securities, however, should be randomly selected and represent different asset classes for the portfolio to effectively diversify risk. Otherwise, one may be better off using an index (e.g., the S&P 500 for a diversified large-cap equity portfolio and other indexes for other asset classes).

Exhibit 13 Diversification with Number of Stocks



Let's begin constructing the optimal portfolio with a portfolio of securities like the S&P 500. Although the S&P 500 is a portfolio of 500 securities, it is a good starting point because it is readily available as a single security for trading. In contrast, it represents only the large corporations that are traded on the US stock markets and, therefore, does not encompass the global market entirely. Because the S&P 500 is the base portfolio, however, we treat it is as the market for the CAPM.

Any security not included in the S&P 500 can be evaluated to determine whether it should be integrated into the portfolio. That decision is based on the α_i of the security, which is calculated using the CAPM with the S&P 500 as the market portfolio. Note that security i may not necessarily be priced incorrectly for it to have a non-zero α_i; α_i can be positive merely because it is not well correlated with the S&P 500 and its return is sufficient for the amount of systematic risk it contains. For example, assume a new stock market, ABC, opens to foreign investors only and is being considered for inclusion in the portfolio. We estimate ABC's model parameters relative to the S&P 500 and find an α_i of approximately 3 percent, with a β_i of 0.60. Because α_i is positive, ABC should be added to the portfolio. Securities with a significantly negative α_i may be short sold to maximize risk-adjusted return. For convenience, however, we will assume that negative positions are not permitted in the portfolio.

In addition to the securities that are correctly priced but enter the portfolio because of their risk–return superiority, securities already in the portfolio (S&P 500) may be undervalued or overvalued based on investor expectations that are incongruent with the market. Securities in the S&P 500 that are overvalued (negative α_i) should be dropped from the S&P 500 portfolio, if it is possible to exclude individual securities, and positions in securities in the S&P 500 that are undervalued (positive α_i) should be increased.

This brings us to the next question: What should the relative weight of securities in the portfolio be? Because we are concerned with maximizing risk-adjusted return, securities with a higher α_i should have a higher weight, and securities with greater nonsystematic risk should be given less weight in the portfolio. A complete analysis of portfolio optimization is beyond the scope of this reading, but the following principles are helpful. The weight in each nonmarket security should be proportional to $\frac{\alpha_i}{\sigma_{ei}^2}$, where the denominator is the nonsystematic variance of security i. The total weight of nonmarket securities in the portfolio is proportional to $\frac{\sum_{i=1}^{N} w_i \alpha_i}{\sum_{i=1}^{N} w_i^2 \sigma_{ei}^2}$. The weight in the market portfolio is a function of $\frac{E(R_m)}{\sigma_m^2}$. The information ratio, $\frac{\alpha_i}{\sigma_{ei}}$ (i.e., alpha divided by nonsystematic risk), measures the abnormal return per unit of risk added by the security to a well-diversified portfolio. The larger the information ratio is, the more valuable the security.

EXAMPLE 11

Optimal Investor Portfolio with Heterogeneous Beliefs

A Japanese investor is holding the Nikkei 225 index, which is her version of the market. She thinks that three stocks, P, Q, and R, which are not in the Nikkei 225, are undervalued and should form a part of her portfolio. She has the following information about the stocks, the Nikkei 225, and the risk-free rate (the information is given as expected return, standard deviation, and beta):

P: 15%, 30%, 1.5

Q: 18%, 25%, 1.2

R: 16%, 23%, 1.1

Nikkei 225: 12%, 18%, 1.0

Risk-free rate: 2%, 0%, 0.0

1. Calculate Jensen's alpha for P, Q, and R.
2. Calculate nonsystematic variance for P, Q, and R.
3. Should any of the three stocks be included in the portfolio? If so, which stock should have the highest weight in the portfolio?

Solution to 1:

Stock P's α: $R_i - [R_f + \beta_i(R_m - R_f)] = 0.15 - (0.02 + 1.5 \times 0.10) = -0.02$

Stock Q's α: $R_i - [R_f + \beta_i(R_m - R_f)] = 0.18 - (0.02 + 1.2 \times 0.10) = 0.04$

Stock R's α: $R_i - [R_f + \beta_i(R_m - R_f)] = 0.16 - (0.02 + 1.1 \times 0.10) = 0.03$

Solution to 2:

Total variance = Systematic variance + Nonsystematic variance. From Section 6, we write the equation as $\sigma_{ei}^2 = \sigma_i^2 - \beta_i^2 \sigma_m^2$.

Stock P's nonsystematic variance = $(0.30 \times 0.30) - (1.5 \times 1.5 \times 0.18 \times 0.18) = 0.09 - 0.0729 = 0.0171$

Stock Q's nonsystematic variance = $(0.25 \times 0.25) - (1.2 \times 1.2 \times 0.18 \times 0.18) = 0.0625 - 0.0467 = 0.0158$

Stock R's nonsystematic variance = $(0.23 \times 0.23) - (1.1 \times 1.1 \times 0.18 \times 0.18) = 0.0529 - 0.0392 = 0.0137$

Solution to 3:

Stock P has a negative α and should not be included in the portfolio, unless a negative position can be assumed through short selling. Stocks Q and R have a positive α; therefore, they should be included in the portfolio with positive weights.

The relative weight of Q is 0.04/0.0158 = 2.53.

The relative weight of R is 0.03/0.0137 = 2.19.

Stock Q will have the largest weight among the nonmarket securities to be added to the portfolio. In relative terms, the weight of Q will be 15.5 percent greater than the weight of R (2.53/2.19 = 1.155). As the number of securities increases, the analysis becomes more complex. However, the contribution of each additional security toward improvement in the risk–return trade-off will decrease and eventually disappear, resulting in a well-diversified portfolio.

SUMMARY

In this reading, we discussed the capital asset pricing model in detail and covered related topics such as the capital market line. The reading began with an interpretation of the CML, uses of the market portfolio as a passive management strategy, and leveraging of the market portfolio to obtain a higher expected return. Next, we discussed systematic and nonsystematic risk and why one should not expect to be compensated for taking on nonsystematic risk. The discussion of systematic and nonsystematic risk was followed by an introduction to beta and return-generating models. This broad topic was then broken down into a discussion of the CAPM and, more specifically, the relationship between beta and expected return. The final section included applications of the CAPM to capital budgeting, portfolio performance evaluation, and security selection. The highlights of the reading are as follows.

- The capital market line is a special case of the capital allocation line, where the efficient portfolio is the market portfolio.
- Obtaining a unique optimal risky portfolio is not possible if investors are permitted to have heterogeneous beliefs because such beliefs will result in heterogeneous asset prices.
- Investors can leverage their portfolios by borrowing money and investing in the market.
- Systematic risk is the risk that affects the entire market or economy and is not diversifiable.
- Nonsystematic risk is local and can be diversified away by combining assets with low correlations.
- Beta risk, or systematic risk, is priced and earns a return, whereas nonsystematic risk is not priced.
- The expected return of an asset depends on its beta risk and can be computed using the CAPM, which is given by $E(R_i) = R_f + \beta_i[E(R_m) - R_f]$.
- The security market line is an implementation of the CAPM and applies to all securities, whether they are efficient or not.
- Expected return from the CAPM can be used for making capital budgeting decisions.

- Portfolios can be evaluated by several CAPM-based measures, such as the Sharpe ratio, the Treynor ratio, M^2, and Jensen's alpha.
- The SML can assist in security selection and optimal portfolio construction.

By successfully understanding the content of this reading, you should feel comfortable decomposing total variance into systematic and nonsystematic variance, analyzing beta risk, using the CAPM, and evaluating portfolios and individual securities.

REFERENCES

Carhart, Mark. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance*, vol. 52, no. 1:57–82.

Fama, Eugene, and Kenneth French. 1992. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, vol. 47, no. 2:427–466.

Lintner, John. 1965a. "Security Prices, Risk, and Maximal Gains from Diversification." *Journal of Finance*, vol. 20, no. 4:587–615.

Lintner, John. 1965b. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics*, vol. 47, no. 1:13–37.

Markowitz, Harry. 1952. "Portfolio Selection." *Journal of Finance*, vol. 7, no. 1:77–91.

Mossin, Jan. 1966. "Equilibrium in a Capital Asset Market." *Econometrica*, vol. 34, no. 4:768–783.

Roll, Richard. 1977. "A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory." *Journal of Financial Economics*, vol. 4, no. 2:129–176.

Ross, Stephen A. 1976. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory*, vol. 13, no. 3:341–360.

Sharpe, William F. 1964. "Capital Asset Prices: A Theory Of Market Equilibrium under Conditions of Risk." *Journal of Finance*, vol. 19, no. 3:425–442.

Treynor, Jack L. 1961. *Market Value, Time, and Risk*. Unpublished manuscript.

Treynor, Jack L. 1962. *Toward a Theory of Market Value of Risky Assets*. Unpublished manuscript.

PRACTICE PROBLEMS

1 The line depicting the total risk and expected return of portfolio combinations of a risk-free asset and any risky asset is the:

A security market line.

B capital allocation line.

C security characteristic line.

2 The portfolio of a risk-free asset and a risky asset has a better risk-return tradeoff than investing in only one asset type because the correlation between the risk-free asset and the risky asset is equal to:

A −1.0.

B 0.0.

C 1.0.

3 With respect to capital market theory, an investor's optimal portfolio is the combination of a risk-free asset and a risky asset with the highest:

A expected return.

B indifference curve.

C capital allocation line slope.

4 Highly risk-averse investors will *most likely* invest the majority of their wealth in:

A risky assets.

B risk-free assets.

C the optimal risky portfolio.

5 The capital market line (CML) is the graph of the risk and return of portfolio combinations consisting of the risk-free asset and:

A any risky portfolio.

B the market portfolio.

C the leveraged portfolio.

6 Which of the following statements *most accurately* defines the market portfolio in capital market theory? The market portfolio consists of all:

A risky assets.

B tradable assets.

C investable assets.

7 With respect to capital market theory, the optimal risky portfolio:

A is the market portfolio.

B has the highest expected return.

C has the lowest expected variance.

8 Relative to portfolios on the CML, any portfolio that plots above the CML is considered:

A inferior.

B inefficient.

C unachievable.

© 2019 CFA Institute. All rights reserved.

9 A portfolio on the capital market line with returns greater than the returns on the market portfolio represents a(n):

A lending portfolio.

B borrowing portfolio.

C unachievable portfolio.

10 With respect to the capital market line, a portfolio on the CML with returns less than the returns on the market portfolio represents a(n):

A lending portfolio.

B borrowing portfolio.

C unachievable portfolio.

11 Which of the following types of risk is *most likely* avoided by forming a diversified portfolio?

A Total risk.

B Systematic risk.

C Nonsystematic risk.

12 Which of the following events is *most likely* an example of nonsystematic risk?

A A decline in interest rates.

B The resignation of chief executive officer.

C An increase in the value of the US dollar.

13 With respect to the pricing of risk in capital market theory, which of the following statements is *most accurate*?

A All risk is priced.

B Systematic risk is priced.

C Nonsystematic risk is priced.

14 The sum of an asset's systematic variance and its nonsystematic variance of returns is equal to the asset's:

A beta.

B total risk.

C total variance.

15 With respect to return-generating models, the intercept term of the market model is the asset's estimated:

A beta.

B alpha.

C variance.

16 With respect to return-generating models, the slope term of the market model is an estimate of the asset's:

A total risk.

B systematic risk.

C nonsystematic risk.

17 With respect to return-generating models, which of the following statements is *most accurate*? Return-generating models are used to directly estimate the:

A expected return of a security.

B weights of securities in a portfolio.

C parameters of the capital market line.

The following information relates to Questions 18–20

An analyst gathers the following information:

Security	Expected Annual Return (%)	Expected Standard Deviation (%)	Correlation between Security and the Market
Security 1	11	25	0.6
Security 2	11	20	0.7
Security 3	14	20	0.8
Market	10	15	1.0

18 Which security has the *highest* total risk?

A Security 1.

B Security 2.

C Security 3.

19 Which security has the *highest* beta measure?

A Security 1.

B Security 2.

C Security 3.

20 Which security has the *least* amount of market risk?

A Security 1.

B Security 2.

C Security 3.

---

21 With respect to capital market theory, the average beta of all assets in the market is:

A less than 1.0.

B equal to 1.0.

C greater than 1.0.

22 The slope of the security characteristic line is an asset's:

A beta.

B excess return.

C risk premium.

23 The graph of the capital asset pricing model is the:

A capital market line.

B security market line.

C security characteristic line.

24 With respect to capital market theory, correctly priced individual assets can be plotted on the:

A capital market line.

B security market line.

C capital allocation line.

25 With respect to the capital asset pricing model, the primary determinant of expected return of an individual asset is the:

A asset's beta.

B market risk premium.

C asset's standard deviation.

26 With respect to the capital asset pricing model, which of the following values of beta for an asset is *most likely* to have an expected return for the asset that is less than the risk-free rate?

A −0.5

B 0.0

C 0.5

27 With respect to the capital asset pricing model, the market risk premium is:

A less than the excess market return.

B equal to the excess market return.

C greater than the excess market return.

The following information relates to Questions 28–31

An analyst gathers the following information:

Security	Expected Standard Deviation (%)	Beta
Security 1	25	1.50
Security 2	15	1.40
Security 3	20	1.60

28 With respect to the capital asset pricing model, if the expected market risk premium is 6% and the risk-free rate is 3%, the expected return for Security 1 is *closest* to:

A 9.0%.

B 12.0%.

C 13.5%.

29 With respect to the capital asset pricing model, if expected return for Security 2 is equal to 11.4% and the risk-free rate is 3%, the expected return for the market is *closest* to:

A 8.4%.

B 9.0%.

C 10.3%.

30 With respect to the capital asset pricing model, if the expected market risk premium is 6% the security with the *highest* expected return is:

A Security 1.

B Security 2.

C Security 3.

31 With respect to the capital asset pricing model, a decline in the expected market return will have the *greatest* impact on the expected return of:

A Security 1.

B Security 2.

C Security 3.

---

32 Three equity fund managers have performance records summarized in the following table:

	Mean Annual Return (%)	Standard Deviation of Return (%)
Manager 1	14.38	10.53
Manager 2	9.25	6.35
Manager 3	13.10	8.23

Given a risk-free rate of return of 2.60%, which manager performed best based on the Sharpe ratio?

A Manager 1

B Manager 2

C Manager 3

33 Which of the following performance measures is consistent with the CAPM?

A *M*-squared.

B Sharpe ratio.

C Jensen's alpha.

34 Which of the following performance measures does *not* require the measure to be compared to another value?

A Sharpe ratio.

B Treynor ratio.

C Jensen's alpha.

35 Which of the following performance measures is *most* appropriate for an investor who is *not* fully diversified?

A *M*-squared.

B Treynor ratio.

C Jensen's alpha.

36 Analysts who have estimated returns of an asset to be greater than the expected returns generated by the capital asset pricing model should consider the asset to be:

A overvalued.

B undervalued.

C properly valued.

37 With respect to capital market theory, which of the following statements *best* describes the effect of the homogeneity assumption? Because all investors have the same economic expectations of future cash flows for all assets, investors will invest in:

A the same optimal risky portfolio.

B the Standard and Poor's 500 Index.

C assets with the same amount of risk.

38 With respect to capital market theory, which of the following assumptions allows for the existence of the market portfolio? All investors:

A are price takers.

B have homogeneous expectations.

C plan for the same, single holding period.

39 The intercept of the best fit line formed by plotting the excess returns of a manager's portfolio on the excess returns of the market is *best* described as Jensen's:

A beta.

B ratio.

C alpha.

40 Portfolio managers who are maximizing risk-adjusted returns will seek to invest *more* in securities with:

A lower values of Jensen's alpha.

B values of Jensen's alpha equal to 0.

C higher values of Jensen's alpha.

41 Portfolio managers, who are maximizing risk-adjusted returns, will seek to invest *less* in securities with:

A lower values for nonsystematic variance.

B values of nonsystematic variance equal to 0.

C higher values for nonsystematic variance.

SOLUTIONS

1 B is correct. A capital allocation line (CAL) plots the expected return and total risk of combinations of the risk-free asset and a risky asset (or a portfolio of risky assets).

2 B is correct. A portfolio of the risk-free asset and a risky asset or a portfolio of risky assets can result in a better risk-return tradeoff than an investment in only one type of an asset, because the risk-free asset has zero correlation with the risky asset.

3 B is correct. Investors will have different optimal portfolios depending on their indifference curves. The optimal portfolio for each investor is the one with highest utility; that is, where the CAL is tangent to the individual investor's highest possible indifference curve.

4 B is correct. Although the optimal risky portfolio is the market portfolio, highly risk-averse investors choose to invest most of their wealth in the risk-free asset.

5 B is correct. Although the capital allocation line includes all possible combinations of the risk-free asset and any risky portfolio, the capital market line is a special case of the capital allocation line, which uses the market portfolio as the optimal risky portfolio.

6 A is correct. The market includes all risky assets, or anything that has value; however, not all assets are tradable, and not all tradable assets are investable.

7 A is correct. The optimal risky portfolio is the market portfolio. Capital market theory assumes that investors have homogeneous expectations, which means that all investors analyze securities in the same way and are rational. That is, investors use the same probability distributions, use the same inputs for future cash flows, and arrive at the same valuations. Because their valuations of all assets are identical, all investors will invest in the same optimal risky portfolio (i.e., the market portfolio).

8 C is correct. Theoretically, any point above the CML is not achievable and any point below the CML is dominated by and inferior to any point on the CML.

9 B is correct. As one moves further to the right of point M on the capital market line, an increasing amount of borrowed money is being invested in the market portfolio. This means that there is negative investment in the risk-free asset, which is referred to as a leveraged position in the risky portfolio.

10 A is correct. The combinations of the risk-free asset and the market portfolio on the CML where returns are less than the returns on the market portfolio are termed 'lending' portfolios.

11 C is correct. Investors are capable of avoiding nonsystematic risk by forming a portfolio of assets that are not highly correlated with one another, thereby reducing total risk and being exposed only to systematic risk.

12 B is correct. Nonsystematic risk is specific to a firm, whereas systematic risk affects the entire economy.

13 B is correct. Only systematic risk is priced. Investors do not receive any return for accepting nonsystematic or diversifiable risk.

14 C is correct. The sum of systematic variance and nonsystematic variance equals the total variance of the asset. References to total risk as the sum of systematic risk and nonsystematic risk refer to variance, not to risk.

15 B is correct. In the market model, $R_i = \alpha_i + \beta_i R_m + e_i$, the intercept, α_i, and slope coefficient, β_i, are estimated using historical security and market returns.

16 B is correct. In the market model, $R_i = \alpha_i + \beta_i R_m + e_i$, the slope coefficient, β_i, is an estimate of the asset's systematic or market risk.

17 A is correct. In the market model, $R_i = \alpha_i + \beta_i R_m + e_i$, the intercept, α_i, and slope coefficient, β_i, are estimated using historical security and market returns. These parameter estimates then are used to predict firm-specific returns that a security may earn in a future period.

18 A is correct. Security 1 has the highest total risk = 0.25 compared to Security 2 and Security 3 with a total risk of 0.20.

19 C is correct. Security 3 has the highest beta value; $1.07 = \dfrac{\rho_{3,m}\sigma_3}{\sigma_m} = \dfrac{(0.80)(20\%)}{15\%}$ compared to Security 1 and Security 2 with beta values of 1.00 and 0.93, respectively.

20 B is correct. Security 2 has the lowest beta value; $0.93 = \dfrac{\rho_{2,m}\sigma_2}{\sigma_m} = \dfrac{(0.70)(20\%)}{15\%}$ compared to Security 1 and 3 with beta values of 1.00 and 1.07, respectively.

21 B is correct. The average beta of all assets in the market, by definition, is equal to 1.0.

22 A is correct. The security characteristic line is a plot of the excess return of the security on the excess return of the market. In such a graph, Jensen's alpha is the intercept and the beta is the slope.

23 B is correct. The security market line (SML) is a graphical representation of the capital asset pricing model, with beta risk on the x-axis and expected return on the y-axis.

24 B is correct. The security market line applies to any security, efficient or not. The CAL and the CML use the total risk of the asset (or portfolio of assets) rather than its systematic risk, which is the only risk that is priced.

25 A is correct. The CAPM shows that the primary determinant of expected return for an individual asset is its beta, or how well the asset correlates with the market.

26 A is correct. If an asset's beta is negative, the required return will be less than the risk-free rate in the CAPM. When combined with a positive market return, the asset reduces the risk of the overall portfolio, which makes the asset very valuable. Insurance is an example of a negative beta asset.

27 B is correct. In the CAPM, the market risk premium is the difference between the return on the market and the risk-free rate, which is the same as the return in excess of the market return.

28 B is correct. The expected return of Security 1, using the CAPM, is 12.0% = 3% + 1.5(6%); $E(R_i) = R_f + \beta_i[E(R_m) - R_f]$.

29 B is correct. The expected risk premium for Security 2 is 8.4%, (11.4% – 3%), indicates that the expected market risk premium is 6%; therefore, since the risk-free rate is 3% the expected rate of return for the market is 9%. That is, using the CAPM, $E(R_i) = R_f + \beta_i[E(R_m) - R_f]$, 11.4% = 3% + 1.4(X%), where X% = (11.4% – 3%)/1.4 = 6.0% = market risk premium.

30 C is correct. Security 3 has the highest beta; thus, regardless of the value for the risk-free rate, Security 3 will have the highest expected return:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f]$$

31 C is correct. Security 3 has the highest beta; thus, regardless of the risk-free rate the expected return of Security 3 will be most sensitive to a change in the expected market return.

32 C is correct. The Sharpe ratio $\left(\widehat{\text{SR}}\right)$ is the mean excess portfolio return per unit of risk, where a higher Sharpe ratio indicates better performance:

$$\widehat{SR}_1 = \frac{\overline{R}_p - \overline{R}_f}{\hat{\sigma}_p} = \frac{14.38 - 2.60}{10.53} = 1.12$$

$$\widehat{SR}_2 = \frac{\overline{R}_p - \overline{R}_f}{\hat{\sigma}_p} = \frac{9.25 - 2.60}{6.35} = 1.05$$

$$\widehat{SR}_3 = \frac{\overline{R}_p - \overline{R}_f}{\hat{\sigma}_p} = \frac{13.10 - 2.60}{8.23} = 1.28$$

33 C is correct. Jensen's alpha adjusts for systematic risk, and *M*-squared and the Sharpe Ratio adjust for total risk.

34 C is correct. The sign of Jensen's alpha indicates whether or not the portfolio has outperformed the market. If alpha is positive, the portfolio has outperformed the market; if alpha is negative, the portfolio has underperformed the market.

35 A is the correct. *M*-squared adjusts for risk using standard deviation (i.e., total risk).

36 B is correct. If the estimated return of an asset is above the SML (the expected return), the asset has a lower level of risk relative to the amount of expected return and would be a good choice for investment (i.e., undervalued).

37 A is correct. The homogeneity assumption refers to all investors having the same economic expectation of future cash flows. If all investors have the same expectations, then all investors should invest in the same optimal risky portfolio, therefore implying the existence of only one optimal portfolio (i.e., the market portfolio).

38 B is correct. The homogeneous expectations assumption means that all investors analyze securities in the same way and are rational. That is, they use the same probability distributions, use the same inputs for future cash flows, and arrive at the same valuations. Because their valuation of all assets is identical, they will generate the same optimal risky portfolio, which is the market portfolio.

39 C is correct. This is because of the plot of the excess return of the security on the excess return of the market. In such a graph, Jensen's alpha is the intercept and the beta is the slope.

40 C is correct. Since managers are concerned with maximizing risk-adjusted returns, securities with a higher value of Jensen's alpha, α_i, should have a higher weight.

41 C is correct. Since managers are concerned with maximizing risk-adjusted returns, securities with greater nonsystematic risk should be given less weight in the portfolio.

Glossary

A priori probability A probability based on logical analysis rather than on observation or personal judgment.

Abandonment option The ability to terminate an investment at some future time if the financial results are disappointing.

Abnormal return The amount by which a security's actual return differs from its expected return, given the security's risk and the market's return.

Absolute advantage A country's ability to produce a good or service at a lower absolute cost than its trading partner.

Absolute dispersion The amount of variability present without comparison to any reference point or benchmark.

Absolute frequency The actual number of observations counted for each unique value of the variable (also called raw frequency).

Accelerated book build An offering of securities by an investment bank acting as principal that is accomplished in only one or two days.

Accelerated methods Depreciation methods that allocate a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Accounting costs Monetary value of economic resources used in performing an activity. These can be explicit, out-of-pocket, current payments, or an allocation of historical payments (depreciation) for resources. They do not include implicit opportunity costs.

Accounting profit Income as reported on the income statement, in accordance with prevailing accounting standards, before the provisions for income tax expense. Also called *income before taxes* or *pretax income*.

Accounts payable Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Accrued expenses Liabilities related to expenses that have been incurred but not yet paid as of the end of an accounting period—an example of an accrued expense is rent that has been incurred but not yet paid, resulting in a liability "rent payable." Also called *accrued liabilities*.

Accrued interest Interest earned but not yet paid.

Acquisition method A method of accounting for a business combination where the acquirer is required to measure each identifiable asset and liability at fair value. This method was the result of a joint project of the IASB and FASB aiming at convergence in standards for the accounting of business combinations.

Action lag Delay from policy decisions to implementation.

Active investment An approach to investing in which the investor seeks to outperform a given benchmark.

Active return The return on a portfolio minus the return on the portfolio's benchmark.

Activity ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory. Also called *asset utilization ratios* or *operating efficiency ratios*.

Add-on rates Bank certificates of deposit, repos, and indexes such as Libor and Euribor are quoted on an add-on rate basis (bond equivalent yield basis).

Addition rule for probabilities A principle stating that the probability that *A* or *B* occurs (both occur) equals the probability that *A* occurs, plus the probability that *B* occurs, minus the probability that both *A* and *B* occur.

Agency bond See *quasi-government bond*.

Agency costs Costs associated with the conflict of interest present between principals and agents when a company is managed by non-owners. Agency costs result from the inherent conflicts of interest between managers, bondholders, and equity owners.

Agency costs of debt Costs arising from conflicts of interest between managers and debtholders.

Agency costs of equity The smaller the stake managers have in the company, the less their share in bearing the cost of excessive perquisite consumption—consequently, the less their desire to give their best efforts in running the company.

Agency RMBS In the United States, securities backed by residential mortgage loans and guaranteed by a federal agency or guaranteed by either of the two GSEs (Fannie Mae and Freddie Mac).

Aggregate demand The quantity of goods and services that households, businesses, government, and non-domestic customers want to buy at any given level of prices.

Aggregate demand curve Inverse relationship between the price level and real output.

Aggregate income The value of all the payments earned by the suppliers of factors used in the production of goods and services.

Aggregate output The value of all the goods and services produced during a specified period.

Aggregate supply The quantity of goods and services producers are willing to supply at any given level of price.

Aggregate supply curve The level of domestic output that companies will produce at each price level.

All-or-nothing (AON) orders An order that includes the instruction to trade only if the trade fills the entire quantity (size) specified.

Allocationally efficient A characteristic of a market, a financial system, or an economy that promotes the allocation of resources to their highest value uses.

Alternative data Non-traditional data types generated by the use of electronic devices, social media, satellite and sensor networks, and company exhaust.

Alternative hypothesis The hypothesis that is accepted if the null hypothesis is rejected.

Alternative investment markets Market for investments other than traditional securities investments (i.e., traditional common and preferred shares and traditional fixed income instruments). The term usually encompasses direct and indirect investment in real estate (including timberland and farmland) and commodities (including precious metals); hedge funds, private equity, and other investments requiring specialized due diligence.

Alternative trading systems Trading venues that function like exchanges but that do not exercise regulatory authority over their subscribers except with respect to the conduct

of the subscribers' trading in their trading systems. Also called *electronic communications networks* or *multilateral trading facilities*.

American depository receipt A US dollar-denominated security that trades like a common share on US exchanges.

American depository share The underlying shares on which American depository receipts are based. They trade in the issuing company's domestic market.

American-style Type of option contract that can be exercised at any time up to the option's expiration date.

Amortisation The process of allocating the cost of intangible long-term assets having a finite useful life to accounting periods; the allocation of the amount of a bond premium or discount to the periods remaining until bond maturity.

Amortised cost The historical cost (initially recognised cost) of an asset, adjusted for amortisation and impairment.

Amortizing bond Bond with a payment schedule that calls for periodic payments of interest and repayments of principal.

Amortizing loans Loans with a payment schedule that calls for periodic payments of interest and repayments of principal.

Analysis of variance (ANOVA) The analysis that breaks the total variability of a dataset (such as observations on the dependent variable in a regression) into components representing different sources of variation. With reference to regression, ANOVA provides the inputs for an *F*-test of the significance of the regression as a whole, as well as the inputs for the coefficient of determination and the standard error of the estimate.

Analytical duration The use of mathematical formulas to estimate the impact of benchmark yield-to-maturity changes on bond prices.

Anchoring and adjustment bias An information-processing bias in which the use of a psychological heuristic influences the way people estimate probabilities.

Annual percentage rate The cost of borrowing expressed as a yearly rate.

Annuity A finite set of level sequential cash flows.

Annuity due An annuity having a first cash flow that is paid immediately.

Anomalies Apparent deviations from market efficiency.

Antidilutive With reference to a transaction or a security, one that would increase earnings per share (EPS) or result in EPS higher than the company's basic EPS—antidilutive securities are not included in the calculation of diluted EPS.

Arbitrage 1) The simultaneous purchase of an undervalued asset or portfolio and sale of an overvalued but equivalent asset or portfolio, in order to obtain a riskless profit on the price differential. Taking advantage of a market inefficiency in a risk-free manner. 2) The condition in a financial market in which equivalent assets or combinations of assets sell for two different prices, creating an opportunity to profit at no risk with no commitment of money. In a well-functioning financial market, few arbitrage opportunities are possible. 3) A risk-free operation that earns an expected positive net profit but requires no net investment of money.

Arbitrage-free pricing The overall process of pricing derivatives by arbitrage and risk neutrality. Also called the *principle of no arbitrage*.

Arbitrageurs Traders who engage in arbitrage. See *arbitrage*.

Arithmetic mean The sum of the observations divided by the number of observations.

Artificial intelligence Computer systems that exhibit cognitive and decision-making ability comparable (or superior) to that of humans.

Ask The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size). See *offer*.

Ask size The maximum quantity of an asset that pertains to a specific ask price from a trader. For example, if the ask for a share issue is $30 for a size of 1,000 shares, the trader is offering to sell at $30 up to 1,000 shares.

Asset allocation The process of determining how investment funds should be distributed among asset classes.

Asset-backed securities A type of bond issued by a legal entity called a *special purpose entity* (SPE) on a collection of assets that the SPE owns. Also, securities backed by receivables and loans other than mortgages.

Asset-based valuation models Valuation based on estimates of the market value of a company's assets.

Asset beta The unlevered beta; which reflects the business risk of the assets; the assets' systematic risk.

Asset class A group of assets that have similar characteristics, attributes, and risk–return relationships.

Asset swap Converts the periodic fixed coupon of a specific bond to an MRR plus or minus a spread.

Asset utilization ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Assets Resources controlled by an enterprise as a result of past events and from which future economic benefits to the enterprise are expected to flow.

Assignment of accounts receivable The use of accounts receivable as collateral for a loan.

Asymmetric information The differential of information between corporate insiders and outsiders regarding the company's performance and prospects. Managers typically have more information about the company's performance and prospects than owners and creditors.

At the money An option in which the underlying's price equals the exercise price.

Auction A type of bond issuing mechanism often used for sovereign bonds that involves bidding.

Autarkic price The price of a good or service in an autarkic economy.

Autarky A state in which a country does not trade with other countries.

Automatic stabilizer A countercyclical factor that automatically comes into play as an economy slows and unemployment rises.

Availability bias An information-processing bias in which people take a heuristic approach to estimating the probability of an outcome based on how easily the outcome comes to mind.

Available-for-sale Under US GAAP, debt securities not classified as either held-to-maturity or held-for-trading securities. The investor is willing to sell but not actively planning to sell. In general, available-for-sale debt securities are reported at fair value on the balance sheet, with unrealized gains included as a component of other comprehensive income.

Average fixed cost Total fixed cost divided by quantity produced.

Average life See *weighted average life*.

Average product Measures the productivity of inputs on average and is calculated by dividing total product by the total number of units for a given input that is used to generate that output.

Average revenue Total revenue divided by quantity sold.

Average total cost Total cost divided by quantity produced.

Average variable cost Total variable cost divided by quantity produced.

Back-testing With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backfill bias A bias introduced if a fund backfills its past performance data to the database when it first starts reporting actual performance.

Backtesting With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backup line of credit A type of credit enhancement provided by a bank to an issuer of commercial paper to ensure that the issuer will have access to sufficient liquidity to repay maturing commercial paper if issuing new paper is not a viable option.

Backwardation A condition in the futures markets in which the spot price exceeds the futures price, the forward curve is downward sloping, and the convenience yield is high.

Balance of payments A double-entry bookkeeping system that summarizes a country's economic transactions with the rest of the world for a particular period of time, typically a calendar quarter or year.

Balance of trade deficit When the domestic economy is spending more on non-domestic goods and services than non-domestic economies are spending on domestic goods and services.

Balance sheet The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet). Also called *statement of financial position* or *statement of financial condition.*

Balance sheet ratios Financial ratios involving balance sheet items only.

Balanced With respect to a government budget, one in which spending and revenues (taxes) are equal.

Balloon payment Large payment required at maturity to retire a bond's outstanding principal amount.

Bar chart A chart for plotting the frequency distribution of categorical data, where each bar represents a distinct category and each bar's height is proportional to the frequency of the corresponding category. In technical analysis, a bar chart that plots four bits of data for each time interval—the high, low, opening, and closing prices. A vertical line connects the high and low prices. A cross-hatch left indicates the opening price and a cross-hatch right indicates the closing price.

Barter economy An economy where economic agents as house-holds, corporations, and governments "pay" for goods and services with another good or service.

Base-rate neglect A type of representativeness bias in which the base rate or probability of the categorization is not adequately considered.

Base rates The reference rate on which a bank bases lending rates to all other customers.

Basic EPS Net earnings available to common shareholders (i.e., net income minus preferred dividends) divided by the weighted average number of common shares outstanding.

Basis point Used in stating yield spreads, one basis point equals one-hundredth of a percentage point, or 0.01%.

Basket of listed depository receipts An exchange-traded fund (ETF) that represents a portfolio of depository receipts.

Bayes' formula The rule for updating the probability of an event of interest—given a set of prior probabilities for the event, information, and information given the event—if you receive new information.

Bearer bonds Bonds for which ownership is not recorded; only the clearing system knows who the bond owner is.

Bearish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from above; this movement is considered bearish. A **death cross** is a bearish crossover based on 50-day and 200-day moving averages.

Behavioral finance A field of finance that examines the psychological variables that affect and often distort the investment decision making of investors, analysts, and portfolio managers.

Behind the market Said of prices specified in orders that are worse than the best current price; e.g., for a limit buy order, a limit price below the best bid.

Benchmark A comparison portfolio; a point of reference or comparison.

Benchmark issue The latest sovereign bond issue for a given maturity. It serves as a benchmark against which to compare bonds that have the same features but that are issued by another type of issuer.

Benchmark rate Typically the yield-to-maturity on a government bond having the same or close to the same time-to-maturity.

Benchmark spread The yield spread over a specific benchmark, usually measured in basis points.

Bernoulli random variable A random variable having the outcomes 0 and 1.

Bernoulli trial An experiment that can produce one of two outcomes.

Best bid The highest bid in the market.

Best effort offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-efforts offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-in-class An ESG implementation approach that seeks to identify the most favorable companies in an industry based on ESG considerations.

Best offer The lowest offer (ask price) in the market.

Beta A measure of the sensitivity of a given investment or portfolio to movements in the overall market.

Bid The price at which a dealer or trader is willing to buy an asset, typically qualified by a maximum quantity.

Bid–ask spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid–offer spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid size The maximum quantity of an asset that pertains to a specific bid price from a trader.

Big Data The vast amount of data being generated by industry, governments, individuals, and electronic devices that arises from both traditional and non-traditional data sources.

Bilateral loan A loan from a single lender to a single borrower.

Bimodal A distribution that has two most frequently occurring values.

Binomial model A model for pricing options in which the underlying price can move to only one of two possible new prices.

Binomial random variable The number of successes in n Bernoulli trials for which the probability of success is constant for all trials and the trials are independent.

Binomial tree The graphical representation of a model of asset price dynamics in which, at each period, the asset moves up with probability p or down with probability $(1 - p)$.

Bitcoin A cryptocurrency using blockchain technology that was created in 2009.

Bivariate correlation See *Pearson correlation*.

Block brokers A broker (agent) that provides brokerage services for large-size trades.

Blockchain A type of digital ledger in which information is recorded sequentially and then linked together and secured using cryptographic methods.

Blue chip Widely held large market capitalization companies that are considered financially sound and are leaders in their respective industry or local stock market.

Bollinger Bands A price-based technical analysis indicator consisting of a line representing the moving average, a higher line representing the moving average plus a set number of standard deviations from the average (for the same number of periods as was used to calculate the moving average), and a lower line representing the moving average minus the same number of standard deviations.

Bond Contractual agreement between the issuer and the bondholders.

Bond equivalent yield A calculation of yield that is annualized using the ratio of 365 to the number of days to maturity. Bond equivalent yield allows for the restatement and comparison of securities with different compounding periods.

Bond indenture The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *trust deed*.

Bond market vigilantes Bond market participants who might reduce their demand for long-term bonds, thus pushing up their yields.

Bond yield plus risk premium approach An estimate of the cost of common equity that is produced by summing the before-tax cost of debt and a risk premium that captures the additional yield on a company's stock relative to its bonds. The additional yield is often estimated using historical spreads between bond yields and stock yields.

Bonus issue of shares A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Book building Investment bankers' process of compiling a "book" or list of indications of interest to buy part of an offering.

Book value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. Also called *carrying value*.

Boom An expansionary phase characterized by economic growth "testing the limits" of the economy.

Bootstrap A resampling method that repeatedly draws samples with replacement of the selected elements from the original observed sample. Bootstrap is usually conducted by using computer simulation and is often used to find standard error or construct confidence intervals of population parameters.

Bottom-up analysis An investment selection approach that focuses on company-specific circumstances rather than emphasizing economic cycles or industry analysis.

Box and whisker plot A graphic for visualizing the dispersion of data across quartiles. It consists of a "box" with "whiskers" connected to the box.

Breakdown A breakdown occurs when the price of an asset moves below a support level.

Breakeven point The number of units produced and sold at which the company's net income is zero (Revenues = Total cost); in the case of perfect competition, the quantity at which price, average revenue, and marginal revenue equal average total cost.

Breakout A breakout occurs when the price of an asset moves above a resistance level.

Bridge financing Interim financing that provides funds until permanent financing can be arranged.

Broad money Encompasses narrow money plus the entire range of liquid assets that can be used to make purchases.

Broker 1) An agent who executes orders to buy or sell securities on behalf of a client in exchange for a commission. 2) See *futures commission merchants*.

Broker–dealer A financial intermediary (often a company) that may function as a principal (dealer) or as an agent (broker) depending on the type of trade.

Brokered market A market in which brokers arrange trades among their clients.

Brownfield investment Investing in *existing* infrastructure assets.

Bubble line chart A line chart that uses varying-sized bubbles to represent a third dimension of the data. The bubbles are sometimes color-coded to present additional information.

Budget surplus/deficit The difference between government revenue and expenditure for a stated fixed period of time.

Bullet bond Bond in which the principal repayment is made entirely at maturity.

Bullish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from below; this movement is considered bullish. A **golden cross** is a bullish crossover based on 50-day and 200-day moving averages.

Business cycles Are recurrent expansions and contractions in economic activity affecting broad segments of the economy.

Business risk The risk associated with operating earnings. Operating earnings are uncertain because total revenues and many of the expenditures contributed to produce those revenues are uncertain.

Buy-side firm An investment management company or other investor that uses the services of brokers or dealers (i.e., the client of the sell side firms).

Buyback A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Call An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call market A market in which trades occur only at a particular time and place (i.e., when the market is called).

Call money rate The interest rate that buyers pay for their margin loan.

Call option An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call protection The time during which the issuer of the bond is not allowed to exercise the call option.

Callable bond A bond containing an embedded call option that gives the issuer the right to buy the bond back from the investor at specified prices on pre-determined dates.

Calmar ratio A comparison of the average annual compounded return to the maximum drawdown risk.

Candlestick chart A price chart with four bits of data for each time interval. A candle indicates the opening and closing price for the interval. The body of the candle is shaded if the opening price was higher than the closing price, and the body is white (or clear) if the opening price was lower than the closing price. Vertical lines known as wicks or shadows extend from the top and bottom of the candle to indicate, respectively, the high and low prices for the interval.

Cannibalization Cannibalization occurs when an investment takes customers and sales away from another part of the company.

Cap rate A metric by which real estate managers are often judged; the annual rent actually earned (net of any vacancies) divided by the price originally paid for the property.

Capacity The ability of the borrower to make its debt payments on time.

Capital account A component of the balance of payments account that measures transfers of capital.

Capital allocation line (CAL) A graph line that describes the combinations of expected return and standard deviation of return available to an investor from combining the optimal portfolio of risky assets with the risk-free asset.

Capital allocation process The process that companies use for decision making on capital investments—those projects with a life of one year or longer.

Capital asset pricing model (CAPM) An equation describing the expected return on any asset (or portfolio) as a linear function of its beta relative to the market portfolio.

Capital consumption allowance A measure of the wear and tear (depreciation) of the capital stock that occurs in the production of goods and services.

Capital deepening investment Increases the stock of capital relative to labor.

Capital expenditure Expenditure on physical capital (fixed assets).

Capital light Capital light businesses require little incremental investment in fixed assets or working capital to enable revenue growth.

Capital loss ratio The percentage of capital in deals that have been realized below cost, net of any recovered proceeds, divided by total invested capital.

Capital market expectations An investor's expectations concerning the risk and return prospects of asset classes.

Capital market line (CML) The line with an intercept point equal to the risk-free rate that is tangent to the efficient frontier of risky assets; represents the efficient frontier when a risk-free asset is available for investment.

Capital market securities Securities with maturities at issuance longer than one year.

Capital markets Financial markets that trade securities of longer duration, such as bonds and equities.

Capital rationing A capital rationing environment assumes that the company has a fixed amount of funds to invest.

Capital restrictions Controls placed on foreigners' ability to own domestic assets and/or domestic residents' ability to own foreign assets.

Capital stock The accumulated amount of buildings, machinery, and equipment used to produce goods and services.

Capital structure The mix of debt and equity that a company uses to finance its business; a company's specific mixture of long-term financing.

Carry The net of the costs and benefits of holding, storing, or "carrying" an asset.

Carrying amount The amount at which an asset or liability is valued according to accounting principles.

Carrying value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. For a bond, the purchase price plus (or minus) the amortized amount of the discount (or premium).

Cartel Participants in collusive agreements that are made openly and formally.

Cash collateral account Form of external credit enhancement whereby the issuer immediately borrows the credit-enhancement amount and then invests that amount, usually in highly rated short-term commercial paper.

Cash conversion cycle A financial metric that measures the length of time required for a company to convert cash invested in its operations to cash received as a result of its operations; equal to days of inventory on hand + days of sales outstanding – number of days of payables. Also called *net operating cycle*.

Cash flow additivity principle The principle that dollar amounts indexed at the same point in time are additive.

Cash flow from operating activities The net amount of cash provided from operating activities.

Cash flow from operations The net amount of cash provided from operating activities.

Cash flow yield The internal rate of return on a series of cash flows.

Cash market securities Money market securities settled on a "same day" or "cash settlement" basis.

Cash markets See *spot markets*.

Cash prices See *spot prices*.

Cash-settled forwards See *non-deliverable forwards*.

Catch-up clause A clause in an agreement that favors the GP. For a GP who earns a 20% performance fee, a catch-up clause allows the GP to receive 100% of the distributions above the hurdle rate *until* she receives 20% of the profits generated, and then every excess dollar is split 80/20 between the LPs and GP.

Categorical data Values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize (also called **qualitative data**).

CBOE Volatility Index (VIX) A measure of near-term market volatility as conveyed by S&P 500 stock index option prices.

Central bank funds market The market in which deposit-taking banks that have an excess reserve with their national central bank can lend money to banks that need funds for maturities ranging from overnight to one year. Called the Federal or Fed funds market in the United States.

Central bank funds rate Interest rate at which central bank funds are bought (borrowed) and sold (lent) for maturities ranging from overnight to one year. Called Federal or Fed funds rate in the United States.

Central banks The dominant bank in a country, usually with official or semi-official governmental status.

Central limit theorem The theorem that states the sum (and the mean) of a set of independent, identically distributed random variables with finite variances is normally distributed, whatever distribution the random variables follow.

Certificate of deposit An instrument that represents a specified amount of funds on deposit with a bank for a specified maturity and interest rate. CDs are issued in various denominations and can be negotiable or non-negotiable.

Change in polarity principle A tenet of technical analysis that states that once a support level is breached, it becomes a resistance level. The same holds true for resistance levels: Once breached, they become support levels.

Change of control put A covenant giving bondholders the right to require the issuer to buy back their debt, often at par or at some small premium to par value, in the event that the borrower is acquired.

Character The quality of a debt issuer's management.

Chartist An individual who uses charts or graphs of a security's historical prices or levels to forecast its future trends.

Chi-square test of independence A statistical test for detecting a potential association between categorical variables.

Classical cycle Refers to fluctuations in the level of economic activity when measured by GDP in volume terms.

Classified balance sheet A balance sheet organized so as to group together the various assets and liabilities into subcategories (e.g., current and noncurrent).

Clawback A requirement that the general partner return any funds distributed as incentive fees until the limited partners have received back their initial investment and a percentage of the total profit.

Clearing The process by which the exchange verifies the execution of a transaction and records the participants' identities.

Clearing instructions Instructions that indicate how to arrange the final settlement ("clearing") of a trade.

Clearinghouse An entity associated with a futures market that acts as middleman between the contracting parties and guarantees to each party the performance of the other.

Closed economy An economy that does not trade with other countries; an *autarkic economy*.

Closed-end fund A mutual fund in which no new investment money is accepted. New investors invest by buying existing shares, and investors in the fund liquidate by selling their shares to other investors.

Cluster sampling A procedure that divides a population into subpopulation groups (clusters) representative of the population and then randomly draws certain clusters to form a sample.

Clustered bar chart See *grouped bar chart*.

Co-investing In co-investing, the investor invests in assets *indirectly* through the fund but also possesses rights (known as co-investment rights) to invest *directly* in the same assets. Through co-investing, an investor is able to make an investment *alongside* a fund when the fund identifies deals.

Code of ethics An established guide that communicates an organization's values and overall expectations regarding member behavior. A code of ethics serves as a general guide for how community members should act.

Coefficient of determination The percentage of the variation of the dependent variable that is explained by the independent variable. Also referred to as the "R-squared" or "R^2."

Coefficient of variation The ratio of a set of observations' standard deviation to the observations' mean value.

Cognitive cost The effort involved in processing new information and updating beliefs.

Cognitive dissonance The mental discomfort that occurs when new information conflicts with previously held beliefs or cognitions.

Cognitive errors Behavioral biases resulting from faulty reasoning; cognitive errors stem from basic statistical, information-processing, or memory errors.

Coincident economic indicators Turning points that are usually close to those of the overall economy; they are believed to have value for identifying the economy's present state.

Collateral Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Collateral manager Buys and sells debt obligations for and from the CDO's portfolio of assets (i.e., the collateral) to generate sufficient cash flows to meet the obligations to the CDO bondholders.

Collateral trust bonds Bonds secured by securities, such as common shares, other bonds, or other financial assets.

Collateralized bond obligations A structured asset-backed security that is collateralized by a pool of bonds.

Collateralized debt obligation Generic term used to describe a security backed by a diversified pool of one or more debt obligations.

Collateralized loan obligations A structured asset-backed security that is collateralized by a pool of loans.

Collateralized mortgage obligation A security created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collateralized mortgage obligations Securities created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collaterals Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Combination A listing in which the order of the listed items does not matter.

Combination formula (binomial formula) The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is ${}_nC_r = \binom{n}{r} = \dfrac{n!}{(n-r)!r!}$.

Commercial paper A short-term, negotiable, unsecured promissory note that represents a debt obligation of the issuer.

Committed (Regular) lines of credit A bank commitment to extend credit up to a pre-specified amount; the commitment is considered a short-term liability and is usually in effect for 364 days (one day short of a full year).

Committed capital The amount that the limited partners have agreed to provide to the private equity fund.

Commodity swap A swap in which the underlying is a commodity such as oil, gold, or an agricultural product.

Common market Level of economic integration that incorporates all aspects of the customs union and extends it by allowing free movement of factors of production among members.

Common shares A type of security that represent an ownership interest in a company.

Common-size analysis The restatement of financial statement items using a common denominator or reference item that allows one to identify trends and major differences; an example is an income statement in which all items are expressed as a percent of revenue.

Common stock See *common shares*.

Company analysis Analysis of an individual company.

Comparable company A company that has similar business risk, usually in the same industry and preferably with a single line of business.

Comparative advantage A country's ability to produce a good or service at a lower relative cost, or opportunity cost, than its trading partner.

Competitive strategy A company's plans for responding to the threats and opportunities presented by the external environment.

Complement The event not-S, written S^C, given the event S. Note that $P(S) + P(S^C) = 1$.

Complements Goods that tend to be used together; technically, two goods whose cross-price elasticity of demand is negative.

Complete markets Informally, markets in which the variety of distinct securities traded is so broad that any desired payoff in a future state-of-the-world is achievable.

Component cost of capital The rate of return required by suppliers of capital for an individual source of a company's funding, such as debt or equity.

Compounding The process of accumulating interest on interest.

Comprehensive income The change in equity of a business enterprise during a period from nonowner sources; includes all changes in equity during a period except those resulting from investments by owners and distributions to owners; comprehensive income equals net income plus other comprehensive income.

Conditional expected value The expected value of a stated event given that another event has occurred.

Conditional pass-through covered bonds Covered bonds that convert to pass-through securities after the original maturity date if all bond payments have not yet been made and the sponsor is in default.

Conditional probability The probability of an event given (conditioned on) another event.

Conditional variances The variance of one variable, given the outcome of another.

Confidence level The complement of the level of significance.

Confirmation bias A belief perseverance bias in which people tend to look for and notice what confirms their beliefs, to ignore or undervalue what contradicts their beliefs, and to misinterpret information as support for their beliefs.

Confusion matrix A type of contingency table used for evaluating the performance of a classification model.

Conservatism bias A belief perseverance bias in which people maintain their prior views or forecasts by inadequately incorporating new information.

Consolidation The movement of a stock's price within a well-defined range of trading levels for a period of time. The price consolidates between a support level and a resistance level.

Constant-yield price trajectory A graph that illustrates the change in the price of a fixed-income bond over time assuming no change in yield-to-maturity. The trajectory shows the "pull to par" effect on the price of a bond trading at a premium or a discount to par value.

Constituent securities With respect to an index, the individual securities within an index.

Consumer surplus The difference between the value that a consumer places on units purchased and the amount of money that was required to pay for them.

Contango A condition in the futures markets in which the spot price is lower than the futures price, the forward curve is upward sloping, and there is little or no convenience yield.

Contingency provision Clause in a legal document that allows for some action if a specific event or circumstance occurs.

Contingency table A tabular format that displays the frequency distributions of two or more categorical variables simultaneously and is used for finding patterns between the variables. A contingency table for two categorical variables is also known as a two-way table.

Contingent claim A security whose value is based on some other underlying security.

Contingent claims Derivatives in which the payoffs occur if a specific event occurs; generally referred to as options.

Contingent convertible bonds Bonds that automatically convert into equity if a specific event or circumstance occurs, such as the issuer's equity capital falling below the minimum requirement set by the regulators. Also called *CoCos*.

Continuation pattern A type of pattern used in technical analysis to predict the resumption of a market trend that was in place prior to the formation of a pattern.

Continuous data Data that can be measured and can take on any numerical value in a specified range of values.

Continuous random variable A random variable for which the range of possible outcomes is the real line (all real numbers between $-\infty$ and $+\infty$) or some subset of the real line.

Continuous trading market A market in which trades can be arranged and executed any time the market is open.

Continuously compounded return The natural logarithm of 1 plus the holding period return, or equivalently, the natural logarithm of the ending price over the beginning price.

Contra account An account that offsets another account.

Contract rate See *mortgage rate*.

Contraction The period of a business cycle after the peak and before the trough; often called a *recession* or, if exceptionally severe, called a *depression*.

Contraction risk The risk that when interest rates decline, the security will have a shorter maturity than was anticipated at the time of purchase because borrowers refinance at the new, lower interest rates.

Contractionary Tending to cause the real economy to contract.

Contractionary fiscal policy A fiscal policy that has the objective to make the real economy contract.

Contracts for differences See *non-deliverable forwards*.

Contribution margin The amount available for fixed costs and profit after paying variable costs; revenue minus variable costs.

Controlling shareholders A particular shareholder or block of shareholders holding a percentage of shares that gives them significant voting power.

Convenience sampling A procedure of selecting an element from a population on the basis of whether or not it is accessible to a researcher or how easy it is for a researcher to access the element.

Convenience yield A non-monetary advantage of holding an asset.

Conventional bond See *plain vanilla bond*.

Conventional cash flows A conventional cash flow pattern is one with an initial outflow followed by a series of inflows.

Convergence The tendency for differences in output per capita across countries to diminish over time. In technical analysis, the term describes the case when an indicator moves in the same manner as the security being analyzed.

Conversion premium The difference between the convertible bond's price and its conversion value.

Conversion price For a convertible bond, the price per share at which the bond can be converted into shares.

Conversion ratio For a convertible bond, the number of common shares that each bond can be converted into.

Conversion value For a convertible bond, the current share price multiplied by the conversion ratio.

Convertible bond Bond that gives the bondholder the right to exchange the bond for a specified number of common shares in the issuing company.

Convertible preference shares A type of equity security that entitles shareholders to convert their shares into a specified number of common shares.

Convexity adjustment For a bond, one half of the annual or approximate convexity statistic multiplied by the change in the yield-to-maturity squared.

Core inflation Refers to the inflation rate calculated based on a price index of goods and services except food and energy.

Corporate governance The system of internal controls and procedures by which individual companies are managed.

Correlation A measure of the linear relationship between two random variables.

Correlation coefficient A number between −1 and +1 that measures the consistency or tendency for two investments to act in a similar way. It is used to determine the effect on portfolio risk when two assets are combined.

Cost averaging The periodic investment of a fixed amount of money.

Cost of capital The cost of financing to a company; the rate of return that suppliers of capital require as compensation for their contribution of capital.

Cost of carry See *carry*.

Cost of debt The cost of debt financing to a company, such as when it issues a bond or takes out a bank loan; the rate of return required by a company's bondholders.

Cost of equity The rate of return required by a company's common stockholders.

Cost of preferred stock The cost to a company of issuing preferred stock; the dividend yield that a company must commit to pay preferred stockholders.

Cost structure The mix of a company's variable costs and fixed costs.

Counterparty risk The risk that the other party to a contract will fail to honor the terms of the contract.

Coupon rate The interest rate promised in a contract; this is the rate used to calculate the periodic interest payments.

Cournot assumption Assumption in which each firm determines its profit-maximizing production level assuming that the other firms' output will not change.

Covariance A measure of the co-movement (linear association) between two random variables.

Covariance matrix A matrix or square array whose entries are covariances; also known as a variance–covariance matrix.

Covenants The terms and conditions of lending agreements that the issuer must comply with; they specify the actions that an issuer is obligated to perform (affirmative covenant) or prohibited from performing (negative covenant).

Covered bond Debt obligation secured by a segregated pool of assets called the cover pool. The issuer must maintain the value of the cover pool. In the event of default, bondholders have recourse against both the issuer and the cover pool.

Covered bonds Senior debt obligations issued by a financial institution and backed by a segregated pool of assets. Covered bond investors have recourse to both the underlying assets and the financial institution.

Credit analysis The evaluation of credit risk; the evaluation of the creditworthiness of a borrower or counterparty.

Credit default swap (CDS) A type of credit derivative in which one party, the credit protection buyer who is seeking credit protection against a third party, makes a series of regularly scheduled payments to the other party, the credit protection seller. The seller makes no payments until a credit event occurs.

Credit derivatives A contract in which one party has the right to claim a payment from another party in the event that a specific credit event occurs over the life of the contract.

Credit enhancements Provisions that may be used to reduce the credit risk of a bond issue.

Credit-linked coupon bond Bond for which the coupon changes when the bond's credit rating changes.

Credit-linked note (CLN) Fixed-income security in which the holder of the security has the right to withhold payment of the full amount due at maturity if a credit event occurs.

Credit migration risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *downgrade risk*.

Credit risk The risk of loss caused by a counterparty's or debtor's failure to make a promised payment. Also called *default risk*.

Credit spread option An option on the yield spread on a bond.

Credit tranching A structure used to redistribute the credit risk associated with the collateral; a set of bond classes created to allow investors a choice in the amount of credit risk that they prefer to bear.

Creditworthiness The perceived ability of the borrower to pay its debt obligations in a timely manner; it represents the ability of a company to withstand adverse impacts on its cash flows.

Critical values Values of the test statistic at which the decision changes from fail to reject the null hypothesis to reject the null hypothesis.

Cross-default Covenant or contract clause that specifies a borrower is considered in default if they default on another debt obligation.

Cross-default provisions Provisions whereby events of default, such as non-payment of interest on one bond, trigger default on all outstanding debt; implies the same default probability for all issues.

Cross-price elasticity of demand The percentage change in quantity demanded for a given percentage change in the price of another good; the responsiveness of the demand for Product A that is associated with the change in price of Product B.

Cross-sectional analysis Analysis that involves comparisons across individuals in a group over a given time period or at a given point in time.

Cross-sectional data A list of the observations of a specific variable from multiple observational units at a given point in time. The observational units can be individuals, groups, companies, trading markets, regions, etc.

Crossing networks Trading systems that match buyers and sellers who are willing to trade at prices obtained from other markets.

Crowding out The thesis that government borrowing may divert private sector investment from taking place.

Cryptocurrency An electronic medium of exchange that lacks physical form.

Cryptography An algorithmic process to encrypt data, making the data unusable if received by unauthorized parties.

Cumulative absolute frequency Cumulates (i.e., adds up) in a frequency distribution the absolute frequencies as one moves from the first bin to the last bin.

Cumulative distribution function A function giving the probability that a random variable is less than or equal to a specified value.

Cumulative frequency distribution chart A chart that plots either the cumulative absolute frequency or the cumulative relative frequency on the y-axis against the upper limit of the interval and allows one to see the number or the percentage of the observations that lie below a certain value.

Cumulative preference shares Preference shares for which any dividends that are not paid accrue and must be paid in full before dividends on common shares can be paid.

Cumulative relative frequency A sequence of partial sums of the relative frequencies in a frequency distribution.

Cumulative voting A voting process whereby shareholders can accumulate and vote all their shares for a single candidate in an election, as opposed to having to allocate their voting rights evenly among all candidates.

Currencies Monies issued by national monetary authorities.

Currency option bonds Bonds that give bondholders the right to choose the currency in which they want to receive interest payments and principal repayments.

Currency swap A swap in which each party makes interest payments to the other in different currencies.

Current account A component of the balance of payments account that measures the flow of goods and services.

Current assets Assets that are expected to be consumed or converted into cash in the near future, typically one year or less. *Also called liquid assets.*

Current cost With reference to assets, the amount of cash or cash equivalents that would have to be paid to buy the same or an equivalent asset today; with reference to liabilities, the undiscounted amount of cash or cash equivalents that would be required to settle the obligation today.

Current government spending With respect to government expenditures, spending on goods and services that are provided on a regular, recurring basis including health, education, and defense.

Current liabilities Short-term obligations, such as accounts payable, wages payable, or accrued liabilities, that are expected to be settled in the near future, typically one year or less.

Current yield The sum of the coupon payments received over the year divided by the flat price; also called the *income* or *interest yield* or *running yield*.

Curve duration The sensitivity of the bond price (or the market value of a financial asset or liability) with respect to a benchmark yield curve.

Customs union Extends the free trade area (FTA) by not only allowing free movement of goods and services among members, but also creating a common trade policy against nonmembers.

CVaR Conditional VaR, a tail loss measure. The weighted average of all loss outcomes in the statistical distribution that exceed the VaR loss.

Cyclical A cyclical company is one whose profits are strongly correlated with the strength of the overall economy.

Cyclical companies Companies with sales and profits that regularly expand and contract with the business cycle or state of economy.

Daily settlement See *mark to market* and *marking to market.*

Dark pools Alternative trading systems that do not display the orders that their clients send to them.

Data A collection of numbers, characters, words, and text—as well as images, audio, and video—in a raw or organized format to represent facts or information.

Data mining The practice of determining a model by extensive searching through a dataset for statistically significant patterns. Also called *data snooping.*

Data science An interdisciplinary field that brings computer science, statistics, and other disciplines together to analyze and produce insights from Big Data.

Data snooping The practice of determining a model by extensive searching through a dataset for statistically significant patterns.

Data table see **two-dimensional rectangular array**.

Day order An order that is good for the day on which it is submitted. If it has not been filled by the close of business, the order expires unfilled.

Days of inventory on hand An activity ratio equal to the number of days in the period divided by inventory turnover over the period.

Dealers A financial intermediary that acts as a principal in trades.

Dealing securities Securities held by banks or other financial intermediaries for trading purposes.

Death cross See **Bearish crossover**.

Debentures Type of bond that can be secured or unsecured.

Debt rating An independent measure of the quality and safety of a company's debt based upon an analysis of the company's ability to pay the promised cash flows in full and on time. It includes an analysis of any indentures.

Debt-rating approach A method for estimating a company's before-tax cost of debt based on the yield on comparably rated bonds for maturities that closely match that of the company's existing debt.

Debt tax shield The tax benefit from interest paid on debt being tax deductible from income, equal to the marginal tax rate multiplied by the value of the debt.

Debt-to-assets ratio A solvency ratio calculated as total debt divided by total assets.

Debt-to-capital ratio A solvency ratio calculated as total debt divided by total debt plus total shareholders' equity.

Debt-to-equity ratio A solvency ratio calculated as total debt divided by total shareholders' equity.

Deciles Quantiles that divide a distribution into 10 equal parts.

Declaration date The day that the corporation issues a statement declaring a specific dividend.

Decreasing returns to scale When a production process leads to increases in output that are proportionately smaller than the increase in inputs.

Deductible temporary differences Temporary differences that result in a reduction of or deduction from taxable income in a future period when the balance sheet item is recovered or settled.

Deep learning Machine learning using neural networks with many hidden layers.

Deep learning nets Machine learning using neural networks with many hidden layers.

Default probability The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default risk*.

Default risk The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default probability*.

Default risk premium An extra return that compensates investors for the possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount.

Defensive companies Companies with sales and profits that have little sensitivity to the business cycle or state of the economy.

Defensive interval ratio A liquidity ratio that estimates the number of days that an entity could meet cash needs from liquid assets; calculated as (cash + short-term marketable investments + receivables) divided by daily cash expenditures.

Deferred coupon bond Bond that pays no coupons for its first few years but then pays a higher coupon than it otherwise normally would for the remainder of its life. Also called *split coupon bond*.

Deferred income A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred tax assets A balance sheet asset that arises when an excess amount is paid for income taxes relative to accounting profit. The taxable income is higher than accounting profit and income tax payable exceeds tax expense. The company expects to recover the difference during the course of future operations when tax expense exceeds income tax payable.

Deferred tax liabilities A balance sheet liability that arises when a deficit amount is paid for income taxes relative to accounting profit. The taxable income is less than the accounting profit and income tax payable is less than tax expense. The company expects to eliminate the liability over the course of future operations when income tax payable exceeds tax expense.

Defined benefit pension plans Plans in which the company promises to pay a certain annual amount (defined benefit) to the employee after retirement. The company bears the investment risk of the plan assets.

Defined contribution pension plans Individual accounts to which an employee and typically the employer makes contributions during their working years and expect to draw on the accumulated funds at retirement. The employee bears the investment and inflation risk of the plan assets.

Deflation Negative inflation.

Degree of confidence The probability that a confidence interval includes the unknown population parameter.

Degree of financial leverage (DFL) The ratio of the percentage change in net income to the percentage change in operating income; the sensitivity of the cash flows available to owners when operating income changes.

Degree of operating leverage (DOL) The ratio of the percentage change in operating income to the percentage change in units sold; the sensitivity of operating income to changes in units sold.

Degree of total leverage The ratio of the percentage change in net income to the percentage change in units sold; the sensitivity of the cash flows to owners to changes in the number of units produced and sold.

Degrees of freedom The number of independent variables used in defining sample statistics, such as variance, and the probability distributions they measure.

Delta The sensitivity of the derivative price to a small change in the value of the underlying asset.

Demand curve Graph of the inverse demand function. A graph showing the demand relation, either the highest quantity willingly purchased at each price or the highest price willingly paid for each quantity.

Demand function A relationship that expresses the quantity demanded of a good or service as a function of own-price and possibly other variables.

Demand shock A typically unexpected disturbance to demand, such as an unexpected interruption in trade or transportation.

Dependent With reference to events, the property that the probability of one event occurring depends on (is related to) the occurrence of another event.

Dependent variable The variable whose variation about its mean is to be explained by the regression; the left-side variable in a regression equation. Also referred to as the *explained variable*.

Depository bank A bank that raises funds from depositors and other investors and lends it to borrowers.

Depository institutions Commercial banks, savings and loan banks, credit unions, and similar institutions that raise funds from depositors and other investors and lend it to borrowers.

Depository receipt A security that trades like an ordinary share on a local exchange and represents an economic interest in a foreign company.

Depreciation The process of systematically allocating the cost of long-lived (tangible) assets to the periods during which the assets are expected to provide economic benefits.

Derivative pricing rule A pricing rule used by crossing networks in which a price is taken (derived) from the price that is current in the asset's primary market.

Derivatives A financial instrument whose value depends on the value of some underlying asset or factor (e.g., a stock price, an interest rate, or exchange rate).

Descriptive statistics Measures that summarize central tendency and spread variation in the data's distribution.

Diffuse prior The assumption of equal prior probabilities.

Diffusion index Reflects the proportion of the index's components that are moving in a pattern consistent with the overall index.

Diluted EPS The EPS that would result if all dilutive securities were converted into common shares.

Diluted shares The number of shares that would be outstanding if all potentially dilutive claims on common shares (e.g., convertible debt, convertible preferred stock, and employee stock options) were exercised.

Diminishing balance method An accelerated depreciation method, i.e., one that allocates a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Diminishing marginal productivity Describes a state in which each additional unit of input produces less output than previously.

Direct format With reference to the cash flow statement, a format for the presentation of the statement in which cash flow from operating activities is shown as operating cash receipts less operating cash disbursements. Also called *direct method*.

Direct investing Occurs when an investor makes a direct investment in an asset without the use of an intermediary.

Direct method See *direct format.*

Direct taxes Taxes levied directly on income, wealth, and corporate profits.

Direct write-off method An approach to recognizing credit losses on customer receivables in which the company waits until such time as a customer has defaulted and only then recognizes the loss.

Discount To reduce the value of a future payment in allowance for how far away it is in time; to calculate the present value of some future amount. Also, the amount by which an instrument is priced below its face (par) value.

Discount margin See *required margin.*

Discount rates In general, the interest rates used to calculate present values. In the money market, however, a discount rate is a specific type of quoted rate.

Discounted cash flow models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security.

Discouraged worker A person who has stopped looking for a job or has given up seeking employment.

Discrete data Numerical values that result from a counting process; therefore, practically speaking, the data are limited to a finite number of values.

Discrete random variable A random variable that can take on at most a countable number of possible values.

Discriminatory pricing rule A pricing rule used in continuous markets in which the limit price of the order or quote that first arrived determines the trade price.

Diseconomies of scale Increase in cost per unit resulting from increased production.

Dispersion The variability of a population or sample of observations around the central tendency.

Display size The size of an order displayed to public view.

Disposition effect As a result of loss aversion, an emotional bias whereby investors are reluctant to dispose of losers. This reluctance results in an inefficient and gradual adjustment to deterioration in fundamental value.

Distributed ledger A type of database that may be shared among entities in a network.

Distributed ledger technology Technology based on a distributed ledger.

Divergence In technical analysis, a term that describes the case when an indicator moves differently from the security being analyzed.

Diversification ratio The ratio of the standard deviation of an equally weighted portfolio to the standard deviation of a randomly selected security.

Dividend A distribution paid to shareholders based on the number of shares owned.

Dividend discount model (DDM) A present value model that estimates the intrinsic value of an equity share based on the present value of its expected future dividends.

Dividend payout ratio The ratio of cash dividends paid to earnings for a period.

Divisor A number (denominator) used to determine the value of a price return index. It is initially chosen at the inception of an index and subsequently adjusted by the index provider, as necessary, to avoid changes in the index value that are unrelated to changes in the prices of its constituent securities.

Doji In the Japanese terminology used in candlestick charting, the doji signifies that after a full day of trading, the positive price influence of buyers and the negative price influence of sellers exactly counteracted each other—with opening and closing prices that are virtually equal—which suggests that the market under analysis is in balance.

Domestic content provisions Stipulate that some percentage of the value added or components used in production should be of domestic origin.

Double bottom In technical analysis, a reversal pattern that is formed when the price reaches a low, rebounds, and then declines back to the first low level. A double bottom is used to predict a change from a downtrend to an uptrend.

Double coincidence of wants A prerequisite to barter trades, in particular that both economic agents in the transaction want what the other is selling.

Double declining balance depreciation An accelerated depreciation method that involves depreciating the asset at double the straight-line rate. This rate is multiplied by the book value of the asset at the beginning of the period (a declining balance) to calculate depreciation expense.

Double top In technical analysis, a reversal pattern that is formed when an uptrend reverses twice at roughly the same high price level. A double top is used to predict a change from an uptrend to a downtrend.

Down transition probability The probability that an asset's value moves down in a model of asset price dynamics.

Downgrade risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *credit migration risk.*

Downside risk Risk of incurring returns below a specified value.

Downtrend A pattern that occurs when the price of an asset moves lower over a period of time.

Drag on liquidity When receipts lag, creating pressure from the decreased available funds.

Drawdown A percentage peak-to-trough reduction in net asset value.

Dual-currency bonds Bonds that make coupon payments in one currency and pay the par value at maturity in another currency.

DuPont analysis An approach to decomposing return on investment, e.g., return on equity, as the product of other financial ratios.

Duration A measure of the approximate sensitivity of a security to a change in interest rates (i.e., a measure of interest rate risk).

Duration gap A bond's Macaulay duration minus the investment horizon.

Dutch Book Theorem A result in probability theory stating that inconsistent probabilities create profit opportunities.

Early repayment option See *prepayment option*.

Earnings per share The amount of income earned during a period per share of common stock.

Earnings surprise The portion of a company's earnings that is unanticipated by investors and, according to the efficient market hypothesis, merits a price adjustment.

Economic costs All the remuneration needed to keep a productive resource in its current employment or to acquire the resource for productive use; the sum of total accounting costs and implicit opportunity costs.

Economic indicator A variable that provides information on the state of the overall economy.

Economic loss The amount by which accounting profit is less than normal profit.

Economic profit Equal to accounting profit less the implicit opportunity costs not included in total accounting costs; the difference between total revenue (TR) and total cost (TC). Also called *abnormal profit* or *supernormal profit*.

Economic stabilization Reduction of the magnitude of economic fluctuations.

Economic union Incorporates all aspects of a common market and in addition requires common economic institutions and coordination of economic policies among members.

Economies of scale Reduction in cost per unit resulting from increased production.

Effective annual rate The amount by which a unit of currency will grow in a year with interest on interest included.

Effective convexity A *curve convexity* statistic that measures the secondary effect of a change in a benchmark yield curve on a bond's price.

Effective duration The sensitivity of a bond's price to a change in a benchmark yield curve.

Effective interest rate The borrowing rate or market rate that a company incurs at the time of issuance of a bond.

Efficient market A market in which asset prices reflect new information quickly and rationally.

Elastic Said of a good or service when the magnitude of elasticity is greater than one.

Elasticity The percentage change in one variable for a percentage change in another variable; a general measure of how sensitive one variable is to a change in the value of another variable.

Elasticity of demand A measure of the sensitivity of quantity demanded to a change in a product's own price: $\%\Delta Q^D/\%\Delta P$.

Elasticity of supply A measure of the sensitivity of quantity supplied to a change in price: $\%\Delta Q^S/\%\Delta P$.

Electronic communications networks See *alternative trading systems*.

Embedded option Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Embedded options Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Emotional biases Behavioral biases resulting from reasoning influenced by feelings; emotional biases stem from impulse or intuition.

Empirical duration The use of statistical methods and historical bond prices to estimate the price–yield relationship for a specific bond or portfolio of bonds.

Empirical probability The probability of an event estimated as a relative frequency of occurrence.

Employed The number of people with a job.

Endowment bias An emotional bias in which people value an asset more when they hold rights to it than when they do not.

Engagement/active ownership An ESG investment approach that uses shareholder power to influence corporate behavior through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is directed by ESG guidelines.

Enterprise risk management An overall assessment of a company's risk position. A centralized approach to risk management sometimes called firmwide risk management.

Enterprise value A measure of a company's total market value from which the value of cash and short-term investments have been subtracted.

Equal weighting An index weighting method in which an equal weight is assigned to each constituent security at inception.

Equipment trust certificates Bonds secured by specific types of equipment or physical assets.

Equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Equity risk premium The expected return on equities minus the risk-free rate; the premium that investors demand for investing in equities.

Equity swap A swap transaction in which at least one cash flow is tied to the return to an equity portfolio position, often an equity index.

Error term The difference between an observation and its expected value, where the expected value is based on the true underlying population relation between the dependent and independent variables. Also known simply as the *error*.

ESG An acronym that encompasses environmental, social, and governance.

ESG integration An ESG investment approach that focuses on systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions for the purpose of achieving the product's stated investment objectives. Used interchangeably with **ESG investing**.

ESG investing The consideration of environmental, social, and governance factors in the investment process. Used interchangeably with **ESG integration**.

Estimate The particular value calculated from sample observations using an estimator.

Estimated parameters With reference to a regression analysis, the estimated values of the population intercept and population slope coefficients in a regression.

Estimator An estimation formula; the formula used to compute the sample mean and other sample statistics are examples of estimators.

Ethical principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Ethics The study of moral principles or of making good choices. Ethics encompasses a set of moral principles and rules of conduct that provide guidance for our behavior.

Eurobonds Type of bond issued internationally, outside the jurisdiction of the country in whose currency the bond is denominated.

European-style Said of an option contract that can only be exercised on the option's expiration date.

Event Any outcome or specified set of outcomes of a random variable.

Ex-dividend date The first date that a share trades without (i.e., "ex") the dividend.

Excess kurtosis Degree of kurtosis (fatness of tails) relative to the kurtosis of the normal distribution.

Exchanges Places where traders can meet to arrange their trades.

Execution instructions Instructions that indicate how to fill an order.

Exercise The process of using an option to buy or sell the underlying.

Exercise price The fixed price at which an option holder can buy or sell the underlying. Also called *strike price*, *striking price*, or *strike*.

Exercise value The value obtained if an option is exercised based on current conditions. Also known as *intrinsic value*.

Exhaustive Covering or containing all possible outcomes.

Expansion The period of a business cycle after its lowest point and before its highest point.

Expansionary Tending to cause the real economy to grow.

Expansionary fiscal policy Fiscal policy aimed at achieving real economic growth.

Expected inflation The level of inflation that economic agents expect in the future.

Expected loss Default probability times loss severity given default.

Expected return on the portfolio ($E(R_p)$) The weighted average of the expected returns (R_1 to R_n) on the component securities using their respective weights (w_1 to w_n).

Expected value The probability-weighted average of the possible outcomes of a random variable.

Expenses Outflows of economic resources or increases in liabilities that result in decreases in equity (other than decreases because of distributions to owners); reductions in net assets associated with the creation of revenues.

Export subsidy Paid by the government to the firm when it exports a unit of a good that is being subsidized.

Exports Goods and services that an economy sells to other countries.

Extension risk The risk that when interest rates rise, fewer prepayments will occur because homeowners are reluctant to give up the benefits of a contractual interest rate that now looks low. As a result, the security becomes longer in maturity than anticipated at the time of purchase.

Externalities Spillover effects of production and consumption activities onto others who are not directly involved in a particular transaction, activity, or decision.

Externality An effect of a market transaction that is borne by parties other than those who transacted.

Extra dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Extreme value theory A branch of statistics that focuses primarily on extreme outcomes.

Face value The amount of cash payable by a company to the bondholders when the bonds mature; the promised payment at maturity separate from any coupon payment.

Factoring arrangement When a company sells its accounts receivable to a lender (known as a factor) who assumes responsibility for the credit granting and collection process.

Fair value The amount at which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's-length transaction; the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants.

False discovery approach An adjustment in the p-values for tests performed multiple times.

False discovery rate The rate of Type I errors in testing a null hypothesis multiple times for a given level of significance.

Fat-Tailed Describes a distribution that has fatter tails than a normal distribution (also called leptokurtic).

Fed funds rate The US interbank lending rate on overnight borrowings of reserves.

Federal funds rate The US interbank lending rate on overnight borrowings of reserves.

Fiat money Money that is not convertible into any other commodity.

Fiduciary An entity designated to represent the rights and responsibilities of a beneficiary whose assets they are managing, such as a bond trustee acting on behalf of fixed-income investors.

Fiduciary call A combination of a European call and a risk-free bond that matures on the option expiration day and has a face value equal to the exercise price of the call.

FIFO method The first in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the order in which they were placed in inventory.

Fill or kill See *immediate or cancel order*.

Finance lease From the lessee perspective, under US GAAP, a type of lease which is more akin to the purchase of an asset by the lessee. From the lessor perspective, under IFRS, a lease which "transfers substantially all the risks and rewards incidental to ownership of an underlying asset."

Financial account A component of the balance of payments account that records investment flows.

Financial distress Heightened uncertainty regarding a company's ability to meet its various obligations because of diminished earnings power or actual current losses.

Financial flexibility The ability to react and adapt to financial adversity and opportunities.

Financial leverage The extent to which a company can effect, through the use of debt, a proportional change in the return on common equity that is greater than a given proportional change in operating income; also, short for the financial leverage ratio.

Financial leverage ratio A measure of financial leverage calculated as average total assets divided by average total equity.

Financial risk The risk that environmental, social, or governance risk factors will result in significant costs or other losses to a company and its shareholders; the risk arising from a company's obligation to meet required payments under its financing agreements.

Financing activities Activities related to obtaining or repaying capital to be used in the business (e.g., equity and long-term debt).

Fintech Technological innovation in the design and delivery of financial services and products in the financial industry.

Firm commitment offering See *underwritten offering*.

First-degree price discrimination Where a monopolist is able to charge each customer the highest price the customer is willing to pay.

First lien debt Debt secured by a pledge of certain assets that could include buildings, but it may also include property and equipment, licenses, patents, brands, etc.

First mortgage debt Debt secured by a pledge of a specific property.

Fiscal multiplier The ratio of a change in national income to a change in government spending.

Fiscal policy The use of taxes and government spending to affect the level of aggregate expenditures.

Fisher effect The thesis that the real rate of interest in an economy is stable over time so that changes in nominal interest rates are the result of changes in expected inflation.

Fixed charge coverage A solvency ratio measuring the number of times interest and lease payments are covered by operating income, calculated as (EBIT + lease payments) divided by (interest payments + lease payments).

Fixed costs Costs that remain at the same level regardless of a company's level of production and sales.

Fixed-for-floating interest rate swap An interest rate swap in which one party pays a fixed rate and the other pays a floating rate, with both sets of payments in the same currency. Also called *plain vanilla swap* or *vanilla swap*.

Fixed-rate perpetual preferred stock Nonconvertible, noncallable preferred stock that has a fixed dividend rate and no maturity date.

Flag A technical analysis continuation pattern formed by parallel trendlines, typically over a short period.

Flat price The full price of a bond minus the accrued interest; also called the *quoted* or *clean* price.

Float-adjusted market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by adjusting its market capitalization for its market float.

Floaters See *floating-rate notes*.

Floating-rate notes Notes on which interest payments are not fixed but instead vary from period to period depending on the current level of a reference interest rate.

Flotation cost Fees charged to companies by investment bankers and other costs associated with raising new capital.

Foreclosure Allows the lender to take possession of a mortgaged property if the borrower defaults and then sell it to recover funds.

Foreign currency reserves Holding by the central bank of non-domestic currency deposits and non-domestic bonds.

Foreign direct investment Direct investment by a firm in one country (the source country) in productive assets in a foreign country (the host country).

Foreign exchange gains (or losses) Gains (or losses) that occur when the exchange rate changes between the investor's currency and the currency that foreign securities are denominated in.

Foreign portfolio investment Shorter-term investment by individuals, firms, and institutional investors (e.g., pension funds) in foreign financial instruments such as foreign stocks and foreign government bonds.

Forward commitments Class of derivatives that provides the ability to lock in a price to transact in the future at a previously agreed-upon price.

Forward contract An agreement between two parties in which one party, the buyer, agrees to buy from the other party, the seller, an underlying asset at a later date for a price established at the start of the contract.

Forward curve A series of forward rates, each having the same time frame.

Forward market For future delivery, beyond the usual settlement time period in the cash market.

Forward price The fixed price or rate at which the transaction scheduled to occur at the expiration of a forward contract will take place. This price is agreed on at the initiation date of the contract.

Forward rate The interest rate on a bond or money market instrument traded in a forward market. A forward rate can be interpreted as an incremental, or marginal, return for extending the time-to-maturity for an additional time period.

Forward rate agreements A forward contract calling for one party to make a fixed interest payment and the other to make an interest payment at a rate to be determined at the contract expiration.

Fractile A value at or below which a stated fraction of the data lies. Also called quantile.

Fractional reserve banking Banking in which reserves constitute a fraction of deposits.

Framing bias An information-processing bias in which a person answers a question differently based on the way in which it is asked (framed).

Free cash flow The actual cash that would be available to the company's investors after making all investments necessary to maintain the company as an ongoing enterprise (also referred to as free cash flow to the firm); the internally generated funds that can be distributed to the company's investors (e.g., shareholders and bondholders) without impairing the value of the company.

Free cash flow hypothesis The hypothesis that higher debt levels discipline managers by forcing them to make fixed debt service payments and by reducing the company's free cash flow.

Free-cash-flow-to-equity models Valuation models based on discounting expected future free cash flow to equity.

Free float The number of shares that are readily and freely tradable in the secondary market.

Free trade When there are no government restrictions on a country's ability to trade.

Free trade areas One of the most prevalent forms of regional integration, in which all barriers to the flow of goods and services among members have been eliminated.

Frequency distribution A tabular display of data constructed either by counting the observations of a variable by distinct values or groups or by tallying the values of a numerical variable into a set of numerically ordered bins (also called a one-way table).

Frequency polygon A graph of a frequency distribution obtained by drawing straight lines joining successive points representing the class frequencies.

Frequency table A representation of the frequency of occurrence of two discrete variables.

Full price The price of a security with accrued interest; also called the *invoice* or *dirty* price.

Fund investing In fund investing, the investor invests in assets indirectly by contributing capital to a fund as part of a group of investors. Fund investing is available for all major alternative investment types.

Fundamental analysis The examination of publicly available information and the formulation of forecasts to estimate the intrinsic value of assets.

Fundamental value The underlying or true value of an asset based on an analysis of its qualitative and quantitative characteristics. Also called *intrinsic value*.

Fundamental weighting An index weighting method in which the weight assigned to each constituent security is based on its underlying company's size. It attempts to address the disadvantages of market-capitalization weighting by using measures that are independent of the constituent security's price.

Funds of hedge funds Funds that hold a portfolio of hedge funds, more commonly shortened to *funds of funds*.

Fungible Freely exchangeable, interchangeable, or substitutable with other things of the same type. Money and commodities are the most common examples.

Future value (FV) The amount to which a payment or series of payments will grow by a stated future date.

Futures contract A variation of a forward contract that has essentially the same basic definition but with some additional features, such as a clearinghouse guarantee against credit losses, a daily settlement of gains and losses, and an organized electronic or floor trading facility.

Futures price The agreed-upon price of a futures contract.

FX swap The combination of a spot and a forward FX transaction.

G-spread The yield spread in basis points over an actual or interpolated government bond.

Gains Asset inflows not directly related to the ordinary activities of the business.

Game theory The set of tools decision makers use to incorporate responses by rival decision makers into their strategies.

Gamma A numerical measure of how sensitive an option's delta (the sensitivity of the derivative's price) is to a change in the value of the underlying.

Gap opening A gap is an area of a chart where a security's price either rises or falls from the previous day's close with no trading occurring in between. A gap opening is the start of a new trading session with a gap.

GDP deflator A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

General partner (GP) The partner that runs the business and theoretically bears unlimited liability for the business's debts and obligations.

Geometric mean A measure of central tendency computed by taking the nth root of the product of n non-negative values.

Giffen goods Goods that are consumed more as the price of the good rises because it is a very inferior good whose income effect overwhelms its substitution effect when price changes.

Gilts Bonds issued by the UK government.

Global depository receipt A depository receipt that is issued outside of the company's home country and outside of the United States.

Global minimum-variance portfolio The portfolio on the minimum-variance frontier with the smallest variance of return.

Global registered share A common share that is traded on different stock exchanges around the world in different currencies.

Gold standard With respect to a currency, if a currency is on the gold standard a given amount can be converted into a prespecified amount of gold.

Golden Cross See **Bullish crossover**.

Good-on-close An execution instruction specifying that an order can only be filled at the close of trading. Also called *market on close*.

Good-on-open An execution instruction specifying that an order can only be filled at the opening of trading.

Good-till-cancelled order An order specifying that it is valid until the entity placing the order has cancelled it (or, commonly, until some specified amount of time such as 60 days has elapsed, whichever comes sooner).

Goodwill An intangible asset that represents the excess of the purchase price of an acquired company over the value of the net assets acquired.

Government equivalent yield A yield that restates a yield-to-maturity based on a 30/360 day count to one based on actual/actual.

Green bonds Bonds used in green finance whereby the proceeds are earmarked toward environmental-related products.

Green finance A type of finance that addresses environmental concerns while achieving economic growth.

Green loans Any loan instruments made available exclusively to finance or re-finance, in whole or in part, new and/or existing eligible green projects. Green loans are commonly aligned in the market with the Green Loan Principles.

Greenfield investment Investing in infrastructure assets *that are to be constructed*.

Grey market The forward market for bonds about to be issued. Also called "when issued" market.

Gross domestic product The market value of all final goods and services produced within the economy during a given period (output definition) or, equivalently, the aggregate income earned by all households, all companies, and the government within the economy during a given period (income definition).

Gross margin Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit margin The ratio of gross profit to revenues.

Grouped bar chart A bar chart for showing joint frequencies for two categorical variables (also known as a **clustered bar chart**).

Grouping by function With reference to the presentation of expenses in an income statement, the grouping together of expenses serving the same function, e.g. all items that are costs of goods sold.

Grouping by nature With reference to the presentation of expenses in an income statement, the grouping together of expenses by similar nature, e.g., all depreciation expenses.

Growth cycle Refers to fluctuations in economic activity around the long-term potential trend growth level, focusing on how much actual economic activity is below or above trend growth in economic activity.

Growth cyclical A term sometimes used to describe companies that are growing rapidly on a long-term basis but that still experience above-average fluctuation in their revenues and profits over the course of a business cycle.

Growth investors With reference to equity investors, investors who seek to invest in high-earnings-growth companies.

Growth option The ability to make additional investments in a project at some future time if the financial results are strong. Also called *expansion option*.

Growth rate cycle Refers to fluctuations in the growth rate of economic activity.

Guarantee certificate A type of structured financial instrument that provides investors with capital protection. It combines a zero-coupon bond and a call option on some underlying asset.

Haircut See *repo margin*.

Halo effect An emotional bias that extends a favorable evaluation of some characteristics to other characteristics.

Hard-bullet covered bonds Covered bonds for which a bond default is triggered and bond payments are accelerated in the event of sponsor default if payments do not occur according to the original maturity schedule.

Harmonic mean A type of weighted mean computed as the reciprocal of the arithmetic average of the reciprocals.

Head and shoulders pattern In technical analysis, a reversal pattern that is formed in three parts: a left shoulder, a head, and a right shoulder. A head and shoulders pattern is used to predict a change from an uptrend to a downtrend.

Headline inflation Refers to the inflation rate calculated based on the price index that includes all goods and services in an economy.

Heat map A type of graphic that organizes and summarizes data in a tabular format and represents it using a color spectrum.

Hedge funds Private investment vehicles that typically use leverage, derivatives, and long and short investment strategies.

Hedge portfolio A hypothetical combination of the derivative and its underlying that eliminates risk.

Held-to-maturity Debt (fixed-income) securities that a company intends to hold to maturity; these are presented at their original cost, updated for any amortisation of discounts or premiums.

Herding Clustered trading that may or may not be based on information.

Heteroskedasticity A violation of the simple linear regression assumption of homoskedasticity; with heteroskedasticity, the variability of the residuals is not equal across all observations.

Hidden order An order that is exposed not to the public but only to the brokers or exchanges that receive it.

High-frequency trading A form of algorithmic trading that makes use of vast quantities of data to execute trades on ultra-high-speed networks in fractions of a second.

High-water mark The highest value, net of fees, which a fund has reached in history. It reflects the highest cumulative return used to calculate an incentive fee.

Hindsight bias A bias with selective perception and retention aspects in which people may see past events as having been predictable and reasonable to expect.

Histogram A chart that presents the distribution of numerical data by using the height of a bar or column to represent the absolute frequency of each bin or interval in the distribution.

Historical cost In reference to assets, the amount paid to purchase an asset, including any costs of acquisition and/or preparation; with reference to liabilities, the amount of proceeds received in exchange in issuing the liability.

Historical equity risk premium approach An estimate of a country's equity risk premium that is based on the historical averages of the risk-free rate and the rate of return on the market portfolio.

Holder-of-record date The date that a shareholder listed on the corporation's books will be deemed to have ownership of the shares for purposes of receiving an upcoming dividend.

Holding period return The return that an investor earns during a specified holding period; a synonym for total return.

Home bias A preference for securities listed on the exchanges of one's home country.

Homogeneity of expectations The assumption that all investors have the same economic expectations and thus have the same expectations of prices, cash flows, and other investment characteristics.

Homoskedasticity An assumption of simple linear regression such that the variability of the residuals is equal across all observations.

Horizon yield The internal rate of return between the total return (the sum of reinvested coupon payments and the sale price or redemption amount) and the purchase price of the bond.

Horizontal analysis Common-size analysis that involves comparing a specific financial statement with that statement in prior or future time periods; also, cross-sectional analysis of one company with another.

Horizontal demand schedule Implies that at a given price, the response in the quantity demanded is infinite.

Hostile takeover An attempt by one entity to acquire a company without the consent of the company's management.

Household A person or a group of people living in the same residence, taken as a basic unit in economic analysis.

Human capital The accumulated knowledge and skill that workers acquire from education, training, or life experience and the corresponding present value of future earnings to be generated by said skilled individual.

Hurdle rate The minimum rate of return on investment that a fund must reach before a GP receives carried interest.

Hypothesis A proposed explanation or theory that can be tested.

Hypothesis testing The process of testing of hypotheses about one or more populations using statistical inference.

I-spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Iceberg order An order in which the display size is less than the order's full size.

If-converted method A method for accounting for the effect of convertible securities on earnings per share (EPS) that specifies what EPS would have been if the convertible securities had been converted at the beginning of the period, taking account of the effects of conversion on net income and the weighted average number of shares outstanding.

Illusion of control bias A bias in which people tend to believe that they can control or influence outcomes when, in fact, they cannot.

Immediate or cancel order An order that is valid only upon receipt by the broker or exchange. If such an order cannot be filled in part or in whole upon receipt, it cancels immediately. Also called *fill or kill*.

Impact investing Investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.

Impact lag The lag associated with the result of actions affecting the economy with delay.

Implicit price deflator for GDP A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

Implicit selection bias One type of selection bias introduced through the presence of a threshold that filters out some unqualified members.

Implied forward rates Calculated from spot rates, an implied forward rate is a breakeven reinvestment rate that links the return on an investment in a shorter-term zero-coupon bond to the return on an investment in a longer-term zero-coupon bond.

Implied volatility The volatility that option traders use to price an option, implied by the price of the option and a particular option-pricing model.

Import license Specifies the quantity of a good that can be imported into a country.

Imports Goods and services that a domestic economy (i.e., house-holds, firms, and government) purchases from other countries.

In the money Options that, if exercised, would result in the value received being worth more than the payment required to exercise.

Income Increases in economic benefits in the form of inflows or enhancements of assets, or decreases of liabilities that result in an increase in equity (other than increases resulting from contributions by owners).

Income elasticity of demand A measure of the responsiveness of demand to changes in income, defined as the percentage change in quantity demanded divided by the percentage change in income.

Income tax paid The actual amount paid for income taxes in the period; not a provision, but the actual cash outflow.

Income tax payable The income tax owed by the company on the basis of taxable income.

Increasing marginal returns When the marginal product of a resource increases as additional units of that input are employed.

Increasing returns to scale When a production process leads to increases in output that are proportionately larger than the increase in inputs.

Incremental cash flow The cash flow that is realized because of a decision; the changes or increments to cash flows resulting from a decision or action.

Indenture Legal contract that describes the form of a bond, the obligations of the issuer, and the rights of the bondholders. Also called the *trust deed*.

Independent With reference to events, the property that the occurrence of one event does not affect the probability of another event occurring. With reference to two random variables X and Y, they are independent if and only if $P(X,Y) = P(X)P(Y)$.

Independent projects Independent projects are capital investments whose cash flows are independent of each other.

Independent variable A variable used to explain the dependent variable in a regression; a right-side variable in a regression equation. Also referred to as the *explanatory variable*.

Independently and identically distributed With respect to random variables, the property of random variables that are independent of each other but follow the identical probability distribution.

Index-linked bond Bond for which coupon payments and/or principal repayment are linked to a specified index.

Indexing An investment strategy in which an investor constructs a portfolio to mirror the performance of a specified index.

Indicator variable A variable that takes on only one of two values, 0 or 1, based on a condition. In simple linear regression, the slope is the difference in the dependent variable for the two conditions. Also referred to as a *dummy variable*.

Indifference curve A curve representing all the combinations of two goods or attributes such that the consumer is entirely indifferent among them.

Indirect format With reference to cash flow statements, a format for the presentation of the statement which, in the operating cash flow section, begins with net income then shows additions and subtractions to arrive at operating cash flow. Also called *indirect method*.

Indirect method See *indirect format*.

Indirect taxes Taxes such as taxes on spending, as opposed to direct taxes.

Industry A group of companies offering similar products and/or services.

Industry analysis The analysis of a specific branch of manufacturing, service, or trade.

Inelastic Said of a good or service when the magnitude of elasticity is less than one. Insensitive to price changes.

Inferior goods A good whose consumption decreases as income increases.

Inflation The percentage increase in the general price level from one period to the next; a sustained rise in the overall level of prices in an economy.

Inflation-linked bond Type of index-linked bond that offers investors protection against inflation by linking the bond's coupon payments and/or the principal repayment to an index of consumer prices. Also called *linkers*.

Inflation premium An extra return that compensates investors for expected inflation.

Inflation reports A type of economic publication put out by many central banks.

Inflation uncertainty The degree to which economic agents view future rates of inflation as difficult to forecast.

Information cascade The transmission of information from those participants who act first and whose decisions influence the decisions of others.

Information-motivated traders Traders that trade to profit from information that they believe allows them to predict future prices.

Informationally efficient market A market in which asset prices reflect new information quickly and rationally.

Initial coin offering An unregulated process whereby companies raise capital by selling crypto tokens to investors in exchange for fiat money or another agreed-upon cryptocurrency.

Initial margin The amount that must be deposited in a clearinghouse account when entering into a futures contract.

Initial margin requirement The margin requirement on the first day of a transaction as well as on any day in which additional margin funds must be deposited.

Initial public offering (IPO) The first issuance of common shares to the public by a formerly private corporation.

Input productivity The amount of output produced by workers in a given period of time—for example, output per hour worked; measures the efficiency of labor.

Intangible assets Assets lacking physical substance, such as patents and trademarks.

Interbank market The market of loans and deposits between banks for maturities ranging from overnight to one year.

Intercept The expected value of the dependent variable when the independent variable in a simple linear regression is equal to zero.

Interest Payment for lending funds.

Interest coverage A solvency ratio calculated as EBIT divided by interest payments.

Interest-only mortgage A loan in which no scheduled principal repayment is specified for a certain number of years.

Interest rate A rate of return that reflects the relationship between differently dated cash flows; a discount rate.

Interest rate effect The effect through which price level changes, through demand for money, impact interest rate, which in turn impacts investment and consumption.

Interest rate swap A swap in which the underlying is an interest rate. Can be viewed as a currency swap in which both currencies are the same and can be created as a combination of currency swaps.

Intermarket analysis A field within technical analysis that combines analysis of the major categories of securities—namely, equities, bonds, currencies, and commodities—to identify market trends and possible inflections in trends.

Internal rate of return The discount rate that makes net present value equal 0; the discount rate that makes the present value of an investment's costs (outflows) equal to the present value of the investment's benefits (inflows).

Internet of Things A network arrangement of structures and devices whereby the objects on the network are able to interact and share information.

Interpolated spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Interquartile range The difference between the third and first quartiles of a dataset.

Interval With reference to grouped data, a set of values within which an observation falls.

Intrinsic value See *exercise value.*

Inventory investment Net change in business inventory.

Inventory turnover An activity ratio calculated as cost of goods sold divided by average inventory.

Inverse demand function A restatement of the demand function in which price is stated as a function of quantity.

Inverse floater A type of leveraged structured financial instrument. The cash flows are adjusted periodically and move in the opposite direction of changes in the reference rate.

Inverse transformation method A method using randomly generated numbers from the continuous uniform distribution to generate random observations from any distribution.

Investing activities Activities associated with the acquisition and disposal of property, plant, and equipment; intangible assets; other long-term assets; and both long-term and short-term investments in the equity and debt (bonds and loans) issued by other companies.

Investment banks Financial intermediaries that provide advice to their mostly corporate clients and help them arrange transactions such as initial and seasoned securities offerings.

Investment policy statement (IPS) A written planning document that describes a client's investment objectives and risk tolerance over a relevant time horizon, along with constraints that apply to the client's portfolio.

Investment property Property used to earn rental income or capital appreciation (or both).

Jackknife A resampling method that repeatedly draws samples by taking the original observed data sample and leaving out one observation at a time (without replacement) from the set.

January effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January. Also called *turn-of-the-year effect.*

Joint frequencies The entry in the cells of the contingency table that represent the joining of one variable from a row and the other variable from a column to count observations.

Joint probability The probability of the joint occurrence of stated events.

Joint probability function A function giving the probability of joint occurrences of values of stated random variables.

Judgmental Sampling A procedure of selectively handpicking elements from the population based on a researcher's knowledge and professional judgment.

Key rate duration A method of measuring the interest rate sensitivities of a fixed-income instrument or portfolio to shifts in key points along the yield curve.

Keynesians Economists who believe that fiscal policy can have powerful effects on aggregate demand, output, and employment when there is substantial spare capacity in an economy.

Kurtosis The statistical measure that indicates the combined weight of the tails of a distribution relative to the rest of the distribution.

Labor force The portion of the working age population (over the age of 16) that is employed or is available for work but not working (unemployed).

Labor productivity The quantity of goods and services (real GDP) that a worker can produce in one hour of work.

Lagging economic indicators Turning points that take place later than those of the overall economy; they are believed to have value in identifying the economy's past condition.

Law of demand The principle that as the price of a good rises, buyers will choose to buy less of it, and as its price falls, they will buy more.

Law of diminishing marginal returns The observation that a variable factor's marginal product must eventually fall as more of it is added to a fixed amount of the other factors.

Law of diminishing returns The smallest output that a firm can produce such that its long run average costs are minimized.

Law of one price The condition in a financial market in which two equivalent financial instruments or combinations of financial instruments can sell for only one price. Equivalent to the principle that no arbitrage opportunities are possible.

Lead underwriter The lead investment bank in a syndicate of investment banks and broker–dealers involved in a securities underwriting.

Leading economic indicators Turning points that usually precede those of the overall economy; they are believed to have value for predicting the economy's future state, usually near-term.

Lease A contract that conveys the right to use an asset for a period of time in exchange for consideration.

Legal tender Something that must be accepted when offered in exchange for goods and services.

Lender of last resort An entity willing to lend money when no other entity is ready to do so.

Leptokurtic Describes a distribution that has fatter tails than a normal distribution (also called fat-tailed).

Lessee The party obtaining the use of an asset through a lease.

Lessor The owner of an asset that grants the right to use the asset to another party.

Letter of credit Form of external credit enhancement whereby a financial institution provides the issuer with a credit line to reimburse any cash flow shortfalls from the assets backing the issue.

Level of significance The probability of a Type I error in testing a hypothesis.

Leverage In the context of corporate finance, leverage refers to the use of fixed costs within a company's cost structure. Fixed costs that are operating costs (such as depreciation or rent) create operating leverage. Fixed costs that are financial costs (such as interest expense) create financial leverage.

Leveraged buyout A transaction whereby the target company's management team converts the target to a privately held company by using heavy borrowing to finance the purchase of the target company's outstanding shares.

Liabilities Present obligations of an enterprise arising from past events, the settlement of which is expected to result in an outflow of resources embodying economic benefits; creditors' claims on the resources of a company.

LIFO layer liquidation With respect to the application of the LIFO inventory method, the liquidation of old, relatively low-priced inventory; happens when the volume of sales rises above the volume of recent purchases so that some sales are made from relatively old, low-priced inventory. Also called *LIFO liquidation*.

LIFO method The last in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the reverse order the items were placed in inventory (i.e., inventory produced or acquired last are assumed to be sold first).

LIFO reserve The difference between the reported LIFO inventory carrying amount and the inventory amount that would have been reported if the FIFO method had been used (in other words, the FIFO inventory value less the LIFO inventory value).

Likelihood The probability of an observation, given a particular set of conditions.

Limit down A limit move in the futures market in which the price at which a transaction would be made is at or below the lower limit.

Limit order Instructions to a broker or exchange to obtain the best price immediately available when filling an order, but in no event accept a price higher than a specified (limit) price when buying or accept a price lower than a specified (limit) price when selling.

Limit order book The book or list of limit orders to buy and sell that pertains to a security.

Limit up A limit move in the futures market in which the price at which a transaction would be made is at or above the upper limit.

Limitations on liens Meant to put limits on how much secured debt an issuer can have.

Limited partners Partners with limited liability. Limited partnerships in hedge and private equity funds are typically restricted to investors who are expected to understand and to be able to assume the risks associated with the investments.

Limited partnership agreement (LPA) A legal document that outlines the rules of the partnership and establishes the framework that ultimately guides the fund's operations throughout its life.

Lin-log model A regression model in which the independent variable is in logarithmic form.

Line chart A type of graph used to visualize ordered observations. In technical analysis, a plot of price data, typically closing prices, with a line connecting the points.

Linear interpolation The estimation of an unknown value on the basis of two known values that bracket it, using a straight line between the two known values.

Linear regression Regression that models the straight-line relationship between the dependent and independent variables. Also known as *least squares regression* and *ordinary least squares regression*.

Linear scale A scale in which equal distances correspond to equal absolute amounts. Also called an *arithmetic scale*.

Linker See *inflation-linked bond*.

Liquid market Said of a market in which traders can buy or sell with low total transaction costs when they want to trade.

Liquidation To sell the assets of a company, division, or subsidiary piecemeal, typically because of bankruptcy; the form of bankruptcy that allows for the orderly satisfaction of creditors' claims after which the company ceases to exist.

Liquidity The extent to which a company is able to meet its short-term obligations using cash flows and those assets that can be readily transformed into cash.

Liquidity premium An extra return that compensates investors for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

Liquidity ratios Financial ratios measuring the company's ability to meet its short-term obligations to creditors.

Liquidity risk The risk that a financial instrument cannot be purchased or sold without a significant concession in price due to the size of the market.

Liquidity trap A condition in which the demand for money becomes infinitely elastic (horizontal demand curve) so that injections of money into the economy will not lower interest rates or affect real activity.

Load fund A mutual fund in which, in addition to the annual fee, a percentage fee is charged to invest in the fund and/or for redemptions from the fund.

Loan-to-value ratio The amount of the property's mortgage divided by the purchase price

Locked limit A condition in the futures markets in which a transaction cannot take place because the price would be beyond the limits.

Lockup period The minimum holding period before investors are allowed to make withdrawals or redeem shares from a fund.

Log-lin model A regression model in which the dependent variable is in logarithmic form.

Log-log model A regression model in which both the dependent and independent variables are in logarithmic form. Also known as the *double-log model*.

Logarithmic scale A scale in which equal distances represent equal proportional changes in the underlying quantity.

London interbank offered rate (Libor) Collective name for multiple rates at which a select set of banks believes they could borrow unsecured funds from other banks in the London interbank market for different currencies and different borrowing periods ranging from overnight to one year.

Long The buyer of a derivative contract. Also refers to the position of owning a derivative.

Long-lived assets Assets that are expected to provide economic benefits over a future period of time, typically greater than one year. Also called *long-term assets*.

Long position A position in an asset or contract in which one owns the asset or has an exercisable right under the contract.

Long-run average total cost The curve describing average total cost when no costs are considered fixed.

Look-ahead bias A bias caused by using information that was unavailable on the test date.

Loss aversion The tendency of people to dislike losses more than they like comparable gains.

Loss-aversion bias A bias in which people tend to strongly prefer avoiding losses as opposed to achieving gains.

Loss severity Portion of a bond's value (including unpaid interest) an investor loses in the event of default.

Losses Asset outflows not directly related to the ordinary activities of the business.

Lower bound The lowest possible value of an option.

M^2 A measure of what a portfolio would have returned if it had taken on the same total risk as the market index.

M^2 alpha Difference between the risk-adjusted performance of the portfolio and the performance of the benchmark.

Macaulay duration The approximate amount of time a bond would have to be held for the market discount rate at purchase to be realized if there is a single change in interest rate. It indicates the point in time when the coupon reinvestment and price effects of a change in yield-to-maturity offset each other.

Machine learning Computer based techniques that seek to extract knowledge from large amounts of data by "learning" from known examples and then generating structure or predictions. ML algorithms aim to "find the pattern, apply the pattern."

Macroeconomics The branch of economics that deals with aggregate economic quantities, such as national output and national income.

Maintenance covenants Covenants in bank loan agreements that require the borrower to satisfy certain financial ratio tests while the loan is outstanding.

Maintenance margin The minimum amount that is required by a futures clearinghouse to maintain a margin account and to protect against default. Participants whose margin balances drop below the required maintenance margin must replenish their accounts.

Maintenance margin requirement The margin requirement on any day other than the first day of a transaction.

Management buy-in Leveraged buyout in which the current management team is being replaced and the acquiring team will be involved in managing the company.

Management buyout A leveraged buyout event in which a group of investors consisting primarily of the company's existing management purchases at least controlling interest in its outstanding shares. At the extreme, they may purchase all shares and take the company private.

Management fee A fee based on assets under management or committed capital, as applicable—also called a *base fee*.

MAR ratio A variation of the Calmar ratio that uses a full investment history and the average drawdown.

Margin The amount of money that a trader deposits in a margin account. The term is derived from the stock market practice in which an investor borrows a portion of the money required to purchase a certain amount of stock. In futures markets, there is no borrowing so the margin is more of a down payment or performance bond.

Margin bond A cash deposit required by the clearinghouse from the participants to a contract to provide a credit guarantee. Also called a *performance bond*.

Margin call A request for the short to deposit additional funds to bring their balance up to the initial margin.

Margin loan Money borrowed from a broker to purchase securities.

Marginal cost The cost of producing an additional unit of a good.

Marginal frequencies The sums determined by adding joint frequencies across rows or across columns in a contingency table.

Marginal product Measures the productivity of each unit of input and is calculated by taking the difference in total product from adding another unit of input (assuming other resource quantities are held constant).

Marginal propensity to consume The proportion of an additional unit of disposable income that is consumed or spent; the change in consumption for a small change in income.

Marginal propensity to save The proportion of an additional unit of disposable income that is saved (not spent).

Marginal revenue The change in total revenue divided by the change in quantity sold; simply, the additional revenue from selling one more unit.

Marginal value curve A curve describing the highest price consumers are willing to pay for each additional unit of a good.

Mark to market The revaluation of a financial asset or liability to its current market value or fair value.

Mark-to-market Refers to the current expected fair market value for which a given security would likely be available for purchase or sale if traded in current market conditions.

Market anomaly Change in the price or return of a security that cannot directly be linked to current relevant information known in the market or to the release of new information into the market.

Market bid–ask spread The difference between the best bid and the best offer.

Market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its market capitalization by the total market capitalization (sum of the market capitalization) of all securities in the index. Also called *value weighting*.

Market discount rate The rate of return required by investors given the risk of the investment in a bond; also called the *required yield* or the *required rate of return*.

Market float The number of shares that are available to the investing public.

Market liquidity risk The risk that the price at which investors can actually transact—buying or selling—may differ from the price indicated in the market.

Market model A regression equation that specifies a linear relationship between the return on a security (or portfolio) and the return on a broad market index.

Market multiple models Valuation models based on share price multiples or enterprise value multiples.

Market-on-close An execution instruction specifying that an order can only be filled at the close of trading.

Market order Instructions to a broker or exchange to obtain the best price immediately available when filling an order.

Market-oriented investors With reference to equity investors, investors whose investment disciplines cannot be clearly categorized as value or growth.

Market rate of interest The rate demanded by purchasers of bonds, given the risks associated with future cash payment obligations of the particular bond issue.

Market reference rate (MRR) Collective name for a set of rates covering different currencies for different maturities, ranging from overnight to one year.

Market risk The risk that arises from movements in interest rates, stock prices, exchange rates, and commodity prices.

Market value The price at which an asset or security can currently be bought or sold in an open market.

Marketable limit order A buy limit order in which the limit price is placed above the best offer, or a sell limit order in which the limit price is placed below the best bid. Such orders generally will partially or completely fill right away.

Markowitz efficient frontier The graph of the set of portfolios offering the maximum expected return for their level of risk (standard deviation of return).

Matching principle The accounting principle that expenses should be recognized in the same period in which the associated revenue is recognized.

Matrix pricing Process of estimating the market discount rate and price of a bond based on the quoted or flat prices of more frequently traded comparable bonds.

Maturity premium An extra return that compensates investors for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended.

Maturity structure A factor explaining the differences in yields on similar bonds; also called *term structure*.

Mean absolute deviation With reference to a sample, the mean of the absolute values of deviations from the sample mean.

Mean square error (MSE) The sum of squares error divided by the degrees of freedom, $n - k - 1$; in a simple linear regression, $n - k - 1 = n - 2$.

Mean square regression (MSR) The sum of squares regression divided by the number of independent variables k; in a simple linear regression, $k = 1$.

Mean–variance analysis An approach to portfolio analysis using expected means, variances, and covariances of asset returns.

Measure of central tendency A quantitative measure that specifies where data are centered.

Measure of value A standard for measuring value; a function of money.

Measures of location Quantitative measures that describe the location or distribution of data. They include not only measures of central tendency but also other measures, such as percentiles.

Median The value of the middle item of a set of items that has been sorted into ascending or descending order (i.e., the 50th percentile).

Medium of exchange Any asset that can be used to purchase goods and services or to repay debts; a function of money.

Medium-term note A corporate bond offered continuously to investors by an agent of the issuer, designed to fill the funding gap between commercial paper and long-term bonds.

Mental accounting bias An information-processing bias in which people treat one sum of money differently from another equal-sized sum based on which mental account the money is assigned to.

Menu costs A cost of inflation in which businesses constantly have to incur the costs of changing the advertised prices of their goods and services.

Mesokurtic Describes a distribution with kurtosis equal to that of the normal distribution, namely, kurtosis equal to three.

Microeconomics The branch of economics that deals with markets and decision making of individual economic units, including consumers and businesses.

Minimum efficient scale The smallest output that a firm can produce such that its long-run average total cost is minimized.

Minimum-variance portfolio The portfolio with the minimum variance for each given level of expected return.

Minority shareholders Particular shareholders or a block of shareholders holding a small proportion of a company's outstanding shares, resulting in a limited ability to exercise control in voting activities.

Minsky moment Named for Hyman Minksy. A point in a business cycle when, after individuals become overextended in borrowing to finance speculative investments, people start realizing that something is likely to go wrong and a panic ensues, leading to asset sell-offs.

Modal interval With reference to grouped data, the interval containing the greatest number of observations (i.e., highest frequency).

Mode The most frequently occurring value in a distribution.

Modern portfolio theory (MPT) The analysis of rational portfolio choices based on the efficient use of risk.

Modified duration A measure of the percentage price change of a bond given a change in its yield-to-maturity.

Momentum oscillator A graphical representation of market sentiment that is constructed from price data and calculated so that it oscillates either between a low and a high or around some number.

Monetarists Economists who believe that the rate of growth of the money supply is the primary determinant of the rate of inflation.

Monetary policy Actions taken by a nation's central bank to affect aggregate output and prices through changes in bank reserves, reserve requirements, or its target interest rate.

Monetary transmission mechanism The process whereby a central bank's interest rate gets transmitted through the economy and ultimately affects the rate of increase of prices.

Monetary union An economic union in which the members adopt a common currency.

Money A generally accepted medium of exchange and unit of account.

Money convexity For a bond, the annual or approximate convexity multiplied by the full price.

Money creation The process by which changes in bank reserves translate into changes in the money supply.

Money duration A measure of the price change in units of the currency in which the bond is denominated given a change in its yield-to-maturity.

Money market The market for short-term debt instruments (one-year maturity or less).

Money market securities Fixed-income securities with maturities at issuance of one year or less.

Money multiplier Describes how a change in reserves is expected to affect the money supply; in its simplest form, 1 divided by the reserve requirement.

Money neutrality The thesis that an increase in the money supply leads in the long-run to an increase in the price level, while leaving real variables like output and employment unaffected.

Money-weighted return The internal rate of return on a portfolio, taking account of all cash flows.

Moneyness The relationship between the price of the underlying and an option's exercise price.

Monopolistic competition Highly competitive form of imperfect competition; the competitive characteristic is a notably large number of firms, while the monopoly aspect is the result of product differentiation.

Monopoly In pure monopoly markets, there are no substitutes for the given product or service. There is a single seller, which exercises considerable power over pricing and output decisions.

Monte Carlo simulation A technique involving the use of computer software to represent the operation of a complex financial system. A characteristic feature is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Moral principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Mortgage-backed securities Debt obligations that represent claims to the cash flows from pools of mortgage loans, most commonly on residential property.

Mortgage loan A loan secured by the collateral of some specified real estate property that obliges the borrower to make a predetermined series of payments to the lender.

Mortgage pass-through security A security created when one or more holders of mortgages form a pool of mortgages and sell shares or participation certificates in the pool.

Mortgage rate The interest rate on a mortgage loan; also called *contract rate* or *note rate*.

Moving average The average of the closing price of a security over a specified number of periods. With each new period, the average is recalculated.

Moving-average convergence/divergence oscillator A momentum oscillator that is based on the difference between short-term and long-term moving averages of a security's price.

Multi-factor model A model that explains a variable in terms of the values of a set of factors.

Multi-market indexes Comprised of indexes from different countries, designed to represent multiple security markets.

Multi-step format With respect to the format of the income statement, a format that presents a subtotal for gross profit (revenue minus cost of goods sold).

Multilateral trading facilities See *alternative trading systems*.

Multinational corporation A company operating in more than one country or having subsidiary firms in more than one country.

Multinomial formula (general formula for labeling problems) The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by $\dfrac{n!}{n_1!n_2!\ldots n_k!}$.

Multiple of invested capital (MOIC) A simplified calculation that measures the total value of all distributions and residual asset values relative to an initial total investment; also known as a *money multiple*.

Multiple testing problem The risk of getting statistically significant test results when performing a test multiple times.

Multiplication rule for counting If one task can be done in n_1 ways, and a second task, given the first, can be done in n_2 ways, and a third task, given the first two tasks, can be done in n_3 ways, and so on for k tasks, then the number of ways the k tasks can be done is $(n_1)(n_2)(n_3) \ldots (n_k)$.

Multiplication rule for expected value of the product of uncorrelated random variables This rule is $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated.

Multiplication rule for independent events The rule that when two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B.

Multiplication rule for probability The rule that the joint probability of events A and B equals the probability of A given B times the probability of B.

Multiplier models Valuation models based on share price multiples or enterprise value multiples.

Multivariate distribution A probability distribution that specifies the probabilities for a group of related random variables.

Multivariate normal distribution A probability distribution for a group of random variables that is completely defined by the means and variances of the variables plus all the correlations between pairs of the variables.

Muni See *municipal bond*.

Municipal bond A type of non-sovereign bond issued by a state or local government in the United States. It very often (but not always) offers income tax exemptions.

Mutual fund A comingled investment pool in which investors in the fund each have a pro-rata claim on the income and value of the fund.

Mutually exclusive Indicates that only one event can occur at a time.

Mutually exclusive projects Mutually exclusive projects compete directly with each other. For example, if Projects A and B are mutually exclusive, you can choose A or B but you cannot choose both.

***n* Factorial** For a positive integer n, the product of the first n positive integers; 0 factorial equals 1 by definition. n factorial is written as $n!$.

Narrow money The notes and coins in circulation in an economy, plus other very highly liquid deposits.

Nash equilibrium When two or more participants in a non-coop-erative game have no incentive to deviate from their respective equilibrium strategies given their opponent's strategies.

National income The income received by all factors of production used in the generation of final output. National income equals gross domestic product (or, in some countries, gross national product) minus the capital consumption allowance and a statistical discrepancy.

Natural language processing Computer programs developed to analyze and interpret human language.

Natural rate of unemployment Effective unemployment rate, below which pressure emerges in labor markets.

Negative screening An ESG investment style that focuses on the exclusion of certain sectors, companies, or practices in a fund or portfolio on the basis of specific ESG criteria.

Net book value The remaining (undepreciated) balance of an asset's purchase cost. For liabilities, the face value of a bond minus any unamortized discount, or plus any unamortized premium.

Net exports The difference between the value of a country's exports and the value of its imports (i.e., value of exports minus imports).

Net income The difference between revenue and expenses; what remains after subtracting all expenses (including depreciation, interest, and taxes) from revenue.

Net present value The present value of an investment's cash inflows (benefits) minus the present value of its cash outflows (costs).

Net profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also called *profit margin* or *return on sales*.

Net realisable value Estimated selling price in the ordinary course of business less the estimated costs necessary to make the sale.

Net revenue Revenue after adjustments (e.g., for estimated returns or for amounts unlikely to be collected).

Net tax rate The tax rate net of transfer payments.

Neural networks Computer programs based on how our own brains learn and process information.

Neutral rate of interest The rate of interest that neither spurs on nor slows down the underlying economy.

No-load fund A mutual fund in which there is no fee for investing in the fund or for redeeming fund shares, although there is an annual fee based on a percentage of the fund's net asset value.

Node Each value on a binomial tree from which successive moves or outcomes branch.

Nominal data Categorical values that are not amenable to being organized in a logical order. An example of nominal data is the classification of publicly listed stocks into sectors.

Nominal GDP The value of goods and services measured at current prices.

Nominal risk-free interest rate The sum of the real risk-free interest rate and the inflation premium.

Non-accelerating inflation rate of unemployment Effective unemployment rate below which pressure emerges in labor markets.

Non-agency RMBS In the United States, securities issued by private entities that are not guaranteed by a federal agency or a GSE.

Non-Bank lender Unlike typical banks, which make loans and take deposits, these lenders only make loans. Although they lend money to companies, they usually specialize in issuing mortgages to first-time homebuyers and parties looking to refinance an existing mortgage.

Non-cumulative preference shares Preference shares for which dividends that are not paid in the current or subsequent periods are forfeited permanently (instead of being accrued and paid at a later date).

Non-current assets Assets that are expected to benefit the company over an extended period of time (usually more than one year).

Non-current liabilities Obligations that broadly represent a probable sacrifice of economic benefits in periods generally greater than one year in the future.

Non-cyclical A company whose performance is largely independent of the business cycle. Also known as defensive.

Non-deliverable forwards Cash-settled forward contracts, used predominately with respect to foreign exchange forwards. Also called *contracts for differences*.

Non-financial risks Risks that arise from sources other than changes in the external financial markets, such as changes in accounting rules, legal environment, or tax rates.

Non-participating preference shares Preference shares that do not entitle shareholders to share in the profits of the company. Instead, shareholders are only entitled to receive a fixed dividend payment and the par value of the shares in the event of liquidation.

Non-probability sampling A sampling plan dependent on factors other than probability considerations, such as a sampler's judgment or the convenience to access data.

Non-recourse loan A loan in which the lender does not have a shortfall claim against the borrower, so the lender can look only to the property to recover the outstanding mortgage balance.

Non-renewable resources Finite resources that are depleted once they are consumed, such as oil and coal.

Non-sovereign bond A bond issued by a government below the national level, such as a province, region, state, or city.

Nonconventional cash flows In a nonconventional cash flow pattern, the initial outflow is not followed by inflows only, but the cash flows can flip from positive (inflows) to negative (outflows) again or even change signs several times.

Nonparametric test A test that is not concerned with a parameter or that makes minimal assumptions about the population from which a sample comes.

Nonsystematic risk Unique risk that is local or limited to a particular asset or industry that need not affect assets outside of that asset class.

Normal distribution A continuous, symmetric probability distribution that is completely described by its mean and its variance.

Normal goods Goods that are consumed in greater quantities as income increases.

Normal profit The level of accounting profit needed to just cover the implicit opportunity costs ignored in accounting costs.

Notching Ratings adjustment methodology where specific issues from the same borrower may be assigned different credit ratings.

Note rate See *mortgage rate*.

Notice period The length of time (typically 30 to 90 days) in advance that investors may be required to notify a fund of their intent to redeem some or all of their investment.

Notional principal An imputed principal amount.

Null hypothesis The hypothesis that is tested.

Numerical data Values that represent measured or counted quantities as a number. Also called **quantitative data**.

Objective probabilities Probabilities that generally do not vary from person to person; includes a priori and empirical probabilities.

Observation The value of a specific variable collected at a point in time or over a specified period of time.

Odds against *E* The reciprocal of odds for *E*.

Odds for *E* The probability of *E* divided by 1 minus the probability of *E*.

Off-the-run Seasoned government bonds are off-the-run securities; they are not the most recently issued or the most actively traded.

Offer The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size).

Official interest rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks. Also called *official policy rate* or *policy rate*.

Official policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Oligopoly Market structure with a relatively small number of firms supplying the market.

On-the-run The most recently issued and most actively traded sovereign securities.

One-dimensional array The simplest format for representing a collection of data of the same data type.

One-sided hypothesis test A test in which the null hypothesis is rejected only if the evidence indicates that the population parameter is greater than or less than the hypothesized parameter; occurs when the alternative hypothesis is stated either as greater than or less than the hypothesized population parameter.

Open economy An economy that trades with other countries.

Open-end fund A mutual fund that accepts new investment money and issues additional shares at a value equal to the net asset value of the fund at the time of investment.

Open interest The number of outstanding contracts in a clearinghouse at any given time. The open interest figure changes daily as some parties open up new positions, while other parties offset their old positions.

Open market operations The purchase or sale of bonds by the national central bank to implement monetary policy. The bonds traded are usually sovereign bonds issued by the national government.

Operating activities Activities that are part of the day-to-day business functioning of an entity, such as selling inventory and providing services.

Operating breakeven The number of units produced and sold at which the company's operating profit is zero (revenues = operating costs).

Operating cash flow The net amount of cash provided from operating activities.

Operating efficiency ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Operating lease A type of lease which is more akin to the rental of the underlying asset.

Operating leverage The use of fixed costs in operations.

Operating profit A company's profits on its usual business activities before deducting taxes. Also called *operating income*.

Operating profit margin A profitability ratio calculated as operating income (i.e., income before interest and taxes) divided by revenue. Also called *operating margin*.

Operating risk The risk attributed to the operating cost structure, in particular the use of fixed costs in operations; the risk arising from the mix of fixed and variable costs; the risk that a company's operations may be severely affected by environmental, social, and governance risk factors.

Operational independence A bank's ability to execute monetary policy and set interest rates in the way it thought would best meet the inflation target.

Operational risk The risk that arises from inadequate or failed people, systems, and internal policies, procedures, and processes, as well as from external events that are beyond the control of the organization but that affect its operations.

Operationally efficient Said of a market, a financial system, or an economy that has relatively low transaction costs.

Opportunity cost The value that investors forgo by choosing a particular course of action; the value of something in its best alternative use.

Optimal capital structure The capital structure at which the value of the company is maximized.

Option A financial instrument that gives one party the right, but not the obligation, to buy or sell an underlying asset from or to another party at a fixed price over a specific period of time. Also referred to as *contingent claim* or *option contract.*

Option-adjusted price The value of the embedded option plus the flat price of the bond.

Option-adjusted spread OAS = Z-spread – Option value (in basis points per year).

Option-adjusted yield The required market discount rate whereby the price is adjusted for the value of the embedded option.

Option contract See *option*.

Option premium The amount of money a buyer pays and seller receives to engage in an option transaction.

Order A specification of what instrument to trade, how much to trade, and whether to buy or sell.

Order-driven markets A market (generally an auction market) that uses rules to arrange trades based on the orders that traders submit; in their pure form, such markets do not make use of dealers.

Order precedence hierarchy With respect to the execution of orders to trade, a set of rules that determines which orders execute before other orders.

Ordinal data Categorical values that can be logically ordered or ranked.

Ordinary annuity An annuity with a first cash flow that is paid one period from the present.

Ordinary shares Equity shares that are subordinate to all other types of equity (e.g., preferred equity). Also called *common stock* or *common shares.*

Organized exchange A securities marketplace where buyers and seller can meet to arrange their trades.

Other comprehensive income Items of comprehensive income that are not reported on the income statement; comprehensive income minus net income.

Out-of-sample test A test of a strategy or model using a sample outside the period on which the strategy or model was developed.

Out of the money Options that, if exercised, would require the payment of more money than the value received and therefore would not be currently exercised.

Outcome A possible value of a random variable.

Over-the-counter (OTC) market A decentralized market where buy and sell orders initiated from various locations are matched through a communications network.

Overbought A market condition in which market sentiment is thought to be unsustainably bullish.

Overcollateralization Form of internal credit enhancement that refers to the process of posting more collateral than needed to obtain or secure financing.

Overconfidence bias A bias in which people demonstrate unwarranted faith in their own intuitive reasoning, judgments, and/or cognitive abilities.

Overfitting An undesirable result from fitting a model so closely to a dataset that it does not perform well on new data.

Oversold A market condition in which market sentiment is thought to be unsustainably bearish.

Own price The price of a good or service itself (as opposed to the price of something else).

Own-price elasticity of demand The percentage change in quantity demanded for a percentage change in good's own price, holding all other things constant.

Owners' equity The excess of assets over liabilities; the residual interest of shareholders in the assets of an entity after deducting the entity's liabilities. Also called *shareholders' equity* or *shareholders' funds*.

***p*-Value** The smallest level of significance at which the null is rejected.

Paired comparisons test See *test of the mean of the differences.*

Panel data A mix of time-series and cross-sectional data that contains observations through time on characteristics of across multiple observational units.

Par curve A sequence of yields-to-maturity such that each bond is priced at par value. The bonds are assumed to have the same currency, credit risk, liquidity, tax status, and annual yields stated for the same periodicity.

Par value The amount of principal on a bond.

Parallel shift A parallel yield curve shift implies that all rates change by the same amount in the same direction.

Parameter A descriptive measure computed from or used to describe a population of data, conventionally represented by Greek letters.

Parametric test Any test (or procedure) concerned with parameters or whose validity depends on assumptions concerning the population generating the sample.

Pari passu Covenant or contract clause that ensures a debt obligation is treated the same as the borrower's other senior debt instruments and is not subordinated to similar obligations.

Partial duration See *key rate duration.*

Participating preference shares Preference shares that entitle shareholders to receive the standard preferred dividend plus the opportunity to receive an additional dividend if the company's profits exceed a pre-specified level.

Pass-through rate The coupon rate of a mortgage pass-through security.

Passive investment A buy and hold approach in which an investor does not make portfolio changes based on short-term expectations of changing market or security performance.

Payable date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payment date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payments system The system for the transfer of money.

Pearson correlation A parametric measure of the relationship between two variables.

Pecking order theory The theory that managers consider how their actions might be interpreted by outsiders and thus order their preferences for various forms of corporate financing. Forms of financing that are least visible to outsiders (e.g., internally generated funds) are most preferable to managers, and those that are most visible (e.g., equity) are least preferable.

Peer company See *comparable company*.

Peer group A group of companies engaged in similar business activities whose economics and valuation are influenced by closely related factors.

Pennant A technical analysis continuation pattern formed by trendlines that converge to form a triangle, typically over a short period.

Per capita real GDP Real GDP divided by the size of the population, often used as a measure of a country's average standard of living.

Per unit contribution margin The amount that each unit sold contributes to covering fixed costs—that is, the difference between the price per unit and the variable cost per unit.

Percentiles Quantiles that divide a distribution into 100 equal parts that sum to 100.

Perfect capital markets Markets in which, by assumption, there are no taxes, transaction costs, or bankruptcy costs and in which all investors have equal ("symmetric") information.

Perfect competition A market structure in which the individual firm has virtually no impact on market price, because it is assumed to be a very small seller among a very large number of firms selling essentially identical products.

Perfectly elastic When the quantity demanded or supplied of a given good is infinitely sensitive to a change in the value of a specified variable (e.g., price).

Perfectly inelastic When the quantity demanded or supplied of a given good is completely insensitive to a change in the value of a specified variable (e.g., price).

Performance bond See *margin bond.*

Performance evaluation The measurement and assessment of the outcomes of investment management decisions.

Performance fee Fees paid to the general partner from the limited partner(s) based on realized net profits.

Period costs Costs (e.g., executives' salaries) that cannot be directly matched with the timing of revenues and which are thus expensed immediately.

Periodicity The assumed number of periods in the year; typically matches the frequency of coupon payments.

Permanent differences Differences between tax and financial reporting of revenue (expenses) that will not be reversed at some future date. These result in a difference between the company's effective tax rate and statutory tax rate and do not result in a deferred tax item.

Permissioned networks Networks that are fully open only to select participants on a DLT network.

Permissionless networks Networks that are fully open to any user on a DLT network.

Permutation An ordered listing.

Permutation formula The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is ${}_nP_r = \dfrac{n!}{(n-r)!}$.

Perpetual bonds Bonds with no stated maturity date.

Perpetuity A perpetual annuity, or a set of never-ending level sequential cash flows, with the first cash flow occurring one period from now. A bond that does not mature.

Personal income A broad measure of household income that includes all income received by households, whether earned or unearned; measures the ability of consumers to make purchases.

Pet projects Investments in which influential managers want the corporation to invest. Often, unfortunately, pet projects are selected without undergoing normal capital allocation analysis.

Plain vanilla bond Bond that makes periodic, fixed coupon payments during the bond's life and a lump-sum payment of principal at maturity. Also called *conventional bond*.

Platykurtic Describes a distribution that has relatively less weight in the tails than the normal distribution (also called thin-tailed).

Point estimate A single numerical estimate of an unknown quantity, such as a population parameter.

Policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Population All members of a specified group.

Portfolio company In private equity, the company in which the private equity fund is investing.

Portfolio demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments.

Portfolio planning The process of creating a plan for building a portfolio that is expected to satisfy a client's investment objectives.

Position The quantity of an asset that an entity owns or owes.

Positive screening The process of including sectors or companies based on specific ESG criteria, typically ESG performance relative to industry peers

Posterior probability An updated probability that reflects or comes after new information.

Potential GDP The level of real GDP that can be produced at full employment; measures the productive capacity of the economy.

Power of a test The probability of correctly rejecting the null—that is, rejecting the null hypothesis when it is false.

Precautionary money balances Money held to provide a buffer against unforeseen events that might require money.

Preference shares A type of equity interest which ranks above common shares with respect to the payment of dividends and the distribution of the company's net assets upon liquidation. They have characteristics of both debt and equity securities. Also called *preferred stock*.

Preferred stock See *preference shares*.

Premium In the case of bonds, premium refers to the amount by which a bond is priced above its face (par) value. In the case of an option, the amount paid for the option contract.

Prepaid expense A normal operating expense that has been paid in advance of when it is due.

Prepayment option Contractual provision that entitles the borrower to prepay all or part of the outstanding mortgage principal prior to the scheduled due date when the principal must be repaid. Also called *early repayment option*.

Prepayment penalty mortgages Mortgages that stipulate a monetary penalty if a borrower prepays within a certain time period after the mortgage is originated.

Prepayment risk The uncertainty that the timing of the actual cash flows will be different from the scheduled cash flows as set forth in the loan agreement due to the borrowers' ability to alter payments, usually to take advantage of interest rate movements.

Present value (PV) The present discounted value of future cash flows: For assets, the present discounted value of the future net cash inflows that the asset is expected to generate; for liabilities, the present discounted value of the future net cash outflows that are expected to be required to settle the liabilities.

Present value models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security. Also called *discounted cash flow models*.

Pretax margin A profitability ratio calculated as earnings before taxes divided by revenue.

Price elasticity of demand Measures the percentage change in the quantity demanded, given a percentage change in the price of a given product.

Price index Represents the average prices of a basket of goods and services.

Price limits Limits imposed by a futures exchange on the price change that can occur from one day to the next.

Price multiple A ratio that compares the share price with some sort of monetary flow or value to allow evaluation of the relative worth of a company's stock.

Price priority The principle that the highest priced buy orders and the lowest priced sell orders execute first.

Price relative A ratio of an ending price over a beginning price; it is equal to 1 plus the holding period return on the asset.

Price return Measures *only* the price appreciation or percentage change in price of the securities in an index or portfolio.

Price return index An index that reflects *only* the price appreciation or percentage change in price of the constituent securities. Also called *price index*.

Price-setting option The option to adjust prices when demand varies from what is forecast.

Price stability In economics, refers to an inflation rate that is low on average and not subject to wide fluctuation.

Price takers Producers that must accept whatever price the market dictates.

Price to book value A valuation ratio calculated as price per share divided by book value per share.

Price to cash flow A valuation ratio calculated as price per share divided by cash flow per share.

Price to earnings ratio (P/E ratio or P/E) The ratio of share price to earnings per share.

Price to sales A valuation ratio calculated as price per share divided by sales per share.

Price value of a basis point A version of money duration, it is an estimate of the change in the full price of a bond given a 1 basis point change in the yield-to-maturity.

Price weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its price by the sum of all the prices of the constituent securities.

Priced risk Risk for which investors demand compensation for bearing (e.g., equity risk, company-specific factors, macroeconomic factors).

Primary bond market A market in which issuers first sell bonds to investors to raise capital.

Primary capital markets (primary markets) The market where securities are first sold and the issuers receive the proceeds.

Primary dealer Financial institution that is authorized to deal in new issues of sovereign bonds and that serves primarily as a trading counterparty of the office responsible for issuing sovereign bonds.

Primary market The market where securities are first sold and the issuers receive the proceeds.

Prime brokers Brokers that provide services that commonly include custody, administration, lending, short borrowing, and trading.

Principal The amount of funds originally invested in a project or instrument; the face value to be paid at maturity.

Principal–agent relationship A relationship in which a principal hires an agent to perform a particular task or service; also known as an *agency relationship*.

Principal amount Amount that an issuer agrees to repay the debtholders on the maturity date.

Principal business activity The business activity from which a company derives a majority of its revenues and/or earnings.

Principal value Amount that an issuer agrees to repay the debtholders on the maturity date.

Principle of no arbitrage See *arbitrage-free pricing*.

Prior probabilities Probabilities reflecting beliefs prior to the arrival of new information.

Priority of claims Priority of payment, with the most senior or highest ranking debt having the first claim on the cash flows and assets of the issuer.

Private equity fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *venture capital fund*.

Private equity funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *venture capital funds*.

Private equity securities Securities that are not listed on public exchanges and have no active secondary market. They are issued primarily to institutional investors via non-public offerings, such as private placements.

Private investment in public equity (PIPE) An investment in the equity of a publicly traded firm that is made at a discount to the market value of the firm's shares.

Private placement Typically, a non-underwritten, unregistered offering of securities that are sold only to an investor or a small group of investors. It can be accomplished directly between the issuer and the investor(s) or through an investment bank.

Probability A number between 0 and 1 describing the chance that a stated event will occur.

Probability density function A function with non-negative values such that probability can be described by areas under the curve graphing the function.

Probability distribution A distribution that specifies the probabilities of a random variable's possible outcomes.

Probability function A function that specifies the probability that the random variable takes on a specific value.

Probability sampling A sampling plan that allows every member of the population to have an equal chance of being selected.

Probability tree diagram A diagram with branches emanating from nodes representing either mutually exclusive chance events or mutually exclusive decisions.

Production-flexibility options The options to alter production when demand varies from what is forecast.

Production function Provides the quantitative link between the levels of output that the economy can produce and the inputs used in the production process.

Productivity The amount of output produced by workers during a given period—for example, output per hour worked measures the efficiency of labor.

Profession An occupational group that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition.

Profit The return that owners of a company receive for the use of their capital and the assumption of financial risk when making their investments.

Profit and loss (P&L) statement A financial statement that provides information about a company's profitability over a stated period of time. Also called the *income statement*.

Profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses.

Profitability ratios Ratios that measure a company's ability to generate profitable sales from its resources (assets).

Project sequencing To defer the decision to invest in a future project until the outcome of some or all of a current investment is known. Investments are sequenced over time, so that making an investment creates the option to invest in future projects.

Promissory note A written promise to pay a certain amount of money on demand.

Property, plant, and equipment Tangible assets that are expected to be used for more than one period in either the production or supply of goods or services, or for administrative purposes.

Prospectus The document that describes the terms of a new bond issue and helps investors perform their analysis on the issue.

Protective put An option strategy in which a long position in an asset is combined with a long position in a put.

Proxy contest Corporate takeover mechanism in which shareholders are persuaded to vote for a group seeking a controlling position on a company's board of directors.

Proxy voting A process that enables shareholders who are unable to attend a meeting to authorize another individual to vote on their behalf.

Public offer See *public offering*.

Public offering An offering of securities in which any member of the public may buy the securities. Also called *public offer*.

Public–private partnership (PPP) An agreement between the public sector and the private sector to finance, build, and operate public infrastructure, such as hospitals and toll roads.

Pull on liquidity When disbursements are paid too quickly or trade credit availability is limited, requiring companies to expend funds before they receive funds from sales that could cover the liability.

Pure discount bonds See *zero-coupon bond*.

Put An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Put–call–forward parity The relationship among puts, calls, and forward contracts.

Put–call parity An equation expressing the equivalence (parity) of a portfolio of a call and a bond with a portfolio of a put and the underlying, which leads to the relationship between put and call prices.

Put/call ratio A technical analysis indicator that evaluates market sentiment based on the volume of put options traded divided by the volume of call options traded for a particular financial instrument.

Put option An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Putable bonds Bonds that give the bondholder the right to sell the bond back to the issuer at a predetermined price on specified dates.

Qualitative data see **categorical data**.

Quantile A value at or below which a stated fraction of the data lies. Also referred to as a fractile.

Quantitative data see **numerical data**.

Quantitative easing An expansionary monetary policy based on aggressive open market purchase operations.

Quantity equation of exchange An expression that over a given period, the amount of money used to purchase all goods and services in an economy, $M \times V$, is equal to monetary value of this output, $P \times Y$.

Quantity theory of money Asserts that total spending (in money terms) is proportional to the quantity of money.

Quartiles Quantiles that divide a distribution into four equal parts.

Quasi-fixed cost A cost that stays the same over a range of production but can change to another constant level when production moves outside of that range.

Quasi-government bond A bond issued by an entity that is either owned or sponsored by a national government. Also called *agency bond*.

Quintiles Quantiles that divide a distribution into five equal parts.

Quota rents Profits that foreign producers can earn by raising the price of their goods higher than they would without a quota.

Quotas Government policies that restrict the quantity of a good that can be imported into a country, generally for a specified period of time.

Quote-driven market A market in which dealers acting as principals facilitate trading.

Quoted interest rate A quoted interest rate that does not account for compounding within the year. Also called *stated annual interest rate*.

Quoted margin The specified yield spread over the reference rate, used to compensate an investor for the difference in the credit risk of the issuer and that implied by the reference rate.

Random number An observation drawn from a uniform distribution.

Random number generator An algorithm that produces uniformly distributed random numbers between 0 and 1.

Random variable A quantity whose future outcomes are uncertain.

Range The difference between the maximum and minimum values in a dataset.

Raw data Data available in their original form as collected.

Real exchange rate effect The effect through which changing price level impacts real exchange rate which in turn impacts net exports and aggregate demand.

Real GDP The value of goods and services produced, measured at base year prices.

Real income Income adjusted for the effect of inflation on the purchasing power of money. Also known as the *purchasing power of income*. If income remains constant and a good's price falls, real income is said to rise, even though the number of monetary units (e.g., dollars) remains unchanged.

Real interest rate Nominal interest rate minus the expected rate of inflation.

Real risk-free interest rate The single-period interest rate for a completely risk-free security if no inflation were expected.

Realizable (settlement) value With reference to assets, the amount of cash or cash equivalents that could currently be obtained by selling the asset in an orderly disposal; with reference to liabilities, the undiscounted amount of cash or cash equivalents expected to be paid to satisfy the liabilities in the normal course of business.

Rebalancing Adjusting the weights of the constituent securities in an index.

Rebalancing policy The set of rules that guide the process of restoring a portfolio's asset class weights to those specified in the strategic asset allocation.

Recession A period during which real GDP decreases (i.e., negative growth) for at least two successive quarters, or a period of significant decline in total output, income, employment, and sales usually lasting from six months to a year.

Recognition lag The lag in government response to an economic problem resulting from the delay in confirming a change in the state of the economy.

Recourse loan A loan in which the lender has a claim against the borrower for any shortfall between the outstanding mortgage balance and the proceeds received from the sale of the property.

Redemption yield See *yield-to-maturity*.

Redemptions Withdrawals of funds by investors, as allowed by the notice period and other terms in the partnership agreement.

Refinancing rate A type of central bank policy rate.

Registered bonds Bonds for which ownership is recorded by either name or serial number.

Regression analysis A tool for examining whether a variable is useful for explaining another variable.

Regression coefficients The intercept and slope coefficient of a regression.

Regret The feeling that an opportunity has been missed; typically, an expression of *hindsight bias*.

Regret-aversion bias An emotional bias in which people tend to avoid making decisions that will result in action out of fear that the decision will turn out poorly.

Relative dispersion The amount of dispersion relative to a reference value or benchmark.

Relative frequency The absolute frequency of each unique value of the variable divided by the total number of observations of the variable.

Relative price The price of a specific good or service in comparison with those of other goods and services.

Relative strength analysis A comparison of the performance of one asset with the performance of another asset or a benchmark, based on changes in the ratio of the two assets' prices over time.

Relative strength index (RSI) A technical analysis momentum oscillator that compares a security's gains with its losses over a set period.

Renewable resources Resources that can be replenished, such as a forest.

Rent Payment for the use of property.

Reorganization Agreements made by a company in bankruptcy under which a company's capital structure is altered and/or alternative arrangements are made for debt repayment; US Chapter 11 bankruptcy. The company emerges from bankruptcy as a going concern.

Replication The creation of an asset or portfolio from another asset, portfolio, and/or derivative.

Repo A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy back the same security from the purchaser at an agreed-on price and future date. The party who sells the

security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repo margin The difference between the market value of the security used as collateral and the value of the loan. Also called *haircut*.

Repo rate The interest rate on a repurchase agreement.

Representativeness bias A belief perseverance bias in which people tend to classify new information based on past experiences and classifications.

Repurchase agreement A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy the same security back from the purchaser at an agreed-on price and future date. The party who sells the security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repurchase date The date when the party who sold the security at the inception of a repurchase agreement buys back the security from the cash lending counterparty.

Repurchase price The price at which the party who sold the security at the inception of the repurchase agreement buys back the security from the cash lending counterparty.

Required margin The yield spread over or under the reference rate such that an FRN is priced at par value on a rate reset date.

Required rate of return See *market discount rate*.

Required yield See *market discount rate*.

Required yield spread The difference between the yield-to-maturity on a new bond and the benchmark rate; additional compensation required by investors for the difference in risk and tax status of a bond relative to a government bond. Sometimes called the *spread over the benchmark*.

Resampling A statistical method that repeatedly draws samples from the original observed data sample for the statistical inference of population parameters.

Reserve accounts Form of internal credit enhancement that relies on creating accounts and depositing in these accounts cash that can be used to absorb losses. Also called *reserve funds*.

Reserve funds See *reserve accounts*.

Reserve requirement The requirement for banks to hold reserves in proportion to the size of deposits.

Residual The difference between an observation and its predicted value, where the predicted value is based on the estimated linear relation between the dependent and independent variables using sample data.

Resistance In technical analysis, a price range in which selling activity is sufficient to stop the rise in the price of a security.

Responsible investing An overall (umbrella) term for several investment strategies—ESG investing, SRI investing, thematic investing, and impact investing—that incorporate ESG analysis in their investment processes.

Restricted payments A bond covenant meant to protect creditors by limiting how much cash can be paid out to shareholders over time.

Retracement In technical analysis, a reversal in the movement of a security's price such that it is counter to the prevailing longer-term price trend.

Return-generating model A model that can provide an estimate of the expected return of a security given certain parameters and estimates of the values of the independent variables in the model.

Return on assets (ROA) A profitability ratio calculated as net income divided by average total assets; indicates a company's net profit generated per dollar invested in total assets.

Return on equity (ROE) A profitability ratio calculated as net income divided by average shareholders' equity.

Return on invested capital A measure of the profitability of a company relative to the amount of capital invested by the equity- and debtholders.

Return on sales An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also referred to as *net profit margin*.

Return on total capital A profitability ratio calculated as EBIT divided by the sum of short- and long-term debt and equity.

Revaluation model Under IFRS, the process of valuing long-lived assets at fair value, rather than at cost less accumulated depreciation. Any resulting profit or loss is either reported on the income statement and/or through equity under revaluation surplus.

Revenue The amount charged for the delivery of goods or services in the ordinary activities of a business over a stated period; the inflows of economic resources to a company over a stated period.

Reversal pattern A type of pattern used in technical analysis to predict the end of a trend and a change in the direction of a security's price.

Reverse repo A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse repurchase agreement A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse stock split A reduction in the number of shares outstanding with a corresponding increase in share price, but no change to the company's underlying fundamentals.

Revolving credit agreements (also known as revolvers) The most reliable form of short-term bank borrowing facilities; they are in effect for multiple years (e.g., three to five years) and can have optional medium-term loan features.

Rho The sensitivity of the option price to the risk-free rate.

Ricardian equivalence An economic theory that implies that it makes no difference whether a government finances a deficit by increasing taxes or issuing debt.

Risk Exposure to uncertainty. The chance of a loss or adverse outcome as a result of an action, inaction, or external event.

Risk averse The assumption that an investor will choose the least risky alternative.

Risk aversion The degree of an investor's inability and unwillingness to take risk.

Risk budgeting The establishment of objectives for individuals, groups, or divisions of an organization that takes into account the allocation of an acceptable level of risk.

Risk exposure The state of being exposed or vulnerable to a risk. The extent to which an organization is sensitive to underlying risks.

Risk governance The top-down process and guidance that directs risk management activities to align with and support the overall enterprise.

Risk management The process of identifying the level of risk an organization wants, measuring the level of risk the organization currently has, taking actions that bring the actual

level of risk to the desired level of risk, and monitoring the new actual level of risk so that it continues to be aligned with the desired level of risk.

Risk management framework The infrastructure, process, and analytics needed to support effective risk management in an organization.

Risk-neutral pricing Sometimes said of derivatives pricing, uses the fact that arbitrage opportunities guarantee that a risk-free portfolio consisting of the underlying and the derivative must earn the risk-free rate.

Risk-neutral probabilities Weights that are used to compute a binomial option price. They are the probabilities that would apply if a risk-neutral investor valued an option.

Risk premium An extra return expected by investors for bearing some specified risk.

Risk shifting Actions to change the distribution of risk outcomes.

Risk tolerance The amount of risk an investor is willing and able to bear to achieve an investment goal.

Risk transfer Actions to pass on a risk to another party, often, but not always, in the form of an insurance policy.

Robo-adviser A machine-based analytical tool or service that provides technology-driven investment solutions through online platforms.

Rule of 72 The principle that the approximate number of years necessary for an investment to double is 72 divided by the stated interest rate.

Running yield See *current yield*.

Safety-first rules Rules for portfolio selection that focus on the risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Sales Generally, a synonym for revenue; "sales" is generally understood to refer to the sale of goods, whereas "revenue" is understood to include the sale of goods or services.

Sales risk Uncertainty with respect to the quantity of goods and services that a company is able to sell and the price it is able to achieve; the risk related to the uncertainty of revenues.

Sample A subset of a population.

Sample correlation coefficient A standardized measure of how two variables in a sample move together. It is the ratio of the sample covariance to the product of the two variables' standard deviations.

Sample covariance A measure of how two variables in a sample move together.

Sample excess kurtosis A sample measure of the degree of a distribution's kurtosis in excess of the normal distribution's kurtosis.

Sample mean The sum of the sample observations divided by the sample size.

Sample selection bias Bias introduced by systematically excluding some members of the population according to a particular attribute—for example, the bias introduced when data availability leads to certain observations being excluded from the analysis.

Sample-size neglect A type of representativeness bias in which financial market participants incorrectly assume that small sample sizes are representative of populations (or "real" data).

Sample skewness A sample measure of the degree of asymmetry of a distribution.

Sample standard deviation The positive square root of the sample variance.

Sample statistic A quantity computed from or used to describe a sample.

Sample variance The sum of squared deviations around the mean divided by the degrees of freedom.

Sampling The process of obtaining a sample.

Sampling distribution The distribution of all distinct possible values that a statistic can assume when computed from samples of the same size randomly drawn from the same population.

Sampling error The difference between the observed value of a statistic and the estimate resulting from using subsets of the population.

Sampling plan The set of rules used to select a sample.

Say on pay A process whereby shareholders may vote on executive remuneration (compensation) matters.

Scatter plot A chart in which two variables are plotted along the axis and points on the chart represent pairs of the two variables. In regression, the dependent variable is plotted on the vertical axis and the independent variable is plotted along the horizontal axis. Also known as a scattergram and a *scatter diagram.*

Scatter plot matrix A tool for organizing scatter plots between pairs of variables, making it easy to inspect all pairwise relationships in one combined visual.

Scenario analysis Analysis that shows the changes in key financial quantities that result from given (economic) events, such as the loss of customers, the loss of a supply source, or a catastrophic event; a risk management technique involving examination of the performance of a portfolio under specified situations. Closely related to stress testing.

Screening The application of a set of criteria to reduce a set of potential investments to a smaller set having certain desired characteristics.

Seasoned offering An offering in which an issuer sells additional units of a previously issued security.

Second-degree price discrimination When the monopolist charges different per-unit prices using the quantity purchased as an indicator of how highly the customer values the product.

Second lien A secured interest in the pledged assets that ranks below first lien debt in both collateral protection and priority of payment.

Secondary bond markets Markets in which existing bonds are traded among investors.

Secondary market The market where securities are traded among investors.

Secondary precedence rules Rules that determine how to rank orders placed at the same time.

Sector A group of related industries (GICS definition).

Sector indexes Indexes that represent and track different economic sectors—such as consumer goods, energy, finance, health care, and technology—on either a national, regional, or global basis.

Secured ("asset-based") loan A loan that is backed by specific, secured company assets.

Secured bonds Bonds secured by assets or financial guarantees pledged to ensure debt repayment in case of default.

Secured debt Debt in which the debtholder has a direct claim—a pledge from the issuer—on certain assets and their associated cash flows.

Securitization A process that involves moving assets into a special legal entity, which then uses the assets as guarantees to secure a bond issue.

Securitized assets Assets that are typically used to create asset-backed bonds; for example, when a bank securitizes a pool of loans, the loans are said to be securitized.

Security characteristic line A plot of the excess return of a security on the excess return of the market.

Security market index A portfolio of securities representing a given security market, market segment, or asset class.

Security market line (SML) The graph of the capital asset pricing model.

Security selection The process of selecting individual securities; typically, security selection has the objective of generating superior risk-adjusted returns relative to a portfolio's benchmark.

Self-attribution bias A bias in which people take too much credit for successes (*self-enhancing*) and assign responsibility to others for failures (*self-protecting*).

Self-control bias A bias in which people fail to act in pursuit of their long-term, overarching goals because of a lack of self-discipline.

Self-investment limits With respect to investment limitations applying to pension plans, restrictions on the percentage of assets that can be invested in securities issued by the pension plan sponsor.

Sell-side firm A broker/dealer that sells securities and provides independent investment research and recommendations to their clients (i.e., buy-side firms).

Semi-strong-form efficient market A market in which security prices reflect all publicly known and available information.

Semiannual bond basis yield An annual rate having a periodicity of two; also known as a *semiannual bond equivalent yield*.

Semiannual bond equivalent yield See *semiannual bond basis yield*.

Seniority ranking Priority of payment of various debt obligations.

Sensitivity analysis Analysis that shows the range of possible outcomes as specific assumptions are changed.

Separately managed account (SMA) An investment portfolio managed exclusively for the benefit of an individual or institution.

Serial maturity structure Structure for a bond issue in which the maturity dates are spread out during the bond's life; a stated number of bonds mature and are paid off each year before final maturity.

Settlement The process that occurs after a trade is completed, the securities are passed to the buyer, and payment is received by the seller.

Settlement date Date when the buyer makes cash payment and the seller delivers the security.

Settlement price The official price, designated by the clearinghouse, from which daily gains and losses will be determined and marked to market.

Share repurchase A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Shareholder activism Strategies used by shareholders to attempt to compel a company to act in a desired manner.

Shareholder engagement The process whereby companies engage with their shareholders.

Shareholders' equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Sharpe ratio The average return in excess of the risk-free rate divided by the standard deviation of return; a measure of the average excess return earned per unit of standard deviation of return.

Shelf registration A type of public offering that allows the issuer to file a single, all-encompassing offering circular that covers a series of bond issues.

Short The seller of an asset or derivative contract. Also refers to the position of being short an asset or derivative contract.

Short position A position in an asset or contract in which one has sold an asset one does not own, or in which a right under a contract can be exercised against oneself.

Short-run average total cost The curve describing average total cost when some costs are considered fixed.

Short selling A transaction in which borrowed securities are sold with the intention to repurchase them at a lower price at a later date and return them to the lender.

Shortfall risk The risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Shutdown point The point at which average revenue is equal to the firm's average variable cost.

Side letters Side agreements created between the GP and specific LPs. These agreements exist *outside* the LPA. These agreements provide additional terms and conditions related to the investment agreement.

Simple interest The interest earned each period on the original investment; interest calculated on the principal only.

Simple linear regression (SLR) A regression that summarizes the relation between the dependent variable and a single independent variable.

Simple random sample A subset of a larger population created in such a way that each element of the population has an equal probability of being selected to the subset.

Simple random sampling The procedure of drawing a sample to satisfy the definition of a simple random sample.

Simple yield The sum of the coupon payments plus the straight-line amortized share of the gain or loss, divided by the flat price.

Simulation Computer-generated sensitivity or scenario analysis that is based on probability models for the factors that drive outcomes.

Simulation trial A complete pass through the steps of a simulation.

Single-step format With respect to the format of the income statement, a format that does not subtotal for gross profit (revenue minus cost of goods sold).

Sinking fund arrangement Provision that reduces the credit risk of a bond issue by requiring the issuer to retire a portion of the bond's principal outstanding each year.

Situational influences External factors, such as environmental or cultural elements, that shape our behavior.

Skewed Not symmetrical.

Skewness A quantitative measure of skew (lack of symmetry); a synonym of skew. It is computed as the average cubed deviation from the mean standardized by dividing by the standard deviation cubed.

Slope coefficient The coefficient of an independent variable that represents the average change in the dependent variable for a one-unit change in the independent variable.

Small country A country that is a price taker in the world market for a product and cannot influence the world market price.

Smart beta Involves the use of simple, transparent, rules-based strategies as a basis for investment decisions.

Smart contract A computer program that is designed to self-execute on the basis of pre-specified terms and conditions agreed to by parties to a contract.

Socially responsible investing An investment approach that excludes investments in companies or industries that deviate from an organization's beliefs and sometimes includes investments with favorable environmental or social profiles.

Soft-bullet covered bonds Covered bonds for which bond default and payment acceleration of bond cash flows may be delayed upon sponsor default until a new final maturity date is reached.

Solvency With respect to financial statement analysis, the ability of a company to fulfill its long-term obligations.

Solvency ratios Ratios that measure a company's ability to meet its long-term obligations.

Solvency risk The risk that an organization does not survive or succeed because it runs out of cash, even though it might otherwise be solvent.

Sovereign bond A bond issued by a national government. Also called "Sovereign."

Spearman rank correlation coefficient A measure of correlation applied to ranked data.

Special dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Special purpose entity A non-operating entity created to carry out a specified purpose, such as leasing assets or securitizing receivables; can be a corporation, partnership, trust, or limited liability partnership formed to facilitate a specific type of business activity. Also called *special purpose vehicle*, *special purpose company*, or *variable interest entity*.

Special purpose vehicle See *special purpose entity*.

Specific identification method An inventory accounting method that identifies which specific inventory items were sold and which remained in inventory to be carried over to later periods.

Speculative demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments. Also called *portfolio demand for money*.

Speculative money balances Monies held in anticipation that other assets will decline in value.

Split coupon bond See *deferred coupon bond*.

Sponsored A type of depository receipt in which the foreign company whose shares are held by the depository has a direct involvement in the issuance of the receipts.

Spot curve A sequence of yields-to-maturity on zero-coupon bonds. Sometimes called *zero* or *strip curve* (because coupon payments are "stripped" off the bonds).

Spot markets Markets in which assets are traded for immediate delivery.

Spot prices The price of an asset for immediately delivery.

Spot rates A sequence of market discount rates that correspond to the cash flow dates; yields-to-maturity on zero-coupon bonds maturing at the date of each cash flow.

Spread In general, the difference in yield between different fixed-income securities. Often used to refer to the difference between the yield-to-maturity and the benchmark.

Spread over the benchmark See *required yield spread*.

Spread risk Bond price risk arising from changes in the yield spread on credit-risky bonds; reflects changes in the market's assessment and/or pricing of credit migration (or downgrade) risk and market liquidity risk.

Spurious correlation Refers to: 1) correlation between two variables that reflects chance relationships in a particular dataset; 2) correlation induced by a calculation that mixes each of two variables with a third variable; and 3) correlation between two variables arising not from a direct relation between them but from their relation to a third variable.

Stacked bar chart An alternative form for presenting the frequency distribution of two categorical variables, where bars representing the sub-groups are placed on top of each other to form a single bar. Each sub-section is shown in a different color to represent the contribution of each sub-group, and the overall height of the stacked bar represents the marginal frequency for the category.

Stackelberg model A prominent model of strategic decision making in which firms are assumed to make their decisions sequentially.

Stagflation The combination of a high inflation rate with a high level of unemployment and a slowdown of the economy.

Staggered boards Election process whereby directors are typically divided into multiple classes that are elected separately in consecutive years—that is, one class every year.

Stakeholder management The identification, prioritization, and understanding of the interests of stakeholder groups and managing the company's relationships with these groups.

Stakeholders Individuals or groups of individuals who may be affected either directly or indirectly by a decision and thus have an interest, or stake, in the decision.

Standard deviation The positive square root of the variance; a measure of dispersion in the same units as the original data.

Standard error of the estimate A measure of the fit of a regression line, calculated as the square root of the mean square error. Also known as the *standard error of the regression* and the *root mean square error*.

Standard error of the forecast A measure of the uncertainty associated with a forecasted value of the dependent variable that depends on the standard error of the estimate, the variability of the independent variable, the deviation of the forecasted independent variable from the mean in the regression, and the number of observations.

Standard error of the slope coefficient The standard error of the slope, which in a simple linear regression is the ratio of the model's standard error of the estimate (s_e) to the square root of the variation of the independent variable.

Standard normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Standardizing A transformation that involves subtracting the mean and dividing the result by the standard deviation.

Standards of conduct Behaviors required by a group; established benchmarks that clarify or enhance a group's code of ethics.

Standing limit orders A limit order at a price below market and which therefore is waiting to trade.

Stated annual interest rate A quoted interest rate that does not account for compounding within the year. Also called *quoted interest rate*.

Statement of changes in equity (statement of owners' equity) A financial statement that reconciles the beginning-of-period and end-of-period balance sheet values of

shareholders' equity; provides information about all factors affecting shareholders' equity. Also called *statement of owners' equity*.

Statement of financial condition The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of financial position The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of operations A financial statement that provides information about a company's profitability over a stated period of time.

Static trade-off theory of capital structure A theory pertaining to a company's optimal capital structure. The optimal level of debt is found at the point where additional debt would cause the costs of financial distress to increase by a greater amount than the benefit of the additional tax shield.

Statistic A summary measure of a sample of observations.

Statistically significant A result indicating that the null hypothesis can be rejected; with reference to an estimated regression coefficient, frequently understood to mean a result indicating that the corresponding population regression coefficient is different from zero.

Status quo bias An emotional bias in which people do nothing (i.e., maintain the status quo) instead of making a change.

Statutory voting A common method of voting where each share represents one vote.

Step-up coupon bond Bond for which the coupon, fixed or floating, increases by specified margins at specified dates.

Stochastic oscillator A momentum indicator that compares a particular closing price of a security to a range of the security's prices over a certain period of time.

Stock dividend A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Stock split An increase in the number of shares outstanding with a consequent decrease in share price, but no change to the company's underlying fundamentals.

Stop-loss order See *stop order*.

Stop order An order in which a trader has specified a stop price condition. Also called *stop-loss order*.

Store of value The quality of tending to preserve value.

Store of wealth Goods that depend on the fact that they do not perish physically over time, and on the belief that others would always value the good.

Straight-line method A depreciation method that allocates evenly the cost of a long-lived asset less its estimated residual value over the estimated useful life of the asset.

Straight voting A shareholder voting process in which shareholders receive one vote for each share owned.

Strategic analysis Analysis of the competitive environment with an emphasis on the implications of the environment for corporate strategy.

Strategic asset allocation The set of exposures to IPS-permissible asset classes that is expected to achieve the client's long-term objectives given the client's investment constraints.

Stratified random sampling A procedure that first divides a population into subpopulations (strata) based on classification criteria and then randomly draws samples from each stratum in sizes proportional to that of each stratum in the population.

Street convention A yield measure that neglects weekends and holidays; the internal rate of return on cash flows assuming payments are made on the scheduled dates, even when the scheduled date falls on a weekend or holiday.

Stress testing A specific type of scenario analysis that estimates losses in rare and extremely unfavorable combinations of events or scenarios.

Strong-form efficient market A market in which security prices reflect all public and private information.

Structural (or cyclically adjusted) budget deficit The deficit that would exist if the economy was at full employment (or full potential output).

Structural subordination Arises in a holding company structure when the debt of operating subsidiaries is serviced by the cash flow and assets of the subsidiaries before funds can be passed to the holding company to service debt at the parent level.

Structured data Data that are highly organized in a predefined manner, usually with repeating patterns.

Structured financial instrument A financial instrument that shares the common attribute of repackaging risks. Structured financial instruments include asset-backed securities, collateralized debt obligations, and other structured financial instruments such as capital protected, yield enhancement, participation, and leveraged instruments.

Subjective probability A probability drawing on personal or subjective judgment.

Subordinated debt A class of unsecured debt that ranks below a firm's senior unsecured obligations.

Subordination Form of internal credit enhancement that relies on creating more than one bond tranche and ordering the claim priorities for ownership or interest in an asset between the tranches. The ordering of the claim priorities is called a senior/subordinated structure, where the tranches of highest seniority are called senior followed by subordinated or junior tranches. Also called *credit tranching*.

Substitutes Said of two goods or services such that if the price of one increases the demand for the other tends to increase, holding all other things equal (e.g., butter and margarine).

Sum of squares error (SSE) The sum of the squared deviations of (1) the value of the dependent variable and (2) the value of the dependent variable based on the estimated regression line. Also referred to as the *residual sum of squares*.

Sum of squares regression (SSR) The sum of the squared deviations of (1) the value of the dependent variable based on the estimated regression line and (2) the mean of the dependent variable.

Sum of squares total (SST) The sum of the squared deviations of the dependent variable from its mean; the variation of the dependent variable. Also referred to as the *total sum of squares*.

Sunk cost A cost that has already been incurred.

Supervised learning A machine learning approach that makes use of labeled training data.

Supply shock A typically unexpected disturbance to supply.

Support In technical analysis, a price range in which buying activity is sufficient to stop the decline in the price of a security.

Support tranches Classes or tranches in CMOs that protect PAC tranches from prepayment risk.

Supranational bond A bond issued by a supranational agency such as the World Bank.

Surety bond Form of external credit enhancement whereby a rated and regulated insurance company guarantees to reimburse bondholders for any losses incurred up to a maximum amount if the issuer defaults.

Survey approach An estimate of the equity risk premium that is based on estimates provided by a panel of finance experts.

Survivorship bias Relates to the inclusion of only current investment funds in a database. As such, the returns of funds that are no longer available in the marketplace (have been liquidated) are excluded from the database. In addition, *backfill bias* is another problem, whereby certain surviving hedge funds may be added to databases and various hedge fund indexes only after they are initially successful and start to report their returns.

Sustainability linked loans These are any types of loan instruments and/or contingent facilities (such as bonding lines, guarantee lines, or letters of credit) that incentivize the borrower's achievement of ambitious, pre-determined sustainability performance objectives.

Sustainable growth rate The rate of dividend (and earnings) growth that can be sustained over time for a given level of return on equity, keeping the capital structure constant and without issuing additional common stock.

Sustainable investing A term used in a similar context to responsible investing, but its key focus is on factoring in sustainability issues while investing.

Sustainable rate of economic growth The rate of increase in the economy's productive capacity or potential GDP.

Swap contract An agreement between two parties to exchange a series of future cash flows.

Syndicated loan A loan from a group of lenders to a single borrower.

Syndicated offering A bond issue underwritten by a group of investment banks.

Systematic risk Risk that affects the entire market or economy; it cannot be avoided and is inherent in the overall market. Systematic risk is also known as non-diversifiable or market risk.

Systematic sampling A procedure of selecting every *k*th member until reaching a sample of the desired size. The sample that results from this procedure should be approximately random.

Tactical asset allocation The decision to deliberately deviate from the strategic asset allocation in an attempt to add value based on forecasts of the near-term relative performance of asset classes.

Tactical asset allocation (TAA) A portfolio strategy that shifts the percentages of assets held in various asset classes (or categories) to take advantage of market opportunities. Allocation shifts can occur within an asset class or across asset classes.

Tag cloud see **word cloud**.

Target capital structure A company's chosen proportions of debt and equity.

Target independent A bank's ability to determine the definition of inflation that they target, the rate of inflation that they target, and the horizon over which the target is to be achieved.

Target semideviation A measure of downside risk, calculated as the square root of the average of the squared deviations of observations below the target (also called target downside deviation).

Tariffs Taxes that a government levies on imported goods.

Tax base The amount at which an asset or liability is valued for tax purposes.

Tax expense An aggregate of an entity's income tax payable (or recoverable in the case of a tax benefit) and any changes in deferred tax assets and liabilities. It is essentially the income tax payable or recoverable if these had been determined based on accounting profit rather than taxable income.

Tax loss carry forward A taxable loss in the current period that may be used to reduce future taxable income.

Taxable income The portion of an entity's income that is subject to income taxes under the tax laws of its jurisdiction.

Taxable temporary differences Temporary differences that result in a taxable amount in a future period when determining the taxable profit as the balance sheet item is recovered or settled.

Technical analysis A form of security analysis that uses price and volume data, often displayed graphically, in decision making.

Technology The process a company uses to transform inputs into outputs.

Tender offer Corporate takeover mechanism that involves shareholders selling their interests directly to the group seeking to gain control.

Tenor The time-to-maturity for a bond or derivative contract. Also called *term to maturity*.

Term maturity structure Structure for a bond issue in which the bond's notional principal is paid off in a lump sum at maturity.

Term structure See *maturity structure*.

Term structure of credit spreads The relationship between the spreads over the "risk-free" (or benchmark) rates and times-to-maturity.

Term structure of yield volatility The relationship between the volatility of bond yields-to-maturity and times-to-maturity.

Terminal stock value The expected value of a share at the end of the investment horizon—in effect, the expected selling price. Also called *terminal value*.

Terminal value The expected value of a share at the end of the investment horizon—in effect, the expected selling price.

Terms of trade The ratio of the price of exports to the price of imports, representing those prices by export and import price indexes, respectively.

Test of the mean of the differences A statistical test for differences based on paired observations drawn from samples that are dependent on each other.

Text analytics The use of computer programs to analyze and derive meaning from typically large, unstructured text- or voice-based datasets.

Thematic investing An ESG investment approach that focuses on investing in themes or assets specifically relating to ESG factors, such as clean energy, green technology, or sustainable agriculture.

Thin-Tailed Describes a distribution that has relatively less weight in the tails than the normal distribution (also called platykurtic).

Third-degree price discrimination When the monopolist segregates customers into groups based on demographic or other characteristics and offers different pricing to each group.

Time-period bias The possibility that when we use a time-series sample, our statistical conclusion may be sensitive to the starting and ending dates of the sample.

Time-series data A sequence of observations for a single observational unit of a specific variable collected over time and at discrete and typically equally spaced intervals of time (such as daily, weekly, monthly, annually, or quarterly).

Time tranching The creation of classes or tranches in an ABS/MBS that possess different (expected) maturities.

Time value The difference between the market price of the option and its intrinsic value.

Time value decay Said of an option when, at expiration, no time value remains and the option is worth only its exercise value.

Time value of money The principles governing equivalence relationships between cash flows with different dates.

Time-weighted rate of return The compound rate of growth of one unit of currency invested in a portfolio during a stated measurement period; a measure of investment performance that is not sensitive to the timing and amount of withdrawals or additions to the portfolio.

Tokenization The process of representing ownership rights to physical assets on a blockchain or distributed ledger.

Top-down analysis An investment selection approach that begins with consideration of macroeconomic conditions and then evaluates markets and industries based upon such conditions.

Total comprehensive income The change in equity during a period resulting from transaction and other events, other than those changes resulting from transactions with owners in their capacity as owners.

Total cost The summation of all costs, for which costs are classified as fixed or variable.

Total factor productivity A scale factor that reflects the portion of growth unaccounted for by explicit factor inputs (e.g., capital and labor).

Total fixed cost The summation of all expenses that do not change as the level of production varies.

Total invested capital The sum of market value of common equity, book value of preferred equity, and face value of debt.

Total probability rule A rule explaining the unconditional probability of an event in terms of probabilities of the event conditional on mutually exclusive and exhaustive scenarios.

Total probability rule for expected value A rule explaining the expected value of a random variable in terms of expected values of the random variable conditional on mutually exclusive and exhaustive scenarios.

Total return Measures the price appreciation, or percentage change in price of the securities in an index or portfolio, plus any income received over the period.

Total return index An index that reflects the price appreciation or percentage change in price of the constituent securities plus any income received since inception.

Total return swap A swap in which one party agrees to pay the total return on a security. Often used as a credit derivative, in which the underlying is a bond.

Total variable cost The summation of all variable expenses.

Tracking error The standard deviation of the differences between a portfolio's returns and its benchmark's returns; a synonym of active risk.

Tracking risk The standard deviation of the differences between a portfolio's returns and its benchmarks returns. Also called *tracking error*.

Trade creation When regional integration results in the replacement of higher cost domestic production by lower cost imports from other members.

Trade credit A spontaneous form of credit in which a purchaser of the goods or service is financing its purchase by delaying the date on which payment is made.

Trade diversion When regional integration results in lower-cost imports from non-member countries being replaced with higher-cost imports from members.

Trade payables Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Trade protection Government policies that impose restrictions on trade, such as tariffs and quotas.

Trade surplus (deficit) When the value of exports is greater (less) than the value of imports.

Trading securities Under US GAAP, a category of debt securities held by a company with the intent to trade them. Also called *held-for-trading securities*.

Traditional investment markets Markets for traditional investments, which include all publicly traded debts and equities and shares in pooled investment vehicles that hold publicly traded debts and/or equities.

Transactions money balances Money balances that are held to finance transactions.

Transfer payments Welfare payments made through the social security system that exist to provide a basic minimum level of income for low-income households.

Transparency Said of something (e.g., a market) in which information is fully disclosed to the public and/or regulators.

Treasury stock method A method for accounting for the effect of options (and warrants) on earnings per share (EPS) that specifies what EPS would have been if the options and warrants had been exercised and the company had used the proceeds to repurchase common stock.

Tree-Map Another graphical tool for displaying categorical data. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group.

Trend A long-term pattern of movement in a particular direction.

Treynor ratio A measure of risk-adjusted performance that relates a portfolio's excess returns to the portfolio's beta.

Triangle pattern In technical analysis, a continuation chart pattern that forms as the range between high and low prices narrows, visually forming a triangle.

Trimmed mean A mean computed after excluding a stated small percentage of the lowest and highest observations.

Trimodal A distribution that has the three most frequently occurring values.

Triple bottom In technical analysis, a reversal pattern that results when the price forms three troughs at roughly the same price level. A triple bottom is used to predict a change from a downtrend to an uptrend.

Triple top In technical analysis, a reversal pattern that results when the price forms three peaks at roughly the same price level. A triple top is used to predict a change from an uptrend to a downtrend.

True yield The internal rate of return on cash flows using the actual calendar, including weekends and bank holidays.

Trust deed The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *bond indenture*.

Turn-of-the-year effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January.

Two-dimensional rectangular array A popular form for organizing data for processing by computers or for presenting data visually. It is comprised of columns and rows to hold multiple variables and multiple observations, respectively (also called a **data table**).

Two-fund separation theorem The theory that all investors regardless of taste, risk preferences, and initial wealth will hold a combination of two portfolios or funds: a risk-free asset and an optimal portfolio of risky assets.

Two-sided hypothesis test A test in which the null hypothesis is rejected in favor of the alternative hypothesis if the evidence indicates that the population parameter is either smaller or larger than a hypothesized value; occurs when the alternative hypothesis is stated as not equal to the hypothesized population parameters.

Two-way table See *contingency table*.

Two-week repo rate The interest rate on a two-week repurchase agreement; may be used as a policy rate by a central bank.

Type I error The error of rejecting a true null hypothesis; a false positive.

Type II error The error of not rejecting a false null hypothesis; false negative.

Unanticipated (unexpected) inflation The component of inflation that is a surprise.

Uncommitted lines of credit The least reliable form of bank borrowing in which a bank offers, without formal commitment, a line of credit for an extended period of time but reserves the right to refuse any request for its use.

Unconditional probability The probability of an event *not* conditioned on another event.

Underemployed A person who has a job but has the qualifications to work a significantly higher-paying job.

Underlying An asset that trades in a market in which buyers and sellers meet, decide on a price, and the seller then delivers the asset to the buyer and receives payment. The underlying is the asset or other derivative on which a particular derivative is based. The market for the underlying is also referred to as the *spot market*.

Underwriter A firm, usually an investment bank, that takes the risk of buying the newly issued securities from the issuer and then reselling them to investors or to dealers, thus guaranteeing the sale of the securities at the offering price negotiated with the issuer.

Underwritten offering A type of securities issue mechanism in which the investment bank guarantees the sale of the securities at an offering price that is negotiated with the issuer. Also known as *firm commitment offering*.

Unearned revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service. Also called *deferred revenue* or *deferred income*.

Unemployed People who are actively seeking employment but are currently without a job.

Unemployment rate The ratio of unemployed to the labor force.

Unexpected inflation The component of inflation that is a surprise.

Unimodal A distribution with a single value that is most frequently occurring.

Unit elastic An elasticity with a magnitude of negative one. Also called *unitary elastic*.

Unit labor cost The average labor cost to produce one unit of output.

Unit normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Unitranche debt Consists of a hybrid or blended loan structure that combines different tranches of secured and unsecured debt into a single loan with a single, blended interest rate.

Units-of-production method A depreciation method that allocates the cost of a long-lived asset based on actual usage during the period.

Univariate distribution A distribution that specifies the probabilities for a single random variable.

Universal owners Long-term investors, such as pension funds, that have significant assets invested in globally diversified portfolios.

Unlimited funds An unlimited funds environment assumes that the company can raise the funds it wants for all profitable investments simply by paying the required rate of return.

Unsecured debt Debt that gives the debtholder only a general claim on an issuer's assets and cash flow.

Unsponsored A type of depository receipt in which the foreign company whose shares are held by the depository has no involvement in the issuance of the receipts.

Unstructured data Data that do not follow any conventionally organized forms.

Unsupervised learning A machine learning approach that does not make use of labeled training data.

Up transition probability The probability that an asset's value moves up.

Uptrend A pattern that occurs when the price of an asset moves higher over a period of time.

Validity instructions Instructions which indicate when the order may be filled.

Valuation allowance A reserve created against deferred tax assets, based on the likelihood of realizing the deferred tax assets in future accounting periods.

Valuation ratios Ratios that measure the quantity of an asset or flow (e.g., earnings) in relation to the price associated with a specified claim (e.g., a share or ownership of the enterprise).

Value at risk A money measure of the minimum value of losses expected during a specified time period at a given level of probability.

Value investors With reference to equity investors, investors who are focused on paying a relatively low share price in relation to earnings or assets per share.

VaR See *value at risk*.

Variable A characteristic or quantity that can be measured, counted, or categorized and that is subject to change (also called a field, an attribute, or a feature).

Variable costs Costs that fluctuate with the level of production and sales.

Variance The expected value (the probability-weighted average) of squared deviations from a random variable's expected value.

Variation margin Additional margin that must be deposited in an amount sufficient to bring the balance up to the initial margin requirement.

Veblen goods Goods that increase in desirability with increasing price.

Vega A measure of the sensitivity of an option's price to changes in the underlying's volatility.

Venture capital Investments that provide "seed" or startup capital, early-stage financing, or later-stage financing (including mezzanine-stage financing) to companies that are in early development stages and require additional capital for expansion or preparation for an initial public offering.

Venture capital fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *private equity fund*.

Venture capital funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *private equity funds.*

Vertical analysis Common-size analysis using only one reporting period or one base financial statement; for example, an income statement in which all items are stated as percentages of sales.

Vertical demand schedule Implies that some fixed quantity is demanded, regardless of price.

Visual technique The most common and readily available method of initial data assessment. Experts in pattern recognition maintain that the visual (or "eyeball") technique is still the most effective way of searching for recognizable patterns.

Visualization The presentation of data in a pictorial or graphical format for the purpose of increasing understanding and for gaining insights into the data.

Volatility The standard deviation of the continuously compounded returns on the underlying asset.

Voluntarily unemployed A person voluntarily outside the labor force, such as a jobless worker refusing an available vacancy.

Voluntary export restraint A trade barrier under which the exporting country agrees to limit its exports of the good to its trading partners to a specific number of units.

Vote by proxy A mechanism that allows a designated party—such as another shareholder, a shareholder representative, or management—to vote on the shareholder's behalf.

Warrant Attached option that gives its holder the right to buy the underlying stock of the issuing company at a fixed exercise price until the expiration date.

Waterfall Represents the distribution method that defines the order in which allocations are made to LPs and GPs. There are two major types of waterfall: *deal by deal* (or *American*) and *whole of fund* (or *European*).

Weak-form efficient market hypothesis The belief that security prices fully reflect all past market data, which refers to all historical price and volume trading information.

Wealth effect An increase (decrease) in household wealth increases (decreases) consumer spending out of a given level of current income.

Web-based lender A lender that operates primarily on the internet, offering loans in relatively small amounts, typically to small businesses in need of cash.

Weighted average cost method An inventory accounting method that averages the total cost of available inventory items over the total units available for sale.

Weighted average cost of capital A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average cost of capital (WACC) A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average coupon rate Weighting the mortgage rate of each mortgage loan in the pool by the percentage of the mortgage outstanding relative to the outstanding amount of all the mortgages in the pool.

Weighted average life A measure that gives investors an indication of how long they can expect to hold the MBS before it is paid off; the convention-based average time to receipt of all principal repayments. Also called *average life*.

Weighted average maturity Weighting the remaining number of months to maturity for each mortgage loan in the pool by the amount of the outstanding mortgage balance.

Weighted mean An average in which each observation is weighted by an index of its relative importance.

Winsorized mean A mean computed after assigning a stated percentage of the lowest values equal to one specified low value and a stated percentage of the highest values equal to one specified high value.

Word cloud A visual device for representing textual data, which consists of words extracted from a source of textual data. The size of each distinct word is proportional to the frequency with which it appears in the given text (also known as **tag cloud**).

Working capital The difference between current assets and current liabilities.

Working capital management The management of a company's short-term assets (such as inventory) and short-term liabilities (such as money owed to suppliers).

World price The price prevailing in the world market.

Yield duration The sensitivity of the bond price with respect to the bond's own yield-to-maturity.

Yield to maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield to redemption* or *redemption yield*.

Yield-to-maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield-to-redemption* or *redemption yield*.

Yield-to-redemption See *yield-to-maturity*.

Yield-to-worst The lowest of the sequence of yields-to-call and the yield-to-maturity.

Zero-coupon bond A bond that does not pay interest during its life. It is issued at a discount to par value and redeemed at par. Also called *pure discount bond*.

Zero-volatility spread (Z-spread) Calculates a constant yield spread over a government (or interest rate swap) spot curve.

CFA Institute®
CFA Program

PORTFOLIO MANAGEMENT AND ETHICAL AND PROFESSIONAL STANDARDS

CFA® Program Curriculum
2022 • LEVEL I • VOLUME 6

WILEY

CONTENTS

How to Use the CFA Program Curriculum **xv**
Background on the CBOK **xv**
Organization of the Curriculum **xvi**
Features of the Curriculum **xvi**
Designing Your Personal Study Program **xvii**
CFA Institute Learning Ecosystem (LES) **xviii**
Prep Providers **xix**
Feedback **xx**

Portfolio Management

Study Session 18 **Portfolio Management (2)** **3**

Reading 51 **Basics of Portfolio Planning and Construction** **5**
Introduction **6**
Portfolio Planning, the Investment Policy Statement (IPS) and Its Major Components **6**
  The Investment Policy Statement **6**
  Major Components of an IPS **7**
IPS Risk and Return Objectives **8**
  Return Objectives **14**
IPS Constraints: Liquidity, Time Horizon, Tax Concerns, Legal and Regulatory Factors, and Unique Circumstances **15**
  Liquidity Requirements **16**
  Time Horizon **16**
  Tax Concerns **17**
  Legal and Regulatory Factors **18**
  Unique Circumstances and ESG Considerations **18**
Gathering Client Information **20**
Portfolio Construction and Capital Market Expectations **23**
  Capital Market Expectations **24**
The Strategic Asset Allocation **24**
Steps Toward an Actual Portfolio and Alternative Portfolio Organizing Principles **32**
  New Developments in Portfolio Management **36**
ESG Considerations in Portfolio Planning and Construction **37**
Summary **40**
Practice Problems **43**
Solutions **47**

Reading 52 **The Behavioral Biases of Individuals** **49**
Introduction and Categorizations of Behavioral Biases **49**
Categorizations of Behavioral Biases **50**
Cognitive Errors **50**
  Belief Perseverance Biases **50**

Processing Errors 56
Emotional Biases 63
Loss-Aversion Bias 63
Overconfidence Bias 65
Self-Control Bias 66
Status Quo Bias 66
Endowment Bias 67
Regret-Aversion Bias 69
How Behavioral Finance Influences Market Behavior 72
Defining Market Anomalies 72
Momentum 73
Bubbles and Crashes 73
Value 75
Summary 76
Practice Problems 78
Solutions 81

Reading 53 **Introduction to Risk Management** 83
Introduction 83
The Risk Management Process 85
The Risk Management Framework 87
Risk Governance - An Enterprise View 93
An Enterprise View of Risk Governance 93
Risk Tolerance 96
Risk Budgeting 98
Identification of Risk - Financial and Non-Financial Risk 101
Financial Risks 101
Non-Financial Risks 102
Identification of Risk - Interactions Between Risks 107
Measuring and Modifying Risk - Drivers and Metrics 110
Drivers 110
Metrics 112
Methods of Risk Modification - Prevention, Avoidance, and Acceptance 116
Risk Prevention and Avoidance 116
Risk Acceptance: Self-Insurance and Diversification 117
Methods of Risk Modification - Transfer, Shifting, Choosing a Method for Modifying 118
Risk Shifting 120
How to Choose Which Method for Modifying Risk 122
Summary 124
Practice Problems 127
Solutions 130

Reading 54 **Technical Analysis** 133
Introduction 133
Technical Analysis: Principles, Assumptions, and links to Investment Analysis 134
Principles and Assumptions 135
Technical Analysis and Behavioral Finance 136

Technical Analysis and Fundamental Analysis **138**
The Differences in Conducting/Interpreting Technical Analysis in Various Types of Markets **139**
Chart Types **141**
Types of Technical Analysis Charts **143**
Trend, Support, and Resistance **153**
Common Chart Patterns **156**
Reversal Patterns **156**
Continuation Patterns **166**
Technical Indicators: Moving Averages and Bollinger Bands **175**
Technical Indicators **177**
Technical Indicators: Oscillators, Relative Strength, and Sentiment **181**
Rate of Change Oscillator **182**
Relative Strength Index **184**
Stochastic Oscillator **186**
Moving-Average Convergence/Divergence Oscillator (MACD) **188**
Sentiment Indicators **189**
Intermarket Analysis **195**
Principles of Intermarket Analysis **196**
Technical Analysis Applications to Portfolio Management **198**
The Role of the Technical Analyst in Fundamental Portfolio Management **210**
Summary **212**
Practice Problems **215**
Solutions **221**

Reading 55 **Fintech in Investment Management** **225**
Introduction and What is Fintech **225**
What Is Fintech? **226**
Big Data **227**
Sources of Big Data **229**
Big Data Challenges **230**
Advanced Analytical Tools: Artificial Intelligence and Machine Learning **230**
Types of Machine Learning **232**
Data Science: Extracting Information from Big Data **233**
Data Processing Methods **233**
Data Visualization **234**
Selected Applications of Fintech to Investment Management; Text Analytics & Natural Language Processing **236**
Text Analytics and Natural Language Processing **236**
Robo-Advisory Services **237**
Risk Analysis **239**
Algorithmic Trading **240**
Distributed Ledger Technology, and Permissioned and Permissionless Networks **240**
Permissioned and Permissionless Networks **243**
Applications of Distributed Ledger Technology to Investment Management **243**
Cryptocurrencies **243**
Tokenization **244**

Post-Trade Clearing and Settlement 244
Compliance 244
Summary 245
Practice Problems 247
Solutions 249

Ethical and Professional Standards

Study Session 19 **Ethical and Professional Standards** 253

Reading 56 **Ethics and Trust in the Investment Profession** 255
Introduction 255
Ethics 257
Ethics and Professionalism 259
How Professions Establish Trust 260
Professions Are Evolving 262
Professionalism in Investment Management 262
Trust in Investment Management 263
CFA Institute as an Investment Management Professional Body 263
Challenges to Ethical Conduct 265
Ethical vs. Legal Standards 267
Ethical Decision-Making Frameworks 270
The Framework for Ethical Decision-Making 270
Applying the Framework 272
Conclusion 278
Summary 278
Practice Problems 281
Solutions 283

Reading 57 **Code of Ethics and Standards of Professional Conduct** 285
Preface 285
Evolution of the CFA Institute Code of Ethics and Standards of Professional Conduct 286
Standards of Practice Handbook 286
Summary of Changes in the Eleventh Edition 287
CFA Institute Professional Conduct Program 289
Adoption of the Code and Standards 290
Acknowledgments 290
Ethics and the Investment Industry 291
Why Ethics Matters 291
CFA Institute Code of Ethics and Standards of Professional Conduct 295
Preamble 295
The Code of Ethics 296
Standards of Professional Conduct 296
Practice Problems 300
Solutions 302

Reading 58

Guidance for Standards I–VII **305**
Standard I(A): Professionalism - Knowledge of the Law **305**
  Standard I(A) Knowledge of the Law **305**
  Guidance **306**
Standard I(A): Recommended Procedures **310**
  Members and Candidates **310**
  Distribution Area Laws **311**
  Legal Counsel **311**
  Dissociation **311**
  Firms **311**
Standard I(A): Application of the Standard **312**
  Example 1 (Notification of Known Violations): **312**
  Example 2 (Dissociating from a Violation): **312**
  Example 3 (Dissociating from a Violation): **312**
  Example 4 (Following the Highest Requirements): **313**
  Example 5 (Following the Highest Requirements): **313**
  Example 6 (Laws and Regulations Based on Religious Tenets): **313**
  Example 7 (Reporting Potential Unethical Actions): **314**
  Example 8 (Failure to Maintain Knowledge of the Law): **314**
Standard I(B): Professionalism - Independence and Objectivity **315**
  Guidance **315**
Standard I(B): Recommended Procedures **320**
Standard I(B): Application of the Standard **322**
  Example 1 (Travel Expenses): **322**
  Example 2 (Research Independence): **322**
  Example 3 (Research Independence and Intrafirm Pressure): **322**
  Example 4 (Research Independence and Issuer Relationship Pressure): **323**
  Example 5 (Research Independence and Sales Pressure): **323**
  Example 6 (Research Independence and Prior Coverage): **323**
  Example 7 (Gifts and Entertainment from Related Party): **324**
  Example 8 (Gifts and Entertainment from Client): **324**
  Example 9 (Travel Expenses from External Manager): **325**
  Example 10 (Research Independence and Compensation Arrangements): **325**
  Example 11 (Recommendation Objectivity and Service Fees): **326**
  Example 12 (Recommendation Objectivity): **326**
  Example 13 (Influencing Manager Selection Decisions): **327**
  Example 14 (Influencing Manager Selection Decisions): **327**
  Example 15 (Fund Manager Relationships): **327**
  Example 16 (Intrafirm Pressure): **328**
Standard I(C): Professionalism – Misrepresentation **328**
  Guidance **328**
Standard I(C): Recommended Procedures **332**
  Factual Presentations **332**
  Qualification Summary **333**
  Verify Outside Information **333**
  Maintain Webpages **333**
  Plagiarism Policy **333**

Standard I(C): Application of the Standard 334
Example 1 (Disclosure of Issuer-Paid Research): 334
Example 2 (Correction of Unintentional Errors): 334
Example 3 (Noncorrection of Known Errors): 334
Example 4 (Plagiarism): 335
Example 5 (Misrepresentation of Information): 335
Example 6 (Potential Information Misrepresentation): 335
Example 7 (Plagiarism): 336
Example 8 (Plagiarism): 336
Example 9 (Plagiarism): 336
Example 10 (Plagiarism): 337
Example 11 (Misrepresentation of Information): 337
Example 12 (Misrepresentation of Information): 337
Example 13 (Avoiding a Misrepresentation): 338
Example 14 (Misrepresenting Composite Construction): 338
Example 15 (Presenting Out-of-Date Information): 339
Example 16 (Overemphasis of Firm Results): 339
Standard I(D): Professionalism – Misconduct 340
Guidance 340
Standard I(D): Recommended Procedures 341
Standard I(D): Application of the Standard 341
Example 1 (Professionalism and Competence): 341
Example 2 (Fraud and Deceit): 341
Example 3 (Fraud and Deceit): 342
Example 4 (Personal Actions and Integrity): 342
Example 5 (Professional Misconduct): 342
Standard II(A): Integrity of Capital Markets - Material Nonpublic Information 343
Standard II(A) Material Nonpublic Information 343
Guidance 343
Standard II(A): Recommended Procedures 347
Achieve Public Dissemination 347
Adopt Compliance Procedures 347
Adopt Disclosure Procedures 348
Issue Press Releases 348
Firewall Elements 348
Appropriate Interdepartmental Communications 348
Physical Separation of Departments 349
Prevention of Personnel Overlap 349
A Reporting System 349
Personal Trading Limitations 350
Record Maintenance 350
Proprietary Trading Procedures 350
Communication to All Employees 350
Standard II(A): Application of the Standard 351
Example 1 (Acting on Nonpublic Information): 351
Example 2 (Controlling Nonpublic Information): 351
Example 3 (Selective Disclosure of Material Information): 352
Example 4 (Determining Materiality): 352
Example 5 (Applying the Mosaic Theory): 352

- Example 6 (Applying the Mosaic Theory): 353
- Example 7 (Analyst Recommendations as Material Nonpublic Information): 353
- Example 8 (Acting on Nonpublic Information): 353
- Example 9 (Mosaic Theory): 354
- Example 10 (Materiality Determination): 354
- Example 11 (Using an Expert Network): 355
- Example 12 (Using an Expert Network): 355

Standard II(B): Integrity of Capital Markets - Market Manipulation 355
- Guidance 356

Standard II(B): Application of the Standard 357
- Example 1 (Independent Analysis and Company Promotion): 357
- Example 2 (Personal Trading Practices and Price): 357
- Example 3 (Creating Artificial Price Volatility): 358
- Example 4 (Personal Trading and Volume): 358
- Example 5 (“Pump-Priming” Strategy): 358
- Example 6 (Creating Artificial Price Volatility): 359
- Example 7 (Pump and Dump Strategy): 360
- Example 8 (Manipulating Model Inputs): 360
- Example 9 (Information Manipulation): 360

Standard III(A): Duties to Clients - Loyalty, Prudence, and Care 361
- Standard III(A) Loyalty, Prudence, and Care 361
- Guidance 361

Standard III(A): Recommended Procedures 365
- Regular Account Information 365
- Client Approval 365
- Firm Policies 365

Standard III(A): Application of the Standard 366
- Example 1 (Identifying the Client—Plan Participants): 366
- Example 2 (Client Commission Practices): 367
- Example 3 (Brokerage Arrangements): 367
- Example 4 (Brokerage Arrangements): 368
- Example 5 (Client Commission Practices): 368
- Example 6 (Excessive Trading): 368
- Example 7 (Managing Family Accounts): 369
- Example 8 (Identifying the Client): 369
- Example 9 (Identifying the Client): 369
- Example 10 (Client Loyalty): 370
- Example 11 (Execution-Only Responsibilities): 370

Standard III(B): Duties to Clients - Fair Dealing 370
- Guidance 371

Standard III(B): Recommended Procedures 373
- Develop Firm Policies 373
- Disclose Trade Allocation Procedures 375
- Establish Systematic Account Review 375
- Disclose Levels of Service 375

Standard III(B): Application of the Standard 375
- Example 1 (Selective Disclosure): 375
- Example 2 (Fair Dealing between Funds): 376

Example 3 (Fair Dealing and IPO Distribution): 376
Example 4 (Fair Dealing and Transaction Allocation): 377
Example 5 (Selective Disclosure): 377
Example 6 (Additional Services for Select Clients): 377
Example 7 (Minimum Lot Allocations): 378
Example 8 (Excessive Trading): 378
Example 9 (Limited Social Media Disclosures): 378
Example 10 (Fair Dealing between Clients): 379
Standard III(C): Duties to Clients – Suitability 379
Guidance 380
Standard III(C): Recommended Procedures 382
Investment Policy Statement 383
Regular Updates 383
Suitability Test Policies 383
Standard III(C): Application of the Standard 384
Example 1 (Investment Suitability—Risk Profile): 384
Example 2 (Investment Suitability—Entire Portfolio): 384
Example 3 (IPS Updating): 384
Example 4 (Following an Investment Mandate): 385
Example 5 (IPS Requirements and Limitations): 385
Example 6 (Submanager and IPS Reviews): 386
Example 7 (Investment Suitability—Risk Profile): 386
Example 8 (Investment Suitability): 386
Standard III(D): Duties to Clients - Performance Presentation 387
Guidance 387
Standard III(D): Recommended Procedures 388
Apply the GIPS Standards 388
Compliance without Applying GIPS Standards 388
Standard III(D): Application of the Standard 388
Example 1 (Performance Calculation and Length of Time): 389
Example 2 (Performance Calculation and Asset Weighting): 389
Example 3 (Performance Presentation and Prior Fund/Employer): 389
Example 4 (Performance Presentation and Simulated Results): 390
Example 5 (Performance Calculation and Selected Accounts Only): 390
Example 6 (Performance Attribution Changes): 390
Example 7 (Performance Calculation Methodology Disclosure): 391
Example 8 (Performance Calculation Methodology Disclosure): 391
Standard III(E): Duties to Clients - Preservation of Confidentiality 392
Guidance 392
Standard III(E): Recommended Procedures 393
Communicating with Clients 394
Standard III(E): Application of the Standard 394
Example 1 (Possessing Confidential Information): 394
Example 2 (Disclosing Confidential Information): 394
Example 3 (Disclosing Possible Illegal Activity): 395
Example 4 (Disclosing Possible Illegal Activity): 395
Example 5 (Accidental Disclosure of Confidential Information): 395
Standard IV(A): Duties to Employers – Loyalty 396
Standard IV(A) Loyalty 396

Guidance 396
Standard IV(A): Recommended Procedures 400
Competition Policy 400
Termination Policy 400
Incident-Reporting Procedures 400
Employee Classification 400
Standard IV(A): Application of the Standard 401
Example 1 (Soliciting Former Clients): 401
Example 2 (Former Employer’s Documents and Files): 401
Example 3 (Addressing Rumors): 402
Example 4 (Ownership of Completed Prior Work): 402
Example 5 (Ownership of Completed Prior Work): 402
Example 6 (Soliciting Former Clients): 403
Example 7 (Starting a New Firm): 403
Example 8 (Competing with Current Employer): 404
Example 9 (Externally Compensated Assignments): 404
Example 10 (Soliciting Former Clients): 404
Example 11 (Whistleblowing Actions): 405
Example 12 (Soliciting Former Clients): 405
Example 13 (Notification of Code and Standards): 405
Example 14 (Leaving an Employer): 406
Example 15 (Confidential Firm Information): 407
Standard IV(B): Duties to Employers - Additional Compensation Arrangements 407
Guidance 407
Standard IV(B): Recommended Procedures 408
Standard IV(B): Application of the Standard 408
Example 1 (Notification of Client Bonus Compensation): 408
Example 2 (Notification of Outside Compensation): 409
Example 3 (Prior Approval for Outside Compensation): 409
Standard IV(C): Duties to Employers - Responsibilities of Supervisors 410
Guidance 410
Standard IV(C): Recommended Procedures 412
Codes of Ethics or Compliance Procedures 412
Adequate Compliance Procedures 413
Implementation of Compliance Education and Training 413
Establish an Appropriate Incentive Structure 414
Standard IV(C): Application of the Standard 414
Example 1 (Supervising Research Activities): 414
Example 2 (Supervising Research Activities): 415
Example 3 (Supervising Trading Activities): 415
Example 4 (Supervising Trading Activities and Record Keeping): 416
Example 5 (Accepting Responsibility): 416
Example 6 (Inadequate Procedures): 417
Example 7 (Inadequate Supervision): 417
Example 8 (Supervising Research Activities): 418
Example 9 (Supervising Research Activities): 418

Standard V(A): Investment Analysis, Recommendations, and Actions - Diligence and Reasonable Basis **419**
  - Standard V(A) Diligence and Reasonable Basis **419**
  - Guidance **419**

Standard V(A): Recommended Procedures **422**

Standard V(A): Application of the Standard **423**
  - Example 1 (Sufficient Due Diligence): **423**
  - Example 2 (Sufficient Scenario Testing): **424**
  - Example 3 (Developing a Reasonable Basis): **424**
  - Example 4 (Timely Client Updates): **424**
  - Example 5 (Group Research Opinions): **425**
  - Example 6 (Reliance on Third-Party Research): **425**
  - Example 7 (Due Diligence in Submanager Selection): **426**
  - Example 8 (Sufficient Due Diligence): **426**
  - Example 9 (Sufficient Due Diligence): **426**
  - Example 10 (Sufficient Due Diligence): **427**
  - Example 11 (Use of Quantitatively Oriented Models): **427**
  - Example 12 (Successful Due Diligence/Failed Investment): **428**
  - Example 13 (Quantitative Model Diligence): **428**
  - Example 14 (Selecting a Service Provider): **429**
  - Example 15 (Subadviser Selection): **429**
  - Example 16 (Manager Selection): **429**
  - Example 17 (Technical Model Requirements): **430**

Standard V(B): Investment Analysis, Recommendations, and Actions - Communication with Clients and Prospective Clients **431**
  - Guidance **431**

Standard V(B): Recommended Procedures **434**

Standard V(B): Application of the Standard **434**
  - Example 1 (Sufficient Disclosure of Investment System): **434**
  - Example 2 (Providing Opinions as Facts): **435**
  - Example 3 (Proper Description of a Security): **435**
  - Example 4 (Notification of Fund Mandate Change): **435**
  - Example 5 (Notification of Fund Mandate Change): **436**
  - Example 6 (Notification of Changes to the Investment Process): **436**
  - Example 7 (Notification of Changes to the Investment Process): **436**
  - Example 8 (Notification of Changes to the Investment Process): **437**
  - Example 9 (Sufficient Disclosure of Investment System): **437**
  - Example 10 (Notification of Changes to the Investment Process): **437**
  - Example 11 (Notification of Errors): **438**
  - Example 12 (Notification of Risks and Limitations): **438**
  - Example 13 (Notification of Risks and Limitations): **439**
  - Example 14 (Notification of Risks and Limitations): **439**

Standard V(C): Investment Analysis, Recommendations, and Actions - Record Retention **440**
  - Guidance **440**

Standard V(C): Recommended Procedures **441**

Standard V(C): Application of the Standard **442**
  - Example 1 (Record Retention and IPS Objectives and Recommendations): **442**

Example 2 (Record Retention and Research Process): 442
Example 3 (Records as Firm, Not Employee, Property): 442
Standard VI(A): Conflicts of Interest - Disclosure of Conflicts 443
Standard VI(A) Disclosure of Conflicts 443
Guidance 443
Standard VI(A): Recommended Procedures 446
Standard VI(A): Application of the Standard 446
Example 1 (Conflict of Interest and Business Relationships): 446
Example 2 (Conflict of Interest and Business Stock Ownership): 447
Example 3 (Conflict of Interest and Personal Stock Ownership): 447
Example 4 (Conflict of Interest and Personal Stock Ownership): 447
Example 5 (Conflict of Interest and Compensation Arrangements): 448
Example 6 (Conflict of Interest, Options, and Compensation Arrangements): 448
Example 7 (Conflict of Interest and Compensation Arrangements): 449
Example 8 (Conflict of Interest and Directorship): 449
Example 9 (Conflict of Interest and Personal Trading): 449
Example 10 (Conflict of Interest and Requested Favors): 450
Example 11 (Conflict of Interest and Business Relationships): 450
Example 12 (Disclosure of Conflicts to Employers): 451
Standard VI(B): Conflicts of Interest - Priority of Transactions 451
Guidance 452
Standard VI(B): Recommended Procedures 453
Standard VI(B): Application of the Standard 455
Example 1 (Personal Trading): 455
Example 2 (Trading for Family Member Account): 455
Example 3 (Family Accounts as Equals): 456
Example 4 (Personal Trading and Disclosure): 456
Example 5 (Trading Prior to Report Dissemination): 456
Standard VI(C): Conflicts of Interest - Referral Fees 457
Guidance 457
Standard VI(C): Recommended Procedures 457
Standard VI(C): Application of the Standard 458
Example 1 (Disclosure of Referral Arrangements and Outside Parties): 458
Example 2 (Disclosure of Interdepartmental Referral Arrangements): 459
Example 3 (Disclosure of Referral Arrangements and Informing Firm): 459
Example 4 (Disclosure of Referral Arrangements and Outside Organizations): 459
Example 5 (Disclosure of Referral Arrangements and Outside Parties): 460
Standard VII(A): Responsibilities as a CFA Institute Member or CFA Candidate - Conduct as Participants in CFA Institute Programs 460
Standard VII(A) Conduct as Participants in CFA Institute Programs 461
Guidance 461
Standard VII(A): Application of the Standard 463
Example 1 (Sharing Exam Questions): 463
Example 2 (Bringing Written Material into Exam Room): 464
Example 3 (Writing after Exam Period End): 464
Example 4 (Sharing Exam Content): 464
Example 5 (Sharing Exam Content): 465

Example 6 (Sharing Exam Content): 465
Example 7 (Discussion of Exam Grading Guidelines and Results): 465
Example 8 (Compromising CFA Institute Integrity as a Volunteer): 466
Example 9 (Compromising CFA Institute Integrity as a Volunteer): 466
Standard VII(B): Responsibilities as a CFA Institute Member or CFA Candidate - Reference to CFA Institute, the CFA Designation, and the CFA Program 466
Guidance 467
Standard VII(B): Recommended Procedures 469
Standard VII(B): Application of the Standard 470
Example 1 (Passing Exams in Consecutive Years): 470
Example 2 (Right to Use CFA Designation): 470
Example 3 (“Retired” CFA Institute Membership Status): 470
Example 4 (Stating Facts about CFA Designation and Program): 471
Example 5 (Order of Professional and Academic Designations): 471
Example 6 (Use of Fictitious Name): 471
Practice Problems 472
Solutions 482

Reading 59 **Introduction to the Global Investment Performance Standards (GIPS)** 491
Why Were the GIPS Created, Who Can Claim Compliance & Who Benefits from Compliance 492
Who Can Claim Compliance? 493
Who Benefits from Compliance? 493
Composites 494
Fundamentals of Compliance 494
Verification 495
Practice Problems 496
Solutions 497

Reading 60 **Ethics Application** 499
Introduction 499
Professionalism 499
Knowledge of the Law 499
Independence and Objectivity 502
Misrepresentation 503
Misconduct 506
Integrity of Capital Markets 507
Material Nonpublic Information 507
Market Manipulation 509
Duties to Clients 510
Loyalty, Prudence, and Care 510
Fair Dealing 513
Suitability 514
Performance Presentation 516
Preservation of Confidentiality 517
Duties to Employers 518
Loyalty 518
Additional Compensation Arrangements 521

Responsibilities of Supervisors **521**
Investment Analysis, Recommendations, and Actions **524**
Diligence and Reasonable Basis **524**
Communication with Clients and Prospective Clients **525**
Record Retention **527**
Conflicts of Interest **528**
Disclosure of Conflicts **528**
Priority of Transactions **528**
Referral Fees **530**
Responsibilities as a CFA Institute Member or CFA Candidate **531**
Conduct as Participants in CFA Institute Programs **531**
Reference to CFA Institute, the CFA Designation, and the CFA Program **532**

Glossary **G-1**

How to Use the CFA Program Curriculum

Congratulations on your decision to enter the Chartered Financial Analyst (CFA®) Program. This exciting and rewarding program of study reflects your desire to become a serious investment professional. You are embarking on a program noted for its high ethical standards and the breadth of knowledge, skills, and abilities (competencies) it develops. Your commitment should be educationally and professionally rewarding.

The credential you seek is respected around the world as a mark of accomplishment and dedication. Each level of the program represents a distinct achievement in professional development. Successful completion of the program is rewarded with membership in a prestigious global community of investment professionals. CFA charterholders are dedicated to life-long learning and maintaining currency with the ever-changing dynamics of a challenging profession. CFA Program enrollment represents the first step toward a career-long commitment to professional education.

The CFA exam measures your mastery of the core knowledge, skills, and abilities required to succeed as an investment professional. These core competencies are the basis for the Candidate Body of Knowledge (CBOK™). The CBOK consists of four components:

- A broad outline that lists the major CFA Program topic areas (www.cfainstitute.org/programs/cfa/curriculum/cbok);
- Topic area weights that indicate the relative exam weightings of the top-level topic areas (www.cfainstitute.org/programs/cfa/curriculum);
- Learning outcome statements (LOS) that advise candidates about the specific knowledge, skills, and abilities they should acquire from readings covering a topic area (LOS are provided in candidate study sessions and at the beginning of each reading); and
- CFA Program curriculum that candidates receive upon exam registration.

Therefore, the key to your success on the CFA exams is studying and understanding the CBOK. The following sections provide background on the CBOK, the organization of the curriculum, features of the curriculum, and tips for designing an effective personal study program.

BACKGROUND ON THE CBOK

CFA Program is grounded in the practice of the investment profession. CFA Institute performs a continuous practice analysis with investment professionals around the world to determine the competencies that are relevant to the profession, beginning with the Global Body of Investment Knowledge (GBIK®). Regional expert panels and targeted surveys are conducted annually to verify and reinforce the continuous feedback about the GBIK. The practice analysis process ultimately defines the CBOK. The CBOK reflects the competencies that are generally accepted and applied by investment professionals. These competencies are used in practice in a generalist context and are expected to be demonstrated by a recently qualified CFA charterholder.

© 2021 CFA Institute. All rights reserved.

The CFA Institute staff—in conjunction with the Education Advisory Committee and Curriculum Level Advisors, who consist of practicing CFA charterholders—designs the CFA Program curriculum in order to deliver the CBOK to candidates. The exams, also written by CFA charterholders, are designed to allow you to demonstrate your mastery of the CBOK as set forth in the CFA Program curriculum. As you structure your personal study program, you should emphasize mastery of the CBOK and the practical application of that knowledge. For more information on the practice analysis, CBOK, and development of the CFA Program curriculum, please visit www.cfainstitute.org.

ORGANIZATION OF THE CURRICULUM

The Level I CFA Program curriculum is organized into 10 topic areas. Each topic area begins with a brief statement of the material and the depth of knowledge expected. It is then divided into one or more study sessions. These study sessions should form the basic structure of your reading and preparation. Each study session includes a statement of its structure and objective and is further divided into assigned readings. An outline illustrating the organization of these study sessions can be found at the front of each volume of the curriculum.

The readings are commissioned by CFA Institute and written by content experts, including investment professionals and university professors. Each reading includes LOS and the core material to be studied, often a combination of text, exhibits, and in-text examples and questions. End of Reading Questions (EORQs) followed by solutions help you understand and master the material. The LOS indicate what you should be able to accomplish after studying the material. The LOS, the core material, and the EORQs are dependent on each other, with the core material and EORQs providing context for understanding the scope of the LOS and enabling you to apply a principle or concept in a variety of scenarios.

The entire readings, including the EORQs, are the basis for all exam questions and are selected or developed specifically to teach the knowledge, skills, and abilities reflected in the CBOK.

You should use the LOS to guide and focus your study because each exam question is based on one or more LOS and the core material and practice problems associated with the LOS. As a candidate, you are responsible for the entirety of the required material in a study session.

We encourage you to review the information about the LOS on our website (www.cfainstitute.org/programs/cfa/curriculum/study-sessions), including the descriptions of LOS “command words” on the candidate resources page at www.cfainstitute.org.

FEATURES OF THE CURRICULUM

End of Reading Questions/Solutions *All End of Reading Questions (EORQs) as well as their solutions are part of the curriculum and are required material for the exam.* In addition to the in-text examples and questions, these EORQs help demonstrate practical applications and reinforce your understanding of the concepts presented. Some of these EORQs are adapted from past CFA exams and/or may serve as a basis for exam questions.

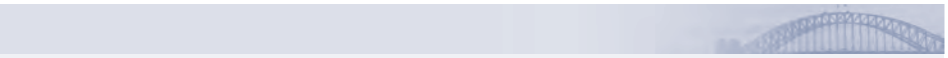
Glossary For your convenience, each volume includes a comprehensive Glossary. Throughout the curriculum, a **bolded** word in a reading denotes a term defined in the Glossary.

Note that the digital curriculum that is included in your exam registration fee is searchable for key words, including Glossary terms.

LOS Self-Check We have inserted checkboxes next to each LOS that you can use to track your progress in mastering the concepts in each reading.

Source Material The CFA Institute curriculum cites textbooks, journal articles, and other publications that provide additional context or information about topics covered in the readings. As a candidate, you are not responsible for familiarity with the original source materials cited in the curriculum.

Note that some readings may contain a web address or URL. The referenced sites were live at the time the reading was written or updated but may have been deactivated since then.



Some readings in the curriculum cite articles published in the *Financial Analysts Journal*®, which is the flagship publication of CFA Institute. Since its launch in 1945, the *Financial Analysts Journal* has established itself as the leading practitioner-oriented journal in the investment management community. Over the years, it has advanced the knowledge and understanding of the practice of investment management through the publication of peer-reviewed practitioner-relevant research from leading academics and practitioners. It has also featured thought-provoking opinion pieces that advance the common level of discourse within the investment management profession. Some of the most influential research in the area of investment management has appeared in the pages of the *Financial Analysts Journal*, and several Nobel laureates have contributed articles.

Candidates are not responsible for familiarity with *Financial Analysts Journal* articles that are cited in the curriculum. But, as your time and studies allow, we strongly encourage you to begin supplementing your understanding of key investment management issues by reading this, and other, CFA Institute practice-oriented publications through the Research & Analysis webpage (www.cfainstitute.org/en/research).

Errata The curriculum development process is rigorous and includes multiple rounds of reviews by content experts. Despite our efforts to produce a curriculum that is free of errors, there are times when we must make corrections. Curriculum errata are periodically updated and posted by exam level and test date online (www.cfainstitute.org/en/programs/submit-errata). If you believe you have found an error in the curriculum, you can submit your concerns through our curriculum errata reporting process found at the bottom of the Curriculum Errata webpage.

DESIGNING YOUR PERSONAL STUDY PROGRAM

Create a Schedule An orderly, systematic approach to exam preparation is critical. You should dedicate a consistent block of time every week to reading and studying. Complete all assigned readings and the associated problems and solutions in each study session. Review the LOS both before and after you study each reading to ensure that

you have mastered the applicable content and can demonstrate the knowledge, skills, and abilities described by the LOS and the assigned reading. Use the LOS self-check to track your progress and highlight areas of weakness for later review.

Successful candidates report an average of more than 300 hours preparing for each exam. Your preparation time will vary based on your prior education and experience, and you will probably spend more time on some study sessions than on others.

You should allow ample time for both in-depth study of all topic areas and additional concentration on those topic areas for which you feel the least prepared.

CFA INSTITUTE LEARNING ECOSYSTEM (LES)

As you prepare for your exam, we will email you important exam updates, testing policies, and study tips. Be sure to read these carefully.

Your exam registration fee includes access to the CFA Program Learning Ecosystem (LES). This digital learning platform provides access, even offline, to all of the readings and End of Reading Questions found in the print curriculum organized as a series of shorter online lessons with associated EORQs. This tool is your one-stop location for all study materials, including practice questions and mock exams.

The LES provides the following supplemental study tools:

Structured and Adaptive Study Plans The LES offers two ways to plan your study through the curriculum. The first is a structured plan that allows you to move through the material in the way that you feel best suits your learning. The second is an adaptive study plan based on the results of an assessment test that uses actual practice questions.

Regardless of your chosen study path, the LES tracks your level of proficiency in each topic area and presents you with a dashboard of where you stand in terms of proficiency so that you can allocate your study time efficiently.

Flashcards and Game Center The LES offers all the Glossary terms as Flashcards and tracks correct and incorrect answers. Flashcards can be filtered both by curriculum topic area and by action taken—for example, answered correctly, unanswered, and so on. These Flashcards provide a flexible way to study Glossary item definitions.

The Game Center provides several engaging ways to interact with the Flashcards in a game context. Each game tests your knowledge of the Glossary terms a in different way. Your results are scored and presented, along with a summary of candidates with high scores on the game, on your Dashboard.

Discussion Board The Discussion Board within the LES provides a way for you to interact with other candidates as you pursue your study plan. Discussions can happen at the level of individual lessons to raise questions about material in those lessons that you or other candidates can clarify or comment on. Discussions can also be posted at the level of topics or in the initial Welcome section to connect with other candidates in your area.

Practice Question Bank The LES offers access to a question bank of hundreds of practice questions that are in addition to the End of Reading Questions. These practice questions, only available on the LES, are intended to help you assess your mastery of individual topic areas as you progress through your studies. After each practice question, you will receive immediate feedback noting the correct response and indicating the relevant assigned reading so you can identify areas of weakness for further study.

Mock Exams The LES also includes access to three-hour Mock Exams that simulate the morning and afternoon sessions of the actual CFA exam. These Mock Exams are intended to be taken after you complete your study of the full curriculum and take practice questions so you can test your understanding of the curriculum and your readiness for the exam. If you take these Mock Exams within the LES, you will receive feedback afterward that notes the correct responses and indicates the relevant assigned readings so you can assess areas of weakness for further study. We recommend that you take Mock Exams during the final stages of your preparation for the actual CFA exam. For more information on the Mock Exams, please visit www.cfainstitute.org.

PREP PROVIDERS

You may choose to seek study support outside CFA Institute in the form of exam prep providers. After your CFA Program enrollment, you may receive numerous solicitations for exam prep courses and review materials. When considering a prep course, make sure the provider is committed to following the CFA Institute guidelines and high standards in its offerings.

Remember, however, that there are no shortcuts to success on the CFA exams; reading and studying the CFA Program curriculum *is* the key to success on the exam. The CFA Program exams reference only the CFA Institute assigned curriculum; no prep course or review course materials are consulted or referenced.

SUMMARY

Every question on the CFA exam is based on the content contained in the required readings and on one or more LOS. Frequently, an exam question is based on a specific example highlighted within a reading or on a specific practice problem and its solution. To make effective use of the CFA Program curriculum, please remember these key points:

1. All pages of the curriculum are required reading for the exam.
2. All questions, problems, and their solutions are part of the curriculum and are required study material for the exam. These questions are found at the end of the readings in the print versions of the curriculum. In the LES, these questions appear directly after the lesson with which they are associated. The LES provides immediate feedback on your answers and tracks your performance on these questions throughout your study.
3. We strongly encourage you to use the CFA Program Learning Ecosystem. In addition to providing access to all the curriculum material, including EORQs, in the form of shorter, focused lessons, the LES offers structured and adaptive study planning, a Discussion Board to communicate with other candidates, Flashcards, a Game Center for study activities, a test bank of practice questions, and online Mock Exams. Other supplemental study tools, such as eBook and PDF versions of the print curriculum, and additional candidate resources are available at www.cfainstitute.org.
4. Using the study planner, create a schedule and commit sufficient study time to cover the study sessions. You should also plan to review the materials, answer practice questions, and take Mock Exams.
5. Some of the concepts in the study sessions may be superseded by updated rulings and/or pronouncements issued after a reading was published. Candidates are expected to be familiar with the overall analytical framework contained in the assigned readings. Candidates are not responsible for changes that occur after the material was written.

FEEDBACK

At CFA Institute, we are committed to delivering a comprehensive and rigorous curriculum for the development of competent, ethically grounded investment professionals. We rely on candidate and investment professional comments and feedback as we work to improve the curriculum, supplemental study tools, and candidate resources.

Please send any comments or feedback to info@cfainstitute.org. You can be assured that we will review your suggestions carefully. Ongoing improvements in the curriculum will help you prepare for success on the upcoming exams and for a lifetime of learning as a serious investment professional.

Portfolio Management

STUDY SESSION

Study Session 18	Portfolio Management (2)

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to explain and demonstrate the use of fundamentals of portfolio and risk management, including return and risk measurement, and portfolio planning and construction.

© 2021 CFA Institute. All rights reserved.

PORTFOLIO MANAGEMENT
STUDY SESSION

18

Portfolio Management (2)

This study session introduces the portfolio planning and construction process, including the development of an investment policy statement (IPS). A discussion of risk management, including the various types and measures of risk, follows, and a risk management framework is provided. Technical analysis, a set of tools that uses asset price, trading volume, and other similar data for making investment decisions, is then examined. The session concludes with coverage on how financial technology (fintech) is impacting areas within the investment industry, such as investment analysis, automated advice, and risk management.

READING ASSIGNMENTS

Reading 51	Basics of Portfolio Planning and Construction by Alistair Byrne, PhD, CFA, and Frank E. Smudde, MSc, CFA
Reading 52	The Behavioral Biases of Individuals by Michael M. Pompian, CFA
Reading 53	Introduction to Risk Management by Don M. Chance, PhD, CFA, and Michael E. Edleson, PhD, CFA
Reading 54	Technical Analysis by Aksel Kibar, CMT, Barry M. Sine, and Robert A. Strong, PhD, CFA
Reading 55	Fintech in Investment Management by Robert Kissell, PhD, and Barbara J. Mack

© 2021 CFA Institute. All rights reserved.

READING

51

Basics of Portfolio Planning and Construction

by Alistair Byrne, PhD, CFA, and Frank E. Smudde, MSc, CFA

Alistair Byrne, PhD, CFA, is at State Street Global Advisors (United Kingdom). Frank E. Smudde, MSc, CFA, is at APG Asset Management (Netherlands).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the reasons for a written investment policy statement (IPS);
☐	**b.** describe the major components of an IPS;
☐	**c.** describe risk and return objectives and how they may be developed for a client;
☐	**d.** explain the difference between the willingness and the ability (capacity) to take risk in analyzing an investor's financial risk tolerance;
☐	**e.** describe the investment constraints of liquidity, time horizon, tax concerns, legal and regulatory factors, and unique circumstances and their implications for the choice of portfolio assets;
☐	**f.** explain the specification of asset classes in relation to asset allocation;
☐	**g.** describe the principles of portfolio construction and the role of asset allocation in relation to the IPS;
☐	**h.** describe how environmental, social, and governance (ESG) considerations may be integrated into portfolio planning and construction.

CFA Institute would like to thank Hardik Sanjay Shah, CFA, for his contributions to the 2022 update of this reading.

© 2021 CFA Institute. All rights reserved.

1

INTRODUCTION

To build a suitable portfolio for a client, investment advisers should first seek to understand the client's investment goals, resources, circumstances, and constraints. Investors can be categorized into broad groups based on shared characteristics with respect to these factors (e.g., various types of individual investors and institutional investors). Even investors within a given type, however, will invariably have a number of distinctive requirements. In this reading, we consider in detail the planning for investment success based on an individualized understanding of the client.

This reading is organized as follows: Section 2 discusses the investment policy statement, a written document that captures the client's investment objectives and the constraints. Section 3 discusses the portfolio construction process, including the first step of specifying a strategic asset allocation for the client. Section 4 concludes and summarizes the reading.

2

PORTFOLIO PLANNING, THE INVESTMENT POLICY STATEMENT (IPS) AND ITS MAJOR COMPONENTS

a describe the reasons for a written investment policy statement (IPS)

b describe the major components of an IPS

Portfolio planning can be defined as a program developed in advance of constructing a portfolio that is expected to define the client's investment objectives. The written document governing this process is the investment policy statement (IPS). The IPS is sometimes complemented by a document outlining policy on responsible investing—the broadest (umbrella) term used to describe principles that typically address one or more environmental, social, and governance themes that an investor requires to be considered when evaluating whether to invest in a particular company, as well as during the period of ownership. Sustainable investing, a term used in a similar context to responsible investing, focuses on factoring in sustainability issues during the investment process. Policies on responsible investing may also be integrated within the IPS itself. In the remainder of this reading, the integration of responsible investing within the IPS will be our working assumption.

2.1 The Investment Policy Statement

The IPS is the starting point of the portfolio management process. Without a full understanding of the client's situation and requirements, it is unlikely that successful results will be achieved. "Success" can be defined as a client achieving his important investment goals using means that he is comfortable with (in terms of risks taken and other concerns). The IPS essentially communicates a plan for achieving investment success.

The IPS is typically developed following a fact-finding discussion with the client. This discussion can include the use of a questionnaire designed to articulate the client's risk tolerance as well as address expectations in connection with specific circumstances. In the case of institutional clients, the fact finding may involve asset–liability management reviews, identification of liquidity needs, and a wide range of tax, legal, and other considerations.

The IPS can take a variety of forms.[1] A typical format will include the client's investment objectives and the constraints that apply to the client's portfolio.

The client's objectives are specified in terms of risk tolerance and return requirements. These elements must be consistent with each other: a client is unlikely to be able to find a portfolio that offers a relatively high expected return without taking on a relatively high level of expected risk. As part of their financial planning, clients may specify specific spending goals, which need to be considered when setting risk tolerance and return requirements.

The constraints section covers factors that need to be taken into account when constructing a portfolio for the client that meets the objectives. The typical categories are liquidity requirements, time horizon, regulatory requirements, tax status, and unique needs. The constraints may be either internal (i.e., set by the client) or external (i.e., set by law or regulation), as we discuss in detail later.

Having a well-constructed IPS for all clients should be standard procedure for an investment manager. The investment manager should build the portfolio with reference to the IPS and be able to refer to it to assess a particular investment's suitability for the client. In some cases, the need for the IPS goes beyond simply being a matter of standard procedure. In certain countries, the IPS (or an equivalent document) is a legal or regulatory requirement. For example, UK pension schemes must have a statement of investment principles under the Pensions Act 1995 (Section 35), and this statement is in essence an IPS. The UK Financial Services Authority also has requirements for investment firms to "know their customers." The European Union's Markets in Financial Instruments Directive ("MiFID") requires firms to assign clients to categories (eligible counterparties, institutional clients, or retail clients), with the category type determining the types of protections and limitations relevant for the client by law.

In the case of an institution, such as a pension plan or university endowment, the IPS may set out the governance arrangements that apply to the investment portfolio. For example, this information could cover the investment committee's approach to appointing and reviewing investment managers for the portfolio, and the discretion that those managers have.

The IPS should be reviewed on a regular basis to ensure that it remains consistent with the client's circumstances and requirements. For example, the UK Pensions Regulator suggests that a pension scheme's statements of investment principles—a form of IPS—should be reviewed at least every three years. The IPS should also be reviewed if the manager becomes aware of a material change in the client's circumstances, as well as on the initiative of the client when her objectives, time horizon, or liquidity needs change.

2.2 Major Components of an IPS

There is no single standard format for an IPS. Many IPS and investment governance documents with a similar purpose (as noted previously), however, include the following sections:

- *Introduction*. This section describes the client.
- *Statement of Purpose*. This section states the purpose of the IPS.

---

[1] In this reading, an IPS is assumed to be a document governing investment management activities covering all or most of a client's financial wealth. In many practical contexts, investment professionals work with investment mandates that cover only parts of a client's wealth or financial risk. Governance documents such as "Limited Partnership Agreements" and "Investment Management Agreements" will govern such mandates. Their contents are to a large degree comparable to the contents of the IPS as described in this reading.

- *Statement of Duties and Responsibilities*. This section details the duties and responsibilities of the client, the custodian of the client's assets, and the investment managers.
- *Procedures*. This section explains the steps to take to keep the IPS current and the procedures to follow to respond to various contingencies.
- *Investment Objectives*. This section explains the client's objectives in investing.
- *Investment Constraints*. This section presents the factors that constrain the client in seeking to achieve the investment objectives.
- *Investment Guidelines*. This section provides information about how policy should be executed (e.g., on the permissible use of leverage and derivatives) and on specific types of assets excluded from investment, if any.
- *Evaluation and Review*. This section provides guidance on obtaining feedback on investment results.
- *Appendices*: (A) Strategic Asset Allocation and (B) Rebalancing Policy. Many investors specify a strategic asset allocation (SAA), also known as the policy portfolio, which is the baseline allocation of portfolio assets to asset classes in view of the investor's investment objectives and the investor's policy with respect to rebalancing asset class weights. This SAA may include a statement of policy concerning hedging risks such as currency risk and interest rate risk.

The sections that are most closely linked to the client's distinctive needs, and probably the most important from a planning perspective, are those dealing with investment objectives and constraints. An IPS focusing on these two elements has been called an IPS in an "objectives and constraints" format.

In the following sections, we discuss the investment objectives and constraints format of an IPS beginning with risk and return objectives. The process of developing the IPS is the basic mechanism for evaluating and trying to improve an investor's overall expected return–risk stance. In a portfolio context, return objectives and expectations must be tailored to be consistent with risk objectives. The risk and return objectives must also be consistent with the constraints that apply to the portfolio. A growing proportion of investors explicitly include non-financial considerations when formulating their investment policies. This approach is often referred to as responsible investing (discussed earlier alongside related terms), which reflects environmental, social, and governance (ESG) considerations. Responsible investing recognizes that ESG considerations may eventually affect the portfolio's financial risk–return profile and may express the investor's societal convictions. In this reading, we discuss responsible investing aspects of investment policy, where relevant.

3

IPS RISK AND RETURN OBJECTIVES

b describe the major components of an IPS

c describe risk and return objectives and how they may be developed for a client

d explain the difference between the willingness and the ability (capacity) to take risk in analyzing an investor's financial risk tolerance

When constructing a portfolio for a client, it is important to ensure that the risk of the portfolio is suitable for the client. The IPS should state clearly the risk tolerance of the client. Risk objectives are specifications for portfolio risk that reflect the client's risk tolerance. Quantitative risk objectives can be absolute, relative, or a combination of the two.

Examples of an absolute risk objective would be a desire not to suffer any loss of capital or not to lose more than a given percentage of capital in any 12-month period. Note that these objectives are unrelated to investment market performance, good or bad, and are absolute in the sense of being self-standing. The fulfillment of such objectives could be achieved by not taking any risk—for example, by investing in an insured bank certificate of deposit at a creditworthy bank. If investments in risky assets are undertaken, however, such statements could be restated as a probability statement to be more operational (i.e., practically useful). For example, the desire not to lose more than 4% of capital in any 12-month period might be restated as an objective that with 95% probability the portfolio not lose more than 4% in any 12-month period. Measures of absolute risk include the variance or standard deviation of returns and **value at risk**.[2]

Some clients may choose to express relative risk objectives, which relate risk relative to one or more benchmarks perceived to represent appropriate risk standards. For example, investments in large-cap UK equities could be benchmarked to an equity market index, such as the FTSE 100 Index. The S&P 500 Index could be used as a benchmark for large-cap US equities; for investments with cash-like characteristics, the benchmark could be an interest rate such as Treasury bill rate. For risk relative to a benchmark, the measure could be **tracking risk**, or **tracking error**.[3] In practice, such risk objectives are used in situations where the total wealth management activities on behalf of a client are divided into partial mandates.

For institutional clients, the benchmark may be linked to some form of liability the institution has. For example, a pension plan must meet the pension payments as they come due, and the risk objective will be to minimize the probability that it will fail to do so. A related return objective might be to outperform the discount rate used in finding the present value of liabilities over a multi-year time horizon.

When a policy portfolio (that is, a specified set of long-term asset class weightings and hedge ratios) is used, the risk objective may be expressed as a desire for the portfolio return to be within a band of plus or minus X% of the benchmark return calculated by assigning an index or benchmark to represent each asset class present in the policy portfolio. Again, this objective may be more usefully interpreted as a statement of probability—for example, a 95% probability that the portfolio return will be within X% of the benchmark return over a stated period. Example 1 reviews this material.

EXAMPLE 1

Types of Risk Objectives

A Japanese institutional investor has a portfolio valued at ¥10 billion. The investor expresses her first risk objective as a desire not to lose more than ¥1 billion in the coming 12-month period. She specifies a second risk objective of achieving returns within 4% of the return to the TOPIX stock market index, which is her benchmark. Based on this information, address the following:

1 **A** Characterize the first risk objective as absolute or relative.

B Give an example of how the risk objective could be restated in a practical manner.

2 **A** Characterize the second risk objective as absolute or relative.

B Identify a measure for quantifying the risk objective.

---

[2] **Value at risk** is a money measure of the minimum value of losses expected during a specified period at a given level of probability.

[3] **Tracking risk** (sometimes called **tracking error**) is the standard deviation of the differences between a portfolio's returns and its benchmark's returns.

Solutions:

1 **A** This is an absolute risk objective.

B This risk objective could be restated in a practical manner by specifying that the 12-month 95% value at risk of the portfolio must be no more than ¥1 billion.

2 **A** This is a relative risk objective.

B This risk objective could be quantified using the tracking risk as a measure. For example, assuming returns follow a normal distribution, an expected tracking risk of 2% would imply a return within 4% of the index return approximately 95% of the time. Remember that tracking risk is stated as a one standard deviation measure.

A client's overall risk tolerance is a function of the client's ability to bear (accept) risk and her "risk attitude," which might be considered as the client's willingness to take risk. For ease of expression, from this point on we will refer to ability to bear risk and willingness to take risk as the two components of risk tolerance. Above-average ability to bear risk and above-average willingness to take risk imply above-average risk tolerance. Below-average ability to bear risk and below-average willingness to take risk imply below-average risk tolerance. These interactions are shown in Exhibit 1.

Exhibit 1 Risk Tolerance

	Ability to Bear Risk	
Willingness to Take Risk	**Below Average**	**Above Average**
Below Average	Below-average risk tolerance	Resolution needed
Above Average	Resolution needed	Above-average risk tolerance

The *ability* to bear risk is measured mainly in terms of objective factors, such as time horizon, expected income, and level of wealth relative to liabilities. For example, an investor with a 20-year time horizon can be considered to have a greater ability to bear risk, other things being equal, than an investor with a 2-year horizon. This difference is because over 20 years, there is more scope for losses to be recovered or other adjustments made to circumstances than there is over 2 years.

Similarly, an investor whose assets are comfortably in excess of their liabilities has more ability to bear risk than an investor whose wealth and expected future expenditure are more closely balanced. For example, a wealthy individual who can sustain a comfortable lifestyle after a very substantial investment loss has a relatively high ability to bear risk. A pension plan that has a large surplus of assets over liabilities has a relatively high ability to bear risk.

The *willingness* to take risk, or risk attitude, is a more subjective factor based on the client's psychology and perhaps also his current circumstances. Although the list of factors related to an individual's risk attitude remains open to debate, it is believed that some psychological factors, such as personality type, self-esteem, and inclination to independent thinking, are correlated with risk attitude. Some individuals are comfortable taking financial and investment risk, whereas others find it distressing. Although there is no single agreed-upon method for measuring risk tolerance, a willingness to take risk may be gauged by discussing risk with the client or by asking the client to complete a psychometric questionnaire. For example, financial planning academic John Grable and collaborators have developed 13-item and 5-item risk attitude questionnaires that have undergone some level of technical validation. The five-item questionnaire is shown in Exhibit 2.

Exhibit 2 A Five-Item Risk Assessment Instrument

1 Investing is too difficult to understand.
   - a Strongly agree
   - b Tend to agree
   - c Tend to disagree
   - d Strongly disagree

2 I am more comfortable putting my money in a bank account than in the stock market.
   - a Strongly agree
   - b Tend to agree
   - c Tend to disagree
   - d Strongly disagree

3 When I think of the word “risk,” the term “loss” comes to mind immediately.
   - a Strongly agree
   - b Tend to agree
   - c Tend to disagree
   - d Strongly disagree

4 Making money in stocks and bonds is based on luck.
   - a Strongly agree
   - b Tend to agree
   - c Tend to disagree
   - d Strongly disagree

5 In terms of investing, safety is more important than returns.
   - a Strongly agree
   - b Tend to agree
   - c Tend to disagree
   - d Strongly disagree

Source: Grable and Joo (2004).

The responses, a), b), c), and d), are coded 1, 2, 3, and 4, respectively, and summed. The lowest score is 5 and the highest score is 20, with higher scores indicating greater risk tolerance. For two random samples drawn from the faculty and staff of large US universities (n = 406), the mean score was 12.86 with a standard deviation of 3.01 and a median (i.e., most frequently observed) score of 13.

Note that a question, such as the first one in Exhibit 2, indicates that risk attitude may be associated with non-psychological factors (such as level of financial knowledge and understanding and decision-making style) as well as psychological factors.

The adviser needs to examine whether a client's ability to accept risk is consistent with the client's willingness to take risk. For example, a wealthy investor with a 20-year time horizon, who is thus able to take risk, may also be comfortable taking risk; in this case the factors are consistent. If the wealthy investor has a low willingness to take risk, there would be a conflict.

The conflict between ability and willingness to take risk can also arise in the institutional context. In addition, different stakeholders within the institution may take different views. For example, the trustees of a well-funded pension plan may desire a low-risk approach to safeguard the funding of the scheme and beneficiaries of the scheme may take a similar view. The sponsor, however, may wish a higher-risk/higher-return approach in an attempt to reduce future funding costs. When a trustee bears a fiduciary responsibility to pension beneficiaries and the interests of the pension sponsor and the pension beneficiaries conflict, the trustee should act in the best interests of the beneficiaries.

When both the ability and willingness to take risk are consistent, the investment adviser's task is the simplest. When ability to take risk is below average and willingness to take risk is above average, the investor's risk tolerance should be assessed as below average overall. When ability to take risk is above average but willingness is below average, the portfolio manager or adviser may seek to counsel the client and explain the conflict and its implications. For example, the adviser could outline the reasons why the client is considered to have a high ability to take risk and explain the likely consequences, in terms of reduced expected return, of not taking risk. The investment adviser, however, should not aim to change a client's willingness to take risk that is not a result of a miscalculation or misperception. Modification of elements of personality is not within the purview of the investment adviser's role. The prudent approach is to reach a conclusion about risk tolerance consistent with the lower of the two factors (ability and willingness) and to document the decisions made.

Example 2 is the first of a set that follows the analysis of an investment client through the preparation of the major elements of an IPS.

EXAMPLE 2

The Case of Henri Gascon: Risk Tolerance

Henri Gascon is an energy trader who works for a major French oil company based in Paris. He is 30 years old and married with one son, aged 5. Gascon has decided that it is time to review his financial situation and consults a financial adviser, who notes the following aspects of Gascon's situation:

- Gascon's annual salary of €250,000 is more than sufficient to cover the family's outgoings.
- Gascon owns his apartment outright and has €1,000,000 of savings.
- Gascon perceives that his job is reasonably secure.
- Gascon has a good knowledge of financial matters and is confident that equity markets will deliver positive returns over the long term.
- In the risk tolerance questionnaire, Gascon strongly disagrees with the statements that "making money in stocks and bonds is based on luck" and "in terms of investing, safety is more important than returns."
- Gascon expects that most of his savings will be used to fund his retirement, which he hopes to start at age 50.

Based only on the information given, which of the following statements is *most* accurate?

A Gascon has a low ability to take risk but a high willingness to take risk.

B Gascon has a high ability to take risk but a low willingness to take risk.

C Gascon has a high ability to take risk and a high willingness to take risk.

Solution:

C is correct. Gascon has a high income relative to outgoings, a high level of assets, a secure job, and a time horizon of 20 years. This information suggests a high *ability* to take risk. At the same time, Gascon is knowledgeable and confident about financial markets and responds to the questionnaire with answers that suggest risk tolerance. This result suggests he also has a high *willingness* to take risk.

EXAMPLE 3

The Case of Jacques Gascon: Risk Tolerance

Marie Gascon is so pleased with the services provided by her financial adviser that she suggests to her brother Jacques that he should also consult the adviser. Jacques thinks it is a good idea. Jacques, a self-employed computer consultant also based in Paris, is 40 years old and divorced with four children, aged between 12 and 16. The financial adviser notes the following aspects of Jacques' situation:

- Jacques' consultancy earnings average €40,000 per annum but are quite volatile.
- Jacques is required to pay €10,000 per year to his ex-wife and children.
- Jacques has a mortgage on his apartment of €100,000 and €10,000 of savings.
- Jacques has a good knowledge of financial matters and expects that equity markets will deliver very high returns over the long term.
- In the risk tolerance questionnaire, Jacques strongly disagrees with the statements "I am more comfortable putting my money in a bank account than in the stock market" and "When I think of the word 'risk', the term 'loss' comes to mind immediately."
- Jacques expects that most of his savings will be required to support his children at university.

Based only on the information given, which statement is correct?

A Jacques has a low ability to take risk but a high willingness to take risk.

B Jacques has a high ability to take risk but a low willingness to take risk.

C Jacques has a high ability to take risk and a high willingness to take risk.

Solution:

A is correct. Jacques does not have a particularly high income, his income is unstable, and he has reasonably high outgoings for his mortgage and maintenance payments. His investment time horizon is approximately two to six years given the ages of his children and his desire to support them at university. This finely balanced financial situation and short time horizon suggests a low ability to take risk. In contrast, his expectations for financial market returns and risk tolerance questionnaire answers suggest a high willingness to take risk. The financial adviser may wish to explain to Jacques how finely balanced his financial situation is and suggest that, despite his desire to take more risk, a relatively cautious portfolio might be the most appropriate approach to take.

3.1 Return Objectives

A client's return objectives can be stated in a number of ways. Similar to risk objectives, return objectives may be stated on an absolute or a relative basis.

As an example of an absolute objective, the client may want to achieve a particular percentage rate of return. This objective could be a nominal rate of return or could be expressed in real (inflation-adjusted) terms.

Alternatively, the return objective can be stated on a relative basis—for example, relative to a benchmark return. The benchmark could be an equity market index, such as the S&P 500 or the FTSE 100, or a cash rate of interest such as Libor. A relative return objective might be stated as, for example, a desire to outperform the benchmark index by one percentage point per year.

Some institutions also set their return objectives relative to a peer group or universe of managers—for example, an endowment aiming for a return that is in the top 50% of returns of similar institutions, or a private equity mandate aiming for returns in the top quartile among the private equity universe. This objective can be problematic when limited information is known about the investment strategies or the return calculation methodology being used by peers, and we must bear in mind the impossibility of *all* institutions being "above average." Furthermore, a good benchmark should be investable—that is, able to be replicated by the investor—and a peer benchmark typically does not meet that criterion.

In each case, the return requirement can be stated before or after fees. Care should be taken that the fee basis used is clear and understood by both the manager and client. The return can also be stated on either a pre- or post-tax basis when the investor is required to pay tax. For a taxable investor, the baseline is to state and analyze returns on an after-tax basis.

The return objective could be a required return—that is, the amount the investor needs to earn to meet a particular future goal—such as a certain level of retirement income.

The manager or adviser must ensure that the return objective is realistic. Care should be taken that client and manager are in agreement on whether the return objective is nominal (which is more convenient for measurement purposes) or real (i.e., inflation-adjusted, which usually relates better to the objective). It must be consistent with the client's risk objective (high expected returns are unlikely to be possible without high levels of risk) and also with the current economic and market environment. For example, 15% nominal returns might be possible when inflation is 10% but will be unlikely when inflation is 3%.

When a client has unrealistic return expectations, the manager or adviser will need to counsel her about what is achievable in the current market environment and within the client's tolerance for risk.

EXAMPLE 4

The Case of Marie Gascon: Return Objectives

Having assessed her risk tolerance, Marie Gascon now begins to discuss her retirement income needs with the financial adviser. She wishes to retire at age 50, which is 20 years from now. Her salary meets current and expected future expenditure requirements, but she does not expect to be able to make any additional pension contributions to her fund. Gascon sets aside €100,000 of her savings as an emergency fund to be held in cash. The remaining €900,000 is invested for her retirement.

Gascon estimates that a before-tax amount of €2,000,000 in today's money will be sufficient to fund her retirement income needs. The financial adviser expects inflation to average 2% per year over the next 20 years. Pension fund contributions and pension fund returns in France are exempt from tax, but pension fund distributions are taxable upon retirement.

1 Which of the following is closest to the amount of money Gascon will have to accumulate in nominal terms by her retirement date to meet her retirement income objective (i.e., expressed in money of the day in 20 years)?

A €900,000

B €2,000,000

C €3,000,000

2 Which of the following is closest to the annual rate of return that Gascon must earn on her pension portfolio to meet her retirement income objective?

A 2.0%

B 6.2%

C 8.1%

Solution to 1:

C is correct. At 2% annual inflation, €2,000,000 in today's money equates to €2,971,895 in 20 years measured in money of the day [€2,000,000 × $(1 + 2\%)^{20}$].

Solution to 2:

B is correct. €900,000 growing at 6.2% per year for 20 years will accumulate to €2,997,318, which is just above the required amount. (The solution of 6.2% comes from €2,997,318/€900,000 = $(1 + X)^{20}$, where X is the required rate of return.)

IPS CONSTRAINTS: LIQUIDITY, TIME HORIZON, TAX CONCERNS, LEGAL AND REGULATORY FACTORS, AND UNIQUE CIRCUMSTANCES

4

b describe the major components of an IPS

e describe the investment constraints of liquidity, time horizon, tax concerns, legal and regulatory factors, and unique circumstances and their implications for the choice of portfolio assets

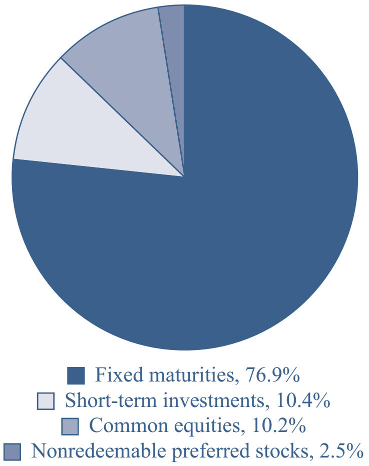
In the following sections, we analyze five major types of constraints on portfolio selection: liquidity, time horizon, tax concerns, legal and regulatory factors, and unique circumstances.

4.1 Liquidity Requirements

The IPS should state what the likely requirements are to withdraw funds from the portfolio. Examples for an individual investor would be outlays for covering healthcare payments or tuition fees. For institutions, it could be spending rules and requirements for endowment funds, the existence of claims coming due in the case of property and casualty insurance, or benefit payments for pension funds and life insurance companies.

When the client does have such a requirement, the manager should allocate part of the portfolio to cover the liability. This part of the portfolio will be invested in assets that are liquid—that is, easily converted to cash—and have low risk when the liquidity need is actually present (e.g., a bond maturing at the time when private education expenses will be incurred), so that their value is known with reasonable certainty. For example, the asset allocation in the insurance portfolios of US insurer Progressive Corporation (see Exhibit 3) shows a large allocation to fixed-income investments (called "Fixed maturities" by the company), some of which are either highly liquid or have a short maturity. These investments enable the company, in the case of automobile insurance, to pay claims for which the timing is unpredictable.

Exhibit 3 Asset Allocation of Progressive Corporation



Source: Progressive Corporation, 2018 Second Quarter Report.

4.2 Time Horizon

The IPS should state the time horizon over which the investor is investing. It may be the period over which the portfolio is accumulating before any assets need to be withdrawn; it could also be the period until the client's circumstances are likely to change. For example, a 55-year-old pension plan investor hoping to retire at age 65 has a 10-year horizon. The portfolio may not be liquidated at age 65, but its structure may need to change, for example, as the investor begins to draw an income from the fund.

The time horizon of the investor will affect the nature of investments used in the portfolio. Illiquid or risky investments may be unsuitable for an investor with a short time horizon because the investor may not have enough time to recover from investment losses, for example. Such investments, however, may be suitable for an investor with a longer horizon, especially if the risky investments are expected to have higher returns.

EXAMPLE 5

Investment Time Horizon

1. Frank Johnson is investing for retirement and has a 20-year horizon. He has an average risk tolerance. Which investment is likely to be the *least* suitable for a major allocation in Johnson's portfolio?
   - **A** Listed equities
   - **B** Private equity
   - **C** US Treasury bills
2. Al Smith has to pay a large tax bill in six months and wants to invest the money in the meantime. Which investment is likely to be the *least* suitable for a major allocation in Smith's portfolio?
   - **A** Listed equities
   - **B** Private equity
   - **C** US Treasury bills

Solution to 1:

C is correct. With a 20-year horizon and average risk tolerance, Johnson can accept the additional risk of listed equities and private equity compared with US Treasury bills.

Solution to 2:

B is correct. Private equity is risky, has no public market, and is the least liquid among the assets mentioned.

4.3 Tax Concerns

Tax status varies among investors. Some investors will be subject to taxation on investment returns and some will not. For example, in many countries, returns to pension funds are exempt from tax. Some investors will face a different tax rate on income (dividends and interest payments) than they do on capital gains (associated with increases in asset prices). Typically, when there is a differential, income is taxed more highly than gains. Gains may be subject to a lower tax rate, or part or all of the gain may be exempt from taxation. Furthermore, income may be taxed as it is earned, whereas gains may be taxed when they are realized. Hence, in such cases there is a time value of money benefit in the deferment of taxation of gains relative to income.

In many cases, the portfolio should reflect the tax status of the client. For example, a taxable investor may wish to hold a portfolio that emphasizes capital gains and receives little income. A taxable investor based in the United States is also likely to consider including US municipal bonds ("munis") in his portfolio because interest income from munis, unlike from Treasuries and corporate bonds, is exempt from taxes. A tax-exempt investor, such as a pension fund, will be relatively indifferent to the form of returns.

4.4 Legal and Regulatory Factors

The IPS should state any legal and regulatory restrictions that constrain how the portfolio is invested.

In some countries, such institutional investors as pension funds are subject to restrictions on portfolio composition. For example, there may be a limit on the proportion of equities or other risky assets in the portfolio or on the proportion of the portfolio that may be invested overseas. The United States has no limits on pension fund asset allocation, but some countries do, examples of which are shown in Exhibit 4. Pension funds also often face restrictions on the percentage of assets that can be invested in securities issued by the plan sponsor, so called **self-investment limits**.

Exhibit 4 Examples of Pension Fund Investment Restrictions

Country	Listed Equity	Real Estate	Government Bonds	Corporate Bonds	Foreign Currency Exposure
Switzerland	50%	30%	100%	100%	Unhedged 30%
Japan	100%	Not permitted	100%	100%	No limits
South Africa	75%	25%	100%	75%	25%

Source: OECD "Survey of Investment Regulations of Pension Funds," July 2018.

When an individual has access to material nonpublic information about a particular security, this situation may also form a constraint. For example, the directors of a public company may need to refrain from trading the company's stock at certain points of the year before financial results are published. The IPS should note this constraint so that the portfolio manager does not inadvertently trade the stock on the client's behalf.

4.5 Unique Circumstances and ESG Considerations

This section of the IPS should cover any other aspect of the client's circumstances, including beliefs and values, that is likely to have a material impact on portfolio composition. A client may have considerations derived from her faith or moral values that could constrain investment choices. For instance, an investor seeking compliance with Shari'a (the Islamic law) will avoid investing in businesses and financial instruments inconsistent with Shari'a, such as casinos and bonds, because Shari'a prohibits gambling and lending money on interest. Similarly, an investor may wish to avoid investments that he believes are inconsistent with his faith. Charitable and pension fund investors may have constituencies that want to express their values in an investment portfolio.

Whether rooted in religious beliefs or not, a client may have personal objections to certain products (e.g., weapons, tobacco, gambling) or practices (e.g., environmental impact of business activities, human impact of government policies, labor standards), which could lead to the exclusion of certain companies, countries, or types of securities (e.g., interest-bearing debt) from the investable universe as well as the client's benchmark. Investing in accordance with such considerations is referred to as socially responsible investing (SRI).

Specific ESG investment approaches can be classified in a variety of ways, and the investment community lacks clear consensus on terminology. We define six generic ESG investment approaches:

- *Negative screening*: Excluding companies or sectors based on business activities or environmental or social concerns;
- *Positive screening*: Including sectors or companies based on specific ESG criteria, typically ESG performance relative to industry peers;
- *ESG integration*: Systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions;
- *Thematic investing*: Investing in themes or assets related to ESG factors;
- *Engagement/active ownership*: Using shareholder power to influence corporate behavior to achieve targeted ESG objectives along with financial returns; and
- *Impact investing*: Investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.

These ESG investment approaches may impact a portfolio manager's investment universe and may also require the investment management firm to put in place a process to systematically incorporate ESG factors into the investment process.

EXAMPLE 6

Ethical Preferences

The BMO Responsible UK Equity Fund is designed for investors who wish to have ethical and ESG principles applied to the selection of their investments. The fund's managers apply both positive (features to be emphasized in the portfolio) and negative (features to be avoided in the portfolio) screening criteria:

Product-Based Screening Criteria

- Alcohol
- Arctic and oil sands
- Coal mining
- Gambling
- Nuclear power generation
- Pornography
- Tobacco
- Weapons

Conduct-Based Screening Criteria

Environmental

- Environmental management

Social

- Animal welfare
- Health and safety
- Human rights and oppressive regimes
- Labor standards

Governance

- Business ethics

[Excerpted from BMO Responsible UK Equity Fund documents; https://www.bmogam.com/gb-en/intermediary/bmo-responsible-uk-equity-2-inc/.]

When the portfolio represents only part of the client's total wealth, there may be aspects or portions of wealth not under the control of the manager that have implications for the portfolio. For example, an employee of a public company whose labor income and retirement income provision are reliant on that company, and who may have substantial investment exposure to the company through employee share options and stock holdings, may decide that his portfolio should not invest additional amounts in that stock. An entrepreneur may be reluctant to see her portfolio invested in the shares of competing businesses or in any business that has risk exposures aligned with her entrepreneurial venture.

A client's income may rely on a particular industry or asset class. Appropriate diversification requires that industry or asset class to be de-emphasized in the client's investments. For example, a stockbroker should consider having a relatively low weighting in equities, as his skills and thus his income-generating ability are worth less when equities do not perform well. Employees should similarly be wary of having concentrated share positions in the equity of the company where they work. If the employer encounters difficulties, not only may its employees lose their jobs but their investment portfolios could also suffer a significant loss of value.

5

GATHERING CLIENT INFORMATION

c describe risk and return objectives and how they may be developed for a client

e describe the investment constraints of liquidity, time horizon, tax concerns, legal and regulatory factors, and unique circumstances and their implications for the choice of portfolio assets

As noted earlier, it is important for portfolio managers and investment advisers to know their clients. For example, in the EU, MiFID II requires financial intermediaries to undertake substantial fact finding. This is required not only in the case of full-service wealth management or in the context of an IPS but also in "lighter" forms of financial intermediation, such as advisory relationships (in which clients make investment decisions after consultation with their investment adviser or broker) or execution-only relationships (in which the client makes investment decisions independently).

An exercise in fact finding about the customer should take place at the beginning of the client relationship. This process will involve gathering information about the client's circumstances as well as discussing the client's objectives and requirements.

Important data to gather from a client should cover family and employment situation as well as financial information. If the client is an individual, it may also be necessary to know about the situation and requirements of the client's spouse or other family members. The health of the client and her dependents is also relevant information. In an institutional relationship, it will be important to know about key stakeholders in the organization and what their perspective and requirements are. Information gathering may be done in an informal way or may involve structured interviews, questionnaires, or analysis of data. Many advisers will capture data electronically and use special systems that record data and produce customized reports.

Good recordkeeping is very important and may be crucial in a case in which any aspect of the client relationship comes into dispute at a later stage.

EXAMPLE 7

Marie Gascon: Description of Constraints

Marie Gascon continues to discuss her investment requirements with her financial adviser. The adviser begins to draft the constraints section of the IPS.

Gascon expects that she will continue to work for the oil company and that her relatively high income will continue for the foreseeable future. Gascon and her husband plan to have no additional children but expect that their son will go to a university at age 18. They expect that their son's education costs can be met out of their salary income.

Gascon's emergency reserve of €100,000 is considered to be sufficient as a reserve for unforeseen expenditures and emergencies. Her retirement savings of €900,000 has been contributed to her defined-contribution pension plan account to fund her retirement. Under French regulation, pension fund contributions are paid from gross income (i.e., income prior to deduction of tax), and pension fund returns are exempt from tax, but pension payments from a fund to retirees are taxed as income to the retiree.

With respect to Gascon's retirement savings portfolio, refer back to Example 2 as needed and address the following:

1 As concerns liquidity,

- **A** a maximum of 50% of the portfolio should be invested in liquid assets.
- **B** the portfolio should be invested entirely in liquid assets because of high spending needs.
- **C** the portfolio has no need for liquidity because there are no short-term spending requirements.

2 The investment time horizon is closest to

- **A** 5 years.
- **B** 20 years.
- **C** 40 years.

3 As concerns taxation, the portfolio

- **A** should emphasize capital gains because income is taxable.
- **B** should emphasize income because capital gains are taxable.
- **C** is tax exempt and thus indifferent between income and capital gains.

4 The principle legal and regulatory factors applying to the portfolio are

- **A** US securities laws.
- **B** European banking laws.
- **C** French pension fund regulations.

5 As concerns unique needs, the portfolio should

- **A** have a high weighting in oil and other commodity stocks.
- **B** be invested only in responsible and sustainable investments.
- **C** not have significant exposure to oil and other commodity stocks.

Solution to 1:

C is correct. The assets are for retirement use, which is 20 years away. Any short-term spending needs will be met from other assets or income.

Solution to 2:

B is correct. The relevant time horizon is to the retirement date, which is 20 years away. The assets may not be liquidated at that point, but a restructuring of the portfolio is to be expected as Gascon starts to draw an income from it.

Solution to 3:

C is correct. Because no tax is paid in the pension fund, it does not matter whether returns come in the form of income or capital gains.

Solution to 4:

C is correct. Management of the portfolio will have to comply with any rules relating to French pension funds.

Solution to 5:

C is correct. Gascon's human capital (i.e., future labor income) is affected by the prospects of the oil industry. If her portfolio has significant exposure to oil stocks, she would be increasing a risk exposure she already has.

Example 8, the final one based on Marie Gascon, shows how the information obtained from the fact-finding exercises might be incorporated into the objectives and constraints section of an IPS.

EXAMPLE 8

Marie Gascon: Outline of an IPS

Following is a simplified excerpt from the IPS the adviser prepares for Marie Gascon, covering objectives and constraints.

Risk Objectives:

- The portfolio may take on relatively high amounts of risk in seeking to meet the return requirements. With a 20-year time horizon and significant assets and income, the client has an above-average ability to take risk. The client is a knowledgeable investor, with an above-average willingness to take risk. Hence, the client's risk tolerance is above average, explaining the aforementioned portfolio risk objective.
- The portfolio should be well diversified with respect to asset classes and concentration of positions within an asset class. Although the client has above-average risk tolerance, his investment assets should be diversified to control the risk of catastrophic loss.

Return Objectives:

- The portfolio's long-term return requirement is 6.2% per year, in nominal terms and net of fees, to meet the client's retirement income goal.

Constraints:

- *Liquidity*: The portfolio consists of pension fund assets, and there is no need for liquidity in the short to medium term.
- *Time Horizon*: The portfolio will be invested with a 20-year time horizon. The client intends to retire in 20 years, at which time an income will be drawn from the portfolio.
- *Tax Status*: Under French law, contributions to the fund are made gross of tax and returns in the fund are tax-free. Hence, the client is indifferent between income and capital gains in the fund.
- *Legal and Regulatory Factors*: Management of the portfolio must comply with French pension fund regulations.
- *Unique Needs*: The client is an executive in the oil industry. The portfolio should strive to minimize additional exposures to oil and related stocks.

PORTFOLIO CONSTRUCTION AND CAPITAL MARKET EXPECTATIONS

6

f explain the specification of asset classes in relation to asset allocation

g describe the principles of portfolio construction and the role of asset allocation in relation to the IPS

Once the IPS has been compiled, the investment manager can construct a suitable portfolio. Strategic asset allocation is a traditional focus of the first steps in portfolio construction. The strategic asset allocation is stated in terms of percentage allocations to asset classes. An **asset class** is a category of assets that have similar characteristics, attributes, and risk–return relationships. The **strategic asset allocation** (SAA) is the set of exposures to IPS-permissible asset classes that is expected to achieve the client's long-term objectives given the client's risk profile and investment constraints. An SAA could include a policy of hedging portfolio risks not explicitly covered by asset class weights. The obvious examples are hedge ratios for foreign currency exposure, or the management of interest rate risk resulting from asset-liability mismatch, and the hedging of inflation risk. So-called "overlay" portfolios of derivatives are often used for this purpose.

The focus on the SAA is the result of a number of important investment principles. One principle is that a portfolio's systematic risk accounts for most of its change in value over the long term. **Systematic risk** is risk related to the economic system (e.g., risk related to business cycle) that cannot be eliminated by holding a diversified portfolio. This risk is different from **nonsystematic risk**, defined as the unique risks of particular assets, which may be avoided by holding other assets with offsetting risks. A second principle is that the returns to groups of similar assets (e.g., long-term debt claims) predictably reflect exposures to certain sets of systematic factors (e.g., for the debt claims, unexpected changes in the interest rate). Thus, the SAA is a means of providing the investor with exposure to the systematic risks of asset classes in proportions that meet the risk and return objectives.

The process of formulating a strategic asset allocation is based on the IPS, already discussed, and capital market expectations.

6.1 Capital Market Expectations

Capital market expectations are the investor's expectations concerning the risk and return prospects of asset classes, however broadly or narrowly the investor defines those asset classes. When associated with the client's investment objectives, the result is the strategic asset allocation that is expected to allow the client to achieve his investment objectives (at least under normal capital market conditions).

Traditionally, capital market expectations are quantified in terms of asset class expected returns, standard deviation of returns, and correlations among pairs of asset classes. Formally, the expected return of an asset class consists of the risk-free rate and one or more risk premium(s) associated with the asset class. Expected returns are in practice developed in a variety of ways, including the use of historical estimates, economic analysis, and various kinds of valuation models. Standard deviations and correlation estimates are frequently based on historical data and risk models.

7

THE STRATEGIC ASSET ALLOCATION

f explain the specification of asset classes in relation to asset allocation

g describe the principles of portfolio construction and the role of asset allocation in relation to the IPS

Traditionally, investors have distinguished cash, equities, bonds (government and corporate), and real estate as the major asset classes. In recent years, this list has been expanded with private equity, hedge funds, high-yield and emerging market bonds, and commodities. In addition, such assets as art and intellectual property rights may be considered asset classes for those investors prepared to take a more innovative approach and to accept some illiquidity. Combining such new asset classes as well as hedge funds and private equity under the header "alternative investments" has become accepted practice.

As the strategic asset allocation is built up by asset classes, the decision about how to define those asset classes is an important one. Defining the asset classes also determines the extent to which the investor controls the risk and return characteristics of the eventual investment portfolio. For example, separating bonds into government bonds and corporate bonds, and then further separating corporate bonds into investment grade and non-investment grade (high yield) and government bonds into domestic and foreign government bonds, creates four bond categories. For these categories, risk–return expectations can be expressed and correlations with other asset classes (and, in an asset–liability management context, with the liabilities) can be estimated. An investment manager who wants to explicitly consider the risk–return characteristics of those bond categories in the strategic asset allocation may choose to treat them as distinct asset classes. Similarly, in equities, some investors distinguish between emerging market and developed market equities, between domestic and international equities, or between large-cap and small-cap equities. In some regulatory environments for institutional investors, asset class definitions are mandatory, thereby forcing investment managers to articulate risk–return expectations (and apply risk management) on the asset classes specified. Conversely, a broader categorization of asset classes leaves the allocation between different categories of bonds and equities, for example, to managers responsible for these asset classes.

When defining asset classes, a number of criteria apply. Intuitively, an asset class should contain relatively homogeneous assets while providing diversification relative to other asset classes. In statistical terms, risk and return expectations should be similar, and paired correlations of assets should be relatively high within an asset class but should be lower versus assets in other asset classes. Also, the asset classes, while being mutually exclusive, should add up to a sufficient approximation of the relevant investable universe. Applying these criteria ensures that the strategic asset allocation process has considered all available investment alternatives.

EXAMPLE 9

Specifying Asset Classes

The strategic asset allocations of many institutional investors make a distinction between domestic equities and international equities or between developed market equities and emerging market equities. Often, equities are separated into different market capitalization brackets, resulting, for example, in an asset class such as domestic small-cap equity.

The correlation matrix in Exhibit 5 shows the paired correlations of monthly returns between different equity asset classes and other asset classes. Specifically, these correlations are measured over the period from December 2000 through August 2018. In addition, the exhibit shows the annualized volatility of monthly returns.

Exhibit 5 Asset Class Correlation Matrix

Correlations	US Equities	Emerging Markets	European Equities	Japanese Equities	US Small-Cap Equities	Commodities	European Gov't. Bonds	US Treasuries	US Credits	US High-Yield Credit
US Equities	1.00	0.78	0.88	0.59	0.89	0.32	0.08	–0.37	0.19	0.66
Emerging Markets Equities	0.78	1.00	0.84	0.64	0.75	0.46	0.21	–0.24	0.34	0.70
European Equities	0.88	0.84	1.00	0.64	0.79	0.43	0.16	–0.28	0.29	0.68
Japanese Equities	0.59	0.64	0.64	1.00	0.57	0.32	0.24	–0.18	0.29	0.52
US Small-Cap Equities	0.89	0.75	0.79	0.57	1.00	0.32	0.09	–0.36	0.19	0.69
Commodities	0.32	0.46	0.43	0.32	0.32	1.00	0.13	–0.18	0.12	0.36
European Gov't. Bonds	0.08	0.21	0.16	0.24	0.09	0.13	1.00	0.45	0.60	0.30
US Treasuries	–0.37	–0.24	–0.28	–0.18	–0.36	–0.18	0.45	1.00	0.58	–0.19
US Credits	0.19	0.34	0.29	0.29	0.19	0.12	0.60	0.58	1.00	0.54
US High-Yield Credit	0.66	0.70	0.68	0.52	0.69	0.36	0.30	–0.19	0.54	1.00
Volatility	14.3%	21.6%	18.4%	15.6%	18.4%	22.3%	4.9%	4.4%	5.5%	9.3%

Correlations and volatilities have been calculated using monthly returns from December 2000 through August 2018, unhedged, in USD.
Source: MSCI Bloomberg, S&P

Based only on the information given, address the following:

1 Contrast the correlations between equity asset classes with the correlations between equity asset classes and US Treasuries.

2 The monthly returns of which equity asset class differ the most from US equities?

Solution to 1:

The matrix reveals very strong correlation between the equity asset classes. For example, the correlation between European equities and US equities is 0.88. The correlation of equities with bonds, however, is much lower. For example, US equities, emerging markets equities, European equities, and Japanese equities all have negative correlation with US government bonds (–0.37, –0.24 and –0.28, and –0.18, respectively). It is worth noting, however, that correlations can vary through time and the values shown may be specific to the sample period used.

Solution to 2:

Among equity asset classes as listed in the table, the correlation between US and Japanese equities is the lowest, at 0.59. By contrast, correlations between US equities and emerging markets, European, and US small cap equities are 0.78 or higher.

Using correlation as a metric, Example 9 tends to indicate that only emerging markets were well differentiated from European equities. So, why do investors still often subdivide equities? Apart from any regulatory reasons, one explanation might be that this decomposition into smaller asset classes corresponds to the way the asset allocation is structured in portfolios. Many investment managers have expertise exclusively in specific areas of the market, such as emerging market equities, US small-cap equity, or international investment-grade credit. Bringing the asset class definitions of the asset allocation in line with investment products actually available in the market may simplify matters from an organizational perspective.

The risk–return profile of the strategic asset allocation depends on the expected returns and risks of the individual asset classes, as well as the correlation between those asset classes. In general, adding assets classes with low correlation improves the risk–return trade-off (more return for similar risk). Typically, the strategic asset allocation for risk-averse investors will have a large weight in government bonds and cash, whereas those with more willingness and ability to take risk will have more of their assets in risky asset classes, such as equities and many types of alternative investments.

It is customary to represent asset classes using benchmarks and universes calculated by providers such as FTSE, MSCI, or Bloomberg. An negative screening or a **best-in-class** policy (discussed previously) limits the number of securities to choose from, potentially impacting the risk and expected return estimates for these asset classes. Some examples of exclusions may be controversial weaponry or tobacco companies, or investments in certain countries. When such exclusions apply, risk and return estimates based on non-traditional ("off-the-shelf") asset class benchmarks may not be applicable. Separate benchmark indices reflecting the exclusions may available from the providers to mitigate this issue.

ABP is the pension fund for the Dutch government sector employees. The fund offers teachers, police officers, members of the military, and other civil servants a defined benefit pension plan, aiming for a pension of 70% of the average career real income for employees. As of the first quarter of 2018, ABP had €405 billion under management. The strategic asset allocation as of this period is shown in Exhibit 6.

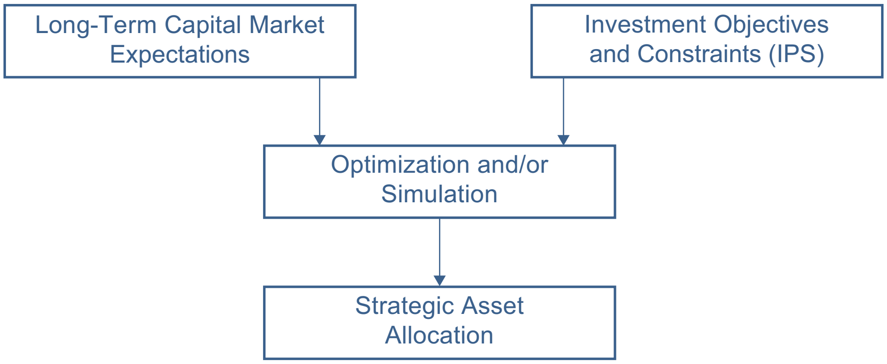
Exhibit 6 Strategic Asset Allocation for ABP

Equity	
Equities, developed countries	27%
Equities, emerging markets	9%
Total equity	***36%***
Alternatives	
Real estate	10%
Private equity	5%
Hedge funds	4%
Commodities	5%
Infrastructure	3%
Total alternatives	***27%***
Fixed-income securities	
Government bonds	13%
Corporate bonds	13%
Inflation-linked bonds	8%
Emerging market bonds	3%
Total fixed income	***37%***
Total	100%

Source: ABP Quarterly Report Q1 2018

A strategic asset allocation results from combining the constraints and objectives articulated in the IPS and long-term capital market expectations regarding the asset classes. The strategic asset allocation or policy portfolio will subsequently be implemented into real portfolios. Exhibit 7 illustrates conceptually how investment objectives and constraints and long-term capital market expectations combine into a policy portfolio.

Exhibit 7 Strategic Asset Allocation Process


Long-Term Capital Market Expectations
Investment Objectives and Constraints (IPS)
Optimization and/or Simulation
Strategic Asset Allocation

In some frameworks used in practice, the asset allocation is an integral part of the investment policy statement. This presentation, however, keeps the asset allocation separate from the investment policy statement because clients' investment objectives and constraints qualitatively differ in nature from capital market expectations, thus requiring different types of analysis, different sources of information, and different review cycles.

The combination of investment objectives/constraints and capital market expectations theoretically occurs using optimization techniques. In this section, we apply mean–variance optimization to a sample set of investment objectives and constraints, using an investment universe with associated market expectations. We assume that investors choosing from a range of asset allocations with similar returns would prefer those with lower risk. Choosing from allocations with similar levels of risk, investors would prefer those with the highest return. Formally, investors' risk and return objectives can be described as a utility function, in which utility increases with higher expected returns and lower risk. This assumption could yield an expected utility equation such as that shown in Equation 1.[4]

$$U_p = E\left(R_p\right) - \lambda\sigma_p^2 \quad (1)$$

where

U_p = the investor's expected utility from the portfolio
$E(R_p)$ = the expected return of the portfolio
σ_p = the standard deviation of returns of the portfolio
λ = a measure of the investor's risk aversion

This utility function expresses a positive relationship between utility and expected portfolio return (i.e., higher expected return increases utility, all else equal) and a negative relationship between utility and volatility of portfolio return as measured by the variance of portfolio returns. The stronger the negative relationship, the greater the investor's risk aversion. The portfolio is understood to represent a particular asset allocation. The asset allocation providing the highest expected utility is the one that is optimal for the investor given his or her risk aversion.

For different values of U_p, a line can be plotted that links those combinations of risk and expected return that produces that level of utility: an indifference curve. An investor would attain equal utility from all risk–return combinations on that curve.

---

[4] Sharpe, Chen, Pinto, and McLeavey (2007).

Capital market expectations, specified in asset classes' expected returns, standard deviations of return, and correlations, translate into an efficient frontier of portfolios. A multi-asset class portfolio's expected return is given by

$$E\left(R_p\right) = \sum_{i=1}^{n} w_i E\left(R_i\right) \tag{2}$$

where w_i equals the weight of asset class i in the portfolio, and its risk is given by

$$\sigma_p = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{p,i} w_{p,j} \mathrm{Cov}\left(R_i, R_j\right)} \tag{3}$$

The covariance between the returns on asset classes i and j is given by the product of the correlation between the two asset classes and their standard deviations of return:

$$\mathrm{Cov}(R_i, R_j) = \rho_{i,j}\sigma_i\sigma_j \tag{4}$$

where

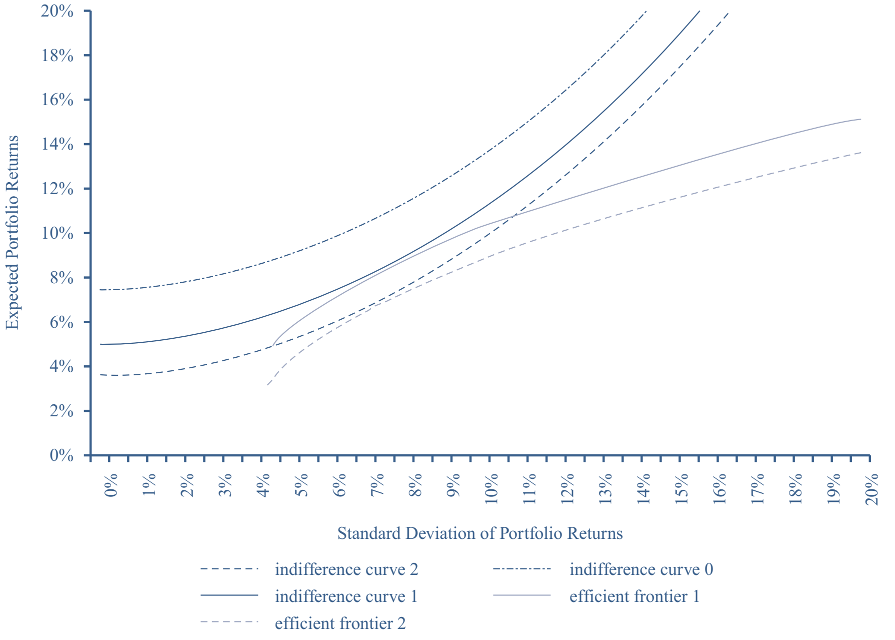
$\mathrm{Cov}(R_i, R_j)$ = the covariance between the return of asset classes i and j

$\rho_{i,j}$ = the correlation between the returns of asset classes i and j

The resulting portfolios can be represented as a scatter of dots in a chart depicting their risk and expected return. Because a portfolio's risk is a positive function of the risk of its assets and the correlations among them, a portfolio consisting of risky assets with low correlation has lower risk than one with similarly risky assets with high correlation. It is therefore possible to construct different portfolios with equal expected returns but with different levels of risk. The line that connects those portfolios with the minimal risk for each level of expected return (above that of the **minimum-variance portfolio**—the portfolio with the minimum variance for each given level of expected return) is the efficient frontier. Clearly, the efficient frontier will move "upward" as more low-correlation assets with sufficient expected return are added to the mix because it lowers the risk in the portfolios for equal expected returns. Similarly, when return expectations increase for asset classes while volatility and correlation assumptions remain unchanged, the efficient frontier will move upward because each portfolio is able to generate higher returns for the same level of risk.

Both the efficient frontier and a range of indifference curves can be plotted in the risk–return space. In Exhibit 8, the dark-colored curves that are concave from below represent efficient frontiers associated with different assumed expected returns. The lighter-colored curves are indifference curves. The point where the efficient frontier intersects with the indifference curve with the highest utility attainable (i.e., the point of tangency) represents the optimal asset allocation for the client/investor. In Exhibit 8, efficient frontier 1 has a point of tangency with indifference curve 1. Higher levels of utility, such as those associated with indifference curve 0, can apparently not be reached with the assets underlying the efficient frontier. It is clear that when capital market expectations change, this change moves the efficient frontier away from its original location. In the chart, this movement is illustrated by efficient frontier 2, which incorporates different capital market expectations. This new efficient frontier has a point of tangency with indifference curve 2, which is associated with a lower level of expected utility. Because the point of tangency represents the strategic asset allocation, it implies the asset allocation should be adjusted. Similarly, should investment objectives or constraints change, the indifference curves will change their shape and location. This change will again move the point of tangency, and hence change the asset allocation.

Exhibit 8 Strategic Asset Allocation Efficient Frontier



This framework describes how investor objectives and capital market expectations should theoretically be reconciled. It will, however, not be the exact procedure that in practice will be followed. First, an IPS does not necessarily translate the client's investment objectives and constraint into a utility function. Rather, an IPS gives threshold levels for risk and expected return, combined with a number of additional constraints that cannot be captured in this model. Second, the model illustrated is a single-period model, whereas in practice, the constraints from the IPS will make it more appropriate to use multi-period models. Multi-period problems can be more effectively addressed using simulation.

EXAMPLE 10

Approaching a SAA for a Private Investor

Rainer Gottschalk recently sold his local home construction company in the south of Germany to a large homebuilder with a nationwide reach. Upon selling his company, he accepted a job as regional manager for that nationwide homebuilder. Gottschalk is now considering his and his family's financial future. He looks forward to his new job—he likes his new role, and the position provides him with income to fulfill his family's short-term and medium-term liquidity needs. Gottschalk feels strongly that he should not invest the proceeds of the sale of his company in real estate because his income already depends on the state of the real estate market. Also, reflecting family values, he feels strongly that his savings should not support the tobacco industry. He therefore wants his equity allocation to exclude any stocks of tobacco product manufacturers or retailers. Gottschalk consults a financial adviser from his bank about how to invest his money to retire in good wealth in 20 years.

The IPS developed by his adviser suggests a return objective of 5%, with a standard deviation of 10%. The bank's asset management division provides Gottschalk and his adviser with the following data (Exhibit 9, Panel 1) on market expectations. The adviser estimates that excluding the tobacco industry from the investment universe affects expected equity returns of European equities by –0.2% and annual standard deviation by +0.1%. The impact on emerging market equities, and on the correlation structure, was considered negligible. Gottschalk accepts the results of these calculations as shown in Exhibit 9, Panel 2.

Exhibit 9 Risk, Return, and Correlation Estimates

			Correlation Matrix		
	Expected Return	Standard Deviation	European Equities	Emerging Mkt Equities	European Govt Bonds
Panel 1					
European equities	6.0%	15.0%	1.00	0.78	–0.08
Emerging market equities	8.0%	20.1%	0.78	1.00	–0.07
European government bonds	2.0%	7.8%	–0.08	–0.07	1.00
Panel 2					
European equities	5.8%	15.1%	1.00	0.78	–0.08
Emerging market equities	8.0%	20.1%	0.78	1.00	–0.07
European government bonds	2.00%	7.8%	–0.08	–0.07	1.00

Standard deviation and correlation calculated over the period March 1999–August 2018. All data in unhedged euros.
Sources: MSCI, Bloomberg

To illustrate the possibilities, the adviser presents Gottschalk with the following plot (Exhibit 10), in which the points forming the shaded curve outline the risk–return characteristics of the portfolios that can be constructed out of the three asset classes. An imaginary line linking the points with the lowest standard deviation for each attainable level of return would be the efficient frontier. The two straight lines show the risk and return objectives. Gottschalk should aim for portfolios that offer an expected return of at least 6% (the straight horizontal line or above) and a standard deviation of return of 12% or lower (the straight vertical line to the left).

Exhibit 10 Efficient Frontier

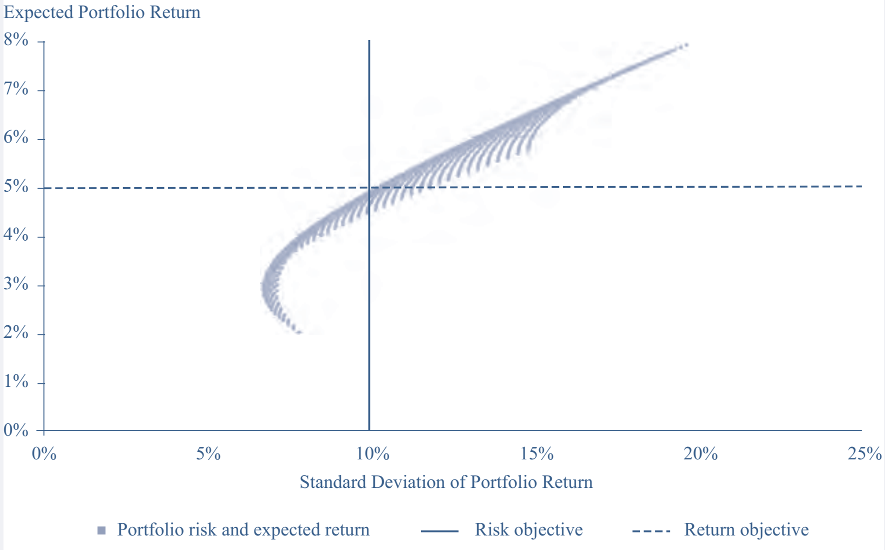


Exhibit 10 shows that no portfolio satisfies the two objectives (return of 5% and standard deviation of 10%) exactly, because the highest expected return that can be attained at a maximum volatility of 10% is 4.9%. This difference, Gottschalk and the adviser agree, is acceptable. The portfolio that would correspond with this expected return consists of 16% European stocks, 38% emerging market equities, and 46% government bonds.

8

STEPS TOWARD AN ACTUAL PORTFOLIO AND ALTERNATIVE PORTFOLIO ORGANIZING PRINCIPLES

g describe the principles of portfolio construction and the role of asset allocation in relation to the IPS

The strategic asset allocation in itself does not yet represent an actual investment portfolio. It is the first step in implementing an investment strategy. For quantitatively oriented portfolio managers, the next step is often risk budgeting.

As used in this reading, **risk budgeting** is the process of deciding on the amount of risk to assume in a portfolio (the overall risk budget) and subdividing that risk over the sources of investment return (e.g., strategic asset allocation, tactical asset allocation, and security selection). Because the decision about the total amount of risk to be taken is made in constructing the IPS, at this stage we are concerned about the subdivision of that risk.

Apart from the exposures to systematic risk factors specified in the strategic asset allocation, the returns of an investment strategy depend on two other sources: tactical asset allocation and security selection. **Tactical asset allocation** is the decision to deliberately deviate from the policy exposures to systematic risk factors (i.e., the policy weights of asset classes) with the intent to add value based on forecasts of the near-term returns of those asset classes. For instance, an investor may decide to temporarily invest more of the portfolio in equities than the SAA prescribes if the investor anticipates that equities will deliver a higher return over the short term than other asset classes. **Security selection** is an attempt to generate higher returns than the asset class benchmark by selecting securities with a higher expected return. For example, an investment manager may decide to add more IBM stock in her portfolio than the weight in her equity benchmark if she expects this stock to do better than the benchmark. To fund this purchase, she may sell another stock expected to do worse than either the benchmark or IBM. Obviously, deciding to deviate from policy weights or to select securities aiming to beat the benchmark creates additional uncertainty about returns. This risk is over and above the risk inherent in the policy portfolio. Hence, an investment policy should set risk limits and desired payoffs for each of these three activities.

Risk budgeting implies that the portfolio manager has to choose, for every asset class, whether to deploy security selection as a return generator. This choice is generally referred to as the choice between active or passive management. Contrary to strategic asset allocation, where exposures to sources of systematic risk are selected and sized, security selection is not rewarded with a long-run payoff to risk. Security selection is a zero-sum game: All investors in an asset class are competing with each other to identify a typically limited number of assets that are misvalued. In total, the gross returns of all market participants average out to the market return (the reward for taking systematic risk). This implies that the average active investor will match the market return and that one investor's gain versus the market return is the other investor's loss versus the market return. Because active managers tend to trade more and have to pay people (including themselves) to generate investment ideas or information leading to such ideas, however, the average active manager will underperform the market, net of costs. This fact does not imply that there are no skillful investment managers who, with some consistency, beat their benchmarks. Neither does it imply that all passive managers will be able to match the benchmark. The higher the turnover of an index, the more trading costs a passive manager will incur, making the task of matching the return of an index more difficult.

The likelihood of adding a significant amount of value from security selection depends on the skills of the manager and the informational efficiency of the market for the asset class his skill relates to. The more efficient an asset class or a subset of that asset class (such as a regional stock, bond, or real estate market or a size category within the stock market), the more skillful an asset manager has to be to add value. Broadly speaking, an efficient market is a market in which prices, on average, very quickly reflect newly available information. That requires a sizeable participation of investors trading risk against expected return, acting on rational expectations, using the same or similar pricing models, and having equal opportunities to access relevant information. Clearly, the market for US large-capitalization equities would be quite efficient. By contrast, some regional bond and equity markets do not have the technical and regulatory systems for information dissemination that are sufficient to serve all investors on a timely basis. Skilled managers should be able to exploit the resulting inefficiencies.

Sometimes the choice between active and passive management is actually made implicitly when the asset class is included in the asset allocation. The markets for some assets—such as those for non-listed real estate and infrastructure assets—are so illiquid that it is very difficult to buy a diversified exposure. As a result, participating in that market is not possible without engaging in security selection.

As the portfolio is constructed and its value changes with the returns of the asset classes and securities in which it is invested, the weights of the asset classes will gradually deviate from the policy weights in the strategic asset allocation. This process is referred to as drift. Periodically, or when a certain threshold deviation from the policy weight (the bandwidth) has been breached, the portfolio should be rebalanced back to the policy weights. The set of rules that guide the process of restoring the portfolio's original exposures to systematic risk factors is known as the **rebalancing policy**. Even absent a formal risk budget, formulating a rebalancing policy is an important element of risk management, as the following example illustrates.

EXAMPLE 11

Strategic and Tactical Asset Allocation for a European Charity

A European charity has an asset allocation at the beginning of the year consisting of the asset classes and weights shown in Exhibit 11.

Exhibit 11 Asset Allocation of a European Charity (beginning of year)

Asset Class	Policy Weight	Corridor (+/–)	Upper Limit	Lower Limit
European equities	30.0%	2.0%	32.0%	28.0%
International equities	15.0%	2.0%	17.0%	13.0%
European government bonds	20.0%	2.0%	22.0%	18.0%
Corporate bonds	20.0%	2.0%	22.0%	18.0%
Cash and money market instruments	15.0%	2.0%	17.0%	13.0%
Total	100.0%			

As Exhibit 11 reveals, the charity has a policy that the asset class weights cannot deviate from the policy weights by more than 2% (the corridor). The resulting upper and lower limits for the asset class weights are shown in the rightmost columns of the table. There are two reasons for asset class actual weights to deviate from policy weights: by deliberate choice (tactical asset allocation or market timing) and as a result of divergence of the returns of the different asset classes (drift). In this example, the asset class weights start the year exactly in line with policy weights.

After half a year, the investment portfolio is as shown in Exhibit 12.

Exhibit 12 Asset Allocation for a European Charity (six months later)

Asset Class	Policy Weight	Corridor (+/–)	Upper Limit	Lower Limit	Period Return	Ending Weight
European equities	30.0%	2.0%	32.0%	28.0%	15.0%	32.4%
International equities	15.0%	2.0%	17.0%	13.0%	10.0%	15.5%
European government bonds	20.0%	2.0%	22.0%	18.0%	0.5%	18.9%
Corporate bonds	20.0%	2.0%	22.0%	18.0%	1.5%	19.1%

Exhibit 12 (Continued)

Asset Class	Policy Weight	Corridor (+/–)	Upper Limit	Lower Limit	Period Return	Ending Weight
Cash and money market instruments	15.0%	2.0%	17.0%	13.0%	1.0%	14.2%
Total	100.0%				6.6%	100.0%

1 Discuss the returns of the portfolio and comment on the main asset weight changes.

Solution to 1:

The investment portfolio generated a return calculated on beginning (policy) weights of 6.55%, rounded to 6.6% (= 0.30 × 15% + 0.15 × 10% + 0.20 × 0.5% + 0.20 × 1.5% + 0.15 × 1.0%), mainly driven by a strong equity market. Bond returns were more subdued, leading to considerable drift in asset class weights. In particular, the European equity weight breached the upper limit of its allowed actual weight.

The investment committee decides against reducing European equities back to policy weight and adding to the fixed income and cash investments toward policy weights. Although this rebalancing would be prudent, the committee decides to engage in tactical asset allocation based on the view that this market will continue to be strong over the course of the year. It decides to just bring European equities back to within its bandwidth (a 32% portfolio weight) and add the proceeds to cash. Exhibit 13 shows the outcome after another half year.

Exhibit 13 Asset Allocation for a European Charity (an additional six months later)

Asset Class	Policy Weight	Starting Weight	Corridor (+/–)	Upper Limit	Lower Limit	Period Return	Ending Weight
European equities	30.0%	32.0%	2.0%	32.0%	28.0%	–9.0%	29.7%
International equities	15.0%	15.5%	2.0%	17.0%	13.0%	–6.0%	14.9%
European government bonds	20.0%	18.9%	2.0%	22.0%	18.0%	4.0%	20.0%
Corporate bonds	20.0%	19.1%	2.0%	22.0%	18.0%	4.0%	20.2%
Cash and money market instruments	15.0%	14.6%	2.0%	17.0%	13.0%	2.0%	15.2%
Total	100.0%					–2.0%	100.0%

The prior decision not to rebalance to policy weights did not have a positive result. Contrary to the investment committee's expectations, both European and international equities performed poorly while bonds recovered. The return of the portfolio was −2.0%.

2 How much of this return can be attributed to tactical asset allocation?

Solution to 2:

Because tactical asset allocation is the deliberate decision to deviate from policy weights, the return contribution from tactical asset allocation equals the difference between the actual return and the return that would have been made if the asset class weights were equal to the policy weights. Exhibit 14 shows this difference to be −0.30%.

Exhibit 14 Returns to Tactical Asset Allocation

Asset Class	Policy Weight I	Starting Weight II	Weights Difference III (= II – I)	Period Return IV	TAA Contribution V(= III × IV)
European equities	30.0%	32.0%	2.0%	−9.0%	−0.18%
International equities	15.0%	15.5%	0.5%	−6.0%	−0.03%
European government bonds	20.0%	18.9%	−1.1%	4.0%	−0.05%
Corporate bonds	20.0%	19.1%	−0.9%	4.0%	−0.04%
Cash and money market instruments	15.0%	14.6%	−0.4%	2.0%	−0.01%
Total	100.0%			−2.0%	−0.30%

The process of executing an investment strategy continues with selecting the appropriate manager(s) for each asset class and allocating funds to them. The investment portfolio management process is then well into the execution stage.

The investment managers' performance will be monitored, as well as the results of the tactical and strategic asset allocation. When asset class weights move outside their corridors, money is transferred from the asset classes that have become too large compared with the SAA to those that fall short. Managers as well as the strategic asset allocation will be reviewed on the basis of the outcome of the monitoring process. In addition, capital market expectations may change, as may the circumstances and objectives of the client. These changes could result in an adjustment of the strategic asset allocation.

8.1 New Developments in Portfolio Management

The portfolio planning and construction framework presented so far relies on a somewhat rigid process. Nonetheless, there are two newer, less structured developments that deserve specific mention.

The first development is the growth in the offering of exchange traded funds, or ETFs, in combination with algorithm-based financial advice (or robo-advice). ETFs are funds that track the performance of an asset class index or sub-index, are easily tradable, and are relatively cheap compared with actively managed funds or managed accounts. The broad array of ETF offerings, covering the main equity and fixed-income indices as well as commodities, enable retail investors to obtain fast, inexpensive, and liquid exposure to asset classes. Robo-advice has further reduced the costs for retail investors to create a well-diversified portfolio.

The second development relates to criticism of asset class return forecasts over relevant time horizons, as well as the perceived instability of asset class correlations and volatilities. Some market participants argue that poor investment portfolio results reflect the sensitivity of modern portfolio theory-based portfolio construction methodologies to small errors in return forecasts or estimated correlations. In response, practitioners developed an investment approach where asset classes were weighted according to risk contribution. This approach is known as *risk parity investing*. Proponents of risk parity investing argue that traditionally constructed portfolios have considerable risk from equities. That is, the typically high (60% or more) weight of equities in institutional portfolios understates the risk impact: equities tend to be much more volatile than fixed income. Opponents of risk parity argue that following the global financial crisis of 2007–2009, favorable results of risk parity portfolios were caused by the long period of decline in interest rates that benefited bond market performance.

ESG CONSIDERATIONS IN PORTFOLIO PLANNING AND CONSTRUCTION

9

h describe how environmental, social, and governance (ESG) considerations may be integrated into portfolio planning and construction

The implementation of a policy on responsible investing affects both strategic asset allocation and implementation of the portfolio construction process. The ESG investment approaches described previously require a set of instructions for investment managers with regard to the selection of securities, the exercise of shareholder rights, and the selection of investment strategies. Examples of issues driving the integration of environmental and social factors in the investment process include scarcity of natural resources, physical impacts of climate change, global economic and demographic trends, diversity and inclusion, and the rise of social media. ESG investment approaches can be implemented with structured, numeric data for many of these issues (e.g., executive salaries and bonuses, carbon footprint, employee turnover, lost time injuries and fatalities, and employee absenteeism). Although companies often are not required to disclose such data, that is changing as many stock exchanges and other regulatory bodies across developed and emerging markets have set up guidelines related to corporate sustainability disclosures for listed companies. In addition, many organizations and regulatory bodies have derived frameworks setting out standards on a number of these issues—examples include the Principles of Responsible Investment, the UN Global Compact, and the OECD Guidelines for Multinational Enterprises. These standards help form the basis of responsible investing policies for asset owners. In turn, asset owners may exclude or engage with companies in accordance with these issues, or demand from their selected investment managers consider these issues in their investment process.

We previously discussed that the limitation in the investment universe from using negative screening policies affects the expected returns and risk. When selecting or instructing active or passive managers, these managers will clearly prefer to see their performance measured against a benchmark that reflects the limited universe. There are benchmarks and investment vehicles (both active and passive) available, particularly in equities, that reflect many commonly excluded companies or sectors. It is also worth noting that with the proliferation of the ESG integration approach, more and more asset owners expect their asset managers to beat the regular benchmarks, because integration of ESG factors into traditional financial analysis and portfolio construction is viewed more as a process enhancement rather than an entirely new way to invest.

EXAMPLE 12

ESG Factors Directly Impacting Portfolio Construction

Based in South Africa, Mountain Materials (Mountain) is a fictitious cement manufacturing company that ranks as one of the largest cement and concrete manufacturers in the world. Mountain operates mostly in South Africa, where environmental regulations have been gradually strengthening since 2015. Because of the large scale of its operations, Mountain is a significant emitter of greenhouse gases (GHGs). During 2019, by setting a carbon price on the country's largest GHG emitters, South Africa launched a new, crucial endeavor in its efforts to tackle air pollution and climate change. Despite having some ad hoc initiatives to manage its carbon emissions, the company lacks firmwide programs to limit energy use or carbon emissions, thereby remaining exposed to increased costs to offset excess emissions. The average price on carbon across seven pilot markets in South Africa was between $5 and $15 per ton of carbon dioxide. In addition, the company’s performance in managing toxic air emissions as well as employee health and safety falls short of industry best practices, leaving Mountain exposed to related risks.

Ved Disha, CFA, is analyzing the effects of the environmental and social factors on Mountain’s financial statements. Exhibit 15 illustrates Disha’s expected internal rate of return (IRR) in the base, bear, and bull case scenarios for Mountain based on his fundamental analysis, and Exhibit 16 illustrates the same scenarios following the integration of these material environmental and social risks.

Exhibit 15 Pre-ESG Integration: Bear/Base/Bull Case Scenario

	Bear Case	Base Case	Bull Case
Revenue growth	0.0%	10.0%	15.0%
Margin improvement	–5.0%	3.0%	5.0%
Cash dividend	1.0%	2.0%	2.0%
Multiple expansion	–10.0%	0.0%	5.0%
IRR	–14.0%	15.0%	27.0%

Exhibit 16 Post-ESG Integration: Bear/Base/Bull Case Scenario

	Bear Case	Base Case	Bull Case
Revenue growth	–5.0%	10.0%	20.0%
Margin improvement	–10.0%	–3.0%	0.0%
Cash dividend	0.0%	0.0%	2.0%

Exhibit 16 (Continued)

	Bear Case	Base Case	Bull Case
Multiple expansion	–5.0%	–2.0%	0.0%
IRR	–20.0%	5.0%	22.0%

Disha assumed that compliance with national and provincial carbon regulations would require the company to increase spending on equipment, resulting in a 1% erosion in operating margin. Moreover, to limit toxic emissions, the company would have to switch to relatively cleaner sources of energy such as gas-based powered plants. This change is expected to further dampen operating margin by 2% because of increased fuel costs. As a result of higher spending, it was assumed that the previously stable cash dividend policy would turn conservative in the short term, and hence Disha reduced the expected dividend from 2% to 0% for the base case. Lastly, there is a downside to multiples as a result of the concerns related to management of health and safety risks, because the company's performance is below that of its peers and capital markets tend to discount the share price in the event of safety incidents. Based on all of these changes, the base case IRR for the cement company case became less attractive. This outlook led Disha to undertake a relatively smaller position of 0.5% (versus 1.5% in the absence of ESG risks) in his portfolio because of the various unmanaged ESG risks at Mountain. Disha also decided to engage with company management to influence better disclosures and management of these environmental and social risks. In this manner, key ESG risks and growth opportunities were integrated with traditional financial analysis to help arrive at a more robust investment decision.

In this example, a high level of unmanaged ESG risks led to a significant change in expected IRR following the ESG integration and hence impacted the position size significantly. It is prudent to note, however, that ESG is one of the many factors that influence investment decision making. Therefore, in many cases, ESG risk and opportunities may have limited effect on a company's financial attractiveness and thereby may not cause a large change in the portfolio. These risks and opportunities have to be analyzed and interpreted on a case-by-case basis.

Shareholder engagement requires good cooperation between investor (client) and investment manager. Engagement efforts are time-consuming, and the interest in such efforts is often that of the clients rather than that of the investment managers. Clients and investment managers must be clear with each other about the exercise of voting rights, filing of shareholder proposals, or entering into conversations with company management. It may be that the engagement and voting is delegated by the client to the investment manager and implemented according to the manager's stewardship policy. Alternatively, the client may instruct some proxy agent to vote on its behalf and according to its own stewardship policies, or the client may instruct voting and maintain dialogue with its investee companies through either individual engagements or collaborative engagements. Collaborative engagement initiatives have gained popularity because it is easier to gain the attention of and encourage positive action from corporations on material ESG issues through collective action. Climate Action 100+ is one such initiative that aims to ensure the world's largest corporate GHG emitters take necessary action on climate change. The initiative aims to engage with more than 100 systemically important carbon emitters, accounting for two-thirds of annual global industrial emissions, alongside more than 60 other companies with significant opportunity to drive the clean energy transition.

Selecting thematic investments, particularly in liquid asset classes, requires finding specialist managers who can identify the right opportunities and manage thematic investment portfolios. In particular, an allocation to thematic investments will bias the total asset class portfolio toward a particular theme, so it is important for the investment manager to demonstrate the impact of the thematic investment on the total risk–return profile of the portfolio. Impact investing specifically selects investment opportunities based on their intention to create a positive environmental and social impact.

The effort and costs associated with limiting the investment universe as part of responsible investing may suggest a negative impact on investment returns. Responsible investing proponents argue, however, that potential improvements in governance, as well as the avoidance of material risks by companies that screen, favorably improve returns. Significant empirical research has been conducted on the performance of ESG factors in equities, including the return differences of ESG equity portfolios relative to mainstream equity portfolios. Academic research remains mixed on the impact of ESG factors on portfolio returns. Nevertheless, ESG investing continues to see strong adoption, with nearly US$31 trillion of AUM dedicated toward responsible investment mandates at the start of 2018. The ESG integration approach that integrates material qualitative and quantitative environmental, social, and governance factors into traditional security and industry analysis as well as portfolio construction is now widely adopted across mainstream funds and not just limited to client-specific separate accounts.

SUMMARY

In this reading, we have discussed construction of a client's investment policy statement, including discussion of risk and return objectives and the various constraints that will apply to the portfolio. We have also discussed the portfolio construction process, with emphasis on the strategic asset allocation decisions that must be made.

- The IPS is the starting point of the portfolio management process. Without a full understanding of the client's situation and requirements, it is unlikely that successful results will be achieved.
- The IPS can take a variety of forms. A typical format will include the client's investment objectives and also list the constraints that apply to the client's portfolio.
- The client's objectives are specified in terms of risk tolerance and return requirements.
- The constraints section covers factors that need to be considered when constructing a portfolio for the client that meets the objectives. The typical constraint categories are liquidity requirements, time horizon, regulatory requirements, tax status, and unique needs.
- Clients may have personal objections to certain products or practices, which could lead to the exclusion of certain companies, countries, or types of securities from the investable universe as well as the client's benchmark. Such considerations are often referred to as ESG (environmental, social, governance).

- ESG considerations can be integrated into an investment policy by negative screening, positive screening, ESG integration, thematic investing, engagement/active ownership, and impact investing.
- Risk objectives are specifications for portfolio risk that reflect the risk tolerance of the client. Quantitative risk objectives can be absolute, relative, or a combination of the two.
- The client's overall risk tolerance is a function of both the client's ability to accept risk and the client's "risk attitude," which can be considered the client's willingness to take risk.
- The client's return objectives can be stated on an absolute or a relative basis. As an example of an absolute objective, the client may want to achieve a particular percentage rate of return. Alternatively, the return objective can be stated on a relative basis—for example, relative to a benchmark return.
- The liquidity section of the IPS should state what the client's requirements are to draw cash from the portfolio.
- The time horizon section of the IPS should state the time horizon over which the investor is investing. This horizon may be the period during which the portfolio is accumulating before any assets need to be withdrawn.
- Tax status varies among investors, and a client's tax status should be stated in the IPS.
- The IPS should state any legal or regulatory restrictions that constrain the investment of the portfolio.
- The unique circumstances section of the IPS should cover any other aspect of a client's circumstances that is likely to have a material impact on portfolio composition. Certain ESG implementation approaches may be discussed in this section.
- Asset classes are the building blocks of an asset allocation. An asset class is a category of assets that have similar characteristics, attributes, and risk–return relationships. Traditionally, investors have distinguished cash, equities, bonds, and real estate as the major asset classes.
- A strategic asset allocation results from combining the constraints and objectives articulated in the IPS and capital market expectations regarding the asset classes.
- As time goes on, a client's asset allocation will drift from the target allocation, and the amount of allowable drift as well as a rebalancing policy should be formalized.
- In addition to taking systematic risk, an investment committee may choose to take tactical asset allocation risk or security selection risk. The amount of return attributable to these decisions can be measured.
- ESG considerations may be integrated into the portfolio planning and construction process. ESG implementation approaches require a set of instructions for investment managers with regard to the selection of securities, the exercise of shareholder rights, and the selection of investment strategies.

REFERENCES

George, Binu, and Hardik Shah. 2018. “*ESG: Improving Your Risk-Adjusted Returns in Emerging Markets.*”

Grable, John E., and Soo-Hyun Joo. 2004. “Environmental and Biopsychosocial Factors Associated with Financial Risk Tolerance.” Financial Counseling and Planning 15 (1): 73–82.

Sharpe, William F., Peng Chen, Jerald E. Pinto, and Dennis W. McLeavey. 2007. “Asset Allocation.” In Managing Investment Portfolios: A Dynamic Process. 3rd ed. New York: Wiley.

PRACTICE PROBLEMS

1 Which of the following is *least* important as a reason for a written investment policy statement (IPS)?

A The IPS may be required by regulation.

B Having a written IPS is part of best practice for a portfolio manager.

C Having a written IPS ensures the client's risk and return objectives can be achieved.

2 Which of the following *best* describes the underlying rationale for a written investment policy statement (IPS)?

A A written IPS communicates a plan for trying to achieve investment success.

B A written IPS provides investment managers with a ready defense against client lawsuits.

C A written IPS allows investment managers to instruct clients about the proper use and purpose of investments.

3 A written investment policy statement (IPS) is *most* likely to succeed if:

A it is created by a software program to assure consistent quality.

B it is a collaborative effort of the client and the portfolio manager.

C it reflects the investment philosophy of the portfolio manager.

4 The section of the investment policy statement (IPS) that provides information about how policy may be executed, including restrictions and exclusions, is *best* described as the:

A *Investment Objectives.*

B *Investment Guidelines.*

C *Statement of Duties and Responsibilities.*

5 Which of the following is *least* likely to be placed in the appendices to an investment policy statement (IPS)?

A *Rebalancing Policy.*

B *Strategic Asset Allocation.*

C *Statement of Duties and Responsibilities.*

6 Which of the following typical topics in an investment policy statement (IPS) is *most* closely linked to the client's "distinctive needs"?

A *Procedures.*

B *Investment Guidelines.*

C *Statement of Duties and Responsibilities.*

7 An investment policy statement that includes a return objective of outperforming the FTSE 100 by 120 basis points is *best* characterized as having a(n):

A relative return objective.

B absolute return objective.

C arbitrage-based return objective.

8 Risk assessment questionnaires for investment management clients are *most* useful in measuring:

A value at risk.

© 2011 CFA Institute. All rights reserved.

B ability to take risk.

C willingness to take risk.

9 Which of the following is *best* characterized as a relative risk objective?

A Value at risk for the fund will not exceed US$3 million.

B The fund will not underperform the DAX by more than 250 basis points.

C The fund will not lose more than €2.5 million in the coming 12-month period.

10 In preparing an investment policy statement, which of the following is *most* difficult to quantify?

A Time horizon.

B Ability to accept risk.

C Willingness to accept risk.

11 After interviewing a client in order to prepare a written investment policy statement (IPS), you have established the following:

- The client has earnings that vary dramatically between £30,000 and £70,000 (pre-tax) depending on weather patterns in Britain.
- In three of the previous five years, the after-tax income of the client has been less than £20,000.
- The client's mother is dependent on her son (the client) for approximately £9,000 per year support.
- The client's own subsistence needs are approximately £12,000 per year.
- The client has more than 10 years' experience trading investments including commodity futures, stock options, and selling stock short.
- The client's responses to a standard risk assessment questionnaire suggest he has above average risk tolerance.

The client is *best* described as having a:

A low ability to take risk, but a high willingness to take risk.

B high ability to take risk, but a low willingness to take risk.

C high ability to take risk and a high willingness to take risk.

12 After interviewing a client in order to prepare a written investment policy statement (IPS), you have established the following:

- The client has earnings that have exceeded €120,000 (pre-tax) each year for the past five years.
- She has no dependents.
- The client's subsistence needs are approximately €45,000 per year.
- The client states that she feels uncomfortable with her lack of understanding of securities markets.
- All of the client's current savings are invested in short-term securities guaranteed by an agency of her national government.
- The client's responses to a standard risk assessment questionnaire suggest she has low risk tolerance.

The client is *best* described as having a:

A low ability to take risk, but a high willingness to take risk.

B high ability to take risk, but a low willingness to take risk.

C high ability to take risk and a high willingness to take risk.

13 A client who is a 34-year old widow with two healthy young children (aged 5 and 7) has asked you to help her form an investment policy statement. She has been employed as an administrative assistant in a bureau of her national government for the previous 12 years. She has two primary financial goals—her retirement and providing for the college education of her children. This client's time horizon is *best* described as being:

A long term.

B short term.

C medium term.

14 The timing of payouts for property and casualty insurers is unpredictable ("lumpy") in comparison with the timing of payouts for life insurance companies. Therefore, in general, property and casualty insurers have:

A lower liquidity needs than life insurance companies.

B greater liquidity needs than life insurance companies.

C a higher return objective than life insurance companies.

15 A client who is a director of a publicly listed corporation is required by law to refrain from trading that company's stock at certain points of the year when disclosure of financial results are pending. In preparing a written investment policy statement (IPS) for this client, this restriction on trading:

A is irrelevant to the IPS.

B should be included in the IPS.

C makes it illegal for the portfolio manager to work with this client.

16 Consider the pairwise correlations of monthly returns of the following asset classes:

	Brazilian Equities	East Asian Equities	European Equities	US Equities
Brazilian equities	1.00	0.70	0.85	0.76
East Asian equities	0.70	1.00	0.91	0.88
European equities	0.85	0.91	1.00	0.90
US equities	0.76	0.88	0.90	1.00

Based solely on the information in the above table, which equity asset class is *most* sharply distinguished from US equities?

A Brazilian equities.

B European equities.

C East Asian equities.

17 Returns on asset classes are *best* described as being a function of:

A the failure of arbitrage.

B exposure to the idiosyncratic risks of those asset classes.

C exposure to sets of systematic factors relevant to those asset classes.

18 In defining asset classes as part of the strategic asset allocation decision, pairwise correlations within asset classes should generally be:

A equal to correlations among asset classes.

B lower than correlations among asset classes.

C higher than correlations among asset classes.

19 Tactical asset allocation is *best* described as:

A attempts to exploit arbitrage possibilities among asset classes.

B the decision to deliberately deviate from the policy portfolio.

C selecting asset classes with the desired exposures to sources of systematic risk in an investment portfolio.

SOLUTIONS

1 C is correct. Depending on circumstances, a written IPS or its equivalent may be required by law or regulation and a written IPS is certainly consistent with best practices. The mere fact that a written IPS is prepared for a client, however, does not *ensure* that risk and return objectives will in fact be achieved.

2 A is correct. A written IPS is best seen as a communication instrument allowing clients and portfolio managers to mutually establish investment objectives and constraints.

3 B is correct. A written IPS, to be successful, must incorporate a full understanding of the client's situation and requirements. As stated in the reading, "The IPS will be developed following a fact finding discussion with the client."

4 B is correct. The major components of an IPS are listed in Section 2.2 of the reading. *Investment Guidelines* are described as the section that provides information about how policy may be executed, including restrictions on the permissible use of leverage and derivatives and on specific types of assets excluded from investment, if any. *Statement of Duties and Responsibilities* "detail[s] the duties and responsibilities of the client, the custodian of the client's assets, the investment managers, and so forth." *Investment Objectives* is "a section explaining the client's objectives in investing."

5 C is correct. The major components of an IPS are listed in Section 2.2 of the reading. Strategic Asset Allocation (also known as the policy portfolio) and Rebalancing Policy are often included as appendices to the IPS. The *Statement of Duties and Responsibilities*, however, is an integral part of the IPS and is unlikely to be placed in an appendix.

6 B is correct. According to the reading, "The sections of an IPS that are most closely linked to the client's distinctive needs are those dealing with investment objectives and constraints." *Investment Guidelines* "[provide] information about how policy may be executed, including investment constraints." *Procedures* "[detail] the steps to be taken to keep the IPS current and the procedures to follow to respond to various contingencies." *Statement of Duties and Responsibilities* "detail[s] the duties and responsibilities of the client, the custodian of the client's assets, the investment managers, and so forth."

7 A is correct. Because the return objective specifies a target return *relative to* the FTSE 100 Index, the objective is best described as a relative return objective.

8 C is correct. Risk attitude is a subjective factor and measuring risk attitude is difficult. Oftentimes, investment managers use psychometric questionnaires, such as those developed by Grable and Joo (2004), to assess a client's willingness to take risk.

9 B is correct. The reference to the DAX marks this response as a relative risk objective. Value at risk establishes a minimum value of loss expected during a specified time period at a given level of probability. A statement of maximum allowed absolute loss (€2.5 million) is an absolute risk objective.

10 C is correct. Measuring willingness to take risk (risk tolerance, risk aversion) is an exercise in applied psychology. Instruments attempting to measure risk attitudes exist, but they are clearly less objective than measurements of ability to take risk. Ability to take risk is based on relatively objective traits such as expected income, time horizon, and existing wealth relative to liabilities.

11 A is correct. The volatility of the client's income and the significant support needs for his mother and himself suggest that the client has a low ability to take risk. The client's trading experience and his responses to the risk assessment questionnaire indicate that the client has an above average willingness to take risk.

12 B is correct. On the one hand, the client has a stable, high income and no dependents. On the other hand, she exhibits above average risk aversion. Her ability to take risk is high, but her willingness to take risk is low.

13 A is correct. The client's financial objectives are long term. Her stable employment indicates that her immediate liquidity needs are modest. The children will not go to college until 10 or more years later. Her time horizon is best described as being long term.

14 B is correct. The unpredictable nature of property and casualty (P&C) claims forces P&C insurers to allocate a substantial proportion of their investments into liquid, short maturity assets. This need for liquidity also forces P&C companies to accept investments with relatively low expected returns. Liquidity is of less concern to life insurance companies given the greater predictability of life insurance payouts.

15 B is correct. When a client has a restriction in trading, such as this obligation to refrain from trading, the IPS "should note this constraint so that the portfolio manager does not inadvertently trade the stock on the client's behalf."

16 A is correct. The correlation between US equities and Brazilian equities is 0.76. The correlations between US equities and East Asian equities and the correlation between US equities and European equities both exceed 0.76. Lower correlations indicate a greater degree of separation between asset classes. Therefore, using solely the data given in the table, returns on Brazilian equities are most sharply distinguished from returns on US equities.

17 C is correct. Strategic asset allocation depends on several principles. As stated in the reading, "One principle is that a portfolio's systematic risk accounts for most of its change in value over the long run." A second principle is that, "the returns to groups of like assets… predictably reflect exposures to certain sets of systematic factors." This latter principle establishes that returns on asset classes primarily reflect the systematic risks of the classes.

18 C is correct. As the reading states, "an asset class should contain homogeneous assets… paired correlations of securities would be high within an asset class, but should be lower versus securities in other asset classes."

19 B is correct. Tactical asset allocation allows actual asset allocation to deviate from that of the strategic asset allocation (policy portfolio) of the IPS. Tactical asset allocation attempts to take advantage of temporary dislocations from the market conditions and assumptions that drove the policy portfolio decision.

READING

52

The Behavioral Biases of Individuals

by Michael M. Pompian, CFA

Michael M. Pompian, CFA, is at Sunpointe Investments (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** compare and contrast cognitive errors and emotional biases;
☐	**b.** discuss commonly recognized behavioral biases and their implications for financial decision making;
☐	**c.** describe how behavioral biases of investors can lead to market characteristics that may not be explained by traditional finance;

1

INTRODUCTION AND CATEGORIZATIONS OF BEHAVIORAL BIASES

Research has demonstrated that when people face complex decisions, they often rely on basic judgments and preferences to simplify the situation rather than acting completely rationally. Although such approaches are quick and intuitively appealing, they may lead to suboptimal outcomes. In contrast to this body of research, traditional economic and financial theory generally assumes that individuals act rationally by considering all available information in the decision-making process, leading them to optimal outcomes and supporting the efficiency of markets. Behavioral finance challenges these assumptions by incorporating research on how individuals and markets actually behave. In this reading, we explore a foundational concept of behavioral finance: behavioral biases. Investment professionals may be able to improve economic outcomes by understanding these biases, recognizing them in themselves and others, and learning strategies to mitigate them.

The reading proceeds as follows. Section 2 describes and broadly characterizes behavioral biases. Sections 3 and 4 discuss specific behavioral biases within two broad categories: cognitive errors and emotional biases. The discussion includes a description of each bias, potential consequences, and guidance on detecting and mitigating the effects of the bias. Section 5 discusses market anomalies, which are essentially aggregate expressions of individual biases among financial market participants. A summary and practice problems conclude the reading.

© 2021 CFA Institute. All rights reserved.

2 CATEGORIZATIONS OF BEHAVIORAL BIASES

a compare and contrast cognitive errors and emotional biases

In general, behavioral biases come in two forms: faulty cognitive reasoning, known as **cognitive errors**, and those based on feelings or emotions, known as **emotional biases**. Both forms of bias, regardless of their source, may cause decisions to deviate from what is assumed by traditional finance theory.

Cognitive errors can often be corrected or eliminated through better information, education, and advice. Emotional biases, on the other hand, are harder to correct because they stem from impulses and intuitions. They arise spontaneously rather than through conscious effort and may even be undesired to the individual feeling them. Thus, it is often possible only to recognize an emotional bias and adapt to it. The cognitive–emotional distinction will help us determine when and how to adjust for behavioral biases in financial decision making.

Researchers have identified numerous behavioral biases. This reading does not attempt to catalog all of them. Rather, it discusses some of the more publicized and recognized biases within the cognitive–emotional framework. Additionally, we limit our focus to gauging the presence or absence—not the magnitude—of each bias discussed. That is, we will not try to measure how strongly the bias is exhibited, but rather we will describe the behavioral bias, its potential consequences, and the detection of and correction for the behavioral bias. In detecting a bias, we will identify statements or thought processes that may indicate the bias. Diagnostic tests of varying degrees of complexity are available to detect biases but are beyond the scope of this reading.

Finally, the individuals of interest in this reading are "financial market participants" (FMPs) engaged in financial decision making. These include both individual investors and financial services professionals.

3 COGNITIVE ERRORS

b discuss commonly recognized behavioral biases and their implications for financial decision making

We classify cognitive errors into two categories: "belief perseverance bias" and "processing errors."

Belief perseverance is the tendency to cling to one's previously held beliefs by committing statistical, information-processing, or memory errors. The belief perseverance biases discussed are conservatism, confirmation, representativeness, illusion of control, and hindsight.

Processing errors describe how information may be processed and used illogically or irrationally in financial decision making. The processing errors discussed are anchoring and adjustment, mental accounting, framing, and availability.

3.1 Belief Perseverance Biases

Belief perseverance biases result from the mental discomfort that occurs when new information conflicts with previously held beliefs or cognitions, known as **cognitive dissonance**. To resolve this discomfort, people may ignore or modify conflicting information and consider only information that confirms their existing beliefs or thoughts.

3.1.1 *Conservatism Bias*

Conservatism bias is a belief perseverance bias in which people maintain their prior views or forecasts by inadequately incorporating new, conflicting information. In Bayesian terms, they tend to overweight their prior probability of the event and underweight the new information, resulting in revised beliefs about probabilities and outcomes that underreact to the new information.

Consequences of Conservatism Bias As a result of conservatism bias, FMPs may do the following:

- Maintain or be slow to update a view or a forecast, even when presented with new information; and
- Maintain a prior belief rather than deal with the mental stress of updating beliefs given complex data. This behavior relates to an underlying difficulty in processing new information.

Detection of and Guidance for Overcoming Conservatism Bias The effect of conservatism bias may be corrected for or reduced by properly analyzing and weighting new information. The first step is to be aware that a bias exists, especially about information that is technical, abstract, and/or statistical, because the **cognitive cost** involved in processing those forms of information is higher than for other types.

When new information is presented, the FMP should ask such questions such as, "How does this information change my forecast?" or "What effect does this information have on my forecast?" FMPs should conduct careful analysis incorporating the new information and then respond appropriately. This updating of prior beliefs in light of new information is consistent with the tenets of Bayes' Rule, in which updated probabilities are derived by systematically combining previous estimates and new information.

If information is difficult to interpret or understand, FMPs should seek guidance from someone who can either explain how to interpret the information or can explain its implications.

3.1.2 *Confirmation Bias*

Confirmation bias refers to the tendency to look for and notice what confirms prior beliefs and to ignore or undervalue whatever contradicts them. A response to cognitive dissonance, confirmation bias reflects a predisposition to justify to ourselves what we want to believe.

Most experienced private wealth advisers have dealt with a client who conducts some research and insists on adding a particular investment to the portfolio. The client may insist on continuing to hold the investment, even when the adviser recommends otherwise, because the client's follow-up research seeks only information that confirms his belief that the investment is still a good value. The confirmation bias is not limited to individual investors; all FMPs should be wary of the potential confirmation biases within themselves.

EXAMPLE 1

Confirmation Bias

A portfolio manager at Sarter Investment Advisors recommended shares of Real Media Inc., a hypothetical television production and distribution company, largely on the basis of compelling analytical and valuation work from a top equity research analyst. Sarter's clients have owned the shares for several years.

Recently, the shares have underperformed significantly as a result of the company missing analysts' earnings estimates and also in response to executive management turnover. The portfolio manager's colleagues believe this underperformance is a result of Real Media losing market share to a competitor with superior technology and distribution. The competitor is publicly traded, but Sarter's portfolio managers and analysts have not done research on it.

After another poor earnings release from Real Media, the portfolio manager speaks with the equity research analyst whose work was the primary source of the investment. The equity research analyst, who maintains a buy rating on the stock, believes that Real Media is now a more compelling investment than ever because its share price has fallen while its earnings estimates remain unchanged. As a result of the conversation, the portfolio manager feels reassured and holds the position.

The portfolio manager is subject to confirmation bias. Rather than speaking with a research analyst who has a sell rating on the stock or conducting research on the competitor to consider a different perspective, the portfolio manager speaks with someone who has an opinion she already shares.

Consequences of Confirmation Bias As a result of confirmation bias, FMPs may do the following:

- Consider only the positive information about an existing investment while ignoring any negative information about the investment.
- Develop screening criteria while ignoring information that either refutes the validity of the criteria or supports other criteria. As a result, some good investments that do not meet the screening criteria may be ignored, and conversely, some bad investments that do meet the screening criteria may be made.
- Under-diversify portfolios. FMPs may become convinced of the value of a single company's stock. They ignore negative news, and they gather and process only information confirming that the company is a good investment. They build a larger position than appropriate and hold an under-diversified portfolio.
- Hold a disproportionate amount of their investment assets in their employing company's stock, because they believe in their company and are convinced of its favorable prospects. Favorable information is cited, and unfavorable information is ignored. If the employee were to acknowledge unfavorable information, the associated mental discomfort might make work very difficult for the employee.

Detection of and Guidance for Overcoming Confirmation Bias The effect of confirmation bias may be corrected for or reduced by actively seeking out information that challenges existing beliefs.

Another useful step is to corroborate an investment decision. For example, if investment selections are based on criteria confirming an existing belief, such as stocks breaking through 52-week highs, it is usually advisable to corroborate that decision with research from another perspective or source (e.g., fundamental research on the company, industry, or sector).

3.1.3 *Representativeness Bias*

Representativeness bias refers to the tendency to classify new information based on past experiences and classifications. New information may resemble or seem *representative* of familiar elements already classified, but in reality, it can be very different. In these instances, the classification reflex can deceive, producing an incorrect understanding that biases all future thinking about the information. *Base-rate neglect* and *sample-size neglect* are two types of representativeness bias that apply to FMPs.

In **base-rate neglect**, a phenomenon's rate of incidence in a larger population—its base rate—is neglected in favor of specific information. The specific, individual information may be misleading relative to the more appropriate base rate or general information. FMPs often follow this erroneous path because diligent research is often conducted on an individual security or strategy, leading FMPs to overlook or ignore general information about the "class" to which an investment belongs, such as an industry, sector, or geography.

A second type of representativeness bias is **sample-size neglect**, in which FMPs incorrectly assume that small sample sizes are *representative* of populations. Individuals prone to sample-size neglect are quick to treat properties reflected in small samples as properties that accurately describe large pools of data, overweighting the information in the small sample.

EXAMPLE 2

Representativeness Bias

Jacques Verte is evaluating the future prospects of APM Company, a large auto parts manufacturer having some difficulties. During the last 50 years, very few auto part manufacturers have failed, even during periods of difficulty. A number of recent headlines have highlighted APM's business and financial difficulties, however, with some commentators suggesting that APM may go out of business.

1 Which of the following scenarios is more likely? Explain why.

  A APM will solve its difficulties.

  B APM will go out of business.

2 If Verte is subject to representativeness bias, is he more likely to classify APM into A or B? Explain why.

Solution to 1:

Scenario A is more likely. The base rate, based on 50 years of data, is that more auto parts companies survive difficult times than fail. Thus, it is more likely that APM will solve its difficulties than go out of business.

Solution to 2:

If Verte is subject to representativeness bias, he is likely to choose Scenario B, predicting that the company will go out of business because of the headlines he has read. In classifying APM as likely to go out of business, Verte would be guilty of base-rate neglect by ignoring the low base rate of auto parts manufacturers failing even during times of difficulty.

Consequences of Representativeness Bias As a result of representativeness bias, FMPs may do the following:

- Adopt a view or a forecast based almost exclusively on individual, specific information or a small sample; and
- Update beliefs using simple classifications rather than deal with the mental stress of updating beliefs given the high cognitive costs of complex data.

Detection of and Guidance on Overcoming Representativeness Bias When FMPs sense that base-rate or sample-size neglect may be a problem, they should ask the following question: "What is the probability that X (the investment under consideration) belongs to Group A (the group it resembles or is considered representative of) versus Group B (the group it is statistically more likely to belong to)?" This question, or a similar question, will help FMPs think through whether they are failing to consider base-rate probabilities or neglecting the law of small numbers and thus inaccurately assessing a particular situation. It may be necessary to do more research to obtain base-rate information and/or widen the sample size of observations.

3.1.4 *Illusion of Control Bias*

In **illusion of control bias**, people tend to believe that they can control or influence outcomes when, in fact, they cannot. Many researchers have uncovered situations where people perceived themselves as possessing more control than they did, inferred causal connections where none existed, or displayed surprisingly great certainty in their predictions for the outcomes of chance events. A classic example is that people prefer choosing their own lottery numbers over random numbers selected for them.

EXAMPLE 3

Illusion of Control Bias

Adelia Scott is a wealth adviser at Sarter Investment Advisors (Sarter), an investment advisory firm for high-net-worth individuals. Scott meets with a client who has 30% of his account in shares of his employer's stock. The client is not subject to any employee holding requirement.

Prior meeting notes indicate that the client initially agreed to diversify the concentrated position over a five-year period. Scott recommends a faster schedule, however, based on recent research indicating that the company's future growth prospects have considerably worsened as a result of industry trends and macroeconomic conditions.

When presented with this information, the client is reluctant to change his diversification plan, citing the company's history of double-digit growth and his belief that this rate of growth will continue for the foreseeable future. The client remarks, "Trust me, my team and I are not going to let those forecasts you're citing come true."

The client is subject to illusion of control bias. He is unwilling to believe Scott's opinion because he believes that he and his team can control the company's performance and stock price.

Consequences of Illusion of Control As a result of illusion of control bias, FMPs may do the following:

- Inadequately diversify portfolios. Research has found that some investors prefer to invest in companies that they feel they have control over, such as the companies they work for, leading them to hold concentrated positions. In fact, most investors have little or no control over the companies they work for. If the company performs poorly, these investors may experience both the loss of employment and investment losses.

- Trade more than is prudent. Researchers have found that portfolio turnover is negatively correlated with investment returns.
- Construct financial models and forecasts that are overly detailed. FMPs may require detailed models before making an investment decision, believing that forecasts from these models control uncertainty. Although a greater understanding of an investment, issuer, or industry is often useful, increased model complexity does not control the inherent risk and uncertainty of investment outcomes.

Detection of and Guidelines for Overcoming Illusion of Control Bias The first and most basic idea that investors need to recognize is that investing is a probabilistic activity. Even the largest investment management firms have little control over the outcomes of the investments they make. Companies are subject to macroeconomic and industry forces, as well as the actions of competitors, customers, and suppliers.

Second, it is advisable to seek contrary viewpoints. As you contemplate a new investment, take a moment to think about considerations that might weigh against making the investment. Ask yourself: What are the downside risks? What might go wrong? When will I sell? It is often useful to speak with someone who has an opposing view, such as an equity research analyst with a sell rating on the subject stock.

3.1.5 *Hindsight Bias*

Hindsight bias refers to believing past events as having been predictable and reasonable to expect. This behavior results from the obvious fact that outcomes that did occur are more readily evident than outcomes that did not. Similarly, people tend to remember their own predictions of the future as more accurate than they actually were because they are biased by the knowledge of what actually occurred. Poorly reasoned decisions with positive results may be remembered as brilliant tactical moves, and poor results of well-reasoned decisions may be described as avoidable mistakes.

EXAMPLE 4

Hindsight Bias

Beverly Bolo, an analyst at an investment advisory firm, is giving a presentation to clients that, among other topics, explains the firm's investment results during past macroeconomic downturns. In the presentation, Bolo points out that the "occurrence of the last recession was obvious upon inspection of the yield curve and other leading indicators eight months before the downturn started."

Bolo's comment exhibits hindsight bias. Recessions, like any other event, appear obvious in hindsight but are hardly ever accurately predicted. Bolo could augment her remarks by exploring how often these leading indicators suggested that a recession is imminent against how often a recession subsequently occurred.

Consequences of Hindsight Bias As a result of hindsight bias, FMPs may do the following:

- Overestimate the degree to which they correctly predicted an investment outcome, or the predictability of an outcome generally. This bias is closely related to overconfidence bias, which is discussed later in the reading.
- Unfairly assess money manager or security performance. Based on the ability to look back at what has taken place in securities markets, performance is compared against what has happened as opposed to expectations at the time the investment was made.

Detection of and Guidelines for Overcoming Hindsight Bias Once understood, hindsight bias should be recognizable. FMPs should ask such questions as, "Am I rewriting history or being honest with myself about the mistakes I made?"

To guard against hindsight bias, FMPs should carefully record their investment decisions and key reasons for making those decisions in writing at or around the time the decision is made. Consulting these written records rather than memory will often produce a far more accurate examination of past decisions.

3.2 Processing Errors

Processing errors refer to information being processed and used illogically or irrationally. As opposed to belief perseverance biases, processing errors are less related to errors of memory or in assigning and updating probabilities; rather, they relate more closely to flaws in how information itself is processed.

3.2.1 *Anchoring and Adjustment Bias*

Anchoring and adjustment bias refers to relying on an initial piece of information to make subsequent estimates, judgments, and decisions. When required to estimate a value with unknown magnitude, people generally begin with some initial default number—an "anchor"—which they then adjust up or down. Regardless of how the initial anchor was chosen, people tend to adjust their anchors insufficiently and produce end approximations that are, consequently, biased. Anchoring and adjustment bias is closely related to conservatism bias. Bayes Rule again provides guidance for how new information should be incorporated into changing prior beliefs.

EXAMPLE 5

Anchoring and Adjustment Bias

Aiden Smythe is an equity research analyst at a brokerage firm. Smythe covers Industrial Lift Plc, a company that manufactures construction machinery. The company's business is sensitive to macroeconomic conditions, particularly non-residential construction activity. Last year, Industrial Lift reported £1.00 in EPS amid mostly strong non-residential construction activity levels. Smythe is updating his EPS estimate for this year. Non-residential construction activity has severely declined in the last two months, and some economists fear that a recession is likely. As a result, Smythe forecasts that EPS will fall 10% from the prior-year level, publishing a £0.90 estimate for the year.

Smythe's estimate exhibits anchoring and adjustment. Smythe's anchor is the prior year's EPS of £1.00, despite the possibility of a material change in underlying conditions.

Consequences of Anchoring and Adjustment Bias As a result of anchoring and adjustment bias, FMPs may stick too closely to their original estimates when learning new information. This mindset is not limited to downside adjustments; the same phenomenon occurs with upside adjustments as well.

Detection of and Guidelines for Overcoming Anchoring and Adjustment Bias The primary action FMPs can take is consciously asking questions that may reveal an anchoring and adjustment bias. Examples of such questions include, "Am I holding onto this stock based on rational analysis, or am I trying to attain a price that I am anchored to, such as the purchase price or a high water mark?" and "Am I making this forecast based on previously observed quantities or based on future expected conditions?"

It is important to remember that a company's revenues and earnings for a given period reflect conditions in that period. If the conditions in the future differ from the past, revenues and earnings will likely differ as well, sometimes radically. Similarly, security prices reflect investors' perception of the future at a given point in time; a given investor's cost basis, past market levels, and other conditions based in the past are often irrelevant. FMPs should look at the basis for any investment recommendation to see whether it is anchored to previous estimates or some "default" number.

3.2.2 *Mental Accounting Bias*

Mental accounting bias refers to mentally dividing money into "accounts" that influence decisions, even though money is **fungible** (Thaler 1980). Despite traditional finance theory assuming that FMPs consider their entire portfolio holistically in a risk–return context, Statman (1999, 2008) contends that the difficulty individuals have in addressing the interaction of different investments leads to mental accounting: Investors construct portfolios in a layered pyramid format, with each layer addressing a specific financial goal.

EXAMPLE 6

Mental Accounting Bias

Kendra Liu, an individual investor, owns shares in New Horizons Ltd., a pharmaceutical company. A drug in that company's research and development pipeline unexpectedly succeeds in a clinical trial, resulting in the shares doubling overnight. Liu sells half of her position in New Horizons and uses the proceeds to purchase shares of Cutting Edge Ltd., a small-cap biotechnology company. Liu believes that an investment in Cutting Edge is a "high-risk, high-reward bet" that could result in a total loss, but she is comfortable making the investment because she is using the proceeds from the sale of New Horizons Ltd., money that she did not expect to have anyway.

Liu's investment in Cutting Edge exhibits mental accounting bias. She has sorted the gains from New Horizons into a mental account based on its source, even though the money is fungible. Liu should invest the proceeds in a manner consistent with a holistic portfolio strategy.

Consequences of Mental Accounting Bias As a result of mental accounting bias, FMPs may do the following:

- Neglect opportunities to reduce risk by combining assets with low correlations. Offsetting positions across various portfolio layers or mental accounts can lead to suboptimal aggregate performance.
- Irrationally distinguish between returns derived from income and those derived from capital appreciation. Although many investors feel the need to preserve principal, they focus on the idea of spending income that the principal generates. As a result, many FMPs chase income streams, unwittingly risking principal in the process.
- Irrationally bifurcate wealth or a portfolio into investment principal and investment returns. Some FMPs may believe that greater risk can be taken with returns (from either income or capital appreciation) than the principal initially contributed. A common euphemism for this scenario, from the casino world, is "playing with house money."

Detection of and Guidelines for Overcoming Mental Accounting Bias An effective way to detect and overcome mental accounting behavior is to recognize its drawbacks. The primary drawback is that correlations between investments are not considered, leading to unintentional risk taking. FMPs should go through the exercise of combining all of their assets onto one spreadsheet (without headings or account labels) to see the holistic asset allocation. This exercise often produces information that is surprising when seen as whole, such as higher cash balances than expected. The logical next step would be to create a portfolio strategy taking all assets into consideration.

3.2.3 *Framing Bias*

Framing bias is an information-processing bias in which a person answers a question differently based on the way in which it is asked or framed. It is often possible to frame a given decision problem in more than one way.

For example, a situation may be presented within a *gain* context (one in four start-up companies succeed) or within a *loss* context (three out of four start-ups fail). Given the first frame, an FMP may adopt a positive outlook and make venture capital investments. Given the second frame, the FMP might not.

Narrow framing occurs when people evaluate information based on a *narrow frame* of reference—that is, losing sight of the big picture in favor of one or two specific points. For example, an investor might focus solely on a company's executive management team, overlooking or even dismissing other important properties such as industry characteristics, fundamental performance, and valuation.

EXAMPLE 7

Effects of Framing Bias

Decision-making frames are quite prevalent in the context of investor behavior. Risk tolerance questionnaires can demonstrate how framing bias may occur in practice and how FMPs should be aware of its effects.

Suppose an investor is to take a risk tolerance questionnaire for the purpose of determining which "risk category" she is in. The risk category will determine asset allocations and the appropriate types of investments. The following information is provided to each questionnaire taker:

> Over a 10-year period, Portfolio ABC has averaged an annual return of 10% with an annual standard deviation of 16%. Assuming a normal return distribution, in a given year there is a 67% probability that the return will fall within one standard deviation of the mean, a 95% probability that the return will fall within two standard deviations of the mean, and a 99.7% probability that the return will fall within three standard deviations of the mean. Thus, there is a 67% chance that the return earned by Portfolio ABC will be between −6% and 26%, a 95% chance that the return will be between −22% and 42%, and a 99.7% chance that the return will be between −38% and 58%.

The following two questions focus on hypothetical Portfolio ABC, DEF, and XYZ. The risk and return for each portfolio are the same in each of the two questions, but the presentation of information differs.

1 Based on the following chart, which investment portfolio fits your risk tolerance and desire for long-term return?

Portfolio	95% Probability Return Range	10-Year Average Return
XYZ	0.5% to 6.5%	3.5%
DEF	−18.0% to 30.0%	6.0%
ABC	−22.0% to 42.0%	10.0%

2 Based on the following chart, which investment portfolio fits your risk tolerance and desire for long-term return?

Portfolio	10-Year Average Return	Standard Deviation of Returns
XYZ	3.5%	1.5%
DEF	6.0%	12.0%
ABC	10.0%	16.0%

An investor may choose different portfolios when asked Question 1 compared with Question 2. Portfolio XYZ may appear more attractive in the first question, which uses two standard deviations to define the range of returns and show the risk, than in the second, which shows only the standard deviations. Also, in the second question, the returns are presented first and the measure of risk second. Thus, how questions are framed and the order in which they are presented can significantly affect how they are answered. FMPs should be acutely aware of how framing can affect investment choices.

Consequences of Framing Bias As a result of framing bias, FMPs may do the following:

- Misidentify risk tolerances because of how questions about risk tolerance were framed, becoming more risk-averse when presented with a gain frame of reference and more risk-seeking when presented with a loss frame of reference. This misidentification may result in suboptimal portfolios.
- Focus on short-term price fluctuations, which may result in long-run considerations being ignored in the decision-making process.

Detection of and Guidelines for Overcoming Framing Bias Framing bias is detected by asking such questions as, “Is the decision the result of focusing on a net gain or net loss position?” As discussed earlier, an investor who has framed the decision as a potential net loss is more likely to select a riskier investment; if the decision is framed as a potential net gain, however, the investor is more likely to go with a less risky investment. When making decisions, FMPs should try to eliminate any reference to gains and losses already incurred. Instead, they should focus on the future prospects of an investment and try to be as neutral and open-minded as possible when interpreting investment-related situations.

3.2.4 *Availability Bias*

Availability bias is an information-processing bias in which people estimate the probability of an outcome or the importance of a phenomenon based on how easily information is recalled. Various sources of availability bias exist; the four most applicable to FMPs are retrievability, categorization, narrow range of experience, and resonance.

Retrievability: If an answer or idea comes to mind more quickly than another answer or idea, the first answer or idea will likely be chosen as correct even if it is not the reality.

Categorization: When solving problems, people gather information from what they perceive as relevant search sets. Different problems require different search sets, which are often based on familiar categorizations. If it is difficult to come up with a search set, because the object of the search is difficult to characterize, the estimated probability of an event may be biased.

Narrow Range of Experience: When making an estimate, a person may use only a narrow range of experience instead of considering multiple perspectives. For example, assuming a product or service that has launched successfully in one country will be globally successful.

Resonance: People are often biased by how closely a situation parallels their own personal situation.

EXAMPLE 8

Availability Bias

A portfolio manager asks an analyst to research and present a list of companies that have "strong growth potential." The manager suggests looking for companies that sell a product or service different from its competitors—but with a compelling value proposition for customers—and that have a small share of a large market.

The analyst is familiar with technology companies, software in particular, based on prior work experience. The analyst has also seen a lot of news articles covering various software companies that, he believes, fit the criteria. The analyst begins screening among technology companies that have high revenue growth rates for the last two quarters. Although the analyst is aware that other companies in other sectors probably fit the criteria as well, the criteria are qualitative and vague such that they cannot be easily translated as screening input.

The analyst's behavior exhibits availability bias from by considering only technology companies in the search because he is familiar with them. The analyst should consult colleagues and/or external resources to widen his search to include all sectors and for help with creatively specifying screening criteria.

Consequences of Availability Bias As a result of availability bias, FMPs may do the following:

- Limit their investment opportunity set. This limitation may result because they use familiar classification schemes. They may restrict investments to stocks or bonds, securities of one country or one sector, and so on.
- Choose an investment, investment adviser, or mutual fund based on advertising or the quantity of news coverage. For instance, when asked to name potential mutual fund companies to invest with, many people will name only the funds that do extensive advertising. The choice of mutual fund should be based on a variety of factors that make it a good fit given the investor's objectives and risk–return profile.
- Fail to diversify. This failure may occur because they make their choices based on a narrow range of experience. For example, an investor who works for a company in a particular industry may overweight investments in that industry.

Detection of and Guidelines for Overcoming Availability Bias To overcome availability bias, investors need to develop an appropriate investment policy strategy, carefully research and analyze investment decisions before making them, and focus on long-term historical data. Questions such as, "How did you decide which investments to consider? Did you choose investments based on your familiarity with the industry or country? Did you see them in an article or research report?" and "Did you choose your investments because you like the companies' products?" can help identify issues of categorization, narrow range of experience, and resonance as sources of availability bias. Availability bias may cause FMPs to think that events that receive heavy media attention are more important than they actually are.

EXAMPLE 9

The following information relates to Questions 1–5

Luca Gerber recently became the chief investment officer for the Ludwigs Family Charity, a mid-size private foundation in Switzerland. Prior to assuming this role, Gerber was a well-known health care industry analyst. The Ludwigs's family fortune is primarily the result of entrepreneurship. Gerhard Ludwigs founded ABC Innovations (ABC), a biotech company dedicated to small cell lung cancer research. The foundation's portfolio is 15% invested in ABC.

Gerber initially feels that the 15% investment in ABC is high. Upon review, however, he decides it is appropriate based on the Ludwigs's involvement and their past success with similar ventures. Gerber makes a mental note to closely monitor the investment in ABC because he is unfamiliar with small-cap start-up companies. The remaining 85% of the foundation's portfolio is invested in equity of high-quality large-cap pharmaceutical companies. Gerber deems this allocation appropriate and is excited that he can continue to use his superior knowledge of the health care industry.

For the past two years, ABC has been dedicated to Project M, an effort directed at developing a drug to treat relapses in small cell lung cancer. Project M has delayed its Phase Two trials twice. Published results from Phase One trials have raised some concerns regarding the drug. In its last two quarterly investors' conference calls, ABC's CEO was very cautious in discussing expectations for Project M. ABC's stock price decreased by more than 20% during the past six months. Gerber believes that the research setbacks are temporary because of ABC's past success with projects. He expects that ABC will begin Phase Two within a year, and he also believes that once Project M goes into Phase Two, ABC's stock price should reach a new 52-week high of CHF80.

Soon after deciding to hold the stock, Gerber reads an article by ABC's chief scientist that details certain scientific results from Project M. As a conclusion, the article states: "Although we still have some major obstacles to overcome, the Project M team has seen positive signs that a treatment for small cell lung cancer is achievable." Although Gerber has difficulty interpreting the scientific results, he feels reassured after reading the concluding statement.

Today, ABC announces the news that it will no longer pursue Project M, citing early signs of the project's failure. In response to the announcement, the stock price drops by 50%. Gerber is stunned. He reviews the company's history and notes that ABC had made numerous comments on its struggles to solve Project M issues, which make the failure seem predictable at the time.

1 Gerber's assessment of the foundation's 100% allocation to biotechnology and pharmaceutical companies is most likely an example of which bias?

A Framing

B Availability

C Hindsight

2 Gerber's belief that the research setbacks were merely temporary given ABC's success with past projects is most likely an example of which bias?

A Availability

B Mental accounting

C Conservatism

3 Researching the biotechnology industry average probability of success of Phase 2 trials, particularly those that have experienced delays, is a strategy that Gerber could have used to most likely mitigate which behavioral bias?

A Hindsight

B Representativeness

C Framing

4 Gerber's approach to reading the article by ABC's chief scientist could best be described by exhibiting which behavioral bias?

A Representativeness

B Confirmation

C Mental accounting

5 Gerber's conclusion, upon re-examining ABC's history, is most likely an example of which behavioral bias?

A Confirmation

B Hindsight

C Conservatism

Solution:

1 B is correct. A consequence of availability bias is a limited investment opportunity set, based on the narrow range of experience of the FMP. Gerber's prior experience has likely resulted in perceiving a portfolio concentrated in a single sector as appropriate because that sector reflects what he knows.

2 C is correct. Conservatism bias can result in FMPs maintaining or only slowly updating their views when presented with new information, especially when the information is complex. Faced with complex information in the form of clinical trial delays that should have reduced his assessment of Project M's probability of success, Gerber elected to maintain his original views.

3 B is correct. A form of representativeness bias is base-rate neglect, in which general information, such as the rate of incidence of a phenomenon for its "reference class" of phenomena, is ignored in favor of individual, specific information. Gerber did not consider the "base rate" of the success of Phase 2 trials.

4 B is correct. Confirmation bias refers to the tendency to seek confirming information and ignore contradictory information. Gerber did not adequately interpret the scientific results or the broader message of the article, choosing instead to focus on the reassuring concluding message that confirmed his existing beliefs.

5 B is correct. Hindsight bias is the result of selectively interpreting the past using knowledge of the present. Although Gerber should have implemented various mitigation strategies over the life of the investment, it is also true that investing, especially investing in a small-cap biotechnology company, is a probabilistic activity. Had Project M succeeded, Gerber may have been tempted to see the investment as evidence of his investment acumen, which would also be biased. The most useful examination of the

past for Gerber would be investigating whether any mitigating actions or strategies would have been useful and putting them in place for subsequent investments.

EMOTIONAL BIASES

4

b discuss commonly recognized behavioral biases and their implications for financial decision making

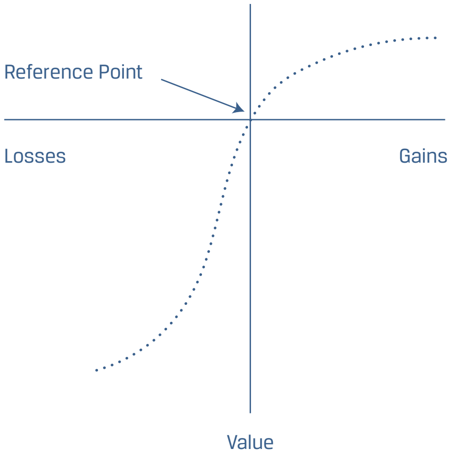
We will now review six emotional biases, their implications for investment decision making, and suggestions for managing the effects of these biases. Emotional biases are harder to correct than cognitive errors because they originate from impulse or intuition rather than conscious calculations. It is often possible only to recognize the bias and adapt to it. The six emotional biases discussed are loss aversion, overconfidence, self-control, status quo, endowment, and regret aversion.

4.1 Loss-Aversion Bias

Loss-aversion bias refers to the tendency to strongly prefer avoiding losses to achieving gains. Rational FMPs should accept more risk to increase gains, not to mitigate losses. Paradoxically, in real life, FMPs instead tend to accept more risk to avoid losses than to achieve gains. Loss aversion leads FMPs to hold their losers to avoid recognizing losses and sell their winners to lock in profits

Kahneman and Tversky (1979) describe loss-averse investor behavior as the evaluation of gains and losses based on a reference point. A utility function that passes through this reference point appears in Exhibit 1. It is S-shaped and asymmetric, implying a greater impact of losses than of gains on utility for the same variation in absolute value. This utility function implies risk-seeking behavior in the domain of losses (below the horizontal axis) and risk avoidance in the domain of gains (above the horizontal axis). An important concept embedded in this utility representation is what has been termed the **disposition effect**: the holding of investments that have experienced losses too long, and the selling of investments that have experienced gains too quickly (i.e., holding on to losers and selling winners). The resulting portfolio may be riskier than the optimal portfolio based on the investor's risk–return objectives.

Exhibit 1 Value Function of Loss Aversion



EXAMPLE 10

Effects of Loss-Aversion Bias

Loss-aversion bias, executed in practice as the disposition effect, is observed often by wealth management practitioners. The classic case of this bias is when an investor opens her monthly account statement and scans the individual investments for winners and losers. Seeing that some investments have lost money and others have gained, the investor is likely to respond by continuing to hold the losing investments. The idea of actually losing money is so painful that the first reaction is to hold the investment until it breaks even.

The investor is acting based on emotions, not cognitive reasoning. In this case, if she did some research, the investor might learn that the company in question is experiencing difficulty and that holding the investment actually adds to the risk in the portfolio (hence the term risk-seeking in the domain of losses).

Conversely, the winners are making money. Loss-averse FMPs tend to sell these investments and realize their gains to avoid seeing gains evaporate. In this case, if the investor did some research, she might learn that the company in question actually improves the portfolio's risk–return profile. By selling the investment, not only is the potential for future losses eliminated but the potential for future gains is also eliminated. Combining the added risk of holding the losers with the elimination of potential gains from selling the winners may make investors' portfolios less efficient than portfolios based on fundamental analysis.

Consequences of Loss Aversion As a result of loss-aversion bias, FMPs may do the following:

- Hold investments in a loss position longer than justified by fundamental analysis, in the hope that they will return to breakeven.
- Sell investments in a gain position earlier than justified by fundamental analysis, out of fear that the gains will erode.

Detection of and Guidelines for Overcoming Loss Aversion A disciplined approach to investment is a good way to alleviate the impact of the loss-aversion bias. It is impossible to make experiencing losses any less painful emotionally, but analyzing investments and realistically considering the probabilities of future losses and gains may help guide the FMP to a rational decision.

4.2 Overconfidence Bias

Overconfidence bias is a bias in which people demonstrate unwarranted faith in their own abilities. Overconfidence may be intensified when combined with **self-attribution bias**, in which people take too much credit for successes (*self-enhancing*) and assign responsibility to others for failures (*self-protecting*). Overconfidence has aspects of both cognitive and emotional errors but is classified as emotional because the bias primarily results from emotion.

Overconfidence bias has two forms: *prediction overconfidence* and *certainty overconfidence*. Both types demonstrate faulty reasoning combined with "gut feel" and hope. Prediction overconfidence occurs when the confidence intervals that FMPs assign to their investment predictions are too narrow. For example, when estimating the future value of a stock, overconfident FMPs will incorporate far too little variation—using a narrower range of expected payoffs and a lower standard deviation of returns—than justified by historical results and fundamental analysis.

Certainty overconfidence occurs when the probabilities that FMPs assign to outcomes are too high. This certainty is often an emotional response rather than a cognitive evaluation.

EXAMPLE 11

Overconfidence Bias

An analyst estimates that the price of oil will increase by 40% over the next 12 months because prevailing prices are lower than many oil producers' cost of production. Unprofitable producers reducing production will eventually put upward pressure on prices so long as oil demand remains stable or increases. Based on this forecast, the analyst recommends several high-yield bonds of oil producers to a portfolio manager.

The portfolio manager asks the analyst for an estimate of downside risk: "How much could we lose if, say, the oil price falls another 10%?" The analyst replies, "That is unrealistic. The current oil price is as low as it can go, and yields on these bonds are as attractive as they will ever be. We must make the investment now. There is no credible downside case."

The analyst is exhibiting overconfidence bias by placing excessive certainty in his prediction and not considering the likelihood or impact of variance from that prediction.

Consequences of Overconfidence Bias As a result of overconfidence bias, FMPs may do the following:

- Underestimate risks and overestimate expected returns.
- Hold poorly diversified portfolios, which may result in significant downside risk.

Detection of and Guidelines for Overcoming Overconfidence Bias FMPs should review their trading records, identify *both* the winners and losers, and calculate portfolio performance over at least two years. Investors with an unfounded belief in their own ability to identify good investments may recall winners and their results but underestimate the number and results of their losers. This review will also identify the amount of trading. Overconfidence is also a cognitive error, and so more-complete information can often help FMPs understand the error of their ways.

It is critical that investors be objective when making and evaluating investment decisions. As an old Wall Street adage states, "Don't confuse brains with a bull market." It is advisable to view the reasoning behind and the results of investments, both winners and losers, as objectively as possible. When did you make money? When did you lose money? Mentally separate your good money-making decisions from your bad ones. Then, review the beneficial decisions and try to discern what, exactly, you did correctly. Did you purchase an investment at a particularly advantageous time based on fundamentals, or did you luck out by timing a market upswing? Similarly, review the decisions that you categorized as poor. Did you invest aptly based on fundamentals and then make an error when it came time to sell, or was the market going through a correction?

When reviewing unprofitable decisions, look for patterns or common mistakes that perhaps you were unaware you were making. Note any such tendencies that you discover, and try to remain mindful of them by brainstorming a rule or reminder such as: "I will do X in the future" or "I will not do Y in the future."

4.3 Self-Control Bias

Self-control bias is a bias in which people fail to act in pursuit of their long-term, overarching goals in favor of short-term satisfaction. For example, individuals pursuing the CFA charter may fail to study sufficiently because of short-term competing demands on their time. Rational behavior suggests that people would do whatever is necessary to achieve their long-term goals, but it often does not happen.

When it comes to money, many people are willing and able to save for the future, but they often have difficulty sacrificing present consumption because of a lack of self-control. This apparent lack of self-control may also be a function of hyperbolic discounting, the human tendency to prefer small payoffs now compared with larger payoffs in the future. Sacrifices in the present require much greater payoffs in the future; otherwise, people will be unwilling to make current sacrifices.

Consequences of Self-Control Bias As a result of self-control bias, FMPs may do the following:

- Save insufficiently for the future, which may, in turn, result in accepting too much risk in portfolios in an attempt to generate higher returns.
- Borrow excessively to finance present consumption.

Detection of and Guidelines for Overcoming Self-Control Bias FMPs should ensure that a proper investment plan is in place and should have a personal budget. Plans need to be in writing, so that they can be reviewed regularly. Similarly, FMPs should look to maintain a strategic asset allocation based on a thorough evaluation.

4.4 Status Quo Bias

Status quo bias is an emotional bias in which people choose to do nothing (i.e., maintain the "status quo") instead of making a change, even when change is warranted. Status quo bias is often discussed in tandem with endowment and regret-aversion biases because an outcome of the biases, maintaining existing positions, is similar. The reasons for maintaining the existing positions differ, however. In the status quo bias, positions are maintained largely because of inertia rather than conscious choice. In endowment and regret-aversion biases, positions are maintained because of conscious, but possibly incorrect, choices.

EXAMPLE 12

The Path of Least Resistance

Using data from three firms, Choi et al. (2001) studied the impact of automatically enrolling employees in a defined contribution pension plan and how default contribution rates and investment options affect participants' behavior.

Automatic enrollment increased employee participation in the defined contribution plan from 26%–43% after six months' tenure and 57%–69% after three years' tenure to >85% for both tenures at all three firms.

Although automatic enrollment increased participation, more than 65% of employees tended to contribute the employer-specified default amount—2% or 3%—and remained in the default investment option. Although this percentage declined slowly over time, even after two years of tenure, more than 40% of participants continued to use the default.

Consequences of Status Quo Bias As a result of status quo bias, FMPs may do the following:

- Unknowingly maintain portfolios with risk characteristics that are inappropriate for their circumstances.
- Fail to explore other opportunities.

Detection of and Guidelines for Overcoming Status Quo Bias Status quo bias may be exceptionally strong and difficult to overcome. Education is essential. FMPs should quantify the risk-reducing and return-enhancing advantages of diversification and proper asset allocation. For example, with a concentrated stock position, showing what can happen to overall wealth levels if the stock collapses may persuade an FMP to diversify.

4.5 Endowment Bias

Endowment bias is an emotional bias in which people value an asset more when they own it than when they do not. Endowment bias is inconsistent with standard economic theory, which asserts that the price a person is *willing to pay* for a good should equal the price at which that person would be *willing to sell* the same good. Psychologists have found, however, that people tend to state minimum selling prices for a good that exceed maximum purchase prices that they are willing to pay for the same good. Effectively, ownership "endows" the asset with added value.

Endowment bias may be the result of several other behavioral biases, such as loss aversion, anchoring and adjustment, and overconfidence. Despite the name, purchased as well as inherited securities can be subject to endowment bias.

EXAMPLE 13

Endowment Bias

Several of an investment analyst's recommended stocks have done well for the past five years, prompting the portfolio manager to ask for a brief update on each, including valuations. For each stock, the analyst estimates that fair value is at least another 40% above the current price. The portfolio manager challenges the analyst by pointing out that the fair value estimates imply valuation multiples that are at least two standard deviations above the five-year average and are well above even the most bullish sell-side analyst's target price. The analyst responds by saying that the market is overlooking these companies' fundamentals. The portfolio manager then asks, "Would you buy these shares today?" The analyst answers, "Probably not."

The analyst is likely exhibiting endowment bias by overestimating the value of shares already owned in the portfolio. This bias is likely the result of having successfully invested in the shares.

Consequences of Endowment Bias Endowment bias may lead FMPs to do the following:

- Fail to sell certain assets and replace them with other assets.
- Continue to hold classes of assets with which they are familiar. FMPs may believe they understand the characteristics of the investments they already own and may be reluctant to purchase assets with which they have less experience. Familiarity adds to owners' perceived value of a security.
- As a result, the FMP may maintain an inappropriate asset allocation. The portfolio may be inappropriate for investors' levels of risk tolerance and financial goals.

Detection of and Guidelines for Overcoming Endowment Bias Many wealth management practitioners have encountered clients who are reluctant to sell securities bequeathed by previous generations. Often in these situations, investors cite feelings of disloyalty associated with the prospect of selling inherited securities, general uncertainty in determining the right choice, and concerns about tax issues. An FMP should ask, "If an equivalent sum to the value of the investments inherited had been received in cash, how would you invest the cash?" Often, the answer is into a very different investment portfolio than the one inherited. It may also be useful to explore the deceased's intent in owning the investment and bequeathing it. "Was the primary intent to leave the specific investment portfolio because it was perceived to be a suitable investment based on fundamental analysis, or was it to leave financial resources to benefit the heirs?" Heirs who affirm the latter conclusion are receptive to considering alternative asset allocations.

An effective way to address endowment bias for purchased securities, when an estimated "sell price" is far higher than any reasonable FMP's estimate of a "buy price" is to ask, "Would you buy this security today at the current price?" A similar question is, "Why are you not buying more of this security today?" Answering these questions can turn the focus away from the past to the present, toward considering the upside from the current price.

4.6 Regret-Aversion Bias

Regret-aversion bias is an emotional bias in which people tend to avoid making decisions out of fear that the decision will turn out poorly. Regret-aversion bias has two dimensions: actions that people take and actions that people *could have* taken. Regret is more intense when the unfavorable outcomes are the result of an action taken versus the result of an action not taken. Thus, no action becomes the default decision.

Consequences of Regret-Aversion Bias As a result of regret-aversion bias, FMPs may do the following:

- Be too conservative in their investment choices as a result of poor outcomes on risky investments in the past. FMPs may wish to avoid the regret of making another bad investment and decide that low-risk instruments are better. This behavior can lead to long-term underperformance and failure to reach investment goals.
- Engage in herding behavior. FMPs may feel safer in popular investments in order to limit potential future regret. It seems safe to be with the crowd, and a reduction in potential emotional pain is perceived. Regret aversion may lead to preference for stocks of well-known companies even in the face of equal risk and return expectations. Choosing the stocks of less-familiar companies is perceived as riskier and involves more personal responsibility and greater potential for regret. As John Maynard Keynes (1936) wrote, "Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally."

Detection of and Guidelines for Overcoming Regret-Aversion Bias FMPs should quantify the risk-reducing and return-enhancing advantages of diversification and proper asset allocation. Regret aversion can cause some FMPs to invest too conservatively or too riskily depending on the current trends. With proper diversification, FMPs will accept the appropriate level of risk in their portfolios depending, of course, on return objectives. To prevent investments from being too conservative, FMPs must recognize that losses happen to everyone and keep in mind the long-term benefits of including risky assets in portfolios. Recognizing that bubbles happen and keeping in mind long-term objectives will prevent a client from making investments that are too risky.

EXAMPLE 14

The following information relates to Questions 6–12

Tiffany Jordan is a hedge fund manager with a history of outstanding performance. For the past 10 years, Jordan's fund has used an equity market–neutral strategy (a long–short strategy that strives to eliminate market risk—i.e., beta should be zero), which has proved effective as a result of Jordan's hard work. An equity market–neutral strategy normally generates large daily trading volume and shifts in individual security positions. Jordan's reputation has grown over the years as her fund has consistently beaten its benchmark. Employee turnover on her team has been high; Jordan tends to be quick to blame and rarely gives credit to team members for success. During the past 12 months, her fund has been significantly underperforming its benchmark.

One of Jordan's junior analysts, Jeremy Tang, is concerned about the underperformance and notes the following:

Observation 1: Certain positions are significantly underwater, have very high risk profiles, and have been held for much longer than normal.

Observation 2: The fund's trading volume has decreased by more than 40% during the past year.

Observation 3: The portfolio is more concentrated in a few sectors than in the past.

Worried that the portfolio may be in violation of the fund's investment policy statement (IPS), Tang brings this concern to Jordan's attention during a regular weekly team meeting. Jordan dismisses Tang's analysis and tells the team not to worry because she knows what she is doing. Jordan indicates that because she believes the pricing misalignment will correct itself, the portfolio would be unable to take advantage of the reversion to the mean if she were to sell certain losing positions. She reassures the team that this strategy has performed well in the past and that the markets will revert, bringing the fund's returns back to normal levels.

Tang tactfully suggests that the team review the fund's IPS together, and Jordan interrupts him and reminds the team that she has memorized the IPS by heart. Tang contemplates his next step. He is concerned that Jordan is displaying behavioral biases that are affecting the fund's performance.

1 By taking credit for successes but assigning blame for failures, Jordan is *most likely* demonstrating:

A loss-aversion bias.

B self-attribution bias.

C illusion of control bias.

2 Which of Tang's observations is *least likely* to be the consequence of Jordan demonstrating loss-aversion bias?

A Observation 1

B Observation 2

C Observation 3

3 Which of Jordan's actions *least* supports that she may be affected by the illusion of control bias?

A Her dismissal of Tang's analysis

B Her routine of holding weekly team meetings

C Her comment on market turnaround and current holdings

4 How does Jordan *most likely* demonstrate loss-aversion bias?

A Telling the team not to worry

B Reducing the portfolio turnover this year

C Deciding to hold the losing positions until they turn around

5 Which of the following emotional biases has Jordan *most likely* exhibited?

A Endowment

B Regret aversion

C Overconfidence

6 Which of the following biases did Jordan *not* demonstrate?

A Self-attribution

B Representativeness

C Illusion of knowledge

7 Which of Tang's findings is *not* a typical consequence of self-control bias?

A Failure to explore other portfolio opportunities

B Asset allocation imbalance problems in the portfolio

C A higher risk profile in the portfolio resulting from pursuit of higher returns

Solution:

1 B is correct. In self-attribution bias, people take credit for successes and assign responsibilities for failure. Jordan claims successful decisions for herself while attributing poor decisions to the team. Her self-esteem affects how she looks at success and failure. Self-attribution and illusion of knowledge biases contribute to overconfidence bias, which Jordan clearly demonstrates later when she tells the team that she knows what she is doing.

2 C is correct. Loss aversion by itself may cause a sector concentration; however, a market-neutral strategy tends to focus on individual stocks without regard to sector. The sector exposure would be mitigated with the balancing of the individual long and short positions.

3 B is correct. Holding weekly team meetings, which indicates a willingness to listen to feedback from others, is not representative of the illusion of control bias. In the illusion of control bias, people believe they can control outcomes to a greater extent than is possible. Individuals exhibiting this bias display great certainty in their predictions of outcomes of chance events and ignore others' viewpoints. Jordan is sure that the market will turn around even though it is out of her control. She chooses not to listen to Tang, who is questioning her viewpoint.

4 C is correct. Jordan's behavior is a classic example of loss aversion: When a loss occurs, she holds on to these positions longer than warranted. By doing so, Jordan has accepted more risk in the portfolio. In loss-aversion bias, people exhibit a strong preference to avoid losses versus achieving gains. One of the consequences of loss aversion bias is that the financial management professional (in this case, Jordan) may hold losing investments in the hope that they will return to breakeven or better.

5 C is correct. Jordan exhibits overconfidence in several ways. She ignores the analysis done by Tang. This may be because Jordan believes she is smarter and more informed than her team members, which is typical of an individual with an illusion of knowledge bias. The certainty she demonstrates that the market will revert is evidence of overconfidence. Her overconfidence is intensified by her self-attribution bias, which is demonstrated through her dealings with her team when she blames them for losses while taking credit for gains. Finally, her portfolio's underperformance against the benchmark is a consequence of overconfidence bias.

6 B is correct. Nowhere in the scenario did it mention that Jordan classified certain information into a personalized category. Representativeness bias is a cognitive bias in which people tend to classify new information based on past experiences and classifications. Jordan is not relating her certainty about the future or her decision to hold losing positions back to anything she has done or experienced in the past.

7 A is correct. Failing to explore other opportunities is a demonstration of status quo bias, not self-control. Self-control bias occurs when individuals deviate from their long-term goals—in this case, the IPS—because of a lack of self-discipline. Jordan is not adhering to the strategy that has succeeded in the past. The consequences of self-control bias include accepting too much risk in the portfolio (C) and asset allocation imbalance problems (B) as Jordan attempts to generate higher returns.

5 HOW BEHAVIORAL FINANCE INFLUENCES MARKET BEHAVIOR

c describe how behavioral biases of investors can lead to market characteristics that may not be explained by traditional finance

Some persistent market patterns run counter to market efficiency. This section focuses on the contributions of behavioral finance to understanding these exceptions to market efficiency, such as momentum, value, bubbles, and crashes, by explaining them as functions of behavioral biases.

5.1 Defining Market Anomalies

Anomalies are apparent deviations from the efficient market hypothesis, identified by persistent abnormal returns that differ from zero and are predictable in direction. Not every deviation is anomalous. Misclassifications tend to stem from three sources: choice of asset pricing model, statistical issues, and temporary disequilibria.

Classifying returns as “abnormal” presupposes a definition of “normal returns,” which generally depends on the asset pricing model used. If a reasonable change in the method of estimating normal returns causes an anomaly to disappear, then it is reasonable to suggest that the anomaly is an illusion. Fama (1998) includes in this category apparently low returns following initial public offerings (called the “IPO puzzle”) and the positive abnormal returns apparent in the 12 months after a stock split. Similarly, when high returns persist on a particular class of securities, or relative to a specific factor in valuation, it might simply be compensation for excess risk rather than an anomaly.

Other apparent anomalies may be explained by the small samples involved, a statistical bias in selection or survivorship, or data mining that overanalyzes data for patterns and treats spurious correlations as relevant. The magnitude of any over- or underperformance also depends critically on the choice of benchmark, which can make it hard to interpret results.

Finally, from time to time, markets can present *temporary disequilibrium behavior*, unusual features that may survive for a period of years but ultimately disappear. Publication of the anomaly, which draws attention to the pattern, usually starts the arbitrage that removes the behavior. For example, the small company January effect, part of the turn-of-the-year effect, does not appear persistent once appropriate adjustment for risk is made. The weekend effect, involving lower stock market returns on Mondays, appears to have diminished in the United States and United Kingdom.

The anomalies discussed in this section reflect behavior that has been identified and analyzed in a number of markets around the world and during different periods. The patterns have been documented in many academic studies, with broadly similar conclusions. Behavioral finance can help provide good explanations by identifying underlying behavioral biases.

5.2 Momentum

Studies have documented, in a range of markets globally, *momentum* or *trending effects* in which future price behavior correlates with that of the recent past (Jegadeesh and Titman 1993; Dimson, Marsh, and Staunton 2008). The positive correlation typically lasts for up to two years before showing a reversal or reversion to the mean, evident in two- to five-year return periods.

EXAMPLE 15 THE MOMENTUM EFFECT: LONDON BUSINESS SCHOOL STUDY

The study involves buying the top 20% of a performance-ranked list of stocks and selling short the bottom 20%. In the 52 years to 2007 in the UK market, the stocks that had outperformed the market most in the previous 12 months went on to generate an annualized return of 18.3%, whereas the market's worst underperformers rose by 6.8% on average. During that period, the market as a whole rose by 13.5% a year. In a subsequent study using data from 2000 to 2007, the momentum effect was also evident in each of the 16 other international markets researched.

The authors noted, "The momentum effect, both in the United Kingdom and globally, has been pervasive and persistent. Though costly to implement on a standalone basis, all investors need to be acutely aware of momentum. Even if they do not set out to exploit it, momentum is likely to be an important determinant of their investment performance."

Source: Dimson, Marsh, and Staunton (2008).

Momentum can be partly explained by availability, hindsight, and loss aversion biases.

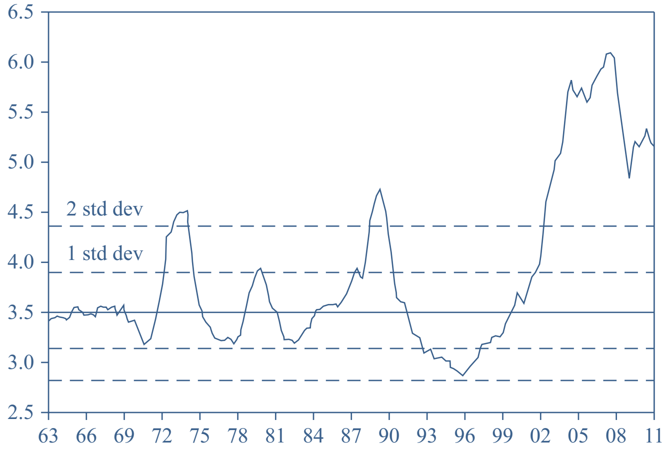
Studies have identified faulty learning models within traders, in which reasoning is based on their recent experience. Behaviorally, this is *availability bias*. In this context, availability bias is also called the *recency effect*, which is the tendency to recall recent events more vividly and give them undue weight. In such models, if the price of an asset rises for a period of time, investors may simply extrapolate this rise to the future. Research points to a tendency for individual private investors to extrapolate trends and to suffer more from recency bias, whereas many investment professionals expect reversion to the mean.

Regret is the feeling that an opportunity has been missed and is typically an expression of *hindsight bias*, which reflects the human tendency to see past events as having been predictable. Regret can be particularly acute when the market is volatile and investors feel they could have predicted the significant market moves, thereby increasing profit or reducing loss. Faced with regret from not owning a mutual fund or stock when it performed well in the previous year, investors may be driven emotionally to remedy this regret. These behavioral factors can explain short-term year-on-year trending and contribute to overtrading.

5.3 Bubbles and Crashes

Although bubbles and crashes have been documented for a long time (Mackay's *Extraordinary Popular Delusions and the Madness of Crowds* was published in 1841), their existence presents a challenge to the concept of market efficiency. Historical examples include the technology bubble of 1999–2000 and the residential property boom of 2005–2007, evident in a range of economies globally including the United States, the United Kingdom, and Australia. Exhibit 3 illustrates the residential property boom in the United Kingdom.

Exhibit 2 UK House Price Average Multiple of Average Family Income



Source: Datastream.

First, note that some bubbles may have rational explanations. Rational investors may expect a future crash but not know its exact timing. For periods of time, there may not be effective arbitrage because of the cost of selling short, unwillingness of investors to bear extended losses, or simply unavailability of suitable instruments. These factors were considerations in past technology and real estate bubbles. Investment managers incentivized on, or accountable for, short-term performance may even rationalize their participation in the bubble in terms of commercial or career risk.

Investors' behaviors and incentives during bubbles are illustrated in Exhibit 4. The manager of Fund A believed he could exit a bubble profitably by selling near the top. The manager of Fund B correctly avoided the bubble, but clients held the manager accountable for short-run performance, which resulted in the fund's closure.

EXAMPLE 16 INVESTOR BEHAVIOR IN BUBBLES

Consider the differing behavior of two managers of major hedge funds during the technology stock bubble of 1998–2000:

The manager of Hedge Fund A was asked why he did not get out of internet stocks earlier even though he knew by December 1999 that technology stocks were overvalued. He replied, "We thought it was the eighth inning, and it was the ninth. I did not think the NASDAQ composite would go down 33% in 15 days." Faced with losses, and despite his previous strong 12-year record, he resigned as Hedge Fund A's manager in April 2000.

The manager of Hedge Fund B refused to invest in technology stocks in 1998 and 1999 because he thought they were overvalued. After strong performance over 17 years, Hedge Fund B was dissolved in 2000 because its returns could not keep up with the returns generated by technology stocks.

Behavioral finance does not yet provide a full explanation for bubbles and crashes, but a number of specific cognitive biases and emotional biases prevalent during such periods can be identified.

In bubbles, investors often exhibit symptoms of *overconfidence*; overtrading, underestimation of risks, failure to diversify, and rejection of contradictory information. With overconfidence, investors are more active and trading volume increases, thus lowering their expected profits. For overconfident investors (active traders), studies have shown that returns are less than returns to either less active traders or the market while risk is higher. At the market level, volatility also often increases in a market with overconfident traders.

The overconfidence and excessive trading that contribute to a bubble are linked to *confirmation bias* and *self-attribution bias*. In a rising market, sales of stocks from a portfolio will typically be profitable, even if winners are being sold too soon. Investors can have faulty learning models that bias their understanding of this profit to take personal credit for success, a form of *hindsight bias*. Selling for a gain appears to validate a good decision in an original purchase and may confer a sense of pride in locking in the profit. This dynamic fuels overconfidence that can lead to poor decisions. Regret aversion can also encourage investors to participate in a bubble, believing they are "missing out" on profit opportunities as stocks appreciate.

As a bubble unwinds, markets may underreact because of anchoring when investors do not sufficiently update their beliefs. The early stages of unwinding a bubble can involve investors in cognitive dissonance, who ignore losses and attempt to rationalize flawed decisions. Eventually, investors capitulate, which accelerates price declines.

5.4 Value

Value stocks are typically characterized by low price-to-earnings ratios, high book-to-market equity, and low price-to-dividend ratios. Growth stock characteristics are generally the opposite of value stock characteristics. For example, growth stocks are characterized by low book-to-market equity, high price-to-earnings ratios, and high price-to-dividend ratios. A number of studies have identified outperformance of value stocks relative to growth stocks over long periods. Fama and French (1998) found that value stocks (high book-to-market equity) outperformed growth stocks (low book-to-market equity) in 12 of 13 major markets during the 1975–1995 period.

Fama and French have also found, however, that the value stock anomaly disappears in a three-factor asset pricing model. This result suggests that size and book-to-market factors are not mispricing but instead represent compensation for risk exposures, such as the greater potential of companies with these characteristics to suffer distress during economic downturns.

A number of other studies have offered behavioral explanations for value anomalies, presenting the anomalies as mispricing rather than compensation for increased risk. These studies recognize the emotional factors involved in appraising stocks. The **halo effect**, for example, extends a favorable evaluation of some characteristics to other characteristics. A company with a good growth record and good previous share price performance might be seen as a good investment, with higher expected returns than its risk characteristics merit. This view is a form of representativeness that can lead investors to extrapolate recent past performance into expected returns. Overconfidence can also be involved in predicting growth rates, potentially leading growth stocks to be overvalued.

Studies have also identified that emotions play a role in estimating risk and expected return of stocks. The impact of emotional biases may be greater with less sophisticated or retail investors, but it has also been identified as a bias in analysts and professional investors. The emotional attraction of a stock can be enhanced by personal experience of products, the value of the brand, marketing expenditures, and the proximity of the headquarters to the analyst or investor. This last issue reflects the **home bias** anomaly, by which portfolios exhibit a strong bias in favor of domestic securities in the context of global portfolios. The effect has also been noted within geographical boundaries, favoring companies headquartered nearer the investor. Home bias may reflect a perceived relative informational advantage, a greater feeling of comfort with the access to company executives that proximity brings (either personal or local brokerage), or a psychological desire to invest in a local community.

To the extent to which less sophisticated investors are influenced by emotions, they may value growth companies more highly. Stock returns of funds that are rated as popular in a *Fortune* magazine survey are found to be subsequently low. A more positive emotional rating in a company leads investors to perceive the company's stock as less risky. Although the capital asset pricing model assumes risk and expected return are positively correlated, many investors behave as if the correlation is negative, expecting higher returns with lower risk.

SUMMARY

Behavioral biases potentially affect the behaviors and decisions of financial market participants. By understanding these biases, financial market participants may be able to moderate or adapt to them and, as a result, improve upon economic outcomes. Behavioral biases may be categorized as either cognitive errors or emotional biases. The type of bias influences whether its impact may be moderated or adapted to.

Among the points made in this reading are the following:

- Individuals do not necessarily act rationally and consider all available information in the decision-making process because they may be influenced by behavioral biases.
- Biases may lead to suboptimal decisions.
- Behavioral biases may be categorized as either cognitive errors or emotional biases. A single bias may have aspects of both, however, with one type of bias dominating.
- Cognitive errors stem from basic statistical, information-processing, or memory errors; cognitive errors typically result from faulty reasoning.
- Emotional biases stem from impulse or intuition and tend to result from reasoning influenced by feelings.
- Cognitive errors are more easily corrected for because they stem from faulty reasoning rather than an emotional predisposition.
- Emotional biases are harder to correct for because they are based on feelings, which can be difficult to change.
- To adapt to a bias is to recognize and accept the bias and to adjust for the bias rather than to attempt to moderate the bias.
- To moderate a bias is to recognize the bias and to attempt to reduce or even eliminate the bias within the individual.
- Cognitive errors can be further classified into two categories: belief perseverance biases and information-processing biases.
- Belief perseverance errors reflect an inclination to maintain beliefs. The belief is maintained by committing statistical, information-processing, or memory errors. Belief perseverance biases are closely related to the psychological concept of cognitive dissonance.

- Belief perseverance biases include conservatism, confirmation, representativeness, illusion of control, and hindsight.
- Information-processing biases result in information being processed and used illogically or irrationally.
- Information-processing biases include anchoring and adjustment, mental accounting, framing, and availability.
- Emotional biases include loss aversion, overconfidence, self-control, status quo, endowment, and regret aversion.
- Understanding and detecting biases is the first step in overcoming the effect of biases on financial decisions. By understanding behavioral biases, financial market participants may be able to moderate or adapt to the biases and, as a result, improve upon economic outcomes.
- Behavioral finance has the potential to explain some apparent deviations from market efficiency (market anomalies).

REFERENCES

Choi, James, David Laibson, Brigitte Madrian, and Andrew Metrick. 2001. *"For Better or For Worse: Default Effects and 401(k) Savings Behavior."* NBER Working Paper, No. 8651. doi:

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2008. *Global Investment Returns Yearbook 2008*. London, UK: ABN AMRO Bank NV.

Fama, Eugene. 1998. "Market Efficiency, Long-Term Returns, and Behavioral Finance." *Journal of Financial Economics* 49 (3): 283–306. doi:

Fama, Eugene F., and Kenneth R. French. 1998. "Value versus Growth: The International Evidence." *Journal of Finance* 53 (6): 1975–99. doi:

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48 (1). doi:

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–91. doi:

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money*. New York: Harcourt, Brace.

Mackay, Charles1841. Extraordinary Popular Delusions and the Madness of Crowds. 1st ed. United Kingdom: Harriman House Publishing. (New edition published September . 2003.)

Statman, Meir. 1999. "Behavioral Finance: Past Battles and Future Engagements." *Financial Analysts Journal* 55 (6): 18–27. doi:

Statman, Meir. 2008. "What Is Behavioral Finance?" In *Handbook of Finance*, ed. Fabozzi, Frank. Hoboken, NJ: John Wiley & Sons.

Thaler, Richard H. 1980. "Toward a Positive Theory of Economic Choice." *Journal of Economic Behavior & Organization* 1 (1): 39–60. doi:

PRACTICE PROBLEMS

1 Under-diversified portfolios are not a potential implication of which of the following behavioral biases?

A Representativeness

B Illusion of control

C Confirmation

2 The advice "Don't confuse brains with a bull market" is aimed at mitigating which of the following behavioral biases?

A Self-control

B Conservatism

C Overconfidence

3 Status quo bias is least similar to which of the following behavioral biases?

A Endowment

B Regret aversion

C Confirmation

4 Which strategy would best mitigate or prevent endowment bias?

A Actively seeking out information that challenges existing beliefs

B When new information is presented, asking "How does this information change my forecast?"

C Asking "Would you buy this security today at the current price?"

5 Jun Park, CFA, works at a hedge fund. Most of Park's colleagues are also CFA charterholders. At an event with recent university graduates, Park comments, "Most CFA charterholders work at hedge funds." Park's remark exhibits which behavioral bias?

A Availability

B Conservatism

C Framing

6 In the 1980s, Japan was viewed by many FMPs as the model economy. Although its growth began to decelerate sharply by 1990, it was not until the mid to late 1990s that FMPs' GDP forecasts were consistently achieved. By taking several years to adapt their forecasts to the lower growth environment, FMPs exhibited which behavioral bias?

A Mental accounting

B Overconfidence

C Conservatism

© 2020 CFA Institute. All rights reserved.

The following information relates to Questions 7–10

Caitríona Daosri is a portfolio manager for an international bank, where she advises high-net-worth clients. Daosri is meeting with a new client, Estêvão Kai, a 40-year-old surgeon with €4 million across various accounts and a salary of €500,000 per annum. Kai explains to Daosri that he has four accounts at four different banks, each with specific sources and uses of funds, as shown in the following table:

Bank Account	Source of Deposits	Use of Funds
1	Salary	Living expenses
2	Bonus	Charitable gifts
3	Portfolio interest	Savings for retirement
4	Portfolio dividends	Mother's living expenses

7 Based on the description of how Kai manages his finances as outlined in Exhibit 1, Kai most likely exhibits the behavioral bias of:

A endowment.

B mental accounting.

C framing.

8 Which of the following is a likely consequence of Kai's approach to managing his finances?

A Concentrated portfolio positions

B Forgone opportunities to reduce risk by combining assets with low correlations

C Excessive trading

9 Which strategy should Daosri use or recommend to Kai?

A Keep written records of investment decisions.

B Ask questions such as, "Is the decision the result of focusing on a net gain or net loss position?"

C Aggregate all accounts and portfolios into a single spreadsheet.

10 Which of the following individual behavioral biases is most strongly associated with market bubbles?

A Overconfidence

B Representativeness

C Framing

11 All of the following are reasons that an apparent deviation from the efficient market hypothesis might not be anomalous except:

A The abnormal returns represent compensation for exposure to risk.

B Changing the asset pricing model makes the deviation to disappear.

C The deviation is well known or documented.

12 The halo effect, which may be evident in FMP's assessments of a company with a history of high revenue growth, is a form of which behavioral bias?

A Endowment

B Representativeness

C Regret aversion

13 Investment managers incentivized or accountable for short-term performance by current and prospective clients is a potentially rational explanation for which of the following?

- **A** Home bias
- **B** Bubbles
- **C** Value stocks outperforming growth stocks

14 Momentum, can be partly explained by the following behavioral biases except:

- **A** availability.
- **B** home bias.
- **C** regret.

15 All of the following are reasons that the historical outperformance of value stocks versus growth stocks may not be anomalous except:

- **A** Abnormal returns represent compensation for risk exposures, such as the heightened risk of value stocks to suffer distress during downturns.
- **B** Companies with strong historical growth rates are viewed as good investments, with higher expected returns than risk characteristics merit.
- **C** The deviation disappears by incorporating a three-factor asset pricing model.

SOLUTIONS

1 A is correct. Under-diversified portfolios are a consequence of both illusion of control and confirmation biases. Researchers have found that some investors prefer to invest in companies that they feel they have control over, such as the companies they work for, leading them to hold concentrated positions. Confirmation bias may lead to FMPs ignoring negative news, paying attention only to information confirming that a company is a good investment, which may result in large positions. Representativeness bias is not typically associated with under-diversified portfolios.

2 C is correct. This advice is specifically aimed at reducing self-attribution bias, a form of overconfidence bias. This bias may result in FMPs taking credit for investment success, as well as assigning responsibility to others for investment failures, when in reality the investment results reflect exogenous market forces.

3 C is correct. Both endowment bias and regret-aversion bias often result in indecision or inertia—a typical outcome of status quo bias, in which people prefer to not make changes even when changes are warranted.

4 C is correct. Endowment bias refers to people attributing additional, unwarranted value to things they possess versus things they do not. This bias is evident in FMPs that systematically and materially overvalue securities in their portfolio versus securities not in their portfolio. The question “Would you buy this security today at the current price?” turns the investor’s attention to assessing the reasonableness of the current price as a buy price rather than solely as a selling price.

5 A is correct. Park is extrapolating his observation based on a narrow range of experience (working at a hedge fund that employs many CFA charterholders) to the entire population of CFA charterholders. Using a narrow range of experience is a form of availability bias.

6 C is correct. Conservatism bias results in maintain or only slowly updating views and forecasts despite the presence of new information. FMPs in the 1990s were reluctant to update forecasts, despite materially different new information for several years.

7 B is correct. Kai has segregated money into four different accounts based on the sources and uses of his funds. Although intuitively appealing, this approach is irrational because money is fungible across the four accounts. Nothing is stopping Kai from collapsing them into a single “account” with a holistic portfolio strategy.

8 B is correct. The most common consequence of mental accounting is neglecting opportunities to reduce risk by combining assets with low correlations, because each account’s asset allocation is examined discretely. Offsetting positions across accounts, or an overall inefficient allocation with respect to risk, can lead to suboptimal aggregate performance.

9 C is correct. Aggregating mental accounts is a logical strategy to combat mental accounting. It is the opposite of disaggregating money into separate accounts.

10 A is correct. The overconfidence and excessive trading that contribute to a bubble are linked to self-attribution bias, a form of overconfidence. In a rising market, sales of stocks from a portfolio will typically be profitable, even if winners are being sold too soon, and FMPs will attribute profits and strong performance to their investment acumen and subsequently underestimate risks.

11 C is correct. Bubbles and crashes are well-known and well-documented phenomena yet represent market anomalies.

12 B is correct. Representativeness refers to the tendency to adopt a view or forecast based on individual information or a small sample, as well to use simple classifications. The halo effect is an example of representativeness, because FMPs extend an overall favorable evaluation to an investment (e.g., a "good company") based on one or few characteristics (e.g., a "visionary CEO").

13 B is correct. Investment managers' incentives—or perhaps more accurately, their perception of their incentives—for short-term performance were named as considerations in the technology and real estate bubbles. Not participating in the bubble presented certain FMPs with commercial or career risk.

14 B is correct. Home bias refers to FMPs preferentially investing in domestic securities, likely reflecting perceived relative informational advantages, a greater feeling of comfort with the access to company executives that proximity brings (either personal or through a local brokerage), or a psychological desire to invest in a local community. Momentum, on the other hand, has been documented in a range of markets around the world, in a time-dependent manner, and reflects some FMPs' availability bias, manifested as a belief that stocks will continue to rise because recently they have only risen, as well as regret aversion by those who invest in past winners because they regret not investing in them in the past.

15 B is correct. This choice describes the halo effect, which does offer a behavioral explanation for the poor performance of growth stocks versus value stocks. Growth stocks are mispriced relative to their risk characteristics, because FMPs focusing on just a few properties, such as a high historical revenue growth rate, while neglecting other characteristics.

READING

53

Introduction to Risk Management

by Don M. Chance, PhD, CFA, and Michael E. Edleson, PhD, CFA

Don M. Chance, PhD, CFA, is at Louisiana State University (USA). Michael E. Edleson, PhD, CFA, is at the University of Chicago (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** define risk management;
☐	**b.** describe features of a risk management framework;
☐	**c.** define risk governance and describe elements of effective risk governance;
☐	**d.** explain how risk tolerance affects risk management;
☐	**e.** describe risk budgeting and its role in risk governance;
☐	**f.** identify financial and non-financial sources of risk and describe how they may interact;
☐	**g.** describe methods for measuring and modifying risk exposures and factors to consider in choosing among the methods.

INTRODUCTION

1

Risk—and risk management—is an inescapable part of economic activity. People generally manage their affairs to be as happy and secure as their environment and resources will allow. But regardless of how carefully these affairs are managed, there is risk because the outcome, whether good or bad, is seldom predictable with complete certainty. There is risk inherent in nearly everything we do, but this reading will focus on economic and financial risk, particularly as it relates to investment management.

All businesses and investors manage risk, whether consciously or not, in the choices they make. At its core, business and investing are about allocating resources and capital to chosen risks. In their decision process, within an environment of uncertainty, these organizations may take steps to avoid some risks, pursue the risks that provide the highest rewards, and measure and mitigate their exposure to these risks as necessary. Risk management processes and tools make difficult business and financial problems easier to address in an uncertain world. Risk is not just a matter

© 2019 CFA Institute. All rights reserved.

of fate; it is something that organizations can actively manage with their decisions, within a risk management framework. Risk is an integral part of the business or investment process. Even in the earliest models of modern portfolio theory, such as mean–variance portfolio optimization and the capital asset pricing model, investment return is linked directly to risk but requires that risk be managed optimally. Proper identification and measurement of risk, and keeping risks aligned with the goals of the enterprise, are key factors in managing businesses and investments. Good risk management results in a higher chance of a preferred outcome—more value for the company or portfolio or more utility for the individual.

Portfolio managers need to be familiar with risk management not only to improve the portfolio's risk–return outcome, but also because of two other ways in which they use risk management at an enterprise level. First, they help to manage their own companies that have their own enterprise risk issues. Second, many portfolio assets are claims on companies that have risks. Portfolio managers need to evaluate the companies' risks and how those companies are addressing them.

This reading takes a broad approach that addresses both the risk management of enterprises in general and portfolio risk management. The principles underlying portfolio risk management are generally applicable to the risk management of financial and non-financial institutions as well.

The concept of risk management is also relevant to individuals. Although many large organizations formally practice risk management, most individuals practice it more informally and some practice it haphazardly, oftentimes responding to risk events after they occur. Although many individuals do take reasonable precautions against unwanted risks, these precautions are often against obvious risks. The more subtle risks are often ignored. Unfortunately, many individuals do not view risk management as a formal, systematic process that would help them achieve not only their financial goals but also the ultimate goal, or maximum utility as economists like to call it, but they should.

Although the primary focus of this reading is on institutions, we will also cover risk management as it applies to individuals. We will show that many common themes underlie risk management—themes that are applicable to both organizations and individuals.

Although often viewed as defensive, risk management is a valuable offensive weapon in the manager's arsenal. In the quest for preferred outcomes, such as higher profit, returns, or share price, management does not usually get to choose the outcomes but does choose the risks it takes in pursuit of those outcomes. The choice of which risks to undertake through the allocation of its scarce resources is the key tool available to management. An organization with a comprehensive risk management culture in place, in which risk is integral to every key strategy and decision, should perform better in the long-term, in good times and bad, as a result of better decision making.

The fact that all businesses and investors engage in risky activities (i.e., activities with uncertain outcomes) raises a number of important questions. The questions that this reading will address include the following:

- What is risk management, and why is it important?
- What risks does an organization (or individual) face in pursuing its objectives?
- How are an organization's goals affected by risk, and how does it make risk management decisions to produce better results?
- How does risk governance guide the risk management process and risk budgeting to integrate an organization's goals with its activities?
- How does an organization measure and evaluate the risks it faces, and what tools does it have to address these risks?

The answers to these questions collectively help to define the process of risk management. This reading is organized along the lines of these questions. Sections 2 and 3 describe the risk management process, and Sections 4–6 discuss risk governance and risk tolerance. Sections 7 and 8 cover the identification of various risks, and Sections 9–11 addresses the measurement and management of risks.

THE RISK MANAGEMENT PROCESS

2

a define risk management;

Risk, broadly speaking, is exposure to uncertainty. Risk is also the concept used to describe all of the uncertain environmental variables that lead to variation in and unpredictability of outcomes. More colloquially, risk is about the chance of a loss or adverse outcome as a result of an action, inaction, or external event.

This last view may make it sound as if risk is something to be avoided. But that is not at all the case. Risk is a key ingredient in the recipe for business or investment success; return without risk is generally a false hope and usually a prescription for falling short of one's goals. Risks taken must be carefully chosen, understood, and well-managed to have a chance at adding value through decisions. Risk and return are the interconnected forces of the financial universe. Many decision makers focus on return, which is not something that is easily controlled, as opposed to risk, or exposure to risk, which may actually be managed or controlled.

Risk exposure is the extent to which the underlying environmental or market risks result in actual risk borne by a business or investor who has assets or liabilities that are sensitive to those risks. It is the state of being exposed or vulnerable to a risk. Risk exposure results from the decisions of an organization or investor to take on risk-sensitive assets and liabilities.

Suppose there is an important announcement in Japan that will result in the yen either appreciating or depreciating by 1%. The range of possible outcomes in real situations is clearly not as simple as the up-or-down 1% case used here, but we will use a simplified example to make an important point. The risk is the uncertain outcome of this event, and the currency risk to a non-Japanese business is the uncertain return or variation in return in domestic currency terms that results from the event. The risk can be described as the range of resulting outcomes and is often thought of in terms of a probability distribution of future returns. Suppose that the underlying amount is ¥1,000,000. The risk exposure of a business may be zero or it could be sizable, depending on whether the business has assets or liabilities tied to this risk—in this case, exposure to that currency. One million yen would, in this example, result in ¥10,000 of risk exposure (1% of ¥1,000,000). Risk management would include, among other things, quantifying and understanding this risk exposure, deciding how and why to have the exposure and how much risk the participant can bear, and possibly mitigating this risk by tailoring the exposure in several ways. The risk management process would inform the decision of whether to operate or invest in this risky currency.

The word "risk" can be confusing because it is used by different people at different times to mean so many different things. Even when used properly, the term has three related but different meanings, which this example illustrates well. Risk can mean, in turn, the underlying uncertainty, the extent of the risky action taken, or the resulting range of risky outcomes to the organization. In this example, the first meaning is the uncertain +1% or –1% movement of the currency. The second meaning is the ¥1,000,000 worth of risky currency, the position taken by the business. The third meaning is the +¥10,000 or –¥10,000 risky outcome that might accrue to the business for having engaged in this risky activity. A common way of more precisely distinguishing among these three "risks" in usage is: *risk driver* for the underlying risk, *risk position* to describe or quantify the risky action taken, and *risk exposure* for the potential valuation change that may result. In the oversimplified example above, the risk exposure is simply the risk position multiplied by the risk driver. In practice the term "risk" is used interchangeably for all three meanings.

> ***Risk management*** *is the process by which an organization or individual defines the level of risk to be taken, measures the level of risk being taken, and adjusts the latter toward the former, with the goal of maximizing the company's or portfolio's value or the individual's overall satisfaction, or utility.*

Said differently, risk management comprises all the decisions and actions needed to *best* achieve organizational or personal objectives while bearing a tolerable level of risk. Risk management is *not* about minimizing risk; it is about actively understanding and embracing those risks that best balance the achievement of goals with an acceptable chance of failure, quantifying the exposure, and continually monitoring and modifying it. A company that shied away from all risk would find that it could not operate. In trying to create wealth, all organizations will find themselves "in the risk business." Risk management is not about avoiding risks any more than a practical diet is about avoiding calories. Risk management is not even about predicting risks. "The Doctrine of No Surprises" is a key mantra among many risk managers, but it does not mean they are expected to predict what will happen. Instead, it means that if an unpredictable event, either positive or negative, happens in an uncertain world, the *impact* of that event on the organization or portfolio would not be a surprise and would have been quantified and considered in advance.

For example, a risk manager of a bank would not have been expected to know that a real estate crisis was going to occur and cause significant defaults on the bank's real estate securities. But a good risk manager would help the bank's management decide how much exposure it should have in these securities by quantifying the potential financial impact of such a crisis destroying, say, 60% of the bank's capital. A good risk management process would include a deep discussion at the governance level about the balance between the likely returns and the unlikely—but sizable—losses and whether such losses are tolerable. Management would ensure that the risk analysis and discussion actively affects their investment decisions, that the potential loss is continuously quantified and communicated, and that it will take actions to mitigate or transfer any portion of the risk exposure that cannot be tolerated.[1] The only surprise here should be the market shock itself; the risk manager should have prepared the organization through stress-testing and scenario analysis, continuously reporting in advance on the potential impact of this sizable risk exposure.

A poor risk management process would have ignored the possibility, though small, of such a significant market event and not quantified the potential loss from exposure to a real estate crisis. As such, the bank's management would have had no idea that more than half of the bank's capital could be at risk, not addressed this risk in any governance/risk appetite discussion, ignored these risks in its investment decisions, and not taken any action to mitigate this risk. In a good risk management process, most of the work is done before an adverse event happens; in a poor risk management process, perhaps just as much work gets done, but it all comes after the event, which is after the damage has been done.

---

[1] For example, hedges may be used to limit loss of capital to 20%.

Good risk management does not prevent losses, but provides a full top-to-bottom framework that rigorously informs the decision-making process—before, during, and after a risk event. Because risks and exposures are dynamic, risk management is a continuous process that is always being reevaluated and revised. If this process is done well, it provides management and staff with the knowledge to navigate as efficiently as possible toward the goals set by the governing body. In turn, this effort increases *ex ante* the value of the business or investment decisions undertaken. Good risk management may allow managers to more quickly or effectively act in the face of a crisis. But *ex post*, even the best risk management may not stop a portfolio from losing money in a market crash nor prevent a business from reduced profits in an economic downturn.

THE RISK MANAGEMENT FRAMEWORK

3

b describe features of a risk management framework;

c define risk governance and describe elements of effective risk governance;

A **risk management framework** flows logically from the definition of risk management that was previously given: It is the infrastructure, process, and analytics needed to support effective risk management in an organization. This process should fully integrate the "risk" and "return" aspects of the enterprise into decisions in support of best achieving its goals within its tolerance for risk. Risk management is not a "one size fits all" solution; it is integral to the enterprise's goals and needs. Thus, it is best achieved through a custom solution. Despite customization, every risk management system or framework should address the following key factors:

- Risk governance
- Risk identification and measurement
- Risk infrastructure
- Defined policies and processes
- Risk monitoring, mitigation, and management
- Communications
- Strategic analysis or integration

Not surprisingly, these factors often overlap in practice. They are defined and discussed here.

Governance is the top-level system of structures, rights, and obligations by which organizations are directed and controlled. Normally performed at the board level, governance is how goals are defined, authority is granted, and top-level decisions are made. The foundation for risk management in the organization is set at the board level as well. **Risk governance** is the top-down process and guidance that directs risk management activities to align with and support the overall enterprise and is addressed in more detail in Sections 4–6. Good governance should include defining an organization's risk tolerance and providing risk oversight. Governance is often driven by regulatory concerns, as well as by the fiduciary role of the governing body. A risk management committee is another facet of governance; it provides top decision makers with a forum for regularly considering risk management issues. To achieve the best results for an organization, risk governance should take an enterprise-wide view. **Enterprise risk management** is an overarching governance approach applied throughout the organization and consistent with its strategy, guiding the risk management framework to focus risk activities on the objectives, health, and value of the *entire* organization.

Risk identification and measurement is the main quantitative core of risk management; but more than that, it must include the qualitative assessment and evaluation of all potential sources of risk and the organization's risk exposures. This ongoing work involves analyzing the environment for relevant risk drivers, which is the common term used for any fundamental underlying factor that results in a risk that is relevant or important to an organization, analyzing the business or portfolio to ascertain risk exposures, tracking changes in those risk exposures, and calculating risk metrics to size these risks under various scenarios and stresses.

Risks are not limited to what is going on in the financial markets. There are many types of risk that can potentially impact a business, portfolio, or individual.

The power of technology has allowed for risk management to be more quantitative and timely. Management can measure and monitor risk, run scenarios, conduct statistical analysis, work with more complex models, and examine more dimensions and risk drivers as well as do it faster. This use of technology needs to be balanced with and supplement—not supplant—experienced business judgment. Technology has made risk infrastructure even more important and beneficial in managing risks.

Risk infrastructure refers to the people and systems required to track risk exposures and perform most of the quantitative risk analysis to allow an assessment of the organization's risk profile. Infrastructure would include risk capture (the important operational process by which a risk exposure gets populated into a risk system), a database and data model, analytic models and systems, a stress or scenario engine, and an ability to generate reports, as well as some amount of skilled and empowered personnel resources dedicated to building and executing the risk framework. With increased reliance on technology, more time and effort must be allotted to test data, models, and results in order to avoid the ironic outcome of the risk of errors coming from within risk systems.

Obviously, the scope of risk infrastructure will be related to the resources, or potential losses, of the organization. Individuals and smaller businesses may rely heavily on an external partner or provider for much of their risk infrastructure and analysis.

Policies and processes are the extension of risk governance into both the day-to-day operation and decision-making processes of the organization. There may be limits, requirements, constraints, and guidelines—some quantitative, some procedural—to ensure risky activities are in line with the organization's predetermined risk tolerance and regulatory requirements. Much of this is just common-sense business practice: updating and protecting data, controlling cash flows, conducting due diligence on investments, handling exceptions and escalations, and making checklists to support important decisions. In a good risk framework, processes would naturally evolve to consider risk at all key decision points, such as investment decisions and asset allocation. Risk management should become an integrated part of the business and not just a policing or regulatory function.

The process of *risk monitoring, mitigation, and management* is the most obvious facet of a risk framework, but also one of the most difficult. Actively monitoring and managing risk requires pulling together risk governance, identification and measurement, infrastructure, and policies and processes and continually reviewing and reevaluating in the face of changing risk exposures and risk drivers. It requires recognizing when risk exposure is not aligned with risk tolerance and then taking action to bring them back into alignment.[2]

---

[2] Risk mitigation and management is discussed in more detail in Sections 9–11.

Communication of critical risk issues must happen continually and across all levels of the organization. Governance parameters, such as risk tolerances and associated constraints, must be clearly communicated to, and understood by, managers. Risk metrics must be reported in a clear and timely manner. Risk issues must be reviewed and discussed as a standard part of decision making. Changes in exposure must be discussed so that action can be taken as appropriate. There should also be a feedback loop with the governance body so that top-level risk guidance can be validated or updated and communicated back to the rest of the organization.

Strategic analysis and integration help turn risk management into an offensive weapon to improve performance. Good risk management is a key to increasing the value of the overall business or portfolio. A risk management framework should provide the tools to better understand the how and why of performance and help sort out which activities are adding value, and which are not. In investing, rigorous analysis can support better investment decisions and improve strategy and risk-adjusted returns.

Exhibit 1 The Risk Management Framework in an Enterprise Context

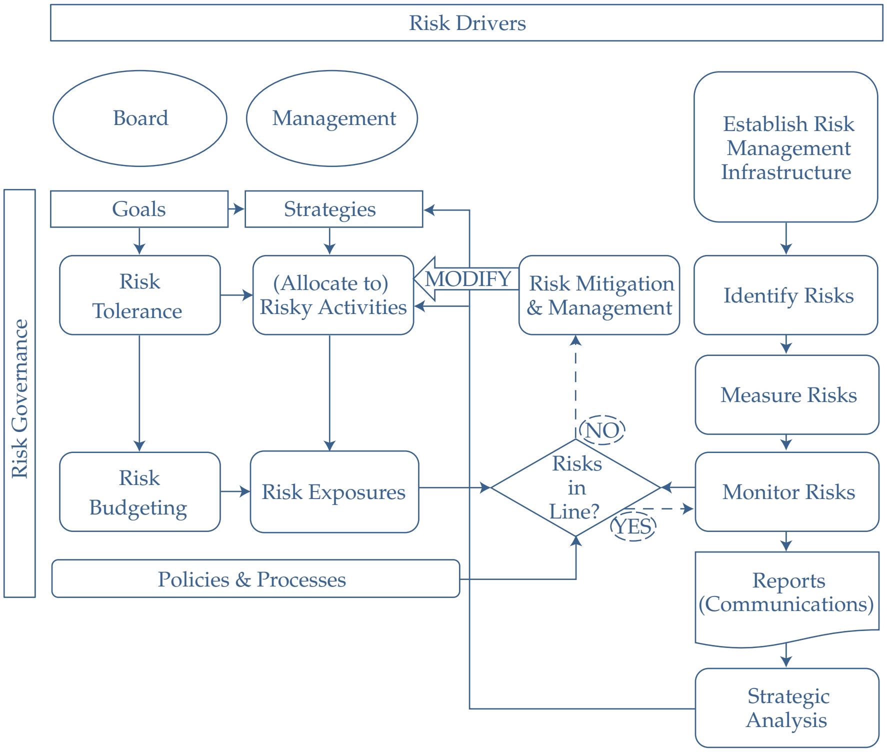

Risk Drivers
Board
Management
Establish Risk Management Infrastructure
Risk Governance
Goals
Strategies
Risk Tolerance
(Allocate to) Risky Activities
MODIFY
Risk Mitigation & Management
Identify Risks
Measure Risks
NO
Risk Budgeting
Risk Exposures
Risks in Line?
YES
Monitor Risks
Policies & Processes
Reports (Communications)
Strategic Analysis

Exhibit 1 illustrates the process of risk management for an enterprise, pulling all the described elements of the risk framework together. Although there are a very high number of risks faced by every organization, most organizations are primarily affected by a small number of key risk drivers, or primary underlying factors that create risk. Along the left side is risk governance, which represents board-level decisions and encompasses and affects the boxes immediately to its right. The governance body, often called a board, defines the goals of the organization and, in turn, decides on its risk tolerance. It may additionally provide guidance on how or where that risk is taken (risk budgeting). The board is also involved in setting high-level policies that will affect most risk management processes. These risk governance activities are a crucial keystone of the risk framework and will be discussed in detail in the next section. When the rest of the risk framework hinges off of these top-down governance elements and is focused on the goals of the entire enterprise (as shown here), the end result is effective enterprise risk management.

The role of management, shown in the middle column, is to plan and execute value-maximizing strategies consistent with their governance guidance. Each management activity in the framework flows not only from management (shown with the arrows) but also from the governance activities on the left. Thus, not only are management's strategies designed to achieve the board's goals, but management also allocates capital to risky activities (its business or investing choices) to execute its strategies consistent with the defined risk tolerance. The risk exposures that result from management's choice of activities should also be aligned with the governing body's risk budget. In addition, management participates actively in setting or implementing policies and establishing procedures that relate to when, how, how much, and by whom each of the other elements of the risk framework are performed.[3]

The rest of the risk management framework comprises a number of important risk activities to help the business achieve all of its strategic and governance goals and mandates. These other elements to implement risk management are shown in the far right column of the illustration. Driven by its need to establish a risk management program to support the enterprise's goals, management would provide the requisite resources for risk management activities by establishing a risk management infrastructure. With risk processes defined and risk infrastructure in place, risks are then identified and measured, which is a regular and continual process of translating risk exposures (produced by the risky activities) into meaningful and generally quantitative risk metrics.

The next major steps—risk monitoring, mitigation, and management—are where much of a firm's day-to-day risk management activity is focused. These activities are split across three boxes in the illustration. Risk levels are continuously monitored, having just been measured.[4] There is a major decision at the monitoring stage: Management must check that all the risks are in line and not outside the limits of the defined risk tolerance or budget.[5] This process involves evaluating the actual risk exposures compared with the organization's risk policies and procedures to ensure that everything is in compliance. If the answer is "no," then risk mitigation and management actions need to be taken to modify risk levels and to bring them back into compliance. There are a variety of methods to accomplish this task, which are addressed in Sections 9–11. Whatever the method, management's allocation of the risk budget to risky activities will be altered by this modification, which includes changing the organization's risk exposures, starting the circle again through the steps on the right, and re-checking to see if risk levels are now consistent with risk policies.

When risks *are* in line with limits, policies, tolerances, mandates, and so on, then the process moves back to continuous monitoring followed by communicating risk levels.[6] This communication, at a minimum, includes reporting key risk metrics on a regular and timely basis across the organization to assist management in its decision-making process and the board in fulfilling its governance duties. Finally, strategic analysis is supported by the risk measurement, reporting, and other steps of the enterprise risk management process. By analyzing all of the enterprise's strategies and risky activities via the risk measurement lens, management can improve its decision-making process and ascertain where to invest its limited capital and risk budget most fruitfully. This step is generally underappreciated and is an inexpensive and beneficial by-product of having built a risk framework. The last two boxes or steps (reports and strategic analysis) represent important feedback loops to inform and improve both governance and the portfolio of risky activities that make up the business.

---

[3] In essence, there could be an arrow from policies and processes to every other box to the right, but these rather obvious relationships are intentionally omitted in the diagram to avoid clutter. Likewise, risk exposures inform nearly all the boxes to the right. Risk management is innately quite interrelated.

[4] Continuous usually does not mean real-time; the frequency of monitoring is based on the resources available, the level of systems support, and the need for risk information in the decision process. At large financial firms, this monitoring will generally be daily; for small businesses and individuals it might be quarterly.

[5] This task is generally delegated to a risk manager; but whatever the title, someone must be accountable for this important check.

[6] While not obvious in the illustration, communication and reporting should happen whether or not risk levels are in compliance; such communications are even more important when risk levels are out of alignment with tolerances.

There are many feedback loops in properly executed risk management. In practice, most of these steps overlap most of the time and are being performed simultaneously. Good risk management ties together all these steps from the highest governance decisions to lower-level specifics, such as models, reports, and operational checklists.

The risk environment is dynamic, and many of our notions of probabilities and likely risk outcomes change in ways we probably could not predict. The risk management framework should be robust enough to anticipate this dynamism—to expect the unexpected. It should be evolutionary—flexible enough to grow with a company or individual and its new challenges.

The complexity of the risk management framework depends on the complexity of an organization's risk exposures and their resources. But that does not mean that smaller organizations or individuals should skip the risk management process; they may simply be able to do less, or have to work with external partners to assist with large portions of the framework, or be less formal about the process. Ultimately, the key principles just covered are still important even to the smallest organization, even if the specific components do not get assembled as described.

THE RISK MANAGEMENT PROCESS FOR AN INDIVIDUAL

Although an individual has neither the resources nor the organizational overhead of a large business, the importance of risk management is not diminished and the risk management framework still applies, albeit most likely in a scaled-down form. Though nearly all of the essential elements of the process illustrated in Exhibit 1 are still useful, the individual can reduce the recipe to six essential ingredients, consistent with the reduced scope of the individual's risk exposures.

The first step for an individual is much like that shown in Exhibit 1 for the most complex organization: the determination of goals or objectives. This step would include most of the elements associated with risk governance, just without all the organizational complexity.

The next functional step involves choosing investments (or other assets) and identifying their risks. Lacking any risk infrastructure, the individual may at this stage already require the services of an investment professional or financial adviser. This step and subsequent steps will probably be executed by the adviser, although the individual principal still needs to stay knowledgeable and involved. In the context of the illustration in Exhibit 1, the individual is effectively their own governance body and the adviser serves the role of management.

The next steps for the individual are equivalent to the heart of the risk management process: risk monitoring and risk mitigation and management. The individual would first evaluate their risk exposure (like the diamond or decision step in the illustration), then consider various alternative approaches to modify the risk if necessary, followed by implementing the risk management solution (insuring, hedging, trading, etc.).

The final functional step for an individual's risk management process would be evaluation and review. This step is parallel to the back-end of the risk management illustration, the boxes at the bottom right. This process may occur with much less frequency for an individual—but it is no less important.

Each individual should simplify the risk management process as required so that they do not end up considering it “too esoteric and complicated to worry about” and thus ignoring risk management altogether. The potential costs of avoiding risk management are essentially the same for an individual as for a large corporation or a hedge fund, although perhaps with less money involved.

At its core, business and investing are about allocating resources and capital to chosen risks. Understanding which risks drive better outcomes should be one of the goals of risk management, and it makes good *risk* management inextricably linked with good management generally. When effective risk management is truly integrated at all levels of the decision-making process and the overall management process, the organization has developed an effective *risk culture*. This culture generally produces better results than just considering risk issues as a separate afterthought, and, in turn, it produces *much* better results than ignoring risk issues altogether in the decision-making process. For individuals, the adoption of a risk culture should result in a personal awareness of the many types of risks, their rewards, the costs, the relationships between them, and the methods of aligning the risks borne with the risks and outcomes desired. This awareness should lead to better investment return and/or smaller losses for the risk taken, resulting in higher satisfaction.

There are a number of other benefits from establishing good risk management: (1) Most obvious is less frequent surprises and a better notion of what the damage would be in the event of a surprise; (2) more decision discipline leading to better consideration of trade-offs and better risk–return relationships; (3) better response and risk mitigation stemming from more awareness and active monitoring, which should trim some of the worst losses; (4) better efficiency and fewer operational errors from policies and procedures, transparency, and risk awareness; (5) better relations, with more trust, between the governing body and management, which generally results in more effective delegation; (6) a better image or reputation because analysts and investors perceive a company as prudent and value-focused. Together, all these benefits should lead to higher value for the enterprise.

EXAMPLE 1

Risk Management and Risk Management Framework

1 Which of the following is *not* a goal of risk management?

 A Measuring risk exposures

 B Minimizing exposure to risk

 C Defining the level of risk appetite

2 Which element of a risk management framework sets the overall context for risk management in an organization?

 A Governance

 B Risk infrastructure

 C Policies and processes

3 Which element of risk management makes up the analytical component of the process?

 A Communication

 B Risk governance

 C Risk identification and measurement

4 Which element of risk management involves action when risk exposures are found to be out of line with risk tolerance?

A Risk governance

B Risk identification and measurement

C Risk monitoring, mitigation, and management

Solution to 1:

B is correct. The definition of risk management includes both defining the level of risk desired and measuring the level of risk taken. Risk management means taking risks actively and in the best, most value-added way possible and is not about minimizing risks.

Solution to 2:

A is correct. Governance is the element of the risk management framework that is the top-level foundation for risk management. Although policies, procedures, and infrastructure are necessary to implement a risk management framework, it is governance that provides the overall context for an organization's risk management.

Solution to 3:

C is correct. Risk identification and measurement is the quantitative part of the process. It involves identifying the risks and summarizing their potential quantitative impact. Communication and risk governance are largely qualitative.

Solution to 4:

C is correct. Risk monitoring, mitigation, and management require recognizing and taking action when these (risk exposure and risk tolerance) are not in line, as shown in the middle of Exhibit 1. Risk governance involves setting the risk tolerance. Risk identification and measurement involves identifying and measuring the risk exposures.

RISK GOVERNANCE - AN ENTERPRISE VIEW

4

c define risk governance and describe elements of effective risk governance;

Risk governance is the foundation for risk management. As defined earlier, it is the top-down process and guidance that directs risk management activities to align with and support the goals of the overall enterprise. It typically emanates from a board of directors with fiduciary obligations and risk oversight and who prescribe goals and authorities. Referring back to the definition of risk management, note that risk management is keenly focused on the risk and value of the overall enterprise.

4.1 An Enterprise View of Risk Governance

In addition to the responsibility for risk oversight, there are two other important areas in which the governing body drives the risk framework. First, it determines the organization's goals, direction, and priorities, which combined serve as a key foundation for enterprise risk management. Recall that enterprise risk management is an overarching governance approach applied across the organization that focuses risk activities on the objectives, health, and value of the whole organization. Second, it spells out the risk appetite or tolerance, meaning which risks are acceptable, which risks are to be mitigated and to what extent, and which risks are unacceptable. Risk governance should also provide a sense of the worst losses that could be tolerated in various scenarios, and management should manage risk accordingly. These considerations should flow naturally into decisions about risk budgeting to guide implementation of an optimal program that is consistent with that risk tolerance.

Risk governance is the impact of the governing body of an organization on the risk management framework. It provides context for and clarity on an organization's value drivers and risk appetite, specifies clear authority to management to execute risk management, and ensures risk oversight to continually determine whether risk management is functioning well and consistent with the organization's value maximization. It is the governing body's job to tie the organizational goals and risk framework together; thus, risk governance happens within an enterprise context. Risk governance and risk oversight also entail compliance with regulatory requirements. Risk governance is a difficult and demanding discipline, and if it is going to flourish in an organization, it needs visible commitment from the top.

Providing clear guidance with sufficient leeway to execute strategy is often a difficult balance. Even more challenging is providing for advance discussion and a clear decision and statement of organizational risk appetite. There is usually substantial discussion about this risk appetite *after* a crisis, but too often there is very little discussion during periods of normalcy, when it would be much more beneficial. Because risk is one of the main strategic tools that management can regulate, it is especially important for governing bodies to openly discuss risk, consider scenarios, understand the impact of negative outcomes on the organization, and make it clear where they are not willing to venture. Much like an automobile that comes with a red zone on some dials to establish boundaries for safe operation, risk governance bodies should likewise establish hypothetical red zones to ensure the safe operation of their enterprise.

Enterprise risk management (focusing risk activities on the objectives, health, and value of the *whole* organization) requires that the entire economic balance sheet of the business be considered, not just the assets or one part of the business in isolation. A narrower view of risk management is unlikely to meet the goal of maximizing the value of the entire enterprise.

Pension fund management provides a classic example of the importance of considering enterprise risk management: "Funds" are the assets and "pension" is the liability. But a true enterprise view requires an even broader outlook. A corporate pension fund's manager might try to maximize only the fund's assets, but this would generally do a disservice to the corporation. The assets and liabilities of a pension fund are both sensitive to market variables, so ignoring the liabilities would be ignoring half the risk. With liabilities that are quite bond-like, a pension fund manager using all equities for maximum growth would potentially make the overall fund insolvent in a market collapse with declining interest rates because, in such a situation, the liabilities would increase substantially in value while the assets fell. Risk tolerance for the assets in isolation would be far different from the risk tolerance of the entire enterprise. One should look beyond just the pension liabilities, which are likely to be a small part of the overall enterprise. Broader still, a true enterprise risk view in this case would *also* consider the parent corporation's business risk profile and not just the pension assets and liabilities. In a market collapse, the overall business might be in a recessionary phase, rendering increasing contributions from the corporation to its pension fund quite painful. Factoring the corporate risk profile into the pension fund investment strategy may cause the risk tolerance to be lower in this case.

Risk governance that focuses on the entire enterprise will result in risk management that is much less likely to be at odds with the goals of the organization and more likely to enhance long-run value. Likewise, consideration of a full spectrum of risks, and not just the most obvious quantitative risks, will result in better risk governance.

The enterprise view of risk management is equally applicable and important to an individual, even if the term “enterprise” is not often used in an individual context.[7] The appropriate set of risks for an individual must be viewed not in isolation, but in consideration of the goals and characteristics of the individual in a holistic view. For example, an adviser may be designing an investment portfolio to maximize a client’s wealth and optimize the risk–return trade-off at some perceived comfortable level of risk.[8] But the client, whose wealth consists not only of financial assets but also of valuable **human capital**, might prefer that risk allocation decisions be made in view of both forms of capital, optimizing her total wealth. For example, a client with a career in real estate would most likely benefit if her financial portfolio is invested in a way that considers her earnings exposure to real estate related risks. Holdings of real estate securities above a certain level, even if optimal from an isolated portfolio perspective, might make this individual less than optimally diversified from a total wealth perspective. In another example, Investor X, who has substantial inflation-adjusted pension benefits, is different from Investor Y, who has a fixed pension benefit, and different still from Investor Z, who has no pension benefit and retires with solely his own investment portfolio. These three investors will require remarkably different investment solutions, not only to deal with inflation but also to deal with the uncertainty surrounding lifespans. Individuals with different goals and characteristics will need differing investment and risk solutions that are best suited to their individual situations. In fact, because of the extremely variable life cycle of an individual and the discrete nature of many individuals’ goals, the enterprise view is even more important to risk management for individuals than it is for institutions.

Risk governance extends into management to include ways to ensure that the risk framework of an organization stays consistent with top-level guidance. One useful approach is to provide a regular forum to discuss the risk framework and key risk issues at the management level. In other words, a risk management committee would be a key element of good risk governance. Its activities could parallel the governance body’s risk deliberations, but at an operational level as opposed to high-level oversight. In this forum, governance overlaps with many of the other aspects of the organization’s risk framework as discussed in Sections 2–3. In fact, if done well, it integrates all of them.

In the same vein, another element of good risk governance is the formal appointment of a responsible executive as chief risk officer (CRO). This officer should be responsible for building and implementing the risk framework for the enterprise and managing the many activities therein. In the same manner that risks are inextricably linked with the core business activities, the CRO is likewise a key participant in the strategic decisions of the enterprise—this position is not solely a policing role. Although the chief executive is responsible for risk as well as all other aspects of an enterprise, it makes no more sense for the CEO to perform the role of the CRO than it would be for the CEO to perform the role of the CFO. Many financial firms now have a CRO in executive management,[9] which had become best practice even in the years prior to the 2008 crisis.

---

[7] Enterprise risk management is an easier concept for an individual; compared with an organization in which deciding, coordinating, and communicating goals can be a big challenge, the scope of risk management efforts for an individual is smaller and more manageable.

[8] Here, the individual is the governing body, setting individual goals and risk appetite; the financial professional or wealth manager is the “management team” executing much of the rest of the risk framework.

[9] Although this is common for financial firms or other large organizations, many less-complex companies will forgo a formal risk structure. The board still maintains its risk governance responsibilities; and it is up to them to work out with management as to how, and to what extent, to meet these responsibilities.

5

RISK TOLERANCE

d explain how risk tolerance affects risk management;

Perhaps the most important element of good risk governance is the **risk tolerance** discussion and decision within the governing body. Business and investment strategy centers on selecting a portfolio of acceptable risk activities that will maximize value and produce the highest returns possible for the given risk level. At the governance level, the duty is generally not to select these activities—a job that usually falls to management—but to establish the organization's risk appetite. Certain risks or levels of risks may be deemed acceptable, other risks deemed unacceptable, and in the middle are risks that may be pursued in a risk-limited fashion. Said differently, risk tolerance identifies the extent to which the organization is willing to experience losses or opportunity costs and to fail in meeting its objectives.

The risk tolerance decision *for an individual* is similar, but not identical, to that of a business enterprise. In traditional finance theory, the individual focuses on maximizing unobservable utility, whereas the business maximizes a generally observable value—the market value or equity price of the company. Although individuals are facing life and certain death on an uncertain timetable, most businesses tend to be relatively short-lived organizations, but with an expectation of immortality. The decisions about risk tolerance from those two very different viewpoints can be expected to differ—for example, risk tolerance in organizations often treats its continued existence as a major consideration. In many ways, the individual's risk tolerance decision is the harder one.

The enterprise risk management perspective is the right lens through which to view the risk appetite question. The risk tolerance decision begins with two different analyses that must be integrated—an "inside" view and an "outside" view. First, what shortfalls within an organization would cause it to fail, or at least fail to achieve some critical goals? Second, what uncertain forces is the organization exposed to? That is, what are its risk drivers? With the answers to these two difficult questions in hand, a board could begin defining dimensions and levels of risk that it finds too uncomfortable to take on. This risk tolerance should be formally chosen and communicated *before* a crisis, and will serve as the high-level guidance for management in its strategic selection of risks. Many organizations will do this *after* a crisis, which is better than not doing it at all but is much like buying insurance after the loss occurs. It is best to take care of it when there seems to be no particularly pressing reason to do so. Similarly, some individuals may not give much thought to their own risk tolerance until after a crisis occurs, when they belatedly decide that the risk was not worth taking.

For example, suppose a Spanish construction equipment manufacturing company's board is determining its risk tolerance. From the inside perspective, it has two main concerns: revenue and liquidity. It determines that it can tolerate a 5%–10% drop in revenue, but that a 20% drop would trigger its debt covenants and put the launch of its new flagship product at risk. Related to this strategy, it needs €40 million of cash flow annually for the next three years for critical capital expenditures and can leave almost none of this cash flow at risk. From the outside perspective, it realizes that there are three main uncertainties or risk drivers over which it has no control: changes in the value of the US dollar, interest rate changes, and market returns on industrial sector equities. Both its business results and its own stock price are strongly correlated with these three risks and could be adversely affected by any of them.

Rather than taking a passive approach as a risk observer, the board in this example uses a top-level analysis to formulate its risk tolerance. In this case, it may decide to limit maximum cash flow variation to €10 million annually and revenue exposure to −10% in a global recession. In addition, it may specify other stated limits, such as the maximum exposure to currency or other risks. This guidance may affect the riskiness of other product strategies that management may pursue. The company may require more expensive financing options to reduce cash flow uncertainty. The governance restrictions may drive risk mitigation programs, such as a hedging strategy, especially for the primary risk drivers that are stated areas of concern.

Governance guidance is important in helping an enterprise target where it should actively pursue risk and where it should mitigate or modify risk. Strategic goals centered on core competencies should be pursued, which leads the company into taking risks that best position the enterprise for success and value creation. Companies sometimes take risks in areas where they have no expertise, which puts their core value creation and their entire organization at peril. A well-functioning risk program would limit or hedge those non-core risks in areas where they have no comparative advantage. Modifying risk is covered in detail in Sections 9–11.

How does a company determine its risk tolerance? There is certainly no formula. Most importantly, a company's goals, its expertise in certain areas, and its strategies will help a board determine which risks the company may pursue and with how much intensity. The ability of a company to respond dynamically to adverse events may allow for a higher risk tolerance. The amount of loss a company can sustain without impairing its status as a going concern should factor into its risk tolerance; some companies are more fragile than others. The competitive landscape matters because both the board's and investors' expectations are usually developed in the context of how a company is positioned in its industry. The government and regulatory landscape is important too, both in their *ex ante* demands on how companies approach risk and in the likely *ex post* reaction in the event of disasters. Quantitative analyses such as scenario analysis, economic models, and sensitivities to macro risk drivers might be used to assess where a board's zone of comfort is bounded. There are other factors that should *not* determine risk tolerance, but in many cases they do. Personal motivations, beliefs, and agendas of board members (the agency problem); company size; whether the market environment seems stable; short-term pressures; and management compensation often affect risk tolerance in ways that might not be in line with the owners' best interests.

Once risk tolerance is determined, the overall risk framework should be geared toward measuring, managing, and communicating compliance with this risk tolerance—getting the risk exposure in line with the enterprise's risk appetite.

This sort of governance exercise not only helps ensure that the organization survives through the worst of times, but also helps ensure a strategic trade-off between risk and return in the decision process, which, in turn, improves potential returns for the given level of risk and value. It is quite easy to find business strategies and investment approaches that produce apparently outsized returns, but they might be at the cost of putting the organization at extreme risk. A somewhat extreme example would be a company selling put options on its own equity, which could produce higher short-term profits but would dramatically increase the chance of the company failing in a steep market decline. Excessive leverage is another risky strategy for boosting short-term profits that may decrease value or lead to failure in the long run. A formal risk governance process with a stated risk tolerance would naturally result in avoidance of many easier, less well-reasoned strategies that entail excessive risk compared with the firm's risk tolerance. Instead, it would lead the strategic discussion into alternative strategies that are more likely to add value while taking reasonable risk within the enterprise's risk tolerance and not simply trade ruin for return. Sincere, good risk governance and risk culture can avoid excessively risky strategies that might put the long-term enterprise value at risk. This approach should produce enhanced value for the enterprise.

6

RISK BUDGETING

e describe risk budgeting and its role in risk governance;

Risk budgeting picks up where risk tolerance leaves off. Whereas risk tolerance focuses on the appetite for risk and what is and is not acceptable, risk budgeting has a more specific focus on how that risk is taken. Risk budgeting quantifies and allocates the tolerable risk by specific metrics; it extends and guides implementation of the risk tolerance decision.

Risk budgeting applies to both business management and portfolio management. Its foundation is the perspective that business or portfolio management involves assembling a number of risk activities or securities, which can be collated into an assemblage of various risk characteristics. For example, a traditional view of a portfolio might be that it is allocated 20% to hedge funds, 30% to private equity, and the remaining 50% is split between stocks and bonds. An alternative risk view of the same portfolio might be 70% driven by global equity returns, 20% by domestic equity returns, with the remaining 10% driven by interest rates. The equity component might be allocated 65% to value and 35% to growth. The portfolio might also have 45% illiquid securities and the remainder liquid. Other allocations can be stated in terms of exposures to inflation, long-term interest rates, currencies, and so on. These multiple dimensions for viewing the allocation of a portfolio are not mutually exclusive: they co-exist. If one is evaluating the risk exposure of a portfolio and trying to keep it in line with a stated risk tolerance, one would be far more concerned with the risk characteristics of the investment assets and portfolio rather than their common classifications of stocks, hedge funds, real estate, private equity, and so on. These terms tell us a little about risk but not enough. Equity is traditionally riskier than hedge funds, but some equities are of quite low risk and some hedge funds are of quite high risk. The risk view may be more meaningful and useful in understanding the portfolio risk than the traditional asset allocation view.

Risk budgeting is any means of allocating a portfolio based on some risk characteristics of the investments. In the purest sense, the term "budget" implies that there is a total risk limit for the organization. Although this approach is not formally required,[10] it would certainly be good practice to have a risk budget that is consistent with the organization's risk tolerance. A risk budget provides a means of implementing the risk tolerance at a strategic level, or in other words, a means of bridging from the high-level governance risk decision to the many management decisions, large and small, that result in the actual risk exposures.

A risk budget can be complex and multi-dimensional, or it can be a simple, one-dimensional risk measure. Even the simplest measure can provide significant benefits in developing an effective risk culture. Four well-known single-dimension measures that are often used are standard deviation, beta, value at risk (VaR), and scenario loss, but there are many others. It is common for some hedge funds to budget risk using standard deviation, managing to a fixed-risk fund target, and evaluating individual investments based on their returns and risks as they affect the *ex ante* standard deviation.

---

[10] One could do risk budgeting even if there were no other risk governance guidance.

More complex forms of risk budgeting use multiple dimensions of risk. One popular approach evaluates risks by their underlying risk classes, such as equity, fixed income, commodity risk, and so on, and then allocates investments by their risk class. Also common are risk factor approaches to risk budgeting, in which exposure to various factors is used to attempt to capture associated risk premiums. An example would be to budget an allocation to give greater emphasis to value stocks based on the belief that they may provide a higher risk-adjusted return than growth stocks. This tactic might be layered over a strategic budget with a certain "beta" as the overall equity risk, supplemented with value and additional factor tilts specified up to some level.

Risk budgeting, although a desirable element of risk governance, cuts across the entire risk management framework, providing a focal point for each of the facets of risk management described in Sections 2 and 3. And although it is true that in practice many organizations operate without a risk budget, it is generally because there has been no specific declaration of their risk tolerance. If a board has a clear understanding of its risk appetite, both the board and management will want some means of implementing a strategic allocation that is consistent with it. Thus, the risk budget becomes a critical overarching construct for the organization's risk framework.

Some individuals may, often through the assistance of a financial planner, engage in some form of risk budgeting, but many do not execute it well or carry it far enough. A classic example of this failure is the tendency of many individuals to invest their financial portfolios in their employers. The risk budget for their total wealth—financial and human capital—is extremely concentrated in one firm and/or one industry. Not surprisingly, such risk budgets typically occur not through formal planning because most formal plans would recognize the problem, but through inaction or inattention.

One major benefit of even the most basic risk budgeting is that it forces risk trade-offs and supports a culture in which risk is considered as a part of all key decisions. Suppose that all the activities a business wants to pursue are in excess of the risk budget. The budgeting of risk should result in an approach, whether explicit or not, of choosing to invest where the return per unit of risk is the highest. Better still, it should also result in a market-benchmarked choice of risk intensity, between possibly doing less of each risky investment or doing more, but with a risk-mitigating hedge. This benefit is extremely important. By choosing between a market hedge or less of a risky investment, one ends up evaluating the investment directly against the market risk–return benchmark. Thus, one is not only comparing risk–return relationships among one's investment choices, but also comparing active versus passive strategies; that is, evaluating investment choices as a whole against the "market return" on a risk-equivalent basis. In other words, one ends up attempting to add active value in each of one's decisions while still staying within the confines of the organization's risk tolerance. The result is even more powerful than merely ensuring that the business is compensated well for the risks they decide to accept. Just having a risk budget in place, forces decision makers to try to add value to the enterprise in every risky decision they make. The risk-budgeting framework makes this consideration innate to the decision process.

EXAMPLE 2

Risk Governance

1 Which of the following approaches is *most* consistent with an enterprise view of risk governance?

- **A** Separate strategic planning processes for each part of the enterprise
- **B** Considering an organization's risk tolerance when developing its asset allocation

- **C** Trying to achieve the highest possible risk-adjusted return on a company's pension fund's assets

2 Which of the following statements about risk tolerance is *most* accurate?

- **A** Risk tolerance is best discussed after a crisis, when awareness of risk is heightened.
- **B** The risk tolerance discussion is about the actions management will take to minimize losses.
- **C** The organization's risk tolerance describes the extent to which the organization is willing to experience losses.

3 Which of the following is *not* consistent with a risk-budgeting approach to portfolio management?

- **A** Limiting the beta of the portfolio to 0.75
- **B** Allocating investments by their amount of underlying risk sources or factors
- **C** Limiting the amount of money available to be spent on hedging strategies by each portfolio manager

4 Who would be the *least* appropriate for controlling the risk management function in a large organization?

- **A** Chief risk officer
- **B** Chief financial officer
- **C** Risk management committee

Solution to 1:

B is correct. The enterprise view is characterized by a focus on the organization as a whole—its goals, value, and risk tolerance. It is not about strategies or risks at the individual business line level.

Solution to 2:

C is correct. Risk tolerance identifies the extent to which the organization is willing to experience losses or opportunity costs and fail in meeting its objectives. It is best discussed before a crisis and is primarily a risk governance or oversight issue at the board level, not a management or tactical one.

Solution to 3:

C is correct. Risk budgeting is any means of allocating a portfolio by some risk characteristics of the investments. This approach could be a strict limit on beta or some other risk measure or an approach that uses risk classes or factors to allocate investments. Risk budgeting does not require nor prohibit hedging, although hedging is available as an implementation tool to support risk budgeting and overall risk governance.

Solution to 4:

B is correct. A chief risk officer or a risk management committee is an individual or group that specializes in risk management. A chief financial officer may have considerable knowledge of risk management, may supervise a CRO, and would likely have some involvement in a risk management committee, but a CFO has broader responsibilities and cannot provide the specialization and attention to risk management that is necessary in a large organization.

IDENTIFICATION OF RISK - FINANCIAL AND NON-FINANCIAL RISK

7

f identify financial and non-financial sources of risk and describe how they may interact;

Having laid the framework for understanding the concept of risk management and risk governance, we now move into the implementation of the process. One of the first important parts of the process is the identification of risks. In this reading, we identify two general categorizations of risks. The first is the set of risks that originate from the financial markets. Accordingly, we refer to this type of risk as **financial risks**. The second group of risks includes those that emanate from outside the financial markets. As such, we refer to these as **non-financial risks**. Although most risks ultimately have monetary consequences, we reserve the term "financial risks" to refer to the risks that arise from events occurring in the financial markets, such as changes in prices or interest rates.[11] In this reading, we will consider the types of financial and non-financial risks faced by organizations and individuals.

7.1 Financial Risks

The risk management industry has come to classify three types of risks as primarily financial in nature. The three primary types of financial risks are market risk, credit risk, and liquidity risk. **Market risk** is the risk that arises from movements in interest rates, stock prices, exchange rates, and commodity prices. This categorization is not to say that these four main factors are the underlying drivers of market risks. Market risks typically arise from certain fundamental economic conditions or events in the economy or industry or developments in specific companies. These are the underlying risk drivers, which we will cover later.

Market risks are among the most obvious and visible risks faced by most organizations and many individuals. The financial markets receive considerable attention in the media, and information on financial market activity is abundant. Institutional investors and many corporations devote considerable resources to processing this information with the objective of optimizing performance. Many individuals also devote considerable attention to market risk, and financial publications and television and radio shows are widely followed in the general population. The state of knowledge in risk management is probably greatest in the area of market risk.

The second primary financial risk is credit risk. **Credit risk** is the risk of loss if one party fails to pay an amount owed on an obligation, such as a bond, loan, or derivative, to another party. In a loan, only one party owes money to the other. In some types of derivatives, only one party owes money to the other, and in other types of derivatives, either party can owe the other. This type of risk is also sometimes called default risk and sometimes counterparty risk. As with market risk, the root source of the risk can arise from fundamental conditions in the economy, industry, or weakness in the market for a company's products. Ultimately, default is an asset-specific risk. Bond and derivatives investors must consider credit risk as one of their primary decision tools.[12] Similar to market risk, credit risk is also a highly visible risk with considerable attention paid to defaults, bankruptcies, and the stresses arising from inadequate cash flow in relation to leverage. Credit risk is a particularly significant risk in that although market prices can go down and bounce back up, defaults and bankruptcies have extremely long-term implications for borrowers.

---

[11] We use the term "financial markets" in a very broad sense. A company may also be exposed to commodity price risk, which we would include as a financial risk.

[12] With certain derivatives (swaps and forwards), either party could be forced to pay off to the other, so each party is concerned about whether its counterparty will pay off, meaning that for some products, credit risk is bilateral.

Although market and credit risk are extremely common risks to institutions, they are also assumed by individuals in their personal investments. One other financial risk, however, is much more common to institutions, although it can be faced by individuals, often unknowingly. This third risk is **liquidity risk**, which is the risk of a significant downward valuation adjustment when selling a financial asset. In order to sell an asset, a party may need to reduce the price to a level that is less than the marked value or the seller's assessment of the asset's true value based on the fundamentals of the asset. In certain market conditions, the seller must make a significant price concession. Having to make price concessions is not necessarily unusual and does not imply a poorly functioning market. Indeed, given no shift in demand, a rightward shift of a supply curve in order to sell a larger quantity is entirely consistent with the notion that a seller must lower the price to sell a greater quantity.

All assets have transaction costs in the market, such as the bid–ask spread. The existence of a sell price that is less than a buy price, however, is not a risk but simply a cost. It is the *uncertainty* of that valuation spread that creates this type of risk. Thus, liquidity risk could also be called "transaction cost risk." The liquidity risk of a $10 stock purchased for $10 is not the risk that one would receive the "bid" price of only $9.99 right after one bought it. That $0.01 spread is a known cost when the stock is purchased, so it is not a risk. The risk is that this spread cost might increase dramatically as a result of either changing market conditions or attempting to maintain a position significantly larger than the normal trading volume for the stock. This problem becomes a serious issue for risk management when the liquidation price falls to less than the seller's estimate of the fundamental value of the asset. Although this risk is often associated with illiquid assets,[13] it really stems from a couple of sources. First, market liquidity varies over time and the market for specific assets may become less liquid; second, as the size of a position increases, the cost and uncertainty associated with liquidating it will increase. In some extreme cases, there may be no price above zero at which the seller can sell the asset.

Of course, one might argue that the cost of illiquidity, and liquidity risk, should thus be part of the investor's assessment of fundamental value, and indeed it is for many analysts. If not, liquidity risk can sometimes be confused with a form of valuation denial in which investors believe that they paid an appropriate price and that the market has not converged to its true value. But less liquidity means a thin market and a lack of investor interest, which may be fertile ground for investment opportunities. Although lack of liquidity can offer benefits, such as the opportunity to buy an asset well before everyone else sees that it is an attractive investment, liquidity risk is generally considered to be a negative factor with which risk managers and indeed all investors must contend.

7.2 Non-Financial Risks

Recall that we refer to financial risks as those arising primarily from events occurring in the financial markets. Although most risks have monetary consequences, there are a number of risks that are typically classified as non-financial in nature. These risks arise from a variety of sources, such as from actions within the organization or from external origins, such as the environment as well as from the relationship between the organization and counterparties, regulators, governments, suppliers, and customers.

---

[13] The illiquid nature of an asset is not itself the risk because that is a direct cost borne immediately upon purchase. Still, uncertainty around the valuation of illiquid assets is a pervasive issue, so it is natural to associate liquidity risk with liquidity characteristics. More importantly, though, the term *liquidity risk* also commonly refers to a much broader set of risks for the organization, which are addressed in the next section.

One important risk of this type is closely related to default risk but deals more with the settling of payments that occur just before a default. This risk is called settlement risk. As an example, suppose Party A enters into a forward contract to purchase ¥200 million of Japanese government bonds from Party B. At expiration if all goes well, Party A would wire the money and Party B would transfer the bonds. Each party fulfills its obligation expecting that the other will do so as well. However, suppose Party A wires the money but Party B does not send the bonds because it has declared bankruptcy. At this point, Party A cannot get the money back, except possibly much later through the potentially slow and cumbersome bankruptcy process.[14] Although the financial consequences are very high, the root source of this risk is the timing of the payment process itself.

Organizations face two types of risks related to the law, and as such, this risk is referred to as legal risk. One risk is simply the risk of being sued over a transaction or for that matter, anything an organization does or fails to do. In financial risk management, however, the major legal concern is that the terms of a contract will not be upheld by the legal system. For example, suppose Bank E enters into a derivatives contract with Party F. Assume that as the underlying changes in price, Party F incurs a loss, whereas there is a corresponding gain to Bank E. But suppose that Party F then identifies a legal issue that it interprets as giving it the right to refuse to pay. If the court upholds Party F's position, Bank E could incur a loss. Litigation always involves uncertainty because even a seemingly weak case can prevail in court.

The following three non-financial risks are related: regulatory risk, accounting risk, and tax risk. They could even be collectively referred to as compliance risk because they all deal with the matter of conforming to policies, laws, rules, and regulations as set forth by governments and authoritative bodies, such as accounting governing boards. Obviously the regulatory, accounting, and tax environment is always subject to change, but the rapid expansion of financial products and strategies in relation to the relatively slow manner in which government and private regulators are able to respond means that laws and regulations are nearly always catching up with the financial world. When these laws and regulations are updated, it can result in significant unexpected costs, back taxes, financial restatements, and penalties.

Another type of non-financial risk is model risk, which is the risk of a valuation error from improperly using a model. This risk arises when an organization uses the wrong model or uses the right model incorrectly. A simple example applicable to both a portfolio manager and a corporate analyst is the assumption of constant dividend growth in the dividend discount model when, in fact, growth is not constant.

Closely related to model risk is tail risk—more events in the tail of the distribution than would be expected by probability models. This risk is a facet of market risk, but it also infects valuations and models when it is ignored or mishandled. Tail risk is known to be especially severe for the normal distribution, which tends to be overused in modeling. As an example, consider the monthly returns on the S&P 500 Index from January 1950 to October 2018. The monthly average return was 0.70%, and the monthly standard deviation was 4.10%. If we rank the monthly returns, we would find that the largest negative return was −21.76%, which occurred in the well-known market crash of October 1987. With a normal distribution, we would find that a return that low would occur only once every 2,199,935 years.[15] The second and third worst monthly returns of –16.94% (October 2008) and –14.58% (August 1998) would occur only once every 6,916 and 654 years respectively. If the normal distribution is a realistic descriptor of returns, results of these magnitudes should *never* have occurred in recorded market history, and yet we have seen three such instances. Interestingly, according to the normal distribution, the largest positive return of 16.30% in October 1974 would occur only once every 888 years. Technically, one could argue that if we go another 2,199,935 years and do not observe a monthly return as low as –21.76%, then the assumption of a normal distribution might seem reasonable, but it seems safe to reject the normal distribution for at least another two million years. Similar comments can apply to the second and third worst returns albeit over shorter periods.

---

[14] This type of risk often arises because of significant time zone differences. Settlement risk is also called Herstatt risk; Herstatt was the name of a German bank that failed in 1974 after receiving "overnight" payments and then defaulting.

[15] This calculation and those that follow are based on determining the probability of the given return or less.

Many quantitative models (e.g., option models) and decision models (e.g., portfolio construction and asset allocation, relying on variances and covariances in analysis and decisions) ignore the existence of fat tails in returns; as a result, market risk is often considered and dealt with in an oversimplified fashion. Tail risk, as the term is used in practice, is important and is discussed separately because financial professionals realize the implicit failure of modeling market risk. More plainly, ignoring tail risk is a form of model risk. And although tail risk might seem more of a financial risk than a non-financial risk, the mistake occurs internally, arising from poor choices made in modeling.

Most of the internal risks faced by an organization are often grouped together and referred to as operational risk. **Operational risk** is the risk that arises inadequate or failed people, systems, and internal policies, procedures, and processes, as well as from external events that are beyond the control of the organization but that affect its operations. Although the factors that give rise to such risks can arise externally, the risks themselves are largely internal to an organization because it would be expected to have its people, systems and internal policies, procedures, and processes functioning effectively regardless of pressures placed on it by external forces.

Employees themselves are major sources of potential internal risks. Banks are keenly aware of the vulnerability to employee theft, given the ease with which so many employees have access to accounts and systems for making entries. But even perfectly honest employees make mistakes, and some can be quite costly. The employee who credits someone's account $100,000 for a $100 deposit may have made an honest mistake, but it is a mistake that could quickly lead to the rapid disappearance of money. In the past, employees up to senior management have been guilty of perpetrating accounting fraud, not necessarily for their own direct benefit but to make the company look better.

In banks and other companies that trade in the financial markets, there is the risk that a trader or portfolio manager will fail to follow laws, rules, or guidelines and put the company at great financial risk. This individual is commonly described as a "rogue trader." Personified by Nick Leeson of Barings Bank, who in 1995 destroyed the 200-year old company by engaging in a series of highly speculative trades to cover up losses, the rogue trader has become a standard concern of risk managers. Although it was never clear if Leeson's trades were truly unauthorized, his legacy left the fear that institutions bear the risk that one trader can imperil the entire organization by making large and highly speculative trades that put the bank's entire capital base at risk. In essence, a rogue trader is a trader who engages in risky transactions without regard for the organization's limits or conforming to its controls.

Organizations are also threatened by business interruptions, such as those caused by extreme weather and natural disasters. Events such as floods, earthquakes, or hurricanes can cause significant damage and temporarily shut down an organization. Although extreme weather and natural disasters are external forces that are completely out of the control of an organization, it does not excuse the organization from having the appropriate internal procedures for managing problems caused by their external environment. Simple and fairly low-cost actions, such as having generators, backup facilities, or providing employees the option to work remotely, can go a long way toward keeping employees working during extreme weather events and when natural disasters strike. Yet, some organization have not heeded inclement weather forecasts. Failing to react to warnings can result in considerable loss.

In a world that is increasingly digital, cyber risk is a major operational risk that spares no organization and that can have significant consequences. Organizations are expected to understand and manage the risk associated with the disruption of or failure related to their information technology (IT) systems. For example, a hacker breaking into a company's IT system and stealing customer or client data is an external threat. Hacking, however, is not simply a random act of mischief. Companies are aware of the threat of hackers, and hackers can break in to a system only if that system is vulnerable. An organization is responsible for ensuring cyber security and establishing sufficiently robust IT safeguards, such as data encryption, to deter hackers from breaking in and either stealing or causing disruption. Cyber-attacks and data breaches can have serious reputational and compliance consequences. For example, all organizations targeting European citizens must comply with the General Data Protection Regulation (GDPR) and notify regulators and data subjects of any data breaches regarding sensitive personal information within 72 hours. Failure to do so can lead to fines of several million euros, including for organizations based outside the European Union. In addition to the threats posed by hackers and viruses, even secure IT systems themselves are a particular source of risk. Programming errors and bugs can create the possibility of costly mistakes.

Terrorism is another form of operational risks that poses a threat to organizations and individuals. The 1993 attacks on the World Trade Center led many companies to recognize that the New York City financial district was a major terrorist target and that, as such, their operations could be shut down by these acts of violence. When the more destructive attacks of 11 September 2001 occurred, many organizations had already established backup facilities sufficiently far away from that area. Of course, such risk is not confined to major financial centers, and indeed, organizations worldwide have begun to take security measures that address this operational risk.

Some of these operational risks are insurable, at least to a modest extent. We will briefly discuss insurance later, but most companies would much prefer to take proactive steps toward prevention than to incur the inconvenience of losses and then have an outside organization compensate them for their losses.

Solvency risk is the risk that the organization does not survive or succeed because it runs out of cash, even though it might otherwise be solvent.[16] This was probably the most underappreciated component of risk prior to the financial crisis of 2008.[17] The collapse of Lehman Brothers was often associated with an excess of leverage, which was certainly a key factor in its failure. But it was solvency risk that forced the company into bankruptcy. Almost overnight, Lehman's liquidity disappeared because most funding sources would no longer willingly bear Lehman's counterparty risk. Even if it had experienced large market gains on the day it went under, it had already been destroyed by solvency risk. Across the entire financial industry, from hedge funds to pension funds, painful but valuable lessons were learned about the critical importance of funds availability and solvency risk, even if all other risks were well-aligned.

---

[16] Solvency risk is often referred to as liquidity risk by industry professionals, even though the expression *liquidity risk* was used earlier to refer to the risk of valuation shock when selling a security. Although the term "liquidity risk" is used in practice in both contexts, in this reading we will refer to the risk relating to the cash position of an organization as "solvency risk."

[17] Bank runs are perhaps the simplest example of solvency risk. An otherwise solvent bank can easily be ruined by a bank run that wipes out its ability to make good on short-term liabilities.

Solvency risk is now viewed as one of the key factors in running a successful hedge fund because investors are extremely sensitive to not recovering their investment in the event of a "run on the fund."

Solvency, in the personal or institutional sense, is the availability of funding to continue to operate without liquidating—or at a less extreme level, to be able to make good on liabilities and meet one's cash flow requirements. Solvency risk is the ultimate example of the importance of taking an enterprise view of risk management. For example, a university's investment officer might have a perfectly well-balanced set of risks in the endowment portfolio when viewed in isolation. But as a part of a university, the portfolio may be affected by a deep recession because the university's professional degree revenue, grant money, and donations will fall at the same time as the portfolio's investment value and cash distributions are in decline. Although the endowment and university may survive, it might be necessary for the endowment to take many emergency actions that impair its value, simply attributable to the overall solvency risk and the ultimate need of the enterprise to not run out of cash.

Solvency risk is easily mitigated, though never eliminated, by a large number of possible safeguards, none of which is free. Many businesses produce short-term higher returns by essentially ignoring solvency risk, but in doing so, they are not managing risk very well. Since the 2008 crisis, most businesses are keenly aware of the consequences of bad solvency management, and have taken such steps as using less leverage, securing more stable sources of financing, investing in models to provide more transparency on solvency risk, incorporating solvency risk at an enterprise level in risk governance, and holding more cash equivalents and assets with less liquidity risk.

Individuals can also face a number of risks of an operational nature. These include hackers breaking into one's computer and the threat of burglary and robbery. One of the most commonly cited risks for individuals is identity theft. For individuals, however, we consider their primary non-financial risks to be related to their life and health as well as other life-changing events.

Obviously, the health of an individual is an extremely important risk. Poor health can result from poor choices in life, but it can also arise from factors that are outside the control of the individual. These risks can result in direct health care expenses, reduced income because of disability, and reduced lifespan or quality of life. People vary widely in the risk management strategies they undertake to control their health, such as in their choices in diet, exercise, preventive health care, and avoidance of undue health risks. Some individuals address only their financial exposure to health risks, and still others do not take proactive steps to address this risk at all.

Closely related to health risk is mortality risk—the risk of dying relatively young—and longevity risk—the risk of outliving one's financial resources. Not only are these risks a primary determinant of the quality of life, they are also critical factors in investment planning. Although it is probably desirable not to know when one will die, financial planning for one's years in retirement is heavily dependent on one's mortality assumption. Insurance companies, defined benefit pension plans, and vendors of retirement annuities need only know the group average mortality. Mortality tables are reasonably accurate, so these institutions have relatively precise estimates of death rates for groups as a whole. Individuals themselves, however, clearly do not know how long they will live. People who use defined contribution plans must therefore build portfolios and control retirement distributions so that their assets outlive them, which is difficult to do when they do not know when they will die. No one wants to outlive their money, but with an increasingly aging population and good health care, this problem is becoming a greater concern.

There are a number of other major non-financial risks that individuals face, which are generally involved with some sort of life-changing disaster. The largest ones—fire, natural disaster, or massive liability stemming from harming others, such as in a car accident—are generally considered “property and casualty” risks and are insured as such.

IDENTIFICATION OF RISK - INTERACTIONS BETWEEN RISKS

8

f identify financial and non-financial sources of risk and describe how they may interact;

In some cases, a risk classified into one category could easily have been classified into another. Indeed, the interactions between risks are numerous. It has been said that market risk begets credit risk, which begets operational risk. That is, given unexpected market moves, one party then owes the other party money. Given the debtor–creditor nature of the relationship, the two parties must have internal operations that process the transactions and pay or collect the money. Thus, whenever there is credit risk, there is settlement risk. If there were no market risk, the other risks in the chain would likely be relatively minor. Legal risk often arises from market or credit risk. Large market moves create losses for one party. There is a long history of parties searching for loopholes in contracts and suing to avoid incurring the loss.

One simple example of an adverse risk interaction is counterparty risk. When trading a derivative contract, it is important to consider the cost of counterparty risk. Suppose Party A buys an out-of-the-money put option with a strike price of ¥1000—a contract theoretically worth ¥100 entitling him to as much as ¥1000 from Counterparty C if an underlying equity index is down. But there is a 2% chance that C could default; and assume that the possibility of default is considered independent of the performance of the equity market. This transaction, with payoffs adjusted for the possibility of default, might price at, say, ¥98 to A. But in reality, the credit risk of C’s default is likely dependent on the equity market return. If the probability distribution of default risk overlaps substantially with that of the market being down, which is a likely scenario, then the risks interact, and the cost of risk is higher. In this example, perhaps the probability of C defaulting is 10% or more when the put option is in the money. So, A’s expected payoff is lower as a result of facing a credit risk that is compounded by market risk. In fact, it is quite likely that in the extreme event—a deep decline in the equity market when A would presumably receive ¥1000—Party A will in all likelihood get nothing. Thus, the investor bears much more risk than initially thought as a result of the failure to consider the interaction of the two risks. And in doing so, Party A overpaid for the contract. This sort of risk interaction is so common in markets that practitioners have given it a very fitting term—“wrong-way risk.” In fact, it was extremely common in the financial crisis of 2008, when holders of many securities based on mortgage credit believed that the risks were well-diversified when in truth, the risks were quite systematic.

Another example of interacting risks was experienced by many banks, funds, and private investment partnerships in 2008, as well as the hedge fund Long-Term Capital Management in 1998. Leverage, which manifested itself in higher market risk, interacted in an extremely toxic manner with liquidity risk and solvency risk and impaired or shuttered many investment firms.[18]

In most adverse financial risk interactions, the whole is much worse than the sum of its parts; the combined risk compounds the individual risks in a non-linear manner. For this example, a 2× levered organization might produce a 2% loss when its unlevered twin or baseline risk bears a 1% loss. If liquidity is a serious issue for the organization, then at a 10% baseline loss, the organization might face some moderate distress from liquidity or funding problems that it ends up losing 25% instead of 20%. It would not be surprising if this organization failed at a 30% baseline loss because of the toxic interplay between levered risk and liquidity problems. This resulting non-linear reaction to risk drivers exists across many risk interactions in many markets, making up-front scenario planning even more valuable to the risk process, a point we will return to later.

Earlier, we briefly described a common example of interacting risks for individuals. Suppose an individual works for a publicly traded company and, through an incentive program, receives shares of the company in her company retirement portfolio or for her personal holdings. Company policies may require that employees hold on to these shares for a number of years. When that time has elapsed, however, many individuals fail to recognize the incredibly concentrated risk they are assuming, so they hold on to their shares. An employee's reasoning for not selling the shares is often that the company she works for has been a solid performing company for many years, so she feels no reason to worry. Moreover, the team spirit often imbued in employees generates pride that can make employees believe that there is no better place in which to work and to invest their money. But if something goes wrong in the company or the industry, the employee may lose her job *and* her savings—an incredibly adverse interaction between market risk and human capital risk. The 2003 collapse of Enron remains a powerful historical example, with many loyal and honest employees losing virtually all of their retirement savings by failing to recognize this risk.

In sum, it is important to recognize that risks do not usually arise independently, but generally interact with one another, a problem that is even more critical in stressed market conditions. The resulting combined risk is practically always non-linear in that the total risk faced is worse than the sum of the risks of the separate components. Most risk models and systems do not directly account for risk interactions, which makes the consequences of the risk interaction even worse. Governance bodies, company management, and financial analysts should be keenly aware of the potential risk and damage of risks in combination, and be aware of the dangers of treating risks as separate and unrelated.

EXAMPLE 3

Financial and Non-Financial Sources of Risk

1 Which of the following is *not* a financial risk?

A Credit risk

B Market risk

---

[18] This example illustrates yet another risk, systemic risk, that is a significant concern to regulators and governments. Stresses and failures in one sector transmit to stresses and failures in other sectors, which can ultimately impact an entire economy. Systemic risk is the ultimate example of interactions among risks.

- **C** Operational risk

2 Which of the following *best* describes an example of interactions among risks?

- **A** A stock in Russia declines at the same time as a stock in Japan declines.
- **B** Political events cause a decline in economic conditions and an increase in credit spreads.
- **C** A market decline makes a derivative counterparty less creditworthy while causing it to owe more money on that derivative contract.

3 Which of the following *best* describes a financial risk?

- **A** The risk of an increase in interest rates.
- **B** The risk that regulations will make a transaction illegal.
- **C** The risk of an individual trading without limits or controls.

4 Which of the following is *not* an example of model risk?

- **A** Assuming the tails of a returns distribution are thin when they are, in fact, fat.
- **B** Using standard deviation to measure risk when the returns distribution is asymmetric.
- **C** Using the one-year risk-free rate to discount the face value of a one-year government bond.

5 Which of the following is the risk that arises when it becomes difficult to sell a security in a highly stressed market?

- **A** Liquidity risk
- **B** Systemic risk
- **C** Wrong-way risk

6 The risks that individuals face based on mortality create which of the following problems?

- **A** The risk of loss of income to their families.
- **B** Covariance risk associated with their human capital and their investment portfolios.
- **C** The interacting effects of solvency risk and the risk of being taken advantage of by an unscrupulous financial adviser.

Solution to 1:

C is correct. Operational risk is the only risk listed that is considered non-financial, even though it may have financial consequences. Credit and market risks derive from the possibility of default and market movements, respectively, and along with liquidity risk, are considered financial risks.

Solution to 2:

C is correct. Although most risks are likely to be interconnected in some way, in some cases the risks an organization is exposed to will *interact* in such a way that a loss (or gain) in one exposure will lead directly to a loss in a different exposure as well, such as with many counterparty contracts. Conditions in A and B are much more directly linked in that market participants fully expect what follows—for example, in B, an outbreak of war in one region of the world could well cause widespread uncertainty; a flight to quality, such as to government-backed securities; and a widening in spreads for credit-risky securities. In C, in contrast, the reduction in creditworthiness following the market decline may be expected, but owing more money on an already existing contract as a result comes from the interaction of risks.

Solution to 3:

A is correct because this risk arises from the financial markets.

Solution to 4:

C is correct. The risk-free rate is generally the appropriate rate to use in discounting government bonds. Although government bonds are generally default free, their returns are certainly risky. Assuming a returns distribution has thin tails when it does not and assuming symmetry in an asymmetric distribution are both forms of model risk.

Solution to 5:

A is correct. Securities vary highly in how liquid they are. Those with low liquidity are those for which either the number of agents willing to invest or the amount of capital these agents are willing to invest is limited. When markets are stressed, these limited number of investors or small amount of capital dry up, leading to the inability to sell the security at any price the seller feels is reasonable. Systemic risk is the risk of failure of the entire financial system and a much broader risk than liquidity risk. Wrong-way risk is the extent to which one's exposure to a counterparty is positively related to the counterparty's credit risk.

Solution to 6:

A is correct. The uncertainty about death creates two risks: mortality risk and longevity risk. The mortality risk (risk of dying relatively young) is manifested by a termination of the income stream generated by the person. In contrast, longevity risk is the risk of outliving one's financial resources.

9

MEASURING AND MODIFYING RISK - DRIVERS AND METRICS

g describe methods for measuring and modifying risk exposures and factors to consider in choosing among the methods.

The core element of risk management is the measurement and modification of risk. One cannot modify risk without measuring it. The primary purpose of measuring risk is to determine whether the risk being taken, or being considered, is consistent with the pre-defined risk tolerance. To understand how risk is measured, it is important to understand the basic elements that drive risk.

9.1 Drivers

This section illustrates the origins of risk. Risk is a part of life itself. None of us knows from one day to the next everything that will happen to us in the next 24 hours, let alone over a longer period. We may get a phone call that a relative is extremely sick, or we may be contacted by a head-hunter about an attractive job possibility. We may learn that we are going to be given an award from a prestigious organization, or we may find that our identity has been stolen. All of us can almost surely name something that happened the previous day that was not anticipated. Most of these happenings are minor and often quickly forgotten. Others are serious. Some are good. Some are bad. Some are unpredictable outcomes of known events, such as whether we get an offer following a job interview or whether a medical test reveals that we are healthy or ill. Some events are completely unanticipated, such as getting a phone call from an old friend we have not talked to in many years or having a flat tire on the drive home. Fortunately, the vast majority of risks in life are minor. The ones that are not minor, however, have the potential to be highly unpredictable and financially, and sometimes physically and emotionally, quite costly.

In a conceptual sense, financial risks are no different from the other risks we face in life. All risks arise from the fact that the future is unknown. Financial risks largely emanate from economic risks, and economic risks emanate from the uncertainties of life.

Financial markets generate prices that fluctuate as investors absorb information about the global and domestic state of the economy, the company's industry, and the idiosyncratic characteristics of the company itself. Global and domestic macroeconomies are driven by the companies that operate within them, but much of the tone as well as the ground rules are set by governments and quasi-governmental agencies, such as central banks. Taxes, regulations, laws, and monetary and fiscal policy establish a legal and economic environment and a set of ground rules that greatly affect the degree and quality of economic activity that takes place. Attempts by governments and central banks of different countries to coordinate economic policies can lead to some degrees of success if harmonized, but if not, they can create an environment in which companies engage in practices designed to seek favorable treatment in some countries and avoid unfavorable treatment in others.[19]

All economies, in turn, are composed of industries. Government policies also affect industries, in some cases encouraging economic activity in some industries while discouraging it in others. Some industries are stable, weathering macroeconomic storms quite well, whereas others are highly cyclical.

The uncertainties of global and domestic macroeconomic and central bank policies create risks for economies and industries that we often treat as systematic. Seemingly minor events, such as filling the position of central bank chairperson, are often viewed by investors as major events, signaling possibly a change in policy that can greatly affect the macroeconomy and possibly certain industries.

Moving down to a more fundamental level, investors face the unsystematic or idiosyncratic risks of individual companies. Modern investment analysis prescribes that diversified portfolios bear no unsystematic risk. We are then led to believe that unsystematic risk does not matter in a well-diversified portfolio. But unsystematic risk does matter to the management of a company. It also matters to poorly diversified investors. And it certainly matters to the financial analysts who cover specific companies. And what would appear to be unsystematic risk can oftentimes actually be systematic. For example, poor credit risk management by a major bank can turn into a global financial crisis if that bank is "too big to fail."

In sum, the basic drivers of risk arise from global and domestic macroeconomies, industries, and individual companies. Risk management can control some of this risk, but it cannot control all of it. For example, the risk manager of a company may be able to reduce the likelihood that his company will default, but he cannot control movements in interest rates. For the latter risk, he must accept that interest rate volatility is a given and that he can only position the company to be able to ensure that its risk exposure is aligned with its objective and risk tolerance. In order to do so, he must first be able to measure the risk.

---

[19] This practice is sometimes called regulatory arbitrage. The policies of certain countries can be more conducive to establishing operations. Examples are the flow of money into countries whose banking laws are less restrictive and more conducive to secrecy and incorporation in or moving a company to a country with lighter regulations or more favorable tax treatment.

9.2 Metrics

The notion of metrics in the context of risk refers to the quantitative measures of risk exposure. The most basic metric associated with risk is probability. Probability is a measure of the relative frequency with which one would expect an outcome, series of outcomes, or range of outcomes to occur. One can speak about the probability of rolling a six in one roll of a die as 1/6, the chance of rain in the next 24 hours as 20%, or the odds of a central bank taking actions to increase interest rates of 50%. These are all probabilities, differing in concept by the fact that the die roll is associated with an objective probability measure, whereas the other two examples are subjective probabilities. It is important to note that probability, in and of itself, is not a sufficient metric of risk. A chance of financial loss of 25% does not tell us everything we need to know. There are other measures of risk that incorporate probability but give us more information.

The standard deviation is a measure of the dispersion in a probability distribution. Although there is a formal mathematical definition of standard deviation, at this point we need only understand the conceptual definition. Standard deviation measures a range over which a certain percentage of the outcomes would be expected to occur. For example, in a normal distribution, about 68% of the time the outcomes will lie within plus or minus one standard deviation of the expected value. Two standard deviations in both directions would cover about 95% of the outcomes, whereas three would encompass 99% of the outcomes. Although standard deviation, or volatility, is widely used in the financial world, it does have significant limitations. In particular, standard deviation may not be an appropriate measure of risk for non-normal distributions. Standard deviation may not exist for return distributions with fat tails.

Moreover, according to modern portfolio theory, the risk captured by an asset's standard deviation overstates the risk of that asset's returns in the context of a diversified portfolio. Investors can easily diversify their holdings, thereby eliminating a portion of the risk in their portfolios by diversifying away the security-specific risk. As a result, most financial valuation theories assert that the ability of investors to eliminate security-specific risk, or non-systematic risk, means that investors should not expect to earn a premium to compensate them for the assumption of this risk. As a consequence, the risk of a security may be better measured by its **beta**, a measure of the sensitivity of a security's returns to the returns on the market portfolio. Beta measures relative risk, meaning how much market risk an asset contributes to a well-diversified portfolio.[20]

Beta describes risk well for a portfolio of equities, but other sources of risk may require other descriptive risk metrics. The risk associated with derivatives is one example of this. Although derivatives are widely used to manage risk, they do so by assuming other risks. Even if the derivative is being used to establish a hedge of an existing exposure to risk, it would still result in the assumption of additional risk because the assumed risk is being used to offset an existing risk. For example, if one purchases a call option denominated in euros to buy Russian rubles, one would be assuming the risk of the ruble/euro exchange rate. Because most derivatives exposures are highly leveraged, it is critical that the risk of derivatives be properly measured. There are several specialized measures of derivatives risk.

---

[20] Earlier, we discussed the fact that unsystematic risk matters to some parties. Here we seem to be saying that it should not matter to anyone. Capital market models almost always assume that investors can diversify quite easily and, as a result, they should not expect to earn a premium for bearing diversifiable risk. This assumption does not mean that everyone's wealth is well-diversified. Investors who do not diversify probably cannot expect to earn a return for bearing diversifiable risk, but it does not mean that these investors should not care about measuring the risk they choose to assume by not diversifying.

The sensitivity of the derivative price to a small change in the value of the underlying asset is called the **delta**. It is perhaps the most important measure of derivatives risk. Yet delta is limited to capturing only small changes in the value of the underlying. Large changes are captured by the concept of **gamma**. Whereas delta is a first-order risk, gamma is considered a second-order risk because it reflects the risk of changes in delta.[21] Some derivatives, such as options, are also sensitive to changes in the volatility of the underlying. This risk is captured by a concept called **vega**, which is a first-order measure of the change in the derivative price for a change in the volatility of the underlying. Derivatives are also sensitive to changes in interest rates, which are reflected in a measure called **rho**. Most options have relatively low sensitivity to interest rates.[22] These, and other mathematically derived derivatives metrics, are collectively referred to as “the Greeks.”

Other asset classes may have their own special metrics to describe risk. One well-known example, **duration**, is a measure of the interest rate sensitivity of a fixed-income instrument. Analogous to delta, it is a first-order risk. The wide variety of financial instrument types and asset classes leads to a proliferation of terminology and risk measures, with most of them having no meaning outside their asset class. As financial organizations and asset risk modeling became more sophisticated and computer power increased, an approach was needed to measure and describe financial risk across the broad spectrum of asset classes. Spurred by the onset of global bank capital regulation, this led to the development of value at risk.

Value at risk or **VaR** is a measure of the size of the tail of the distribution of profits on a portfolio or for an organization. A VaR measure contains three elements: an amount stated in units of currency, a time period, and a probability. For example, assume a London bank determines that its VaR is £3 million at 5% for one day. This statement means that the bank expects to lose a minimum of £3 million in one day 5% of the time. A critical, and often overlooked word, is *minimum*. In this example, the bank expects that its losses will be at least £3 million in one day with 5% probability. In a VaR measure, there is no ultimate maximum that one can state. VaR is thus a minimum extreme loss metric. With a probability of 5% and a measurement period of one day, we can interpret the bank’s VaR as expecting a minimum loss of £3 million once every 20 business days. VaR can also be used to measure credit losses, although the construction of the measure is considerably more difficult given the extreme asymmetry of the risk.

VaR is a simple but controversial measure. There are several ways to estimate VaR, each of which has its own advantages and disadvantages. The different measures can lead to highly diverse estimates. Moreover, VaR is subject to the same model risk as derivative pricing models. VaR is based on a particular assumption about the probability distribution of returns or profits. If that assumption is incorrect, the VaR estimate will be incorrect. VaR also requires certain inputs. If those inputs are incorrect, the VaR estimate will be incorrect. Many critics of VaR have argued that naive users of VaR can be lulled into a false sense of security. A presumably tolerable VaR can give the illusion that the risk is under control, when in fact, it is not. Yet, VaR is accepted as a risk measure by most banking regulators and is approved for disclosure purposes in typical accounting standards. As with any risk measure, one should supplement it with other measures.

---

[21] The notion of a first-order risk versus a second-order risk can be seen by considering the following. Suppose A affects B and B, in turn, affects C. A does not affect C directly but does so only indirectly. A is a first-order risk for B and a second-order risk for C.

[22] Options on interest rates, however, have a high sensitivity to interest rates, but only because interest rates are the underlying, and thus, the source of market risk.

As emphasized earlier, VaR does not tell the maximum loss. The maximum loss is the entire equity of an organization or the entire value of a portfolio, but the statistics used to estimate VaR can be used to gauge average extreme losses. Conditional VaR or **CVaR** is a common tail loss measure, defined as the weighted average of all loss outcomes in the statistical distribution that exceed the VaR loss. Another tail risk metric in the credit risk space that is analogous to CVaR is expected loss given default, which answers the question for a debt security, “If the underlying company or asset defaults, how much do we lose on average?”

VaR focuses on the left tail of the distribution and purports to tell us the expected frequency of extreme negative returns, but it can understate the actual risk. For example, the normal distribution gives us a well-defined measure of extreme negative returns, which are balanced by extreme positive returns. Yet, actual historical return distributions have shown that there are more extreme negative returns than would be expected under the normal distribution. We previously described this concern in the form of tail risk. In response to this concern, statisticians have developed a branch of study that focuses primarily on extreme outcomes, which is called **extreme value theory**, and leads to measures of the statistical characteristics of outcomes that occur in the tails of the distribution. There are mathematical rules that define the statistical properties of such large outcomes, and these rules have been widely used for years in the insurance business. In the past 20 years or so, risk managers have taken to using them to help gauge the likelihood of outcomes that exceed those that would normally be expected.

Two measures in particular that are often used to complement VaR are **scenario analysis** and **stress testing**. These are common sense approaches that ask “If this happens, then how are we affected?” Scenario analysis can be thought of as a package of stress tests, usually with some common underlying theme. A scenario defines a set of conditions or market movements that could occur and would put some pressure on a portfolio. An example might be a sharp increase in interest rates coupled with a significant decline in the value of a currency. The portfolio is then evaluated to determine its expected loss under these scenarios. A different means of posing a scenario analysis is stress testing, which is done by proposing specific asset price moves generally involving extremely large and high pressure scenarios that would occur only rarely but would have the potential for destabilizing the entire organization. The US Federal Reserve and other central banks have begun requiring major banks to stress test their portfolios. Although scenario analysis and stress testing can provide some information, they are, as noted previously for other measures, subject to model risk.

Of course, the measures just mentioned focus primarily on market risk. Credit risk, which is covered in more detail in readings on fixed-income analysis, has long relied heavily on the credit ratings provided by private companies, such as Moody’s Analytics, Standard & Poor’s, and Fitch Ratings. In effect, a large part of credit analysis for many lenders has been outsourced since the early part of the 20th century. Most lenders, however, do not rely exclusively on these rating companies. They do their own analysis, which focuses on the creditor’s liquidity, profitability, and leverage. Liquidity measures, such as the current ratio, may indicate how well a borrower can cover short-term obligations. Solvency ratios, such as cash flow coverage or interest coverage, may reveal whether a borrower generates enough cash or earnings to make its promised interest payments. Profitability measures, such as return on assets, estimate whether a company is sufficiently profitable so that it can easily accommodate debt. Leverage measures, such as the ratio of debt to total assets, reflect whether a company has sufficient equity capital in relation to its debt to absorb losses and negative cash flows without defaulting. Credit analysis also examines the strength and cyclicality of the macroeconomy and the company's industry. Other widely used measures of credit risk include credit VaR, probability of default, expected loss given default, and the probability of a credit rating change.

One of the problems of credit risk measurement is that credit events, such as a ratings downgrade or a default, are relatively rare for a particular organization. Certainly, in the aggregate there are many credit losses, but very few companies that default have a history of defaulting. Without a history to go by, estimating the likelihood of an event that has never actually occurred is extremely difficult. Imagine the challenge of assigning a default probability to Lehman Brothers in 2007. It had been in operation since 1850 and had never defaulted. Yet in 2008, Lehman Brothers, one of the most successful financial companies of all time, filed for bankruptcy. Because of the infrequency of default, risk managers normally attempt to assess default probability by aggregating companies with similar characteristics.[23]

Another useful source of information for risk managers about these rare events is the *ex ante* risk cost that is implied by the market pricing of derivatives. A **credit default swap (CDS)** on an issuing company has an observable price that acts as a signal to a bondholder of the risk cost of a default. Put options, exotic options, insurance contracts, and other financial instruments may contain valuable signals of the cost of rare adverse events, or at least the price of hedging them.

Operational risk is one of the most difficult risks to measure. Consider the operational risk event reported in 2014 in which hackers broke into Home Depot's credit card data base. Assessing the likelihood of such an event and estimating the potential losses would be almost impossible. The threat of litigation alone for years afterward is difficult to quantify. As with credit risk, significant operational risk events are rare but usually quite costly if they do occur. Hence, attempts to quantify the risk usually involve a third party aggregating operational risk events across numerous companies and publishing the statistics.

As mentioned, there are numerous other risks that would likewise be difficult to measure. For example, there is always the possibility of changes in accounting rules, laws, tax rates, and regulatory requirements that can result in significant costs as companies adapt their policies and actions from one regulatory environment to a new one. How would one measure such risks? Moreover, the time period spanned by these risks is extremely long, and in fact, theoretically infinite. Changes in these rules and laws are often motivated by politics. How does one quantify such risks when there are no real numeric measures? Analysis invariably reverts to subjective evaluation of the likelihood of such threats and their potential losses.

As we have described, many risks are measurable, at least on an *ex post* basis. Market-related risks are blessed with large quantities of data, so they are relatively measurable. Credit, operational, and other risks are somewhat rare events. Although it is probably a good thing that such events are indeed rare, their infrequency makes measurement more difficult. Nonetheless, virtually any risk manager will attempt to obtain at least a subjective sense of the expected frequency, likelihood, and cost of these events. With either objective or subjective measurements in mind, risk managers can then proceed to modify these risks accordingly.

---

[23] In some sense, aggregating companies with similar characteristics is what credit ratings do. Companies rated BAA/Baa+ can be quite diverse but all are considered similar with respect to their ability to pay their debts.

10 METHODS OF RISK MODIFICATION - PREVENTION, AVOIDANCE, AND ACCEPTANCE

g describe methods for measuring and modifying risk exposures and factors to consider in choosing among the methods.

The notion of risk modification presumes that an analysis has been conducted in the risk governance stage that defines how much risk is acceptable. Coupled with measurements of the actual risk, as discussed in the previous section, the risk manager then proceeds to align the actual risk with the acceptable risk.

It is important to understand, however, that risk modification is not strictly risk reduction. For example, a portfolio with the strategic objective of maintaining a 50/50 split between equity and cash will naturally find that in a market in which cash outperforms equity, the split between equity and cash will tilt toward cash. Thus, the portfolio becomes less risky. Beyond a certain point, the risk of the portfolio is unacceptably low given the return target. Thus, risk modification would take the form of rebalancing by increasing the risk. For the most part, however, risk management focuses more on reducing the risk. Risk reduction is commonly referred to as hedging. A hedge is a transaction put in place to reduce risk. Some hedges are designed to lead to the ultimate in risk reduction—the complete elimination of risk. Others are simply designed to lower the risk to an acceptable level.[24] For some companies, risk management is primarily concerned with keeping the organization solvent. Regardless of the focus, much of what is done to manage risk is the same. In this section, we will examine four broad categories of risk modification: risk prevention and avoidance, risk acceptance, risk transfer, and risk shifting.

10.1 Risk Prevention and Avoidance

One method of managing risk is taking steps to avoid it altogether; however, avoiding risk may not be as simple as it appears. It is difficult to completely avoid risk, but more importantly, it is unclear that every risk should be completely avoided particularly if there are high costs associated with eliminating the risk. Instead we choose a trade-off between cost and benefits. The actual trade-off may be subject to debate because risk assessment and risk management are subject to variation from one person to another.

We could nearly eliminate the risk of being injured or killed in an automobile accident if we choose to never drive or ride in a car. Like any risk-avoidance strategy, however, there would more than likely be a trade-off in terms of the loss of the benefits provided by the activity. We could try to protect our children from all harm, but that may come at the expense of preparing them poorly for adult life. We could invest our entire retirement savings in cash, but would most likely give up protection against inflation and lose out on the opportunity to benefit from long-term economic growth and the performance of investable assets that benefit from that growth.

Insurance companies rely heavily on the techniques of risk prevention and avoidance. An automobile insurance company would prefer that their policyholders never drive their cars. Although it cannot prohibit them from doing so, it can reward them with lower premiums if they drive less and have safe driving records. A life insurance company would prefer that their policyholders do not smoke, and it can reward nonsmokers with lower rates.

---

[24] For example, in the case of the portfolio with a strategic target of 50/50 equity and cash, if equity outperforms cash, the portfolio will tilt toward equity. At some point, a risk-reducing strategy would then be in order. This type of hedge would reduce the risk but not eliminate it.

Nearly every risk we take has an upside, at least as perceived by the person taking the risk. Some counterexamples might seem to belie this point, but not if viewed from the point of view of the risk taker. One could argue that there are no benefits from smoking, but people who smoke may have the opinion that the pleasure they receive exceeds the costs. Casino gambling incurs the risk of significant financial loss and addiction, but it is risk that is acceptable to the consumers who incur it relative to the perceived benefits they receive. The risks of extreme sports, such as skydiving, would seem to be exceeded by the benefits obtained by participants, and yet participants engage in them with apparently much enjoyment. People undertake all types of risky behaviors because they obtain commensurate benefits. These examples are simply cases in which the decision maker chooses to bear a certain degree of risk. They are conceptually the same as an investor who chooses to accept a relatively high degree of risk. Likewise, those who live their lives engaging in very few risky activities are conceptually the same as the investor who keeps only a modest exposure to risky assets.

In organizations, the decision to avoid risk is generally made at the governance level as a part of setting the risk tolerance. Boards will often decide that there are some business or investment activities simply not worth pursuing based on either the goals of the organization or the perceived risk–return trade-off. These are strategic decisions. Boards may exclude some areas or activities to allow management to focus on choosing risks in other areas where they presumably have a better chance of adding value.

We recap this section by noting that risk prevention and avoidance is simply an element of the decision of how much risk to accept, given the trade-off between the risk of loss and the benefit of gain. This could be a direct benefit or an indirect benefit of avoiding or eliminating a risk. Most decisions in life involve a trade-off between benefits and costs, neither of which is necessarily easy to measure. Thus, risk management is an ongoing process of fine-tuning exposure to achieve the level of risk that is desired in relation to the benefits.

If the risk measurement process shows that the risk exceeds the acceptable level, there are three approaches to managing the risk: self-insuring, risk transfer, and risk shifting.

10.2 Risk Acceptance: Self-Insurance and Diversification

In many cases, from both a risk tolerance and a strategic standpoint, it makes sense to keep a risk exposure—but to do so in the most efficient manner possible. Self-insurance is the notion of bearing a risk that is considered undesirable but too costly to eliminate by external means. In some cases, self-insuring means simply to bear the risk. In other cases, it may involve the establishment of a reserve to cover losses. Choosing to not have health insurance can be an optimal choice for some young, healthy adults without responsibility for children. Setting aside some money to cover potential health costs completes the picture of an individual who completely self-insures. Similarly, a young healthy individual who does not buy life insurance but engages in a systematic, well-conceived savings and investment plan is engaging in self-insurance.

One must be careful with this approach, however, because there is a fine line between self-insurance and denial. To the extent that self-insurance results in risks that are completely in line with the enterprise's risk tolerance, it would be an example of good governance. But if there is a risk that is outside the enterprise's risk tolerance, and management decides to bear that risk anyway, saying it is self-insuring, management is basically ignoring that risk, disregarding and violating its risk tolerance, and practicing bad risk governance. For example, an investment management firm, via its risk tolerance decision, may decide that it cannot bear any investment loss exceeding €1 billion and may apply a variety of risk management tools to limit its market and credit risk accordingly. But suppose that the firm makes no move to limit or insure its risks from fraud or a rogue trader on the grounds that it is "self-insuring" this risk, which could result in a loss as high as €3 billion. By leaving itself open to a loss that far exceeds its stated risk tolerance, management is violating the firm's risk governance.

From the perspective of a business organization, self-insurance is obtained by setting aside sufficient capital to cover losses. The banking industry is a classic example of self-insurance. Although in many countries government insurance may protect depositors, banks self-insure to some extent by maintaining capital and loan loss reserves.

Another form of accepting risk, but doing so in the most efficient manner possible, is diversification. Technically, it is a risk-mitigation technique. But diversification and "the efficient frontier" are so central to modern portfolio analysis that capturing the full benefits of diversification seems the obvious thing for all organizations to pursue—a starting point at which other risk modification could be appended. Although diversification is one form of risk management, it is usually not effective if used in isolation.

In the next two subsections, we discuss how undesired risk can be modified or eliminated by selling the risk to another party. We make two subtle classifications of these methods: risk transfer and risk shifting.

11

METHODS OF RISK MODIFICATION - TRANSFER, SHIFTING, CHOOSING A METHOD FOR MODIFYING

g describe methods for measuring and modifying risk exposures and factors to consider in choosing among the methods.

Risk transfer is the process of passing on a risk to another party, often, but not always, in the form of an insurance policy. Insurance is a legal contract in which one party, the insurer, agrees to cover some percentage of a loss incurred by another party, the insured, from a specific event in return for the payment of a premium by the insured. Insurance as a method of risk modification has been in existence for very long time, and in fact, is even mentioned in the Code of Hammurabi almost 4,000 years ago. Insurance has been widely used in the commercial shipping and farming industries going back hundreds of years. Insurance is almost as old as commerce itself.[25]

From the point of view of the insurer, insurance almost always works on the basis of diversification or pooling of risks. An insurer attempts to sell many policies with risks that have low correlations. The insurer assesses the pooled risks and charges a premium that covers the expected aggregate losses and the insurer's operating costs as well as leaves a profit. Insurers need accurate statistics on aggregate risks, but these are often not difficult to obtain. These actuarial data are widely available on accidents, illnesses, property and casualty damage, and death. In principle, a well-diversified insurer does not care if a single insured party has significantly larger-than-average claims as long as there is no reason to believe that the claims are correlated. There will be other parties that have smaller-than-average claims.

Insurers do have to manage their risks carefully. Some risks can be correlated. In the US Gulf Coast region, property insurance, which includes coverage for loss by hurricanes, is typically more expensive than property insurance in other regions.

---

[25] It is worth noting that the insurance industry has for a long time referred to itself using the term "risk management." A department of risk management in a large organization is often the group that manages the organization's insurance policies. But since around 1990 or so, the term "risk management" has increasingly come to refer to far more than insurance.

Even with a higher premium, an insurer has to avoid providing too much property coverage in an area where a systemic event, such as a hurricane, can occur in order to diversify its risk exposure.

Although insurers carefully assess their risk and charge premiums that they believe accurately reflect expected losses, they nonetheless remain responsible for potentially large claims. Insurers also manage their risk by avoiding writing too many policies with similar and potentially correlated risks and by selling some of the risk to another insurer, a practice known as reinsurance. A company that primarily insures property in the US Midwest, which is highly subject to tornado risk, might be willing to accept some Gulf Coast hurricane risk for a reasonable premium. Insurers often write provisions into contracts to exclude coverage of special cases. For example, a war might nullify insurance coverage in an area. Most insurance policies also contain provisions to guard against moral hazards, such as suicide or destroying one's own property. In the last 20 years or so, some insurance companies have issued bonds that permit them to legally avoid paying principal and/or interest if insurance claims exceed a certain amount. These instruments, known as catastrophe bonds, essentially pass some of the insurance risk on to the investors who buy the bonds.

Most insurance policies do not cover *all* of the risk that is insured. It is common for policies to contain a provision known as a deductible. A deductible is a monetary amount of the loss that will be covered by the insured before any claims are paid. Thus, both the insured and the insurer bear some of the risk, although the insurer usually bears the greater amount. Deductibles serve several purposes. Because insurers incur fixed costs for each claim, deductibles reduce the number of small claims. Deductibles also encourage good risk management by the insured parties. Finally, deductibles offer the insured the opportunity to combine risk transfer with self-insurance and thereby achieve a potentially better trade-off of risk and reward.

As noted, the concept of insurance relies on the diversification or pooling of risks. In a few cases, however, the risks are not easy to pool. For example, suppose a volatile but extremely successful actor is signed to star in a movie. The production company knows that it runs the risk that the actor will engage in behavior that damages the ability of the company to finish the movie. The number of volatile and extremely successful actors for whom policies could be written at the same time is somewhat limited. Thus, an insurer would have to bear that risk without the benefit of diversification.

For example, suppose a television network plans to cover the Olympics but is concerned about a possible cancellation or boycott. It might want an insurance policy to cover it against loss. Specialized coverage is possible through such companies as Lloyd's of London. The approximately 350-year old Lloyd's is famous for covering unusual risks. It does so by organizing groups of investors who are willing to bear a risk for a premium. These groups, called syndicates, are subject to the full extent of losses. In many cases, investors in these syndicates have been required to pay substantial amounts of money to cover losses.[26] These examples illustrate how syndicates work. Although there is only one Olympics to insure, there may also be only one actor to insure. Because the two risks are uncorrelated, a company could write policies on both risks and would achieve some diversification. Moreover, there are other unusual risks that can be covered such that the aggregate pool would represent a highly diverse set of risks that have low correlations.

A very slight variation of insurance is a surety bond. With a surety bond, an insurer promises to pay an insured a certain amount of money if a third party fails to fulfill its obligation. For example, if a party engages the services of another party, the first party is covered if the party obligated to provide the service fails to perform to a satisfactory degree. Surety bonds are widely used in commercial activity when one party bears the risk of the potentially high cost of non-performance by another party. A slight variation of a surety bond is a fidelity bond, which is often used to cover against losses that result from employee dishonesty. Bonds of this type work very similarly to insurance and rely on the pooling of uncorrelated risks.[27] Other similar arrangements include indemnity clauses and hold harmless arrangements, such as when two parties sign a contract and one party agrees to hold the other harmless and/or indemnify the other in the event of loss.

---

[26] NBC insured the 1980 Summer Olympics in Moscow through Lloyd's of London to the extent that if a US boycott occurred, Lloyd's would pay NBC for losses that it incurred by prepaying the Soviet Union for broadcasting rights. The United States did boycott the Olympics and NBC collected on its policy.

The use of insurance by so many as a risk management tool suggests that the cost of risk exceeds the actuarial cost to many individuals and enterprises. *Ex ante* consideration of the cost of a risk in terms of the organization's value or utility ties risk mitigation back to the risk tolerance decision and the most fundamental governance decisions on which value-added strategies to pursue. As an alternative to *ignoring* the cost of risk, the impact on enterprise value should be quite positive.

11.1 Risk Shifting

Whereas risk transfer refers to actions taken that pass the risk on to other parties, **risk shifting** refers to actions that change the distribution of risk outcomes. Risk transfer is often associated with insurance, whereas risk shifting generally involves derivatives as the risk modification vehicle. Although insurance is a form of risk management based on the pooling or diversification of risks, risk shifting diverts some portion of the risk distribution to another market participant who either bears the risk or intermediates that risk by moving it to yet another party. The organization may want to adjust its probability distribution of returns, essentially adjusting the payoff diagram of its risk exposures. An example is a company that is willing to make slightly less profit than it otherwise would if the stock market is up to prevent it from losing too much money if the stock market is down, for example, more than 20% next year. It is adjusting its potential economic outcomes by shifting the probability distribution of its profits conditional on market performance. Risk shifting represents the bulk of hedging and is the most common form of risk modification for financial organizations.

The principal device through which risk shifting is performed is a derivative. We briefly mentioned derivatives earlier in this reading. By definition, a derivative is a financial instrument that derives its price from the price of an underlying asset or rate. Because the price of the underlying and the price of the derivative are so closely related, derivatives can provide essentially the same exposure as the underlying but can do so at lower cost and capital requirements. As such, derivatives permit the efficient shifting of risk across the probability distribution and from one party to another. One can hold the underlying and take an offsetting position in the derivative or vice versa. Whereas insurance can be designed to perform similarly, insurance functions primarily through the pooling of diverse risks. With derivatives, risks are shifted across probability distributions or payoffs and across parties, to leave specific outcomes of the conditional probability distribution with the parties most willing to bear the risk.

There are several types of derivatives, and the manner in which they provide risk shifting varies by type. Derivatives are classified into two categories: forward commitments or contingent claims. Forward commitments are agreements that obligate two parties to undertake a transaction at a future date at a price or rate agreed on when the commitment is made. Forward commitments include such instruments as forward contracts, futures contracts, and swaps. Forward commitments can be used to lock in a future interest rate for a borrower or lender, the purchase or sale price of an asset, or an exchange rate for a currency. Parties who engage in forward commitments do not pay any money at the initiation of the contract. In lieu of any up-front payment from one party to the other, the two parties agree on the terms of the transaction that will be consummated at the end of the contract. Depending on movements in the price or rate of the underlying, one party will ultimately gain from the transaction while the other will lose or, in the less likely case, both parties could breakeven. For example, a corporate treasurer can use a forward contract to lock in the rate at which a foreign cash flow will be converted into the company's domestic currency. Regardless of movements in the exchange rate during the life of the contract, the foreign cash flow will convert to the domestic currency at a rate that is locked in when the contract is initiated. On the opposite side of the transaction, the party can be a speculator who simply bears the risk, or it can be a dealer who intermediates the risk between the hedger and the speculator. We will discuss dealers in more detail in a few paragraphs.

---

[27] In the context of surety and fidelity bonds, the word "bond" does not mean a debt obligation issued by one party, the borrower, and bought by another, the lender. In this context, the word refers to assuring one party that it bears no risk for the actions of a specific other party.

The other type of derivative is a contingent claim, which is commonly known as an option. An option is a contract between two parties that gives one party the right to buy or sell an underlying asset or to pay or receive a known underlying rate. An option takes the form of either a call option, which provides the right to buy the underlying or to pay a known rate, or a put option, which provides the right to sell the underlying or to receive a known rate.

With a forward commitment, both parties are mutually obligated to each other. Because an option grants the right, but not the obligation, to one party, that party has an advantage over the other. Consequently, that party, the buyer of the option, must pay cash, called the premium, to the seller of the option at the start of the contract. Once the premium is paid, the option buyer has no further obligation. He can either exercise the option or he can let the option expire unexercised. In the latter case, the option buyer incurs a loss equal to the premium. If the option is a call and it is exercised, the buyer pays the fixed price or rate and receives the underlying. If the option is a put and it is exercised, the buyer receives the fixed price or rate and delivers the underlying.[28] If the buyer of the option does exercise it, he may achieve a gain that exceeds the premium paid but the gain could also be less than the premium paid, thereby resulting in a net overall loss. An option buyer could be using the option to speculate on an upward move in the underlying if a call or downward move if a put. Alternatively, the option buyer could be hedging. In the example used earlier for forward commitments, the corporate treasurer anticipating an inflow of cash in a foreign currency could buy a put option to sell that currency, thereby converting it into his domestic cash flow at a known fixed rate. The option gives the treasurer the flexibility to not exercise it if the underlying currency rises in value. This flexibility comes at the cost of having to pay a premium at the start of the transaction, thus shifting the financial outcome across the entire probability distribution of that uncertain currency rate. In contrast, with the forward contract, the treasurer does not have to pay cash at the start but is obligated to convert at the agreed-upon rate.

Derivatives can be created in public forums, such as on derivatives exchanges, or privately between two parties. On derivatives exchanges, there are a large number of individual and institutional traders that make markets in derivatives. For private derivatives transactions, there is an extensive market of large bank and non-bank dealers willing to buy and sell derivatives. In both types of markets, these dealers assume the risk being transferred from parties who originate the transactions. These dealers almost always restructure and transfer some portion, if not all, of the risk by finding other parties that are willing to take on that risk. Ultimately, the risk is assumed by some party willing to accept the risk, producing an economically efficient outcome for all parties.

---

[28] Instead of one party delivering the underlying, some options call for settlement in cash of an equivalent amount. Some forward commitments also settle in cash.

11.2 How to Choose Which Method for Modifying Risk

Choosing which risk mitigation method to use—risk prevention and avoidance, self-insuring, risk transfer, or risk shifting—is a critical part of the risk management process. Fortunately, the methods are not mutually exclusive, and many organizations use all methods to some extent. No single method provides a clear-cut advantage over the others. As with all decisions, the trade-off is one of costs versus benefits that are weighed in light of the risk tolerance of the organization or individual.

For example, many companies that have extensive foreign operations and are, therefore, highly exposed to exchange rate risk, hedge that risk using derivatives. Some companies prefer forwards, some prefer swaps, some prefer options, and some use multiple instruments. Some companies attempt to hedge currency risk by setting up operations in foreign countries rather than manufacturing domestically and shipping the goods to foreign countries.[29] Some companies manage their currency risk by attempting to balance currency assets and liabilities. Some airlines hedge the risk of oil price changes and others do not. Some airlines that do hedge this risk do so to a far greater degree than others. Additionally, some prefer the certainty of forwards and swaps, whereas others prefer the flexibility of options, even with the up-front cost that options require. Most insurance companies rely on their actuarial knowledge but supplement it with proactive measures, such as selling risk to other parties.

To the extent possible, most organizations should avoid risks that provide few benefits and potentially extreme costs. Reasonable, low-cost precautions against risks with few benefits should always be taken. Thus, risk prevention and risk avoidance are probably the first choice of measures, especially for risks that lie outside the core competencies of the organization and have little reasonable expectation of adding value. Nonetheless, avoidance may not be the best value for its cost. Moreover, avoiding risk may mean avoiding opportunity. Thus, an organization often cannot simply decide not to take a risk, at least not for all undesirable risks.

Organizations that have large amounts of free cash flow may choose to self-insure some risks, but few organizations have so much cash that they can afford to self-insure all risks. Some risks can potentially imperil the entire capital base. Most companies would, however, prefer to self-insure to the extent possible because self-insurance reduces the costs associated with external monitoring and gives the organization the greatest flexibility. Self-insurance and avoidance should generally be clearly addressed at the governance level and be consistent with stated risk tolerance.

Risk transfer, or the use of insurance, is a widely used risk management tactic, but it may not be suitable for many types of risks. Some risks simply are not insurable, at least not in a cost-effective way. Insurance works best when risks can be pooled, and that is not the case for many types of risks, particularly those that can affect a large number of parties at the same time. The use of risk shifting tools, such as derivatives, may not be available for all types of risks, thus limiting their use in risk mitigation. For financial risks that exceed risk appetite, risk shifting is a very common choice.

The various risk management methods are not equal in terms of the risk reduction and the risk profile that remains. For example, contingent claims, such as insurance, provide the flexibility in the form of offering opportunity to profit in one direction and have a loss reduced in the other, but they require payment of cash up front. In contrast, forward commitments lock in an outcome. In other words, they provide little flexibility, but they require no cash payment up front. The risk profile that exists when a contingent claim hedge is put in place differs significantly from the risk profile that exists when a forward commitment hedge is placed. This process requires significant understanding and discussion at all levels of the organization.

---

[29] Here is another example of the interactions of risks. A decision to manufacture products in a foreign country involves trade-offs between exchange rate risk, political risk, and a variety of other risks germane to that country's economy, not to mention a potentially different degree of operational risk, in the pursuit of higher profits.

To recap, risk takers should identify risks that offer few rewards in light of potential costs and avoid those risks when possible. They should self-insure where it makes sense and diversify to the extent possible. They should consider insurance when risks can be pooled effectively if the cost of the insurance is less than the expected benefit. If derivatives are used, they must consider the trade-off of locking in outcomes with forward commitments versus the flexibility relative to cash cost of contingent claims, which can tailor the desired outcomes or payoffs by shifting the risk. Ultimately, the decision is always one of balancing costs against benefits while producing a risk profile that is consistent with the risk management objectives of the organization.

EXAMPLE 4

Measuring and Modifying Risk

1 From the perspective of an organization, which of the following *best* describes risk drivers?

- **A** The probabilities of adverse events
- **B** The statistical methods that measure risk
- **C** Factors that influence macroeconomies and industries

2 Which of the following concepts directly measures the risk of derivatives?

- **A** Probability
- **B** Delta and gamma
- **C** Beta and standard deviation

3 The *best* definition of value at risk is:

- **A** the expected loss if a counterparty defaults.
- **B** the maximum loss an organization would expect to incur over a holding period.
- **C** the minimum loss expected over a holding period a certain percentage of the time.

4 Which of the following are methods commonly used to supplement VaR to measure the risk of extreme events?

- **A** Standard deviation
- **B** Loss given default
- **C** Scenario analysis and stress testing

5 Which of the following is a true statement about insurable risks?

- **A** Insurable risks are less costly.
- **B** Insurable risks have smaller loss limits.
- **C** Insurable risks are typically diversifiable by the insurer.

Solution to 1:

C is correct. Risks (and risk drivers) arise from fundamental factors in macroeconomies and industries.

Solution to 2:

B is correct. Delta and gamma are measures of the movement in an option price, given a movement in the underlying. The other answers can reflect some elements of derivatives risk, but they are not direct measures of the risk.

Solution to 3:

C is correct. VaR measures a minimum loss expected over a holding period a certain percentage of the time. It is not an expected loss nor does it reflect the maximum possible loss, which is the entire equity of the organization.

Solution to 4:

C is correct. Scenario analysis and stress testing both examine the performance of a portfolio subject to extreme events. The other two answers are metrics used in portfolio analysis but are not typically associated with extreme events.

Solution to 5:

C is correct. Insurance works by pooling risks. It is not necessarily less costly than derivatives nor does it have lower loss limits.

SUMMARY

Success in business and investing requires the skillful selection and management of risks. A well-developed risk management process ties together an organization's goals, strategic competencies, and tools to create value to help it both thrive and survive. Good risk management results in better decision making and a keener assessment of the many important trade-offs in business and investing, helping managers maximize value.

- Risk and risk management are critical to good business and investing. Risk management is *not* only about avoiding risk.
- Taking risk is an active choice by boards and management, investment managers, and individuals. Risks must be understood and carefully chosen and managed.
- Risk exposure is the extent to which an organization's value may be affected through sensitivity to underlying risks.
- Risk management is a process that defines risk tolerance and measures, monitors, and modifies risks to be in line with that tolerance.
- A risk management framework is the infrastructure, processes, and analytics needed to support effective risk management; it includes risk governance, risk identification and measurement, risk infrastructure, risk policies and processes, risk mitigation and management, communication, and strategic risk analysis and integration.
- Risk governance is the top-level foundation for risk management, including risk oversight and setting risk tolerance for the organization.
- Risk identification and measurement is the quantitative and qualitative assessment of all potential sources of risk and the organization's risk exposures.
- Risk infrastructure comprises the resources and systems required to track and assess the organization's risk profile.

- Risk policies and processes are management's complement to risk governance at the operating level.
- Risk mitigation and management is the active monitoring and adjusting of risk exposures, integrating all the other factors of the risk management framework.
- Communication includes risk reporting and active feedback loops so that the risk process improves decision making.
- Strategic risk analysis and integration involves using these risk tools to rigorously sort out the factors that are and are not adding value as well as incorporating this analysis into the management decision process, with the intent of improving outcomes.
- Employing a risk management committee, along with a chief risk officer (CRO), are hallmarks of a strong risk governance framework.
- Governance and the entire risk process should take an enterprise risk management perspective to ensure that the value of the entire enterprise is maximized.
- Risk tolerance, a key element of good risk governance, delineates which risks are acceptable, which are unacceptable, and how much risk the overall organization can be exposed to.
- Risk budgeting is any means of allocating investments or assets by their risk characteristics.
- Financial risks are those that arise from activity in the financial markets.
- Non-financial risks arise from actions within an organization or from external origins, such as the environment, the community, regulators, politicians, suppliers, and customers.
- Financial risks consist of market risk, credit risk, and liquidity risk.
- Market risk arises from movements in stock prices, interest rates, exchange rates, and commodity prices.
- Credit risk is the risk that a counterparty will not pay an amount owed.
- Liquidity risk is the risk that, as a result of degradation in market conditions or the lack of market participants, one will be unable to sell an asset without lowering the price to less than the fundamental value.
- Non-financial risks consist of a variety of risks, including settlement risk, legal risk, regulatory risk, accounting risk, tax risk, model risk, tail risk, and operational risk.
- Operational risk is the risk that arises either from within the operations of an organization or from external events that are beyond the control of the organization but affect its operations. Operational risk can be caused by employees, the weather and natural disasters, vulnerabilities of IT systems, or terrorism.
- Solvency risk is the risk that the organization does not survive or succeed because it runs out of cash to meet its financial obligations.
- Individuals face many of the same organizational risks outlined here but also face health risk, mortality or longevity risk, and property and casualty risk.
- Risks are not necessarily independent because many risks arise as a result of other risks; risk interactions can be extremely non-linear and harmful.
- Risk drivers are the fundamental global and domestic macroeconomic and industry factors that create risk.
- Common measures of risk include standard deviation or volatility; asset-specific measures, such as beta or duration; derivative measures, such as delta, gamma, vega, and rho; and tail measures such as value at risk, CVaR and expected loss given default.

- Risk can be modified by prevention and avoidance, risk transfer (insurance), or risk shifting (derivatives).
- Risk can be mitigated internally through self-insurance or diversification.
- The primary determinants of which method is best for modifying risk are the benefits weighed against the costs, with consideration for the overall final risk profile and adherence to risk governance objectives.

PRACTICE PROBLEMS

1 Risk management in the case of individuals is *best* described as concerned with:

- **A** hedging risk exposures.
- **B** maximizing utility while bearing a tolerable level of risk.
- **C** maximizing utility while avoiding exposure to undesirable risks.

2 Which of the following may be controlled by an investor?

- **A** Risk
- **B** Raw returns
- **C** Risk-adjusted returns

3 The process of risk management includes:

- **A** minimizing risk.
- **B** maximizing returns.
- **C** defining and measuring risks being taken.

4 Risk governance:

- **A** aligns risk management activities with the goals of the overall enterprise.
- **B** defines the qualitative assessment and evaluation of potential sources of risk in an organization.
- **C** delegates responsibility for risk management to all levels of the organization's hierarchy.

5 The factors a risk management framework should address include all of the following *except*:

- **A** communications.
- **B** policies and processes.
- **C** names of responsible individuals.

6 Which of the following is the correct sequence of events for risk governance and management that focuses on the entire enterprise? Establishing:

- **A** risk tolerance, then risk budgeting, and then risk exposures.
- **B** risk exposures, then risk tolerance, and then risk budgeting.
- **C** risk budgeting, then risk exposures, and then risk tolerance.

7 Which of the following *best* describes activities that are supported by a risk management infrastructure?

- **A** Risk tolerance, budgeting, and reporting
- **B** Risk tolerance, measurement, and monitoring
- **C** Risk identification, measurement, and monitoring

8 Effective risk governance in an enterprise provides guidance on all of the following *except*:

- **A** unacceptable risks.
- **B** worst losses that may be tolerated.
- **C** specific methods to mitigate risk for each subsidiary in the enterprise.

9 A firm's risk management committee would be expected to do all of the following *except*:

- **A** approving the governing body's proposed risk policies.

© 2019 CFA Institute. All rights reserved.

B deliberating the governing body's risk policies at the operational level.

C providing top decision-makers with a forum for considering risk management issues.

10 Once an enterprise's risk tolerance is determined, the role of risk management is to:

A analyze risk drivers.

B align risk exposures with risk appetite.

C identify the extent to which the enterprise is willing to fail in meeting its objectives.

11 Which factor should *most* affect a company's ability to tolerate risk?

A A stable market environment

B The beliefs of the individual board members

C The ability to dynamically respond to adverse events

12 Risk budgeting includes all of the following *except*:

A determining the target return.

B quantifying tolerable risk by specific metrics.

C allocating a portfolio by some risk characteristics of the investments.

13 A benefit of risk budgeting is that it:

A considers risk tradeoffs.

B establishes a firm's risk tolerance.

C reduces uncertainty facing the firm.

14 Which of the following risks is *best* described as a financial risk?

A Credit

B Solvency

C Operational

15 Liquidity risk is *most* associated with:

A the probability of default.

B a widening bid–ask spread.

C a poorly functioning market.

16 An example of a non-financial risk is:

A market risk.

B liquidity risk.

C settlement risk.

17 If a company has a one-day 5% Value at Risk of $1 million, this means:

A 5% of the time the firm is expected to lose at least $1 million in one day.

B 95% of the time the firm is expected to lose at least $1 million in one day.

C 5% of the time the firm is expected to lose no more than $1 million in one day.

18 An organization choosing to accept a risk exposure may:

A buy insurance.

B enter into a derivative contract.

C establish a reserve fund to cover losses.

19 The choice of risk-modification method is based on:

A minimizing risk at the lowest cost.

B maximizing returns at the lowest cost.

C weighing costs versus benefits in light of the organization's risk tolerance.

SOLUTIONS

1 B is correct. For individuals, risk management concerns maximizing utility while taking risk consistent with individual's level of risk tolerance.

2 A is correct. Many decision makers focus on return, which is not something that is easily controlled, as opposed to risk, or exposure to risk, which may actually be managed or controlled

3 C is correct. Risks need to be defined and measured so as to be consistent with the organization's chosen level of risk tolerance and target for returns or other outcomes.

4 A is correct. Risk governance is the top-down process that defines risk tolerance, provides risk oversight and guidance to align risk with enterprise goals.

5 C is correct. While risk infrastructure, which a risk management framework must address, refers to the people and systems required to track risk exposures, there is no requirement to actually name the responsible individuals.

6 A is correct. In establishing a risk management system, determining risk tolerance must happen before specific risks can be accepted or reduced. Risk tolerance defines the appetite for risk. Risk budgeting determine how or where the risk is taken and quantifies the tolerable risk by specific metrics. Risk exposures can then be measured and compared against the acceptable risk.

7 C is correct. *Risk infrastructure* refers to the people and systems required to track risk exposures and perform most of the quantitative risk analysis to allow an assessment of the organization's risk profile. The risk management infrastructure identifies, measures, and monitors risks (among other things).

8 C is correct. Risk governance is not about specifying methods to mitigate risk at the business line level. Rather, it is about establishing an appropriate level of risk for the entire enterprise. Specifics of dealing with risk fall under risk management and the risk infrastructure framework.

9 A is correct. The risk management committee is a part of the risk governance structure at the operational level—as such, it does not approve the governing body's policies.

10 B is correct. When risk tolerance has been determined, the risk framework should be geared toward measuring, managing, and complying with the risk tolerance, or aligning risk exposure with risk tolerance. The risk tolerance decision begins by looking at what shortfalls within an organization would cause it to fail to achieve some critical goals and what are the organization's risk drivers.

11 C is correct. If a company has the ability to adapt quickly to adverse events may allow for a higher risk tolerance. There are other factors, such as beliefs of board members and a stable market environment, which may but should not affect risk tolerance.

12 A is correct. Risk budgeting does not include determining the target return. Risk budgeting quantifies and allocates the tolerable risk by specific metrics.

13 A is correct. The process of risk budgeting forces the firm to consider risk tradeoffs. As a result, the firm should choose to invest where the return per unit of risk is the highest.

14 A is correct. A financial risk originates from the financial markets. Credit risk is one of three financial risks identified in the reading: Credit risk is the chance of loss due to an outside party defaulting on an obligation. Solvency risk depends at least in part on factors internal to the organization and operational risk is an *internal* risk arising from the people and processes within the organization.

15 B is correct. Liquidity risk is also called transaction cost risk. When the bid–ask spread widens, purchase and sale transactions become increasingly costly. The risk arises from the uncertainty of the spread.

16 C is correct. Settlement risk is related to default risk, but deals with the timing of payments rather than the risk of default.

17 A is correct. The VaR measure indicates the probability of a loss of at least a certain level in a time period.

18 C is correct. Risk acceptance is similar to self-insurance. An organization choosing to self-insure may set up a reserve fund to cover losses. Buying insurance is a form of risk transfer and using derivatives is a form of risk-shifting, not risk acceptance.

19 C is correct. Among the risk-modification methods of risk avoidance, risk acceptance, risk transfer, and risk shifting none has a clear advantage. One must weigh benefits and costs in light of the firm's risk tolerance when choosing the method to use.

READING

54

Technical Analysis

by Aksel Kibar, CMT, Barry M. Sine, and Robert A. Strong, PhD, CFA

Aksel Kibar, CMT, is at Tech Charts Research & Trading (Bulgaria). Barry M. Sine is at Drexel Hamilton, LLC (USA). Robert A. Strong, PhD, CFA, is at the University of Maine (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** explain principles and assumptions of technical analysis;
☐	**b.** describe potential links between technical analysis and behavioral finance;
☐	**c.** compare principles of technical analysis and fundamental analysis;
☐	**d.** describe and interpret different types of technical analysis charts;
☐	**e.** explain uses of trend, support, and resistance lines;
☐	**f.** explain common chart patterns;
☐	**g.** explain common technical indicators;
☐	**h.** describe principles of intermarket analysis;
☐	**i.** explain technical analysis applications to portfolio management.

INTRODUCTION

1

Technical analysis has been used by traders, analysts, and investors for centuries and has achieved broad acceptance among regulators and the academic community—particularly with regard to its behavioral finance aspects. This reading gives a brief overview of the field, compares technical analysis with other schools of analysis, and describes some of the main tools used in technical analysis. Although technical analysis follows predefined rules and principles, the interpretation of results is generally subjective. That is, although certain aspects, such as the calculation of indicators, follow specific rules, the interpretation of findings is often based on a melding of techniques that

© 2020 CFA Institute. All rights reserved.

suit the style and approach of the individual analyst. In this respect, technical analysis is similar to fundamental analysis, which has specific rules for calculating ratios, for example, but introduces increased subjectivity in the evaluation phase.

2

TECHNICAL ANALYSIS: PRINCIPLES, ASSUMPTIONS, AND LINKS TO INVESTMENT ANALYSIS

a explain principles and assumptions of technical analysis;

b describe potential links between technical analysis and behavioral finance;

c compare principles of technical analysis and fundamental analysis;

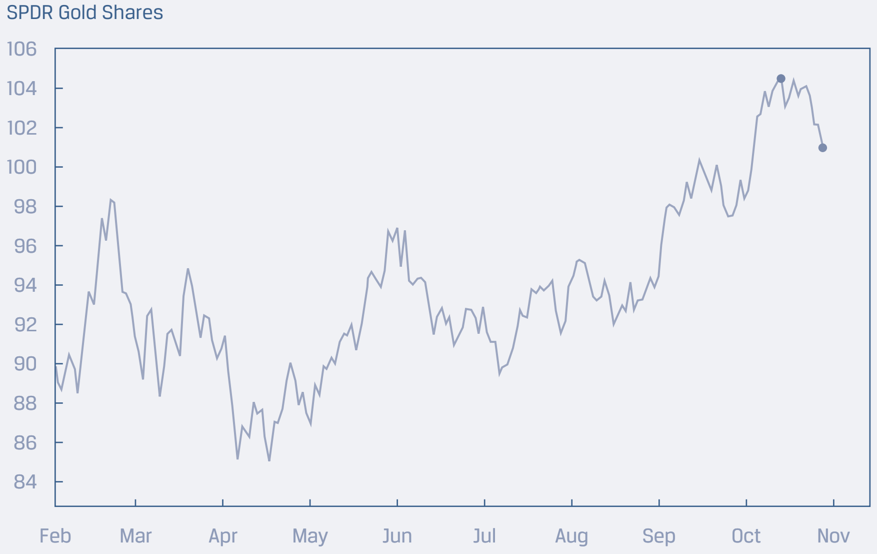
VIGNETTE

Scene 1

You are a portfolio analyst working at a small sovereign wealth fund (SWF) and reporting to the Deputy Chief Investment Officer (CIO). The SWF has recently implemented an investment in GLD for the purposes of improving the risk-adjusted return of the portfolio. This asset allocation decision was based on extensive analysis and agreed on by the CIO, Head of Risk, Head of Alternative Investments, and other members of the Investment Committee. The benefits of adding gold to the portfolio are due to the low correlation gold has with other asset classes in the portfolio.

The Deputy CIO is on a well-deserved sabbatical for several weeks in Antarctica and left you in charge of monitoring the portfolio. You are looking forward to the opportunity to "be in charge" but also hope that there will be no major decisions to make during this time.

Since the investment in GLD is new, you decide to take a look at gold prices and create a basic graph. You prepare a chart that shows the performance of GLD since the start of the year. The purchase price is the blue point, and the current price is the red point. **You realize that the price of GLD is now below the purchase price.**



There is a knock on the door. The Head of Risk comes into your office and says, ***"The price of gold is now below our purchase price. You're in charge of this position, right? Maybe we should sell now to avoid a bigger loss?"***

You respond to the Head of Risk by saying, ***"Let's not panic. After all, the decision to invest in gold was not a short-term decision but a long-term asset allocation decision. We should not fall into the behavioral finance trap of letting fear guide our decisions."***

The Head of Risk leaves your office. You look at the chart again, and you begin to worry. *Maybe we should sell?* You decide to ask the Technical Analyst to prepare a chart and share her thoughts.

Technical analysis is a form of security analysis that uses price and volume data, often graphically displayed, in decision making. Technical analysis can be applied to securities in any freely traded market around the globe. A freely traded market is one in which willing buyers trade with willing sellers without external intervention or impediment. Prices are the result of the interaction of supply and demand in real time. Technical analysis allows us to see a battle between buyers and sellers, along with subtle clues as to which side may be winning. Technical analysis is used on a wide range of financial instruments, including equities, bonds, commodity futures, currency futures, and cryptocurrencies.

The underlying logic of technical analysis is simple:

- Supply and demand determine prices.
- Changes in supply and demand—both in price level and volume—cause changes in prices.
- Past price action can be used to anticipate and project potential future prices with charts and other technical tools.

Basic technical analysis of any financial instrument does not require detailed knowledge of that instrument. As long as the chart represents the action in a freely traded market, a technician does not even need to know the name or type of the security to conduct the analysis. Technical analysis can also be applied over any time frame—from short-term price movements to long-term movements of annual closing prices. Trends that are apparent in short-term charts may also appear over longer time frames. While technical analysis is ideal for short-term trading decisions or tactical asset allocation decisions, long-term chart analysis can add value regarding strategic asset allocation or long-term investment decisions. Looking at recurring technical clues on multiple time intervals (monthly, weekly, daily, 60-minute charts, etc.) can be a particularly useful technique.

2.1 Principles and Assumptions

The three main principles of technical analysis are as follows:

- The market discounts everything.
- Prices move in trends and countertrends.
- Price action creates certain patterns that tend to reoccur and may be cyclical.

The market discounts everything.

One of the biggest assumptions in technical analysis is that price already reflects all known factors impacting a financial instrument. This form of analysis assumes that at any point in time, a stock's price is a reflection of everything that affects the organization, including fundamental factors, such as the balance sheet, the income statement, the cash flow, and the management team. Technical analysis also considers broader economic factors and market psychology to be reflected in pricing.

Prices move in trends and countertrends.

The second assumption in technical analysis is that prices follow trends, which move directionally—upward, downward, sideways, or in a combination of these directions. Once a trend is recognized, the expectation is that any future movement in the price of the asset will follow that trend rather than go against it. A trend in motion is more likely to continue than to reverse. An important corollary is that a trend in motion will continue until it reverses. Indeed, a common saying among technical analysts is, "The trend is your friend."

Price action is repetitive, and certain patterns tend to reoccur.

The repetition of price movements, according to technical analysts, is due to market psychology. News is constantly bombarding assets, and how markets react depends upon the technical setup of the market at the time such news hits. News is always important, but the value of any news depends upon the initial market psychology. Because the behavior of investors repeats itself, price movements can be charted, allowing technicians to recognize patterns.

2.2 Technical Analysis and Behavioral Finance

Technical analysis can be thought of as the study of collective investor psychology or sentiment and thus has a direct connection with behavioral finance. Prices in any freely traded market are set by human beings or their automated proxies (such as computerized trading programs), and a price is set at the equilibrium between supply and demand at any given instant. The role of a good technician is to monitor and evaluate the subtle clues in this battle between supply and demand—much like a doctor looking at a patient for subtle clues to that patient's overall health. Various fundamental theorists have proposed that markets are efficient and rational, but technicians believe that humans are often irrational and emotional and that they tend to behave similarly in similar circumstances. The reason chart patterns have predictive value is that they are graphic representations of human trading activity and human behavior is frequently repeated, especially trading activity that is driven by fear (in market sell-offs) or greed (as evidenced in bubbles—that is, rallies that extend well beyond valuation levels derived from fundamental data). In bubbles, investors, driven by hope and greed, push the price of an asset to irrationally high levels, in the expectation that other buyers will be willing to pay an even higher price for the asset. The resulting chart patterns are analyzed and identified by analysts, traders, and investors for decision-making purposes.

Although fundamental data are key inputs for the determination of value, these data are analyzed by humans, who may be driven, at least partially, by motivations that are not rational. Human behavior is often erratic and driven by emotion in many aspects of one's life, so technicians conclude that it is unreasonable to believe that investing is the one exception where humans always behave rationally. Technicians believe that market trends and patterns reflect this irrational human behavior, which can be seen on price charts in the form of volatile price action, such as major sell-offs or strong buying resulting in parabolic advances. Technical analysis is the study of market trends and patterns. And because trends and patterns tend to repeat themselves, they are potentially identifiable and predictable. So, technicians rely on recognition of patterns that have occurred in the past to project future patterns of security prices.

Another tenet of technical analysis is that the market reflects the collective knowledge and sentiment of the many different participants responsible for the buying and selling activity in a particular security. In a freely traded market, only those market participants who actually buy or sell a security have an impact on its price. And the greater the volume of a participant's trades, the more impact that market participant will tend to have on price. Technical analysis relies on market participants that are taking a directional bet in the market and thereby influencing price and volume. Without trading, there will be no meaningful price fluctuations.

Trades determine volume and price. The impact of a trade occurs instantaneously and frequently anticipates or foreshadows fundamental developments. So, by studying market technical data—price and volume trends—the technician is seeking to understand investor sentiment and to detect any fundamental change that is already transpiring behind the scenes or may shortly transpire. The technician benefits from the wide range of knowledge of market participants and their collective conclusion about a security. In contrast, the fundamental analyst must wait for the release of financial statements to conduct financial statement analysis, so a time lag occurs between the market's activities and the analyst's conclusions. Technicians may therefore be ahead of fundamental analysts in their positioning.

Technical analysts believe that human behavior plays a role in the fluctuation of security prices. Investors with a favorable fundamental view may nonetheless sell a financial instrument for other reasons, including pessimistic investor sentiment, margin calls, and a need for capital. Technicians do not care why market participants are buying or selling, just that they are doing so: These actions move prices, forming trends the technician can profit from. Technical moves happen for both fundamental and behavioral reasons.

Some financial instruments have an associated income stream that contributes to the security's intrinsic value. Bonds have regular coupon payments, and equity shares may have underlying cash flows or dividend streams. A fundamental analyst can adjust these cash flows for risk and use standard time value of money techniques to determine a present value. Other assets, such as bushels of wheat, gallons of crude oil, and ounces of silver, do not have underlying financial statements or an income stream, so valuation models cannot be used to derive their fundamental intrinsic values. For these assets, technical analysis may be particularly valuable. So, whereas fundamental analysis is widely used for fixed-income and equity securities, technical analysis is widely used for commodities, currencies, and futures.

Market participants attempt to anticipate economic developments and enter into trades to profit from them. Technicians believe that security price movements occur before fundamental developments unfold—and certainly before they are reported. This belief may be supported by the fact that stock prices are 1 of the 12 components of the National Bureau of Economic Research's index of leading economic indicators.

Market microstructure

Market microstructure deals with issues of market structure and design, price formation, price discovery, transaction and timing costs, information and disclosure, and investor behavior. It is the process by which buyers find sellers as well as the venues where they meet. The National Bureau of Economic Research (NBER) defines market microstructure as a field of study that is devoted to theoretical, empirical, and experimental research on the economics of security markets. Elements of market microstructure may include minimum tick increments and algorithmic execution interfaces. The field of market microstructure aims to establish connections between activity over the fast-moving short term and the properties that emerge over longer time frames. In this way, market microstructure is a bottom-up approach to understanding financial markets. The recording and analysis of high-quality data that describe the actions and interactions of market participants in the very short term has revealed striking regularities that challenge many theories regarding financial markets.

Market microstructure considers different order types. Every exchange will specify the types of orders that are allowed and how these orders may interact with each other. Some of the important order types are market order, limit order, stop order, good till canceled order, and day order. But hybrid algorithmic orders are also possible and are increasingly prevalent in modern electronic markets (e.g., an order to "participate at 30% of all volume up to a given price limit" or a time-dependent "buy stop close only" order). When placed in the market, orders will have priority depending on such factors as price, time, and size. Buyers with higher bids will have priority over buyers with lower bids. Similarly, sellers with lower asks will have priority over sellers with higher asks. Market orders are executed before limit orders. Orders placed earlier are executed before later orders.

Most trade models assume the presence of both liquidity traders and information traders. Liquidity traders' decisions to buy or sell securities (e.g., for portfolio rebalancing or market making) are assumed to be unrelated to the arrival of new information. Conversely, information traders are cognizant of new information before other market participants during any given trading period. In other words, they have superior knowledge of a security's true equilibrium price. Information traders will buy/sell securities when they can exploit their advantage profitably.

Liquidity traders can be high-frequency traders or market makers on major trading desks. They buy and sell in a tight bid–ask spread with the aim of profiting from small price fluctuations.

Information traders buy and sell because they believe they have sufficient information to predict the next move in a market and thus to beat the market.

2.3 Technical Analysis and Fundamental Analysis

Both technical analysis and fundamental analysis are useful and valid, but they approach the market in different ways. Technicians focus solely on analyzing markets and the trading of financial instruments. Fundamental analysis is a much wider field, encompassing financial and economic analysis as well as analysis of societal and political trends. Technicians analyze market prices. A technician's analysis is derived solely from price and volume data, whereas a fundamental analyst studies a company, incorporating data that are external to the market, and then uses this analysis to predict security price movements. Technical analysis assumes that all of the factors considered by a fundamental analyst are reflected in the price of a financial instrument through buying and selling activity.

A key distinction between technical analysis and fundamental analysis is that the technician has more concrete data, primarily price and volume data, to work with. The financial statements analyzed by fundamental analysts do not contain objective data but rather are the result of numerous estimates and assumptions that have been combined to produce the various line items. Even the cash line on a balance sheet is subject to corporate management's opinion about which assets are liquid enough to be considered "cash." This opinion must be agreed to by auditors and, in many countries, regulators (who sometimes differ with the auditors). Financial statements are subject to restatement because of such issues as changes in accounting assumptions and even fraud. But the price and volume data used in technical analysis are objective. When the data are analyzed, both types of analysis become subjective because judgment is exercised when a technician analyzes a price chart and when a fundamental analyst analyzes an income statement. However, by assigning predefined rules and conditions to their techniques, technical analysts can become more objective.

Fundamental analysis can be considered the more theoretical approach because it seeks to determine the underlying long-term (or intrinsic) value of a security. Technical analysis can be considered the more practical approach because a technician studies the markets and financial instruments as they exist, even if trading activity appears, at times, to be irrational. Technicians seek to project the level at which a financial instrument *will* trade, whereas fundamental analysts seek to predict where it *should* trade.

A drawback of technical analysis is that technicians are limited to studying market movements and do not use other predictive analytical methods, such as interviewing the customers of a subject company to determine future demand for a company's products. Technicians study market trends and are mainly concerned with a security's price trend: Is the security trading up, down, or sideways? Trends are driven by collective investor psychology and generally change gradually, so that the trained technician can spot subtle shifts. However, markets can also change without warning. Additionally, it can take some time for a new trend to become evident. Thus, technicians may make wrong calls and have to change their opinions. Technicians are generally better at identifying market moves when the moves are already underway.

Moreover, trends and patterns must be in place for some time before they are recognizable, so a key shortcoming of technical analysis is that, try as it might to identify clues of an impending trend change, it typically still lags actual price data in such identification. This shortcoming mirrors a key shortcoming of fundamental analysis in that securities often overshoot fundamental fair values in an uptrend and undershoot fundamental fair values in a downtrend. Strictly relying on price targets obtained through fundamental analysis can lead to closing profitable investment positions too early because investors may irrationally bid security prices well above or well below intrinsic value.

Technical analysis and fundamental analysis can be seen as opposing approaches to analyzing securities, but in reality, many investors have experienced success by combining the two techniques. For example, an investor may use fundamental analysis to identify an undervalued stock and then use technical analysis to find specific entry and exit points for that position. Using fundamental and technical analysis together works well, for example, when a security is deemed fundamentally cheap but purchasing the security too early could prove costly. In this instance, technical analysis could help identify an attractive risk/reward entry point and moment in time to express the core fundamental view.

Some technical traders will look at fundamentals to support their trades as well. For example, a trader who is monitoring a **breakout** possibility near an earnings report may look at the fundamentals to get an idea of whether the stock is likely to beat earnings. A breakout refers to when the price of an asset moves above a resistance level.

A good example of a context in which technical analysis can be a superior tool to fundamental analysis is securities fraud. In such cases, fundamental analysts can continue to hold favorable views of a company's equity securities even as its share prices decline.

2.4 The Differences in Conducting/Interpreting Technical Analysis in Various Types of Markets

In general, technical analysis is a trading tool that requires liquidity, so it is best and most easily applied to liquid and deep markets. In other words, the use of technical analysis is limited in illiquid markets, where even modestly sized trades can have an inordinate impact on prices. For example, in considering a thinly traded American Depositary Receipt (ADR), analyzing the more heavily traded local security frequently yields a better analysis—without any potential currency-related distortions. **Gap openings** are more frequent in thinly traded securities and markets. A gap is an area of a chart where a security's price either rises or falls from the previous day's close with no trading occurring in between. A gap opening is the start of a new trading session with a gap.

Technical analysis can be applied to different asset classes, such as equities, commodities, and currencies. The success rate of technical analysis will depend on the nature of the price action in any asset class. Each market has its own character, and the role of the technician is to interpret—but not force—the technical evidence. With regard to equities, in many less-liquid frontier and emerging markets, local retail investors are active traders, whereas developed markets have a higher percentage of institutional investors.

Market participation (institutional vs. retail) can have an impact on technical analysis. The basic difference between retail and institutional investors is that retail investors tend to have less in-depth information (to be more naive) and to be more momentum-centric than institutional investors. As a result, retail investors may depend upon technical analysis and momentum trading somewhat more than institutional investors. Institutional investors also must have enough float/liquidity in a given stock to participate in scale, so are limited in their interest in micro-cap stocks with minimal liquidity. Strong trend periods are the result of inefficiencies, and inefficiencies can be more easily exploited in emerging and frontier markets.

Question 1

Which one of the statements below cannot be stated as a shortcoming of technical analysis?

a Technical analysts might be late in identifying a trend or pattern because they will require enough data to make a conclusion.

b Technical analysts might miss a buying opportunity because the fundamental fair value of a company undershoots in a downtrend.

c Technical analysis will have limited application in the case of an initial public offering (IPO).

Answer: B. The technical analyst will analyze only the price and volume data. If the stock is in a downtrend, the technical analyst will avoid buying it until a new uptrend is established in the analyzed time frame. For that reason, even though a company is fundamentally cheap, the technical analyst will wait for a buy signal that's confirmed by a chart pattern.

A is not the correct answer because being required to wait for confirmation can be viewed as a shortcoming of technical analysis.

C is not the correct answer because an IPO lacks price data, which are required for conducting technical analysis. Not being able to apply technical analysis to securities with limited price history is a shortcoming of technical analysis.

Question 2

Given the behavioral aspects of financial markets, under which conditions and market environments can an investor/analyst apply technical analysis?

a Equity markets with poor liquidity and with limited participation

b Pegged currency markets

c Blue chip technology stocks trading on the NASDAQ Stock Market

Answer: C. For technical analysis to be applied with good results, a market or a stock should have enough participation. Technical analysis measures crowd psychology. An individual's decision making might be difficult to understand and analyze. However, several individuals acting together produce predictable behavior patterns. Crowd behavior is similar for humans around the world. This is why technical analysis can be applied to different markets as long as there is enough human interaction in the form of buying and selling.

A is not the correct answer because in the case of poor liquidity and limited participation, prices are subject to manipulation. Technical analysis is most effective when prices are determined by demand and supply in free markets.

B is not the correct answer because government intervention is the opposite of what we are looking for in terms of free markets. In such cases, technical analysis will have limited scope.

Question 3

Retail and institutional investors behave differently in the way they invest and trade. Which one of the statements below is NOT one of the basic differences between retail and institutional investors?

a Retail investors tend to have less in-depth information when compared to institutional investors.

b Retail investors require more float/liquidity in given stocks in order to participate when compared with institutional investors.

c Retail investors tends to be more momentum-centric than institutional investors.

Answer: B. Retail traders usually trade with smaller size and can more easily buy and sell shares in less-liquid stocks, such as small-cap and micro-cap stocks.

A is incorrect because it is one of the basic differences between retail and institutional investors. Retail investors typically have limited information, whereas institutional investors typically have dedicated research departments and/or access to significant information.

C is incorrect because it is one of the basic differences between retail and institutional investors. Retail investors are more apt to trade based on market movements and intuition, whereas institutional investors follow a more in-depth investment process.

CHART TYPES

3

d describe and interpret different types of technical analysis charts;

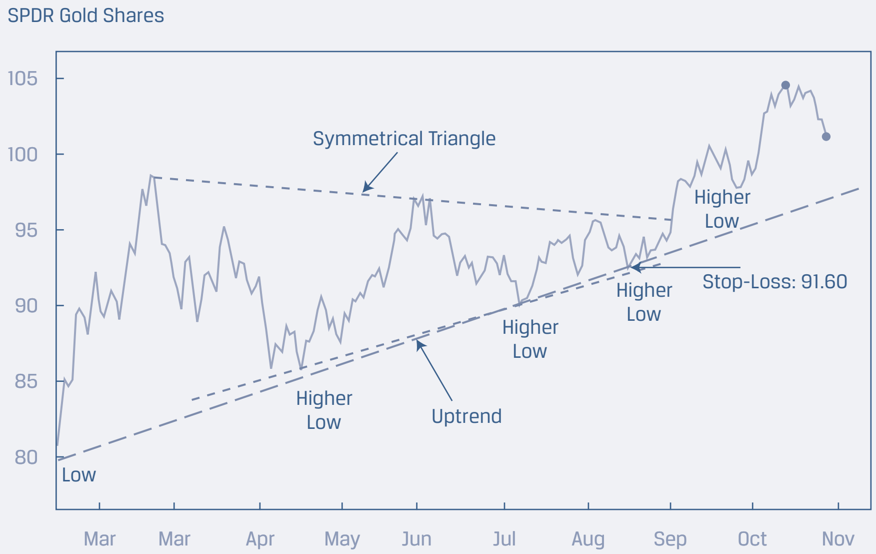
e explain uses of trend, support, and resistance lines;

f explain common chart patterns.

VIGNETTE

Scene 2

The technical analyst tells you that she can happily provide perspective from a technical point of view. **She prepares the following chart:**



She also prepares the following comments:

- Between February and September, the GLD price chart formed a symmetrical triangle chart pattern. The breakout took place in the beginning of September. Following the breakout, the price pulled back to the upper boundary of the chart pattern. The pullback offered an entry opportunity for those who missed the initial breakout in the beginning of September. Following the pullback, the GLD price resumed its upward movement in October.
- Since the beginning of the year, GLD has been in a steady uptrend. The uptrend can be seen in the higher lows. When the higher lows are connected with a trendline, a clear uptrend is visible.
- In late October, the price pulled back and put our long GLD position into an unrealized loss. However, the uptrend was not violated, nor was the chart pattern negated. The minor low at the 91.6 level inside the symmetrical triangle served as the protective stop-loss and the level at which the chart pattern would be negated.

Then she summarizes:

"***So, even though the position shows an unrealized loss, there is no reason to change the positive outlook at this point in time.***"

You breathe a sigh of relief now that you have additional perspective.

The primary tools used in technical analysis are charts and indicators. Charts are a graphical display of price and volume data, and the display may be done in a number of ways. Charts are then subjected to various analyses, including the identification of trends and patterns. Before we discuss several concepts regarding charts, it is important to note that not all price data are suitable for technical analysis. Some instruments (whether equities, futures, commodities, FX, or cryptocurrencies) might not have enough interaction between buyers and sellers to produce meaningful price patterns. Other instruments might become illiquid over time, so that the principles of charting might no longer be applicable.

3.1 Types of Technical Analysis Charts

Charts are an essential component of the technical analyst's toolkit. Charts present information about past price behavior and provide a basis for inferring likely future price behavior. With the help of technology, we have access to financial data through a variety of charts. MetaStock and TradeStation are examples of technical charting and analysis software. A variety of charts can be useful in studying the markets. The selection of the chart to use in technical analysis is determined by the intended purpose of the analysis.

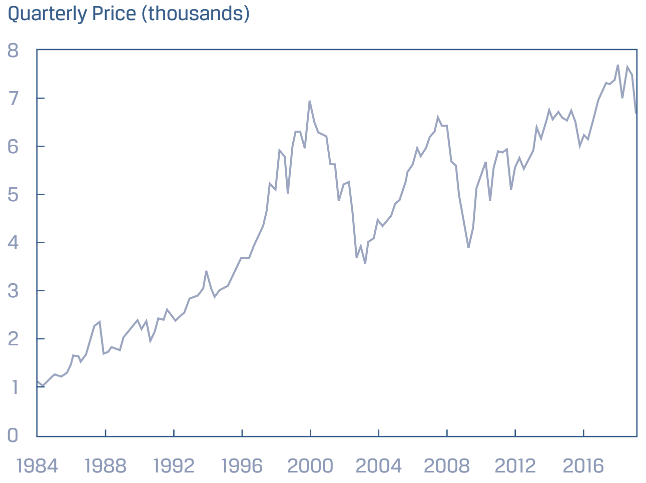
Advanced charting and trading software, utilizing artificial intelligence (AI), can help in simulating trading ideas or strategies as well as in **backtesting** results on historical price data. Backtesting is a method of assessing the viability of a trading strategy by showing how the strategy would play out using historical data.

3.1.1 *Line Chart*

Line charts are familiar to all types of analysts and are a simple graphic display of price trends over time. Line charts are typically drawn with closing prices as the data points. The vertical axis (y-axis) reflects price level, and the horizontal axis (x-axis) represents time. Even though the line chart is the simplest kind of chart, an analyst can quickly glean information from this representation of price data. In fact, line charts might be the most effective tool for analyzing price action because they show the closing price of the day, week, or month. The closing price is regarded by traders and investors as the most important data point, as it reflects the final decision for that period's transactions.

Exhibit 1 shows a quarterly chart of the FTSE 100 Index from 1984 through mid-2018. Up years and down years are clearly evident. The strong rally from 1984 through 1999 and the market decline from late 1999 to late 2002 are also clearly visible. The 2003–07 rally did not exceed the high reached in 1999, which suggests that investors were not willing to pay as high a price for stocks on the London Stock Exchange during that rally as they were in the prior rally. The 2007–09 decline didn't reach the lows of 2002, suggesting that investors viewed the prior recessionary period as a support level. From 2009 through mid-2018, the market was in a general uptrend with some pullback in 2015. This visual information in the price chart provides a broad overview of investor sentiment and can lead to further analysis. Importantly, the analyst can access and analyze this information quickly. Collecting and analyzing the full array of data normally incorporated in fundamental analysis would take much longer.

Exhibit 1 Line Chart: FTSE 100 Quarterly Price Data, 1984–mid-2018



3.1.2 *Bar Chart*

A line chart has one data point per time interval. A **bar chart**, in contrast, has four bits of data in each entry—the high and low prices encountered during the time interval plus the opening and closing prices. Such charts can be constructed for any time period, but they are customarily constructed from daily data. Traders or investors focusing on longer-term analysis can choose to plot their price data on a weekly or monthly scale. How a chart is interpreted will not depend on its periodicity. However, to get the best results, the analyst/investor or trader should decide on the time frame of operation.

As Exhibit 2 shows, a vertical line connects the high and low prices of the day; a cross-hatch to the right indicates the closing price, and a cross-hatch to the left indicates the opening price. The appeal of this kind of chart is that the analyst immediately gets a sense of the nature of that day's trading. A short bar indicates little price movement during the day; that is, the high, low, and close are near the opening price. A long bar indicates a wide divergence between the high and the low for the day.

Exhibit 2 Bar Chart Notation

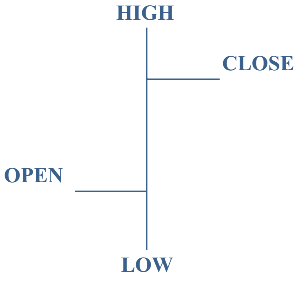


Exhibit 3 shows daily performance of the US dollar (USD) versus the Singapore dollar (SGD) spot FX from early 2018 through late 2018. As we will discuss in the latter sections of this reading, the analyst can identify an orderly, month-long consolidation between the 1.330 and 1.348 levels during the May–June period. Bar charts become more important in the analysis of short-term price actions, where the identification of support and resistance levels makes a major difference in analysis and trading results.

Spikes on the upside and on the downside will form support and resistance levels that can be seen only if the data are plotted as either a bar chart or a candlestick chart (to be discussed shortly). Exhibit 4 shows the same price data plotted as a line chart. With the line chart only, the analyst will not be able to capture the same consolidation period or clearly see the level at which the breakout opportunity will take place. Therefore, it is best to analyze price data that are subject to shorter-term trading by using either bar or candlestick charts.

Exhibit 3 **Bar Chart: US Dollar/Singapore Dollar, January 2018–August 2018 (daily data)**

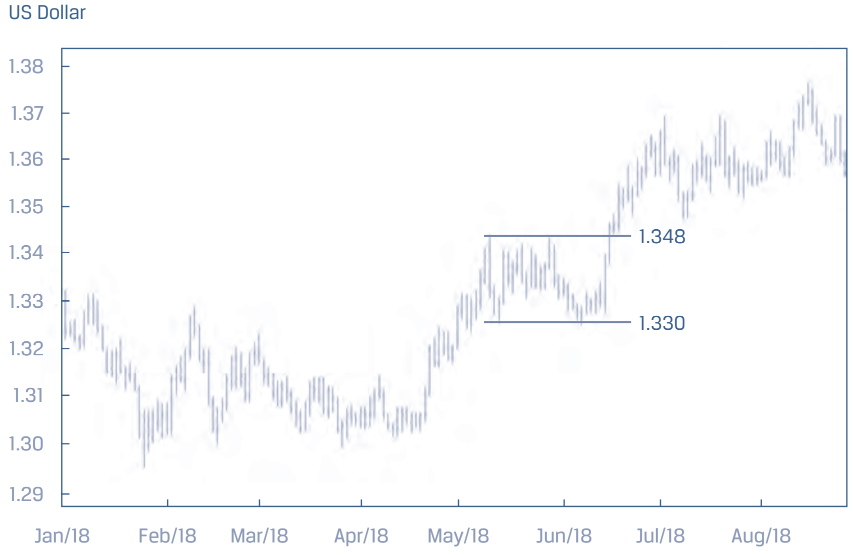
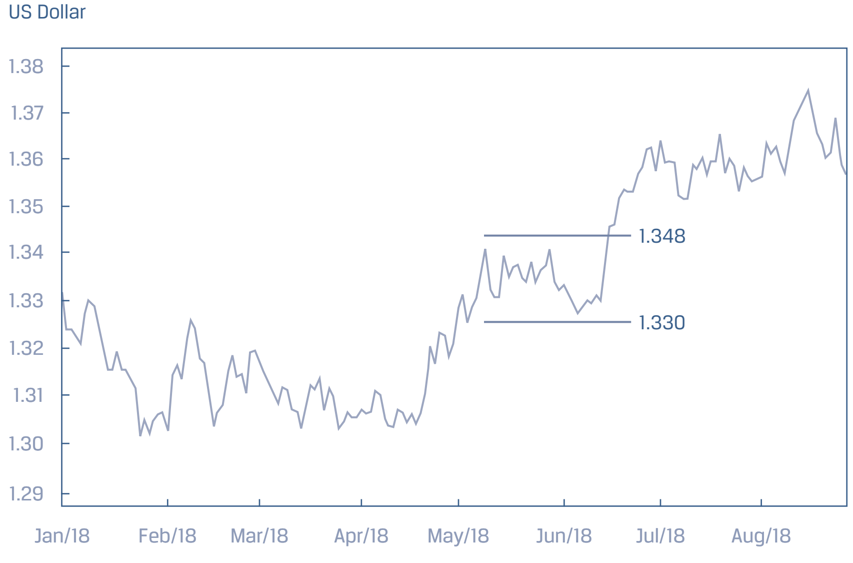


Exhibit 4 **Line Chart: US Dollar/Singapore Dollar, January 2018–August 2018 (daily data)**



3.1.3 *Candlestick Chart*

Candlestick charts trace their roots to Japan, where technical analysis has been in use for centuries. Like a bar chart, a **candlestick chart** provides four prices per data point: the opening and closing prices and the high and low prices during the period. As shown in Exhibit 5, a vertical line represents the range through which the security price traveled during the time period. The line is known as the wick or shadow. The body of the candle is white (or clear) if the opening price was lower than the closing price, and the body of the candle is dark if the opening price was higher than the closing price.

Exhibit 5 Construction of a Candlestick Chart

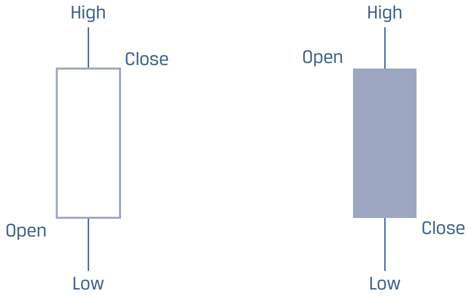
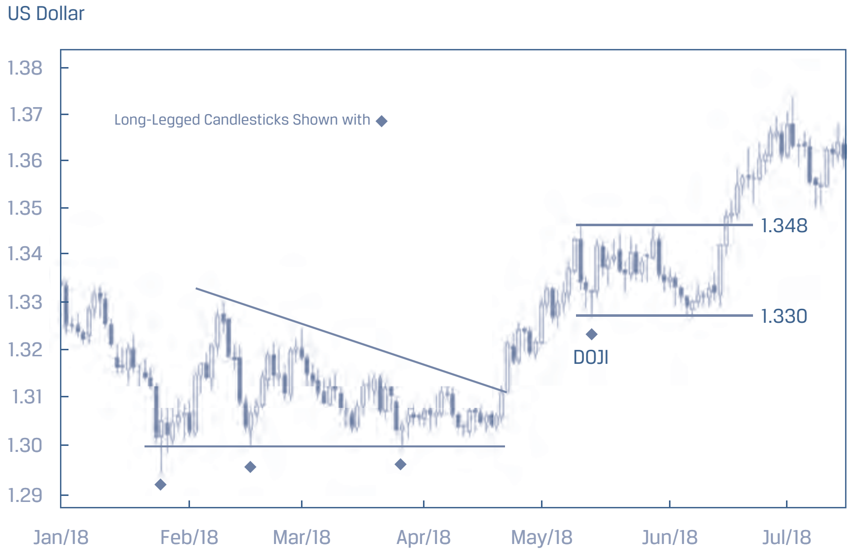


Exhibit 6 shows the US dollar/Singapore dollar daily price chart on a candlestick graph for the period 1 January through 13 July 2018.

Exhibit 6 Candlestick Chart: US Dollar/Singapore Dollar, 1 January–13 July 2018 (daily data)



The advantage of the candlestick chart over the bar chart is that price fluctuations are much more visible on the candlestick chart, which allows for better analysis. The bar chart indicates market volatility only by the height of each bar, but on candlestick charts, the difference between opening and closing prices and their relationship to the highs and lows of the day are clearly apparent. Long-legged candlesticks usually indicate the price finding support after meeting sellers during the day. We will discuss interpretation of candlestick chart patterns in subsequent sections. For now, we can highlight the importance of long shadows. In the Japanese terminology used in candlestick charting, one of the widely traded and analyzed candlestick patterns is called a **doji**. The doji signifies that after a full day of trading, the positive price influence of buyers and the negative price influence of sellers exactly counteracted each other—with opening and closing prices that are virtually equal—which suggests that the market under analysis is in balance. If a doji occurs at the end of a long uptrend or downtrend, it signals that the trend will probably reverse. The doji in mid-May 2018 not only marked a short-term trend reversal from the 1.33 level but also strengthened that level, which later became a support level during June 2018. The added value of such interpretation can be gained only by using candlestick graphs.

3.1.4 *Scale*

For any chart—line, bar, or candlestick—the vertical axis can be constructed with either a **linear scale** (also known as an arithmetic scale) or a **logarithmic scale**, depending on how the analyst wants to view the data. With a logarithmic scale, equal vertical distances on the chart correspond to an equal percentage change. With a linear scale, equal vertical distances on the chart correspond to an equal unit change. A logarithmic scale is appropriate when the analyst is working on longer time frames. A linear scale is better suited for shorter-term price charts. In addition, a logarithmic scale is appropriate when the data move through a range of values representing several orders of magnitude (e.g., from 10 to 10,000); a linear scale is better suited for narrower ranges (e.g., prices from $35 to $50). The difference between a logarithmic price chart and an arithmetic (linear) price chart can be small when you are analyzing a short time frame. However, when you look at longer-term charts (more than two years of price data), you will see major differences.

The horizontal axis shows the passage of time. The appropriate time interval depends on the nature of the underlying data and the specific purpose of the chart. An active trader, for instance, may find 10-minute, 5-minute, or even tick-by-tick data useful, but other technical analysts may prefer daily or weekly data. It is important to note that the shorter the time frame, the more random (and thus less meaningful) price action tends to become.

Consider Exhibits 7 and 8, which both show the yearly history of the Dow Jones Industrial Average (DJIA) from 1928 to 2018. Plotting the index on a linear scale, as in Exhibit 7, makes it difficult to gather much information from the first 60 years of the data series. Analysts can see a slight uptrend but not much else. The eye is drawn to the bull market of the 1980s, the subsequent dot-com bubble, the subprime crisis, and the sustained recovery. When the index is plotted on a logarithmic scale, as in Exhibit 8, however, the data tell a more comprehensive story. The Great Depression of the 1930s stands out, but over the following 80 years, the data follow a relatively stable upward trend.

Exhibit 7 Dow Jones Industrial Average on Linear Scale, 1920–2018 (US$)

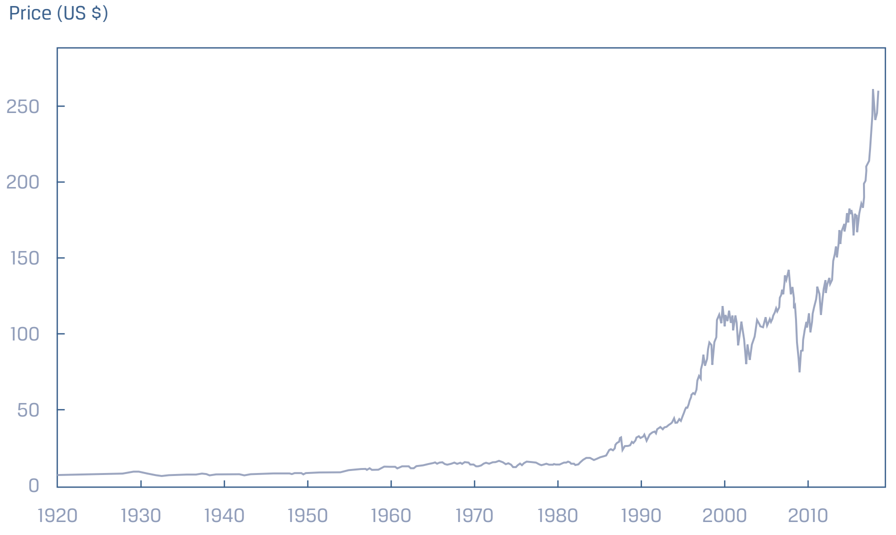
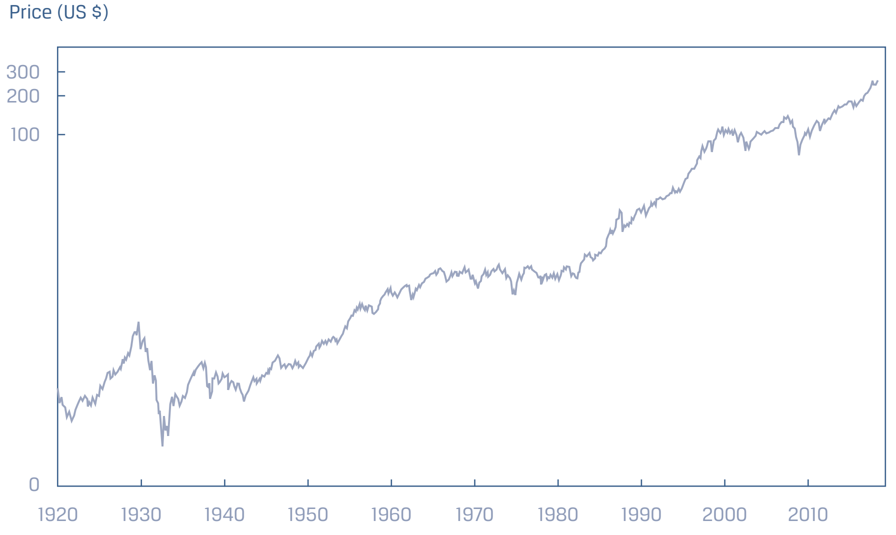


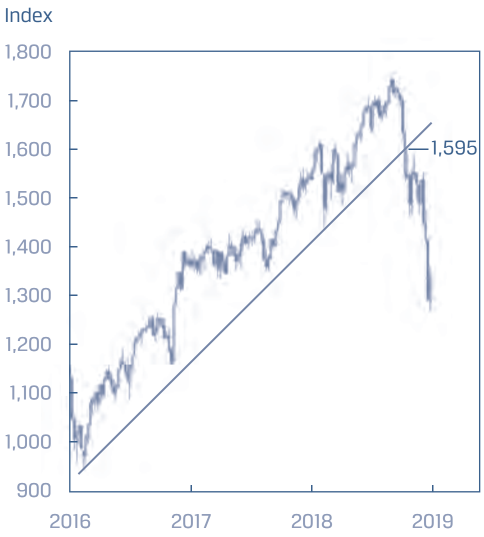
Exhibit 8 Dow Jones Industrial Average on Logarithmic Scale, 1920–2018 (US$)



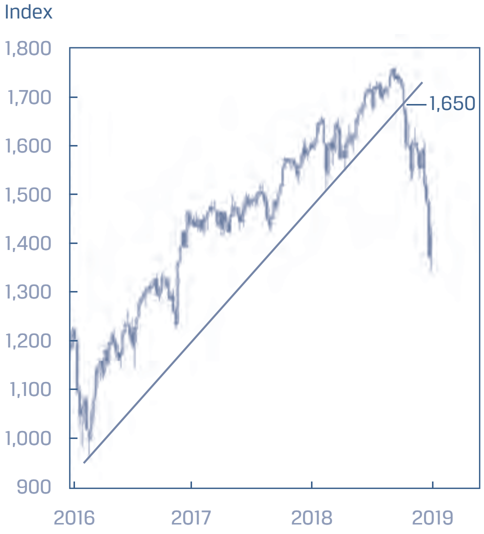
How you construct your scale will impact the trendlines you draw on your price chart. A good example is the price chart of the Russell 2000 Index ETF (IWM). Exhibit 9 shows a price chart of the Russell 2000 Index for the same time period plotted on a linear scale and on a logarithmic scale.

Exhibit 9 Russell 2000 Index Linear and Logarithmic Scales, 2016–2019

A. Linear Scale



B. Logarithmic Scale



The chart on the left is plotted on a linear scale. The chart on the right is plotted on logarithmic scale. The upward-sloping trendline breaks down earlier on the logarithmic scale than on the linear scale. Upward-sloping logarithmic trendlines are broken sooner than upward-sloping linear trendlines. The opposite is true for downtrends. Downward-sloping trendlines on logarithmic scale charts are broken later than downward-sloping trendlines on linear scale charts. The **breakdown** signal—and the resulting change in trend—takes place at the 1,650 level on the logarithmic scale chart, whereas the same signal occurs much lower—at the 1,595 level—on the linear scale price chart. A breakdown occurs when the price of an asset moves below a support level.

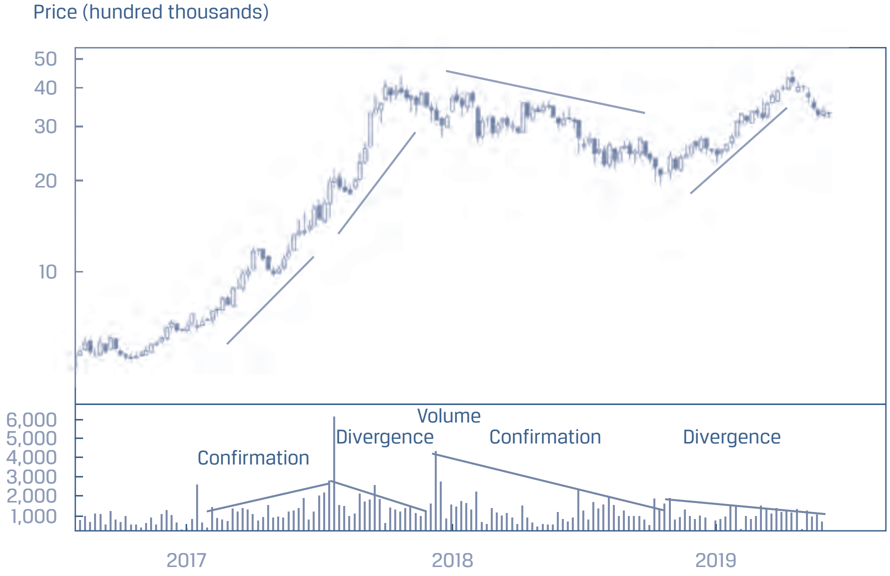
3.1.5 *Volume*

Volume is an important characteristic that is included at the bottom of many charts. Volume is used to assess the strength or conviction of buyers and sellers in determining a security's price. For example, on a daily price chart, a bar chart below the price section will show the volume traded for that day.

Some technicians consider volume information to be crucial. If volume increases during a time frame in which price is also increasing, that combination is considered positive and the two indicators are said to “confirm” each other. The signal would be interpreted to mean that over time, more and more investors are buying the financial instrument, and they are doing so at higher and higher prices. This pattern is considered a positive technical development.

Conversely, if volume and price diverge—for example, if a stock’s price rises while its volume declines—the implication is that fewer and fewer market participants are willing to buy that stock at the new price. If this trend in volume continues, it is expected that the price rally will soon end because demand for the security at higher prices will cease. However, there are cases where a breakout can take place with low volume. Exhibit 10 shows a price chart for Sunac China Holdings with volume displayed separately.

Exhibit 10 Weekly Candlestick Price Chart and Volume Bar Chart: Sunac China Holdings, August 2016–May 2019 (Hong Kong dollars)



3.1.6 *Time Intervals*

Charts can be constructed using any time interval. The same principles of technical analysis apply irrespective of the time interval. Using longer intervals allows the analyst to chart longer time periods, for the simple reason that longer intervals contain fewer data points, so a longer time frame can be presented on the chart. A useful step for many analysts is to begin the analysis of a security with the chart for a long time frame, such as a weekly or monthly chart, and then construct charts with shorter and shorter time intervals, such as daily or hourly charts. Over the years, an increase in computerized trading and high-frequency trading (HFT) has resulted in more frequent false and distorted signals in shorter-term analysis. Traders, investors, and analysts should bear in mind the randomness involved in very short-term (intraday) analysis and trading.

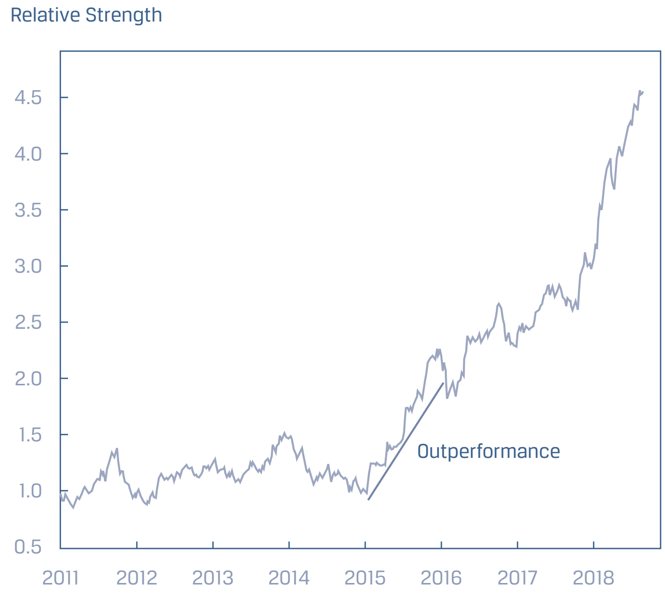
3.1.7 *Relative Strength Analysis*

Relative strength analysis is widely used to compare the performance of a particular asset, such as a common stock, with that of some benchmark—such as, in the case of common stocks, the FTSE 100, the Nikkei 225, or the S&P 500 Index—or with the performance of another security. The intent is to show outperformance or underperformance of the individual issue relative to some other index or asset. Typically, the analyst prepares a line chart of the ratio of two prices, with the asset under analysis as the numerator and with the benchmark or other security as the denominator. A rising line shows the asset is performing better than the benchmark or other stock; a declining line shows the opposite. A flat line shows neutral performance.

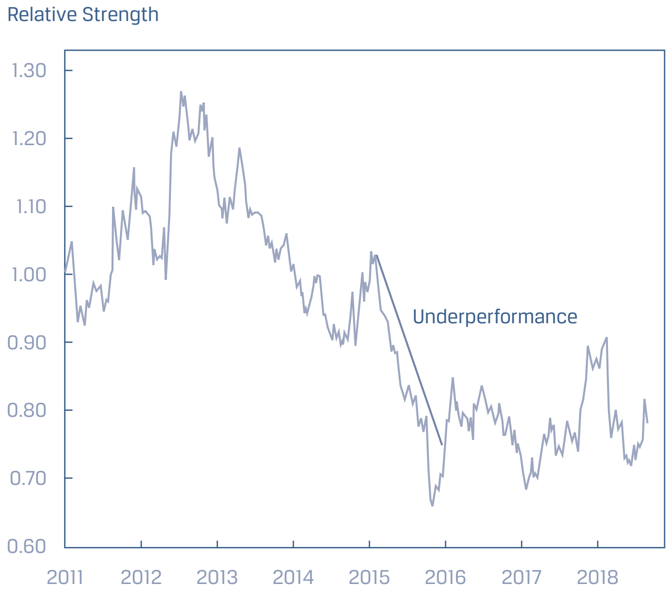
Suppose a private investor wants to understand changing market trends for two investment ideas she has read about. Amazon Inc. (AMZN) is a well-known technology-enabled retail company, and Walmart Inc. (WMT) is a US-based multinational retail company. The investor wants to determine which of the two companies' stocks has been the stronger performer (relative to the S&P 500) over the roughly seven-year period ending August 2018. Exhibit 11 shows relative strength lines for the two stocks between 2011 and 2018. For ease of comparison, the ratio of each retail company's price data to the S&P 500 is indexed to 1.00 at the beginning of 2011. A move from 1.00 to 1.50 on the chart in Panel A indicates a 50% outperformance for Amazon versus the S&P 500. Likewise, a drop from 1.00 to 0.80 on the chart in Panel B indicates a 20% underperformance for Walmart versus the S&P 500.

Exhibit 11 **Relative Strength Analysis of Two Retail Giants: AMZN vs. the S&P 500 and WMT vs. the S&P 500, January 2011–August 2018**

A. Amazon.com Inc. vs. S&P 500



B. Walmart vs. S&P 500



The units on the vertical axis are not significant; the ratio is a function of the relative prices of the assets under consideration. The important information is how the ratio has changed. This type of chart allows an analyst to make a visual determination of that change. As Exhibit 11 illustrates, Amazon was a strong performer starting in 2015, but its relative performance wasn't as strong prior to that year. In contrast, the stock of Walmart lost its leadership in the beginning of 2013, and from that point up to mid-2018, its stock has clearly lagged the performance of the S&P 500.

TREND, SUPPORT, AND RESISTANCE

4

e explain uses of trend, support, and resistance lines;

The concepts of **trend** and **consolidation** are perhaps the most important aspects of technical analysis. A trend is a long-term pattern of movement in a particular direction. When a security is not trending, it is considered to be in a consolidation. Trend analysis is based on the observations that market participants tend to act in herds and that trends tend to stay in place for some time. A security can be considered to be in an uptrend or a downtrend. The timing of buy and sell decisions will depend on how well we are able to differentiate between a consolidation and a trend phase. Not all securities are in a trend. When a security is not trending, it offers opportunity for traders who buy/sell between well-defined ranges. In the latter part of the reading, we will discuss how to interpret each consolidation period with the help of chart patterns. Not every chart will have obvious or clear implications, so the analyst must avoid the temptation to force a conclusion from every chart, which may lead to a wrong interpretation.

A security is in an **uptrend** when the price goes to higher highs and higher lows. In an uptrend, the forces of demand are greater than the forces of supply. So, traders are willing to pay higher and higher prices for the same asset over time. Presumably, the strong demand indicates that investors believe the intrinsic value of the security is increasing. As the price of the security moves up, each subsequent new high is higher than the prior high, and each time there is a **retracement**, which is a reversal in the movement of the security's price, it must stop at a higher low than the prior lows in the trend period. To draw an uptrend line, a technician draws a line connecting the lows of the price chart. Major breakdowns in price, however, when the price drops through and below the trendline, indicate that the uptrend is over and may signal a further decline in the price. Minor breakthroughs below previous lows simply call for the line to be moderately adjusted over time. Time is also a consideration in trends: The longer the security price stays below the trendline, the more meaningful the breakdown is considered to be.

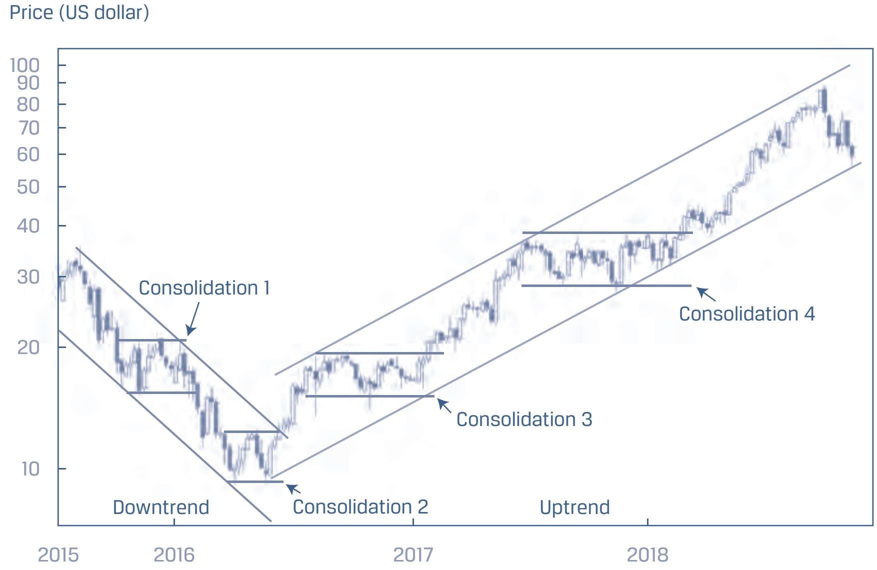
In a **downtrend**, a security makes lower lows and lower highs. As the price of the security moves down, each subsequent new high must be lower than the prior high, and each time there is a retracement, the price must stop at a lower low than the prior lows in the trend period. To draw a downtrend line, a technician draws a line connecting the highs of the price chart. Major breakouts above the downtrend line indicate that the downtrend is over and a rise in the security's price may occur. And as with an uptrend, the longer the security price stays above the trendline, the more meaningful the breakout is considered to be.

In a downtrend, supply overwhelms demand. Over time, sellers are willing to accept lower and lower prices to exit long positions or enter new short positions. This seller behavior generally indicates deteriorating investor sentiment about the asset. However, selling may be prompted by factors not related to the fundamental or intrinsic value of the stock. For example, investors may be forced to sell to meet margin calls in their portfolios. From a purely technical standpoint, the reason is irrelevant. The downtrend is assumed to continue until contrary technical evidence appears. Combining fundamental analysis with technical analysis in such a case, however, might reveal a security that has attractive fundamentals but a currently negative technical position.

A security may trade in a fairly narrow range, moving sideways on the price chart without much upward or downward movement. This pattern indicates a relative balance between supply and demand. A technical analyst might not expect to profit from long or short trades in such securities but might devise profitable option strategies for short-term investors with the ability to accept the risks. For position traders and investors, consolidations are pauses in trends. Such pauses also present an opportunity to assess the strength or weakness of that security. Each consolidation period is eventually followed by a trend period. For buyers of that security, it is important to know how the subsequent trend phase will develop. For sellers of the security, it is important to watch for any change in trend.

Exhibit 12 shows the application of trend and consolidation analysis. It is important to note that each trend period will be followed by a consolidation and each consolidation will give rise to a trend period. Most position traders and investors would like to capture these trend periods. Therefore, differentiation or identification of those consolidation and trend periods becomes an important motive of a **chartist**. A chartist is an individual who uses charts or graphs of a security's historical prices or levels to forecast its future trends.

Exhibit 12 Trend and Consolidation Analysis: Teladoc Health Price Chart, 2015–2019 (US$)



The chart in Exhibit 12 covers roughly four years and would most likely be used by investors with a long time horizon. There are four distinct consolidation periods on the price chart. Identification of those consolidation periods would have helped in making investment decisions. Following its IPO, Teladoc Health started trending lower. In the last quarter of 2015, the stock formed its first consolidation (Consolidation 1). Once the boundaries of the consolidation became clear, the chartist should have asked the question, Which way will the consolidation resolve? If it resolved on the downside, this outcome would suggest that the downtrend would resume. If it resolved on the upside, this outcome would suggest a change in trend, a trend reversal. In the last month of 2015, Teladoc broke down its consolidation to resume its downtrend.

Between March 2016 and May 2016, the stock formed another consolidation (Consolidation 2). This time a breakout on the upside followed. The implication of this breakout was a major trend reversal. It offered a long-term opportunity for investors and traders.

Initial upward thrust pushed Teladoc from around the 12 level to around the 19 level. You will remember that each trend period is followed by a consolidation. Consolidation 3, which took six months to complete, offered an opportunity for traders with short horizons. This was also an opportunity to assess the health of the initial positive change in trend. Once the boundaries of the consolidation became clear, the chartist should have asked the question, Which way will the consolidation resolve? Would Teladoc offer a new breakout opportunity and a continuation of the existing uptrend? The new breakout occurred in January 2017. At this point, we had an established uptrend with higher lows and higher highs.

The continuation of the uptrend took Teladoc from around the 19 level to around the 38 level. At that point, the stock entered into another consolidation (Consolidation 4). From May 2017 until February 2018, the stock remained in a tight consolidation range, which acted as another opportunity for traders with short horizons. These types of consolidations, once they are identified and mature, offer traders a chance to sell at the upper boundary and buy back their shares at the lower boundary. For position traders and investors, the consolidation is a time to assess the strength of the stock. Consolidation 4 was completed in March 2018 with a breakout above the 38 level, and the uptrend resumed higher, reaching the 89 level by the last quarter of 2018.

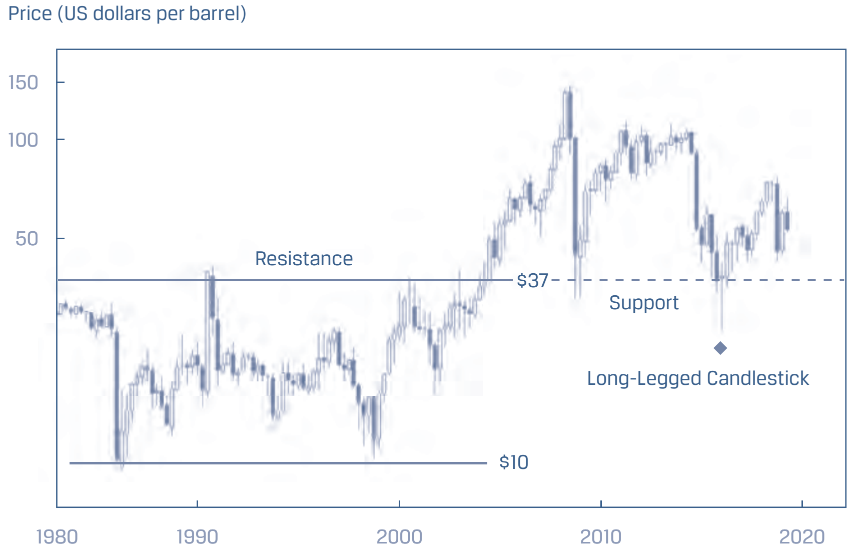
Two concepts related to trend are **support** and **resistance**. Support is defined as a low price range in which buying activity is sufficient to stop the decline in price. It is the opposite of resistance, which is a high price range in which selling activity is sufficient to stop the rise in price. The psychology behind the concepts of support and resistance is that investors have come to a consensus about the price of a security. Support and resistance levels can be sloped lines, as in trendlines or horizontal lines.

A key tenet regarding support and resistance as a part of technical analysis is the **change in polarity principle**, which states that once a support level is breached, it becomes a resistance level. The same holds true for resistance levels; once breached, they become support levels.

Support indicates that at some price level, investors consider a security to be an attractive investment and are willing to buy, even in the wake of a sharp decline. (And resistance indicates that at some level, investors are not willing to buy, even in an uptrend.)

In Exhibit 13, we see an example of support and resistance on a widely followed commodity, WTI Light Crude Oil. For more than two decades, this commodity's price fluctuated between \$10 per barrel and \$37 per barrel. In 2014, the balance between supply and demand changed and pushed the price above the historical resistance at \$37. In four years' time, the commodity was trading as high as the \$147 level. High commodity prices had significant impact on global economies. During the subprime crisis and related sell-off, the WTI Light Crude Oil price fell back to the \$37 level. Previous resistance acted as a strong support. The second time the WTI price tested the \$37 support level was in 2016. Even though the price dipped below the support, the close for the quarter was around the \$38 level. The price chart formed a long-legged candlestick, suggesting the strong support was still valid.

Exhibit 13 Support Level: WTI Light Crude Oil Quarterly Price Chart, 1980–2019 (US$ per barrel)



5

COMMON CHART PATTERNS

f explain common chart patterns;

Chart patterns are formations that appear in price charts that create some type of recognizable shape over time as the single bar lines accumulate. Common patterns appear repeatedly and often lead to similar subsequent price movements. Thus, the identification and analysis of chart patterns is a basic aspect of how technical analysis is used to predict security prices. An important connection to remember is that because patterns form as a result of the behavior of market participants, these patterns are graphical depictions of the collective psychology of the market at a given time. The recurring patterns that appear in charts can be used as a basis for market analysis.

Chart patterns, while most can be analyzed as a type of consolidation, can be divided into two categories: **reversal patterns** and **continuation patterns**. These terms refer to the trend for the security in question prior to the formation of the pattern (consolidation). The most important concept to understand in using chart patterns is that not every chart will lend itself to easy interpretation. Some charts will be clearer than others, and the analyst should not force an interpretation on any chart. This aspect is frequently forgotten by investors, who are so eager to identify and use patterns that they forget the proper application of charts.

5.1 Reversal Patterns

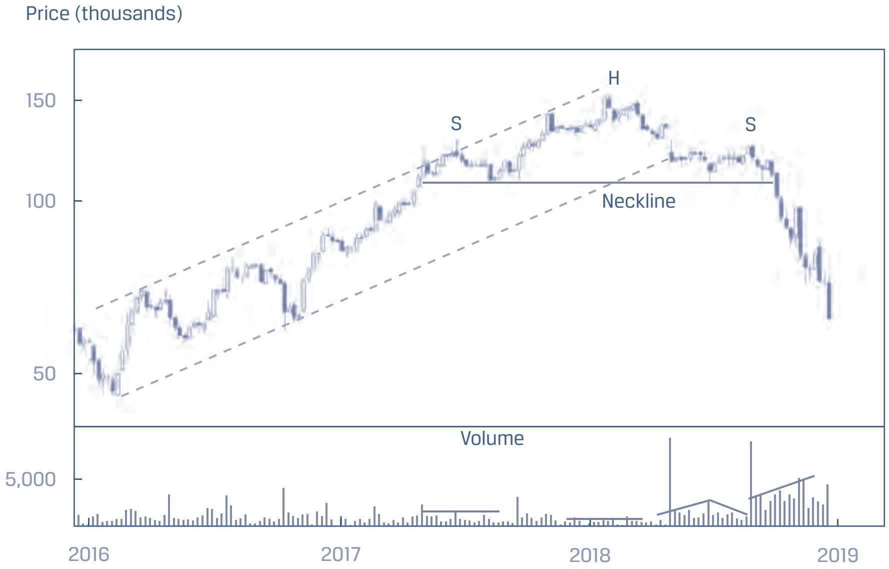
As the name implies, a reversal pattern signals the end of a trend, a change in direction for the financial instrument's price. Evidence that a trend is about to change direction is obviously important, so reversal patterns are noteworthy.

5.1.1 *Head and Shoulders*

Perhaps the most widely recognized reversal pattern is the **head and shoulders pattern**. This pattern consists of three segments. Volume is an important consideration in interpreting this pattern. Because the head and shoulders pattern indicates a trend reversal, it is important to establish the existence of a prior uptrend. Without a prior uptrend to reverse, there cannot be a head and shoulders reversal pattern. Later, we will discuss the *inverse* head and shoulders pattern (preceded by a downtrend).

Exhibit 14 depicts a head and shoulders pattern for Marriott Vacations.

Exhibit 14 Head and Shoulders Pattern: Marriott Vacations Weekly Price Chart, January 2016–December 2018 (US$)



The three parts of the pattern are as follows:

1. Left shoulder: While the security is in an uptrend, the left shoulder forms a peak that marks the high point of the current trend. Often, but not always, this first shoulder may represent the highest volume of the entire pattern. After this peak, a decline ensues to complete the formation of the shoulder.
2. Head: From the low of the left shoulder, an advance begins that exceeds the previous high and marks the top of the head. After this peak, the low of the subsequent decline marks the second point of the neckline. Formation of the head is the first signal that the rally may be coming to an end and that a reversal may be starting. Volume typically will wane a bit into the head high.
3. Right shoulder: The advance from the low of the head forms the right shoulder. This peak is lower than the head (a lower high) and usually in line with the high of the left shoulder. A textbook example should have symmetry between the shoulders in both time and price. The decline from the peak of the right shoulder should break the neckline. It is important that volume be lowest on the second shoulder.

In addition, the following three elements are key for the head and shoulders pattern: neckline, volume, and price target.

- Neckline: The neckline forms by connecting the beginning of the left shoulder and the end of the right shoulder. Depending on the relationship between the two low points, the neckline can slope up, slope down, or be horizontal. A textbook example should have a horizontal chart pattern boundary (neckline) where the lows of both shoulders just touch the horizontal line.
- Volume: As mentioned above, volume plays an important role in the confirmation of a head and shoulders pattern. Ideally, but not always, volume during the advance of the left shoulder should be higher than during the advance of the head. The decrease in volume into the new high of the head then serves as an initial warning sign. The next warning sign comes when volume increases on the decline from the peak of the head, then notably decreases during the advance of the right shoulder. Final confirmation comes when volume increases once again during the decline from the right shoulder.
- Price target: After the price breaks the neckline support, the projected price decline is found by measuring the distance from the neckline to the top of the head. This distance is then subtracted from the neckline to reach a price target. Price targets should be used as guidelines. Price may exceed the chart pattern objective or fall short.

Rarely will an analyst see a perfectly formed head and shoulders pattern; variations include two tops on the shoulders or on the head. The head, however, should rise to a higher price level than either shoulder, whereas the shoulders should be roughly symmetrical. In terms of the neckline price level, the first rally should begin at this level and the left shoulder and head should also decline to roughly this level. But necklines may not always form exactly horizontal lines.

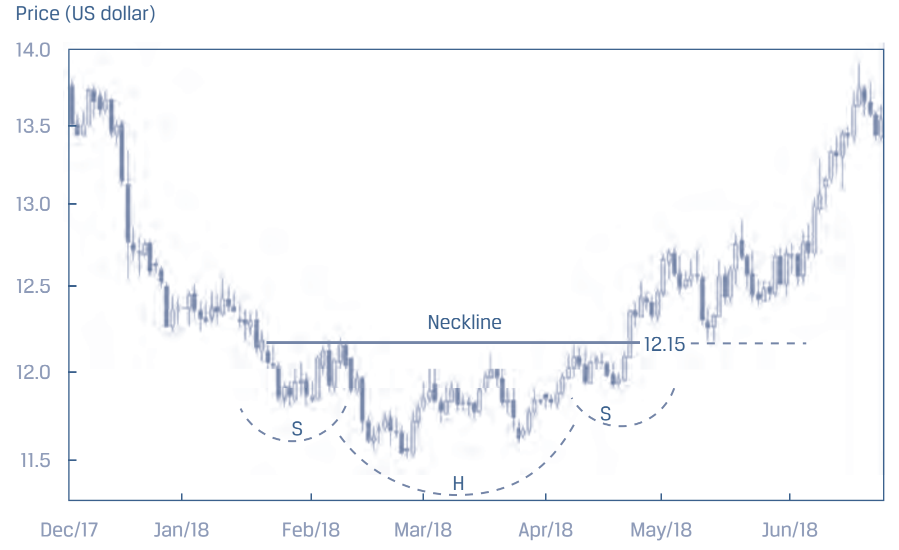
Volume is important in analyzing head and shoulders patterns. A new high in price at the top of the head without a new high in volume signals the presence of fewer bullish market participants. When one indicator is making a new high (or low) but another is not, this situation is called **divergence**.

Once the head and shoulders pattern has formed, the expectation is that the share price will decline through the neckline price. Technicians tend to use filtering rules to make sure that a clear breakdown of the neckline has occurred. These rules may take the form of waiting to trade until the price falls to a meaningful level below the neckline (3% and 5% are commonly used) and/or until the price has remained below the neckline for a specified length of time; when a daily price chart is used, the rule may be several days to a week. Prices commonly rebound to the neckline levels, even after a decline has exceeded the filter levels. Prices generally stop, however, at or around the neckline. The neckline was a support level, and under the change in polarity principle, once a support level is breached, it becomes a resistance level.

5.1.2 *Inverse Head and Shoulders*

The head and shoulders pattern can also form upside down and act as a reversal pattern for a preceding downtrend. Inverse head and shoulders is also referred to as a head and shoulders bottom.

Exhibit 15 Inverse Head and Shoulders Pattern: US Dollar/South African Rand Daily Price Chart, December 2017–July 2018



The three parts of the inverse head and shoulders are as follows:

- Left shoulder: This shoulder appears to show a strong decline, with the slope of the decline greater than that of the prior downtrend. The price movement then reverses back to the level where the rally started, forming a V shape.
- Head: The head is a more pronounced version of the left shoulder. Another decline follows, which takes the price to a lower low than the left shoulder by a significant enough margin that it is clearly evident on the price chart. This second decline also reverses, with price rising to the same level at which the left shoulder began and ended. This price level, the neckline, will also be above the downtrend line formed by connecting the high prices in the downtrend preceding the beginning of the inverse head and shoulders pattern. This pattern is the first signal that the decline may be coming to an end and that a reversal may be near. It is possible for the formation of the head to involve a more complex structure. A head can contain two minor lows, as in the US dollar/South African rand example presented in Exhibit 15.
- Right shoulder: The right shoulder is roughly a mirror image of the left shoulder, signifying less selling enthusiasm. The price declines to roughly the same level as the left shoulder, but the rally reverses at a higher low price than the rally that formed the head.

Volume once again is important to watch and should generally diminish into the final reverse shoulder before ideally exploding higher on the subsequent neckline break.

5.1.3 *Setting Price Targets with the Head and Shoulders Pattern*

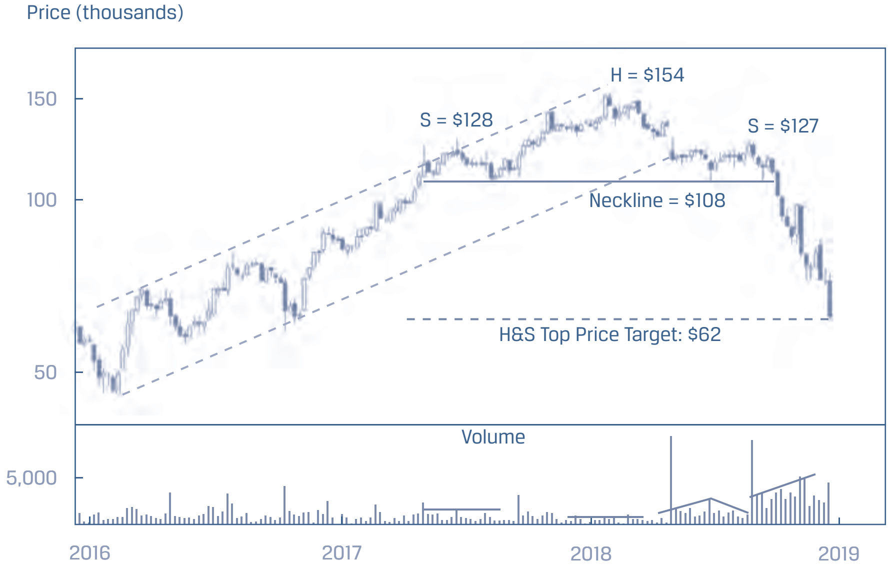
Like all technical patterns, the head and shoulders pattern must be analyzed from the perspective of the security's long-term price trend. The rally before the formation of the pattern must be large enough for there to be something to reverse. The stronger and more pronounced the rally was, the stronger and more pronounced the reversal is likely to be. Similarly, once the neckline is breached, the security is expected to decline by an amount equal to the change in price from the neckline to the top of the head. If the preceding rally started at a price higher than the neckline, however, the correction is unlikely to bring the price lower than the level at the start of the rally. Because a head and shoulders formation is a bearish indicator (i.e., a technician would expect the previously established uptrend to end and a downtrend to commence), a technician would seek to profit by shorting the security under analysis. When attempting to profit from the head and shoulders pattern, a technician will often use the price difference between the head and the neckline to set a price target, which is the price at which the technician anticipates closing the investment position. The price target for the head and shoulders pattern is calculated as follows:

Price target = Neckline – (Head – Neckline).

For example, in Exhibit 16, the high price reached at the top of the head is roughly $154 and the neckline formed at roughly $108, for a difference of $46. So a technician would expect the price to decline to a level $46 below the neckline—that is, to $62:

Price target = $108 – ($154 – $108) = $62.

Exhibit 16 Calculating a Price Target: Marriott Vacations Weekly Price Chart, January 2016–December 2018 (US$)



Price targets should be used as guidelines. With stocks in a downtrend and creating bearish chart patterns, a conservative approach to calculating price objectives is warranted. The conservative approach is to take the percentage difference between the high point of the head and the neckline and project it below the breakdown level (the neckline). Note that with a low-priced stock, such as one with a price of $0.50, there is a possibility of projecting price levels below 0. Because negative values for an asset are not possible, measuring price objectives using a percentage decline is more reasonable.

EXAMPLE 1

Determining a Price Target from a Head and Shoulders Pattern

Danielle Waterhouse is the technical analyst at Kanektok Securities. One of the companies her firm follows is LPA Petroleum. Waterhouse believes that a graph of LPA's share prices over the past six months reveals a classic head and shoulders pattern. The share price peaked at US$108, and she estimates the neckline at US$79. At today's close, the shares traded at US$78. Based on the head and shoulders pattern, what price target should Waterhouse estimate?

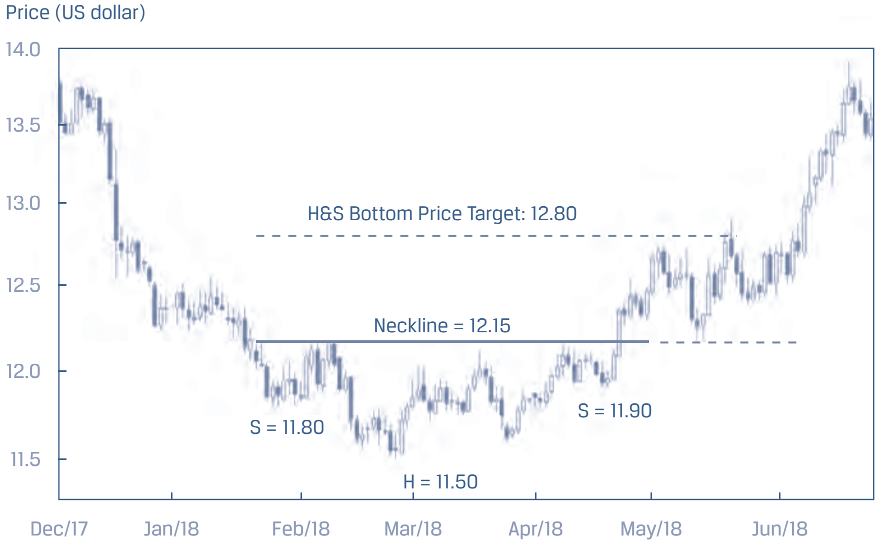
Solution:

Waterhouse estimates the neckline at US$79, which is US$108 minus US$79, or US$29, lower than the head. Her price target is thus US$79 minus US$29, which is US$50. Waterhouse would attempt to sell LPA short at today's price of US$78 and anticipate closing the position at US$50 for a profit of US$28 per share (not accounting for transaction costs).

5.1.4 *Setting Price Targets with the Inverse Head and Shoulders Pattern*

Calculating price targets for inverse head and shoulders patterns is similar to the process for head and shoulders patterns, but because the inverse head and shoulders pattern predicts the end of a downtrend, the technician calculates how high the price is expected to rise once it breaches the neckline. Exhibit 17 illustrates an inverse head and shoulders pattern.

Exhibit 17 Calculating a Price Target for an Inverse Head and Shoulders Pattern: US Dollar/South African Rand Daily Price Chart, December 2017–July 2018



For an inverse head and shoulders pattern, the formula is similar to that for a head and shoulders pattern:

Price target = Neckline + (Neckline – Head).

For example, in the price chart in Exhibit 17, the low price reached at the bottom of the head is roughly 11.5 and the neckline formed at roughly 12.15. The target can thus be calculated as 12.15 + (12.15 – 11.50) = 12.80. In this case, a technician might have taken a long position on 23 April, with the strong daily breakout above the neckline at the 12.15 level, and projected a possible exit near the 12.80 level.

5.1.5 *Double Tops and Bottoms*

A **double top** is formed when an uptrend reverses twice at roughly the same high price level. Typically, volume is lower on the second high than on the first high, signaling a diminishing of demand. The longer the time between the two tops and the deeper the sell-off after the first top, the more significant the pattern is considered to be. Price targets can be calculated from this pattern in a manner similar to the calculation for the head and shoulders pattern. For a double top, price is expected to decline below the low of the valley between the two tops by at least the distance from the valley low to the high of the double top.

EXAMPLE 2

Determining a Price Target from a Double Top Pattern

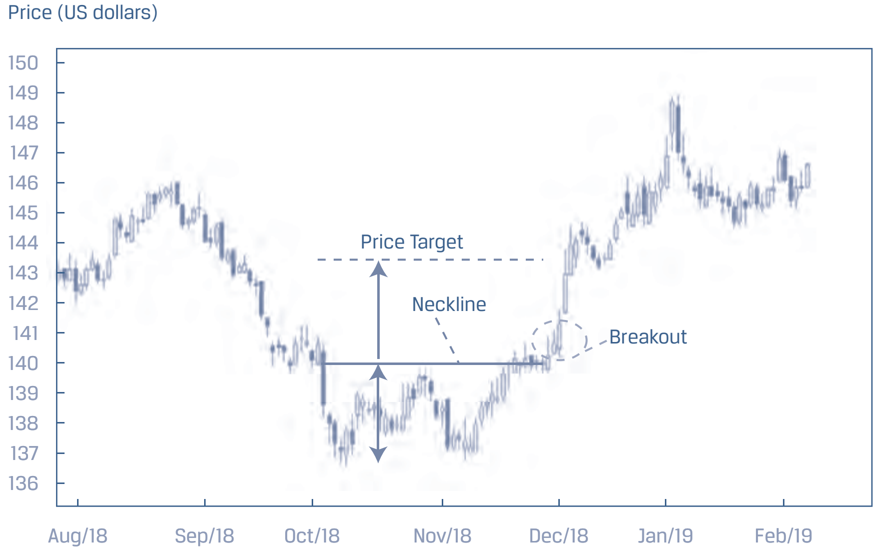
Richard Dupuis is a technician who trades Eurodollar futures for his own account. He analyzes charts based on one-minute intervals, looking for short-term trading opportunities. Eurodollar futures contracts have been trending upward most of the morning, but Dupuis now observes what he believes is a double top pattern: After peaking at US$97.00, the futures contract price falls to US$96.42, climbs again to US$97.02, and then starts a decline. Because of the double top, Dupuis anticipates a reversal from the uptrend to a downtrend. Dupuis decides to open a short position to capitalize on the anticipated trend reversal. What price target should Dupuis estimate for closing the position?

Solution:

Dupuis estimates the price target as US$96.42 – (US$97.02 – US$96.42) = US$95.82.

A **double bottom** is formed when the price reaches a low, rebounds, and then declines again to the first low level. Exhibit 18 depicts a double bottom pattern for US 30-year Treasury bond futures (first-month continuation price). Technicians use the double bottom to predict a change from a downtrend to an uptrend in security prices. The neckline for a double bottom is the horizontal line that touches the minor high between the two major troughs. A breakout takes place once the price breaches the neckline. The distance from the resistance breakout to the trough lows can be added to the area above the resistance breakout to estimate a target.

Exhibit 18 Double Bottom Pattern: US 30-Year Treasury Bond Daily Price Chart, August 2018–February 2019 (first-month continuation futures, US$)

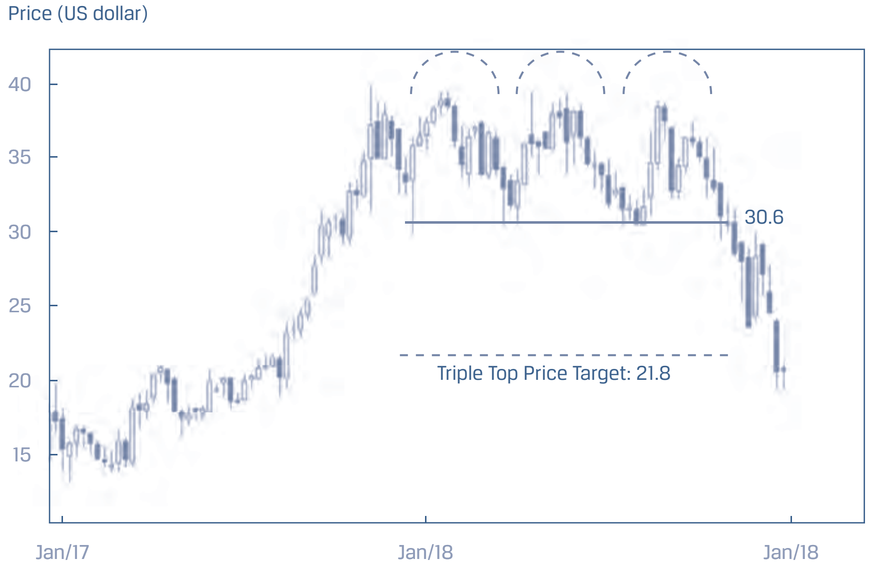


The reason these patterns are significant is that they show that at some price point, investors step in to reverse trends that are underway. For an uptrend, a double top implies that at some price point, enough traders are willing to sell positions (or enter new short positions) that their activities overwhelm and reverse the uptrend created by demand for the shares. A reasonable conclusion is that this price level has been fundamentally derived and that it represents the consensus of investors on the intrinsic value of the security. With double bottoms, if a decline in a security stops at the same price point on two separate occasions, the analyst can conclude that the market consensus is that at that price point, the security is cheap enough to be an attractive investment.

5.1.6 *Triple Tops and Bottoms*

A **triple top** consists of three peaks at roughly the same price level, and a **triple bottom** consists of three troughs at roughly the same price level. A triple top for Odfjell Drilling during 2018 is shown in Exhibit 19.

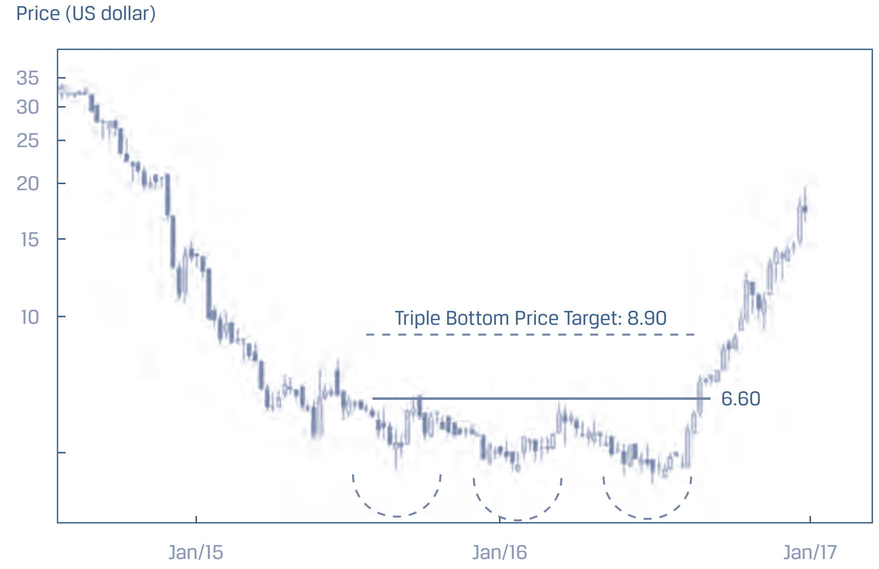
Exhibit 19 Triple Top Pattern: Odfjell Drilling Weekly Price Chart, January 2017–January 2019 (Norwegian krone)



One of the challenges with double top and triple top patterns—and one of the valid criticisms of technical analysis in general—is that an analyst cannot know which pattern will result until after the fact. There is no evidence that market corrections must end with a double bottom or that rallies must end with a double top, and there is no generally accepted technical theory that predicts whether a low will be repeated once or even twice before a reversal occurs. A double bottom is considered to be a more significant pattern than a single bottom because traders have stepped in on two occasions to halt declines. However, traders have no way to determine whether a double top or bottom will be followed by a third top or bottom. The goal of a trader, investor, or analyst is not to predict possible price action but to identify well-defined levels for decision making. It is the breach of these levels—the breakout—that matters most because it completes the lengthy consolidation and gives rise to a new trend period. Triple tops and triple bottoms are rare, but when they occur, they are more significant reversal patterns than double tops or double bottoms. On three separate occasions, traders stepped in to sell or buy shares with enough volume to end a rally or decline underway at the time. Nevertheless, the greater the number of times the price reverses at the same level, and the greater the time interval over which this pattern occurs, the greater the significance of the pattern. Note that a bottoming process, whatever the reversal chart pattern, tends to be more predictable than a topping process. This is because in a bottoming process, market forces are capitulating, whereas in a topping process, irrational exuberance can always take prices to higher levels before ultimately reversing.

Some securities, due to either their volatility or their active trader/investor profile, offer similar well-defined chart pattern opportunities. Exhibit 20 shows a triple bottom chart pattern for Odfjell Drilling prior to the strong uptrend that pushed the stock price from around the 3.5 level to the high teens.

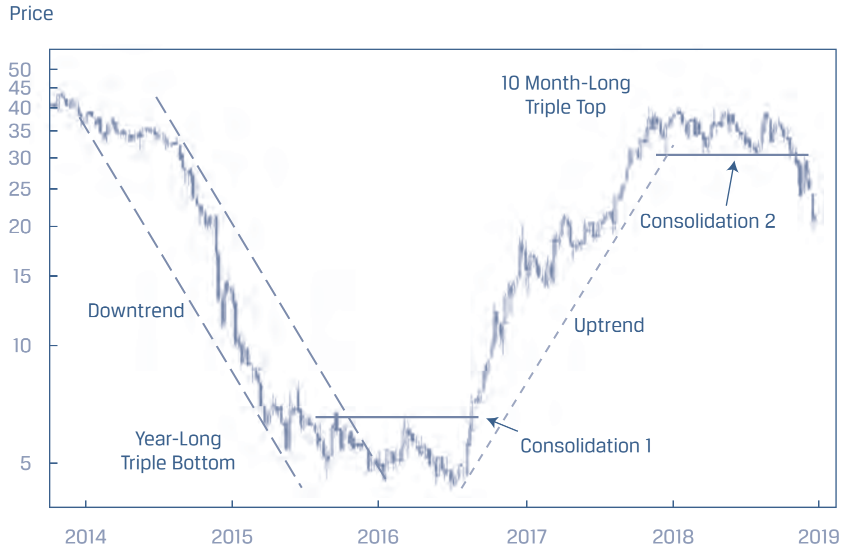
Exhibit 20 Triple Bottom Pattern: Odfjell Drilling Weekly Price Chart, January 2017–January 2019 (Norwegian krone)



When identifying chart patterns, we should remind ourselves of the importance of trends and consolidations. Regardless of the chart pattern in question, the formation of a price range is a *consolidation*, which, once completed with a confirmed breakout in either direction, will be expected to give rise to a trend period. For optimal decision making, traders, investors, and analysts should focus on successfully identifying and differentiating those two phases.

Exhibit 21 shows the Odfjell Drilling price chart between September 2013 and January 2019. The downtrend that started in the last quarter of 2013 reached a bottom in the second half of 2015. A yearlong sideways consolidation offered a trading opportunity for short-term traders between well-defined chart pattern boundaries. Traders, investors, and analysts with a slightly longer horizon should be asking: What type of price action can follow after the yearlong consolidation? If the chart pattern acts as a triple bottom with a breakout above the chart pattern boundary, can it reverse the two-year-long downtrend? In the second half of 2016, Odfjell Drilling completed the triple bottom with a strong weekly breakout and started a new uptrend. The triple bottom reversed the existing trend from down to up.

Exhibit 21 Trend and Consolidation: Odfjell Drilling Weekly Price Chart, September 2013–January 2019 (Norwegian krone)



Similar analysis was required from the technician in the last quarter of 2017. Consolidation 2 was a triple top and acted as a top reversal. The uptrend was followed by a consolidation that reversed the existing trend from up to down.

5.2 Continuation Patterns

A **continuation pattern** is used to predict the resumption of a market trend that was in place prior to the formation of a pattern. From a supply and demand standpoint, a continuation pattern often indicates a change in ownership from one group of investors to another. For example, if a positive trend was in place prior to a pattern and then one group of investors begins selling, the negative impact on price is quickly offset by other investors buying, so the forces of supply and demand go back and forth in terms of their impact on price. But neither has an overwhelming advantage. This type of pattern is often called “a healthy correction” because the long-term market trend does not change and because while one set of investors is seeking to exit, they are replaced by another set of investors willing to take their positions at roughly the same share price. Continuation patterns can take various forms, including triangles, rectangles, and flags, which are detailed in the rest of this section.

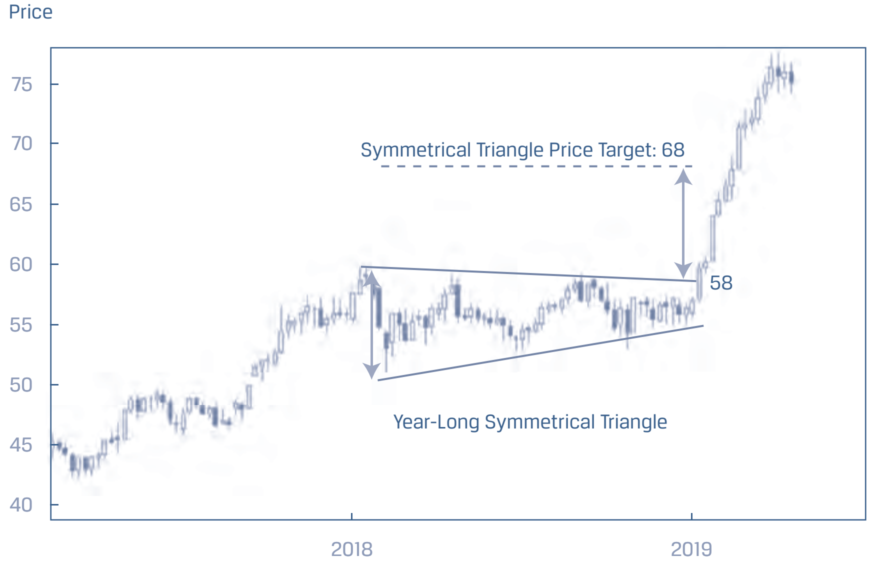
5.2.1 *Triangles*

Triangle patterns are a type of continuation pattern. They come in three forms: symmetrical triangles, ascending triangles, and descending triangles. A triangle pattern forms as the range between high and low prices narrows, visually forming a triangle. In the older terminology, triangles were referred to as “coils” or “springs” because a triangle was considered analogous to a spring being wound up tighter and tighter, storing energy that would at some point be released. In a triangle, a trendline connects the highs and another trendline connects the lows. As the distance between the highs and lows narrows, the trendlines meet, forming a triangle. In a daily price chart, a triangle pattern usually forms over a period of several weeks. In a weekly price chart, a triangle pattern can extend over several months.

In a symmetrical triangle, the trendline formed by the highs slopes down and the trendline formed by the lows slopes up, both at roughly the same angle, forming a symmetrical pattern.

Exhibit 22 contains a symmetrical triangle formed on the price chart of Diös Fastigheter AB during 2018. This triangle indicates that buyers are becoming more bullish while, simultaneously, sellers are becoming more bearish, so buyers and sellers are moving toward a market-clearing point of consensus. Because the sellers are often dominated by long investors exiting positions (as opposed to short sellers creating new short positions), the pressure to sell diminishes once the sellers have sold the security. Thus, the pattern ends with an ongoing move in the same direction as the trend that preceded it, either up or down.

Exhibit 22 **Symmetrical Triangle: Diös Fastigheter AB Weekly Price Chart, March 2017–March 2019 (Swedish krona)**



The possible price target is derived by calculating the difference in price from the two trendlines at the start of the triangle. Once the pattern is completed and the price breaches one of the trendlines that forms the triangle, the analyst expects the price to move by at least the amount of the calculated difference in price from the two trendlines at the start of the triangle above or below the trendline. Typically, price breaks out of a triangle pattern between halfway and three-quarters of the way through the pattern. The longer the triangle pattern persists, the more volatile and sustained the subsequent price movement is likely to be.

In an ascending triangle, as shown in Exhibit 23, the trendline connecting the high prices is horizontal and the trendline connecting the low prices forms an uptrend. This pattern means that market participants are selling the stock at the same price level over a period of time and putting a halt to rallies at the same price point, but buyers are getting more and more bullish and stepping in at increasingly higher prices to halt sell-offs instead of waiting for further price declines. An ascending triangle typically forms in an uptrend. However, there are several cases where an ascending triangle can form as a bottom reversal. The bullish implication of the pattern is more important than where the pattern forms on a price chart. The horizontal line represents sellers taking profits at around the same price point. The fact that the breakout pushes through the horizontal boundary is usually a bullish signal suggesting that buyers have overcome the selling pressure around the horizontal boundary. Irrespective of where the ascending triangle is identified, occasionally as a reversal chart pattern or more frequently as a continuation chart pattern, it should be considered a *consolidation* with bullish implications, and the breakout above the well-defined horizontal boundary should alert traders, investors, and analysts to act accordingly.

Exhibit 23 Ascending Triangle Pattern

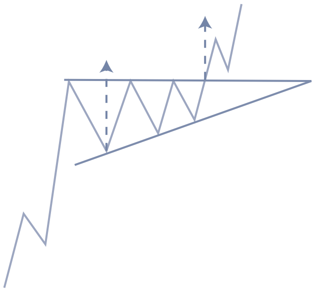
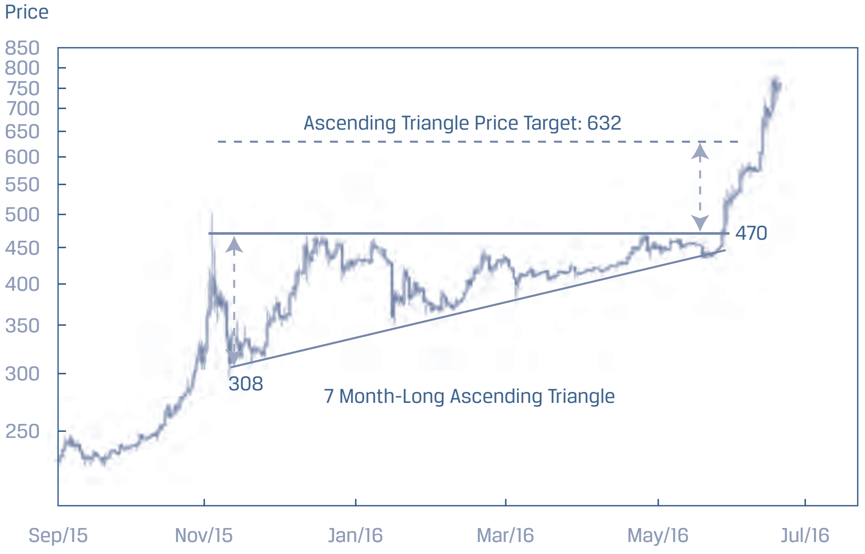


Exhibit 24 contains an ascending triangle formed on the price chart for bitcoin versus the US dollar during 2016. This ascending triangle indicates that buyers are becoming more bullish while sellers are acting around the same price level. The buyers are able to bid the price higher at every correction attempt. Eventually, the sellers lose the battle and the price breaks through the horizontal boundary. The prior trend was up, and the seven-month-long ascending triangle was a consolidation. A breakout above the 470 level completed the consolidation and resulted in a new uptrend. The ascending triangle acted as a bullish continuation chart pattern. The calculation for a chart pattern price target is similar to that for a symmetrical triangle. The possible price target is derived by calculating the difference in price from the two trendlines at the start of the triangle. Once the pattern is broken and the price breaks through one of the trendlines that form the triangle, the analyst will expect the price to move by at least the amount of the breakthrough above or below the trendline.

Exhibit 24 Ascending Triangle: Bitcoin/US Dollar Daily Price Chart, September 2015–June 2016 (US$)



In a descending triangle, as shown in Exhibit 25, the low prices form a horizontal trendline and the high prices form a series of lower and lower highs. Typically, a descending triangle will form in a downtrend. However, there are several situations in which a descending triangle acts as a top reversal. The bearish implication of this pattern is more important than where the pattern forms on a price chart. At some point during the rebounds from the horizontal support, sellers appear with enough supply to halt the countertrend recoveries each time they occur. Sellers push the price to lower levels at each rebound, thereby gaining the upper hand against the buyers who bid the price around the same level. The downward-sloping upper boundary of the descending triangle gives the pattern its bearish bias. Irrespective of where the descending triangle is identified, occasionally as a reversal chart pattern or more frequently as a continuation chart pattern, it should be considered a *consolidation* with bearish implications, and the breakdown below the well-defined horizontal boundary should alert traders, investors, and analysts to act accordingly.

Exhibit 25 Descending Triangle Pattern

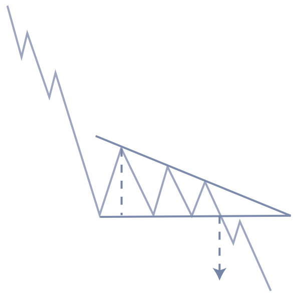
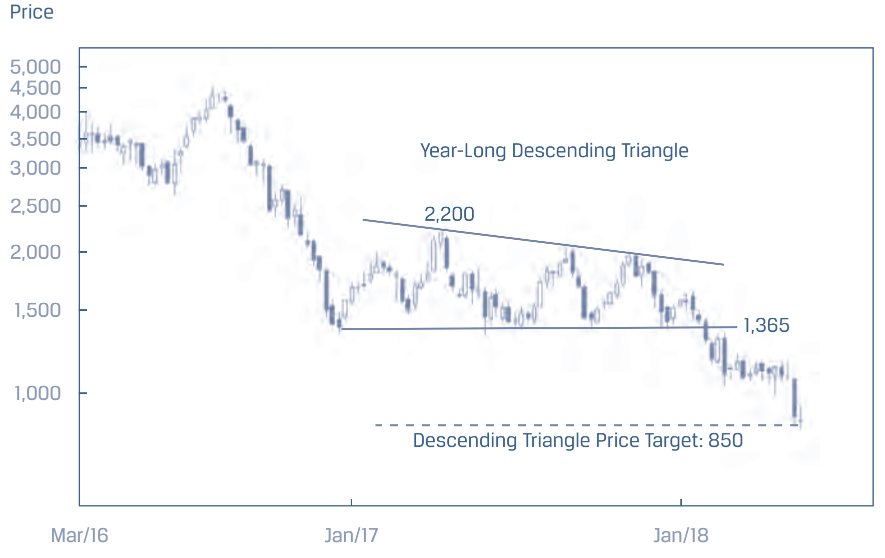


Exhibit 26 shows a descending triangle formed on the price chart of Sibanye Gold Ltd. during 2017. This descending triangle indicates that sellers are becoming more bearish while buyers are acting around the same price level. The sellers are able to push the price lower at every recovery attempt. Eventually, the buyers lose the battle and the price breaks down the horizontal boundary. The prior trend was down, and the yearlong descending triangle was a consolidation. The breakdown below the 1,365 level completed the consolidation and resulted in a downtrend. The descending triangle acted as a bearish continuation pattern. The calculation for a descending triangle pattern price target is slightly different with price charts in a downtrend, partly because of the zero boundary. It is important to note that chart pattern price objectives should be used as guidelines only. Price may reach the projected target, or it may fall short. It is always best to pick the most conservative approach. The possible price target is derived by calculating the percentage change from the two trendlines at the start of the triangle. Sibanye Gold formed a minor high at the 2,200 level. The horizontal boundary was around 1,365. This difference translates roughly into a 38% drop. Once the pattern is completed with a breach of the horizontal support, we can extend the price objective by another 38% on the downside. Please note that we are not taking the absolute price change. A 38% drop from the lower boundary of the descending triangle is equivalent to the 850 level.

Following breakdowns or breakouts, the price can pull back to the chart pattern boundary in an attempt to retest it. Sibanye Gold rebounded to test the previous support at the 1,365 level. The previous support became the new resistance—that is, there was a change in polarity.

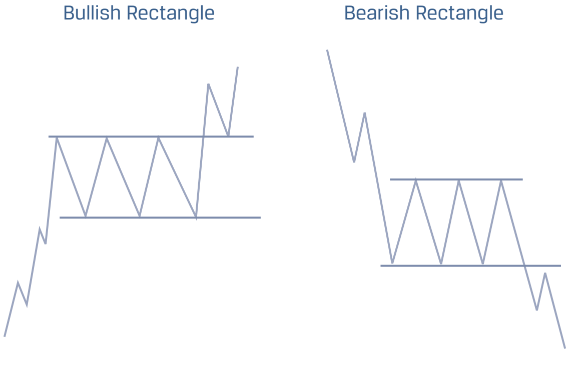
Exhibit 26 Descending Triangle: Sibanye Gold Ltd. Weekly Price Chart, March 2016–May 2018 (South African rand)



5.2.2 *Rectangle Pattern*

A rectangle usually develops as a continuation pattern formed by two parallel trendlines, one connecting the high prices during the pattern and the other connecting the lows. Exhibit 27 shows two rectangle patterns. Like other patterns, the rectangle pattern is a graphical representation of what has been occurring in terms of collective market sentiment. The horizontal resistance line that forms the top of the rectangle shows that investors are repeatedly selling shares at a specific price level, bringing rallies to an end. The horizontal support line forming the bottom of the rectangle indicates that traders are repeatedly making large enough purchases at the same price level to reverse declines. The support level in a bullish rectangle is natural because the long-term trend in the market is bullish. The resistance line may simply represent investors taking profits. Conversely, in a bearish rectangle, the support level may represent investors buying the security. Again, the technician is not concerned with why a pattern has formed, only with the likely next price movement once the price breaks out of the pattern.

Exhibit 27 Rectangle Patterns



Rectangle patterns, due their configuration, are much easier to identify on price charts than other classical chart patterns. Two horizontal boundaries and a clearly visible consolidation period on a price chart allows the analyst to detect the pause in trend periods with ease. This looks similar to a triple top pattern, but a rectangle pattern is a pause in a trend while a triple top pattern is a reversal of a trend.

Exhibit 28 shows a bullish rectangle pattern formed on the price chart of Vitrolife AB between 2017 and 2018. A 15-month-long sideways consolidation formed after a steady uptrend. After several tests of pattern boundaries both on the upside and on the downside, Vitrolife AB cleared the upper boundary that was acting as a strong resistance at the 142 level. The breakout was followed by a pullback, a retesting of the previous resistance. Several weeks of pullback eventually found support at the chart pattern boundary, and the uptrend resumed. The price objective was met in a matter of a few months. It is important to note that Vitrolife AB started a new strong uptrend following the lengthy consolidation. As a result, the price objective was exceeded during the uptrend. The longer the consolidation takes, the stronger the breakout is. The price objective is calculated by taking the width of the rectangle in absolute price terms and adding it to the breakout level.

Exhibit 28 Bullish Rectangle: Vitrolife AB Weekly Price Chart, December 2016–April 2019 (Swedish krona)

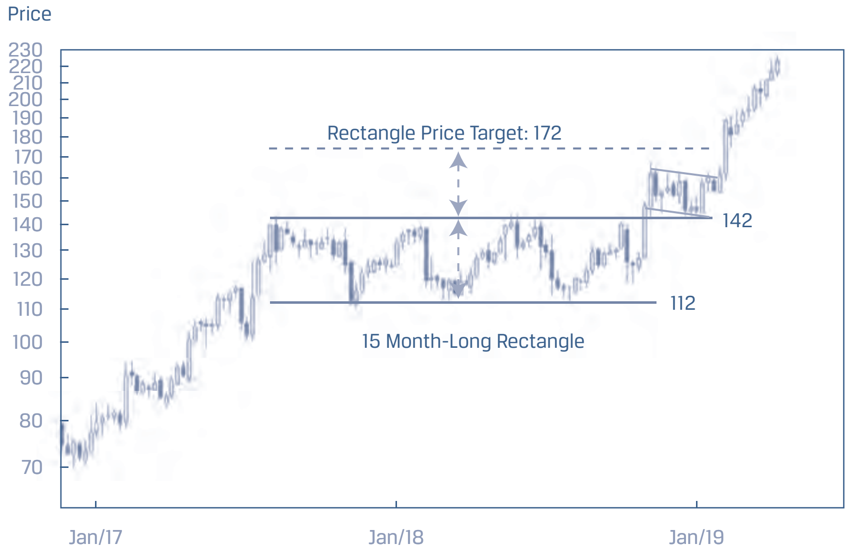
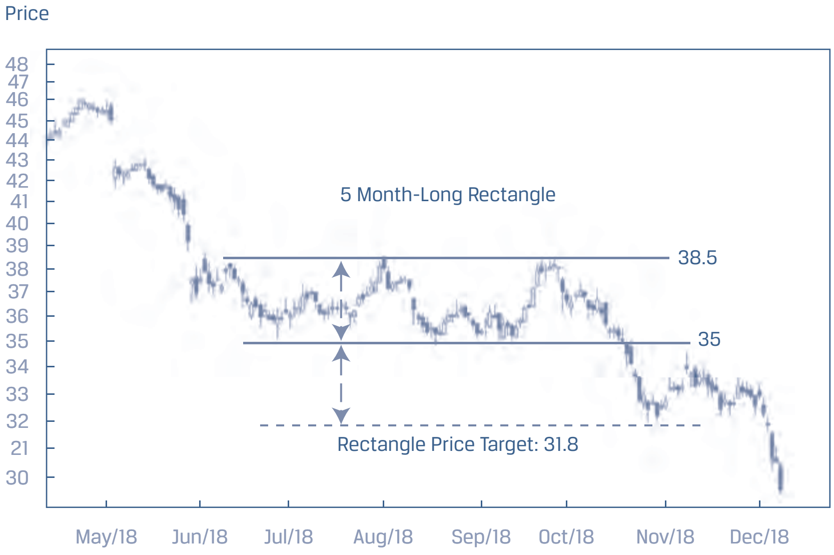


Exhibit 29 shows a bearish rectangle pattern formed on the price chart of Société Générale SA during 2018. A five-month-long sideways consolidation formed after a steady downtrend. After several tests of pattern boundaries, both on the upside and on the downside, Société Générale breached the lower boundary of the rectangle pattern that was acting as a strong support at the 35 level. This breakdown was followed by a pullback, a retesting of the previous support. A short pullback found resistance at the pattern boundary, and the downtrend resumed. The price objective was exceeded during the downtrend. The price objective is calculated by taking the width of the rectangle in percentage terms and projecting it lower from the breakdown level. The difference between the 38.5 and 35 levels translates into a 9% drop. Extending the price objective by 9% from the breakdown level of 35 gives us 31.85.

Exhibit 29 **Bearish Rectangle: Société Générale SA Daily Price Chart, April 2018–December 2018 (Price in Euro)**



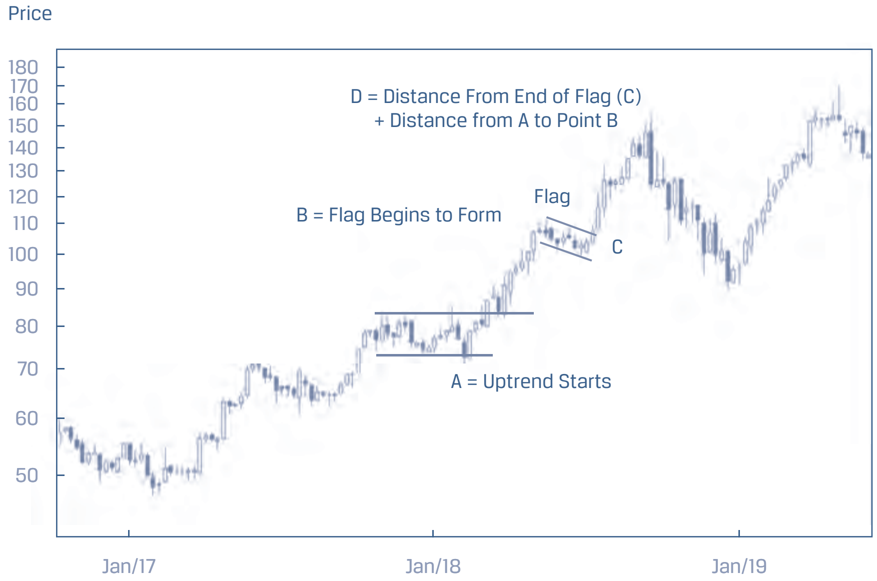
5.2.3 *Flags and Pennants*

Flags and **pennants** are considered minor continuation patterns because they form over short periods of time—for example, on a daily price chart, typically over a week. However, flags or pennants may form on weekly price charts over slightly longer durations. Flags and pennants are similar and have the same uses. A flag is formed by parallel trendlines, in the same way that most countries' flags are rectangular and create a parallelogram. Typically, the trendlines slope in a direction opposite to the trend up to that time; for example, in an uptrend, the flag's trendlines slope down. A pennant formation is similar except that the trendlines converge to form a triangle. The key difference between a triangle and pennant is that a pennant is a short-term formation whereas a triangle is a long-term formation.

The expectation for both flags and pennants is that the trend will continue after the pattern in the same direction it was going prior to the pattern. The price is expected to change by at least the same amount as the price change from the start of the trend to the formation of the flag or pennant. In Exhibit 30, an uptrend begins at point A, which is €70.5. At point B, which is €111, a flag begins to form. The distance from point A to point B is €40.5. The flag ends at point C, which is €104.5. The price target is €104.5 plus €40.5, which is €145, the line labeled D.

As one added note, volume should ideally be high in the left "flagpole" part of a flag pattern, diminish as the flag forms, and then increase again on the subsequent breakout from the flag pattern. It is not unusual to see chart "gaps" near flag patterns, particularly during the continuation breakout period.

Exhibit 30 **Flag Formation: Nemetschek SE, October 2016–May 2019 (Price in Euro)**



Question 4

Which charting style takes into consideration *only* the closing prices?

a Candlestick charts
b Line charts
c Bar charts

B is correct. Line charts connect the closing prices of each period. Both candlestick charts and bar charts include each period's opening, high, low, and closing prices.

Neither A nor C is the correct answer because both bar charts and candlestick charts take into consideration opening, high, low, and closing values.

Question 5

A trader is looking at long-term charts and wants to exit some of his long positions. In doing so, he wants to confirm major trendline breaks to decide when to exit the positions. Which price scale will generate an earlier breakdown alert?

a Linear scale
b Logarithmic scale
c Scale will not impact this

B is correct. A trendline plotted on a logarithmic scale will be steeper in an uptrend when compared to the same trendline plotted on a linear scale. As a result, a reversal will breach the trendline earlier on a logarithmic scale than on a linear scale.

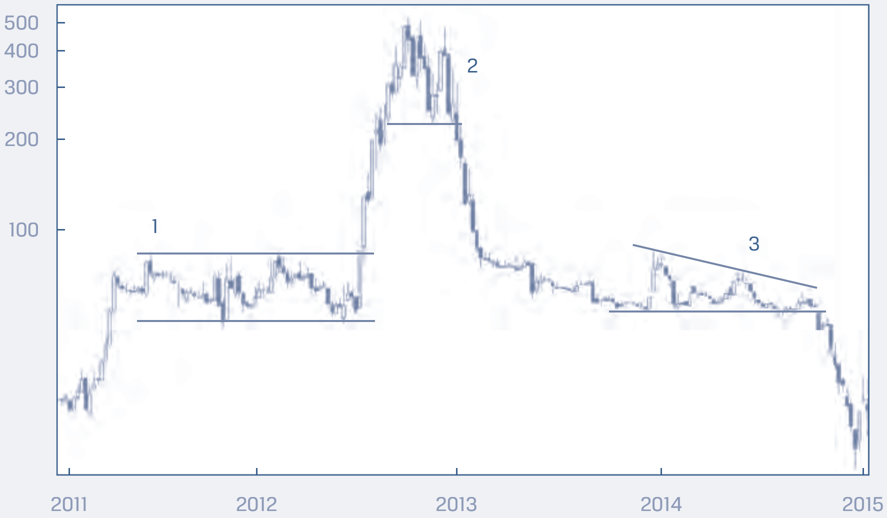
A is not correct because a trendline plotted on a linear scale will be less steep in an uptrend than the same trendline plotted on a logarithmic scale. As a result, a reversal will breach the trendline later on a linear scale than on a logarithmic scale.

C is not correct because scale will indeed impact the timing of either a breakdown below or a breakout above a trendline.

Question 6

The chart below shows the price of Amana Insurance on Saudi Arabia's Tadawul stock exchange between 2011 and 2015. Over the analyzed period, the stock formed three major classic chart patterns. Each one of those patterns (labeled 1, 2, and 3 on the chart) resulted in a strong trend period following the breakout. Which choice below lists the correct chart patterns in the correct order?

Weekly



a Double bottom, rectangle, symmetrical triangle

b Rectangle, double top, descending triangle

c Rectangle, head and shoulder top, ascending triangle

B is correct.

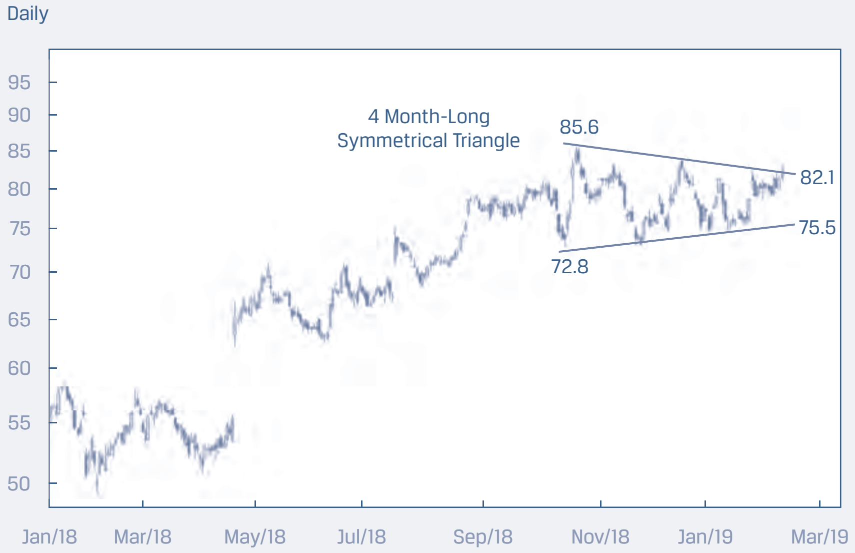
The first chart pattern, which formed between 2011 and 2012, is a rectangle. A rectangle is a continuation pattern. Prior to the rectangle, the stock had been in an uptrend. The rectangle acted as a "rest" period in the strong uptrend. A breakout from the rectangle was followed by another strong upward move. In the second half of 2012, the price formed a double top. A double top is a bearish reversal and should come after an advance. The third chart pattern is a descending triangle. A descending triangle is a bearish continuation pattern. Prior to a descending triangle, the price should be in a downtrend. A descending triangle is usually a "rest" period in a steady downtrend. Following the breakdown, the stock resumed its downtrend.

A is incorrect because chart pattern 1 is not a double bottom. For a double bottom, the prior trend should be downward. In this case, the price moved upward prior to the consolidation. A double bottom is a reversal chart pattern that reverses an existing downtrend. Chart pattern 2 is not a rectangle. Chart pattern 3 is not a symmetrical triangle. A symmetrical triangle has two converging boundaries.

C is incorrect because chart pattern 2 is not a head and shoulders top. For a head and shoulders top, there need to be three peaks, with the head being the highest and the two shoulders being lower than the head. Both head and shoulders and double top are reversal chart patterns, but in this case, pattern 2 can be classified as a double top due to its two peaks. Chart pattern 3 is not an ascending triangle. An ascending triangle has a horizontal upper boundary and an upward-sloping lower boundary. Both symmetrical triangles and ascending triangles can act as continuation patterns, but in this case, chart pattern 3 has a downward-sloping upper boundary and a horizontal lower boundary and is clearly a descending triangle.

Question 7

The chart below shows the Ericsson daily price chart between January 2018 and February 2019. The stock is listed on the Stockholm Stock Exchange, and the price is quoted in Swedish krona. Between October 2018 and February 2019, the Ericsson price chart formed a four-month-long symmetrical triangle. The stock is breaking out of a symmetrical triangle chart pattern, and the analyst wants to calculate the price target. Calculate the price target given the reference levels for the chart pattern boundaries.



a 98.4

b 94.9

c 88.3

B is correct.

The price target is calculated by taking the width of the widest point in the symmetrical triangle and adding it to the breakout level at the upper boundary of the symmetrical triangle.

$$85.6 - 72.8 = 12.8$$

$$12.8 + 82.1 = 94.9$$

A is incorrect because it is not correct to add the width of the widest point (85.6 – 72.8 = 12.8) to the minor high where the symmetrical triangle started forming (12.8 + 85.6 = 98.4).

C is incorrect because it is not correct to add the width of the widest point (85.6 – 72.8 = 12.8) to the level of the lower chart pattern boundary (12.8 + 75.5 = 88.3).

TECHNICAL INDICATORS: MOVING AVERAGES AND BOLLINGER BANDS

6

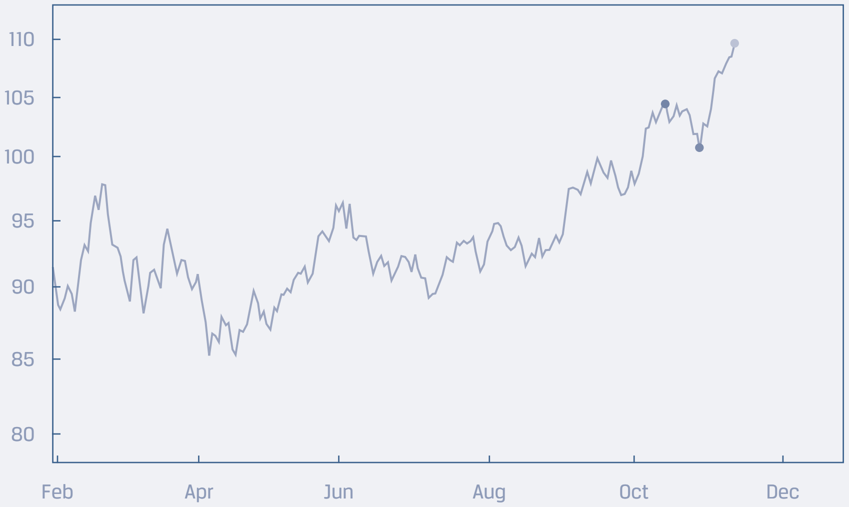
g explain common technical indicators;

VIGNETTE

Scene 3

A few days pass. You update your basic price chart for GLD and note that the price has been rising. The current price is the yellow point.

SPDR Gold Shares ETF – GLD (daily)



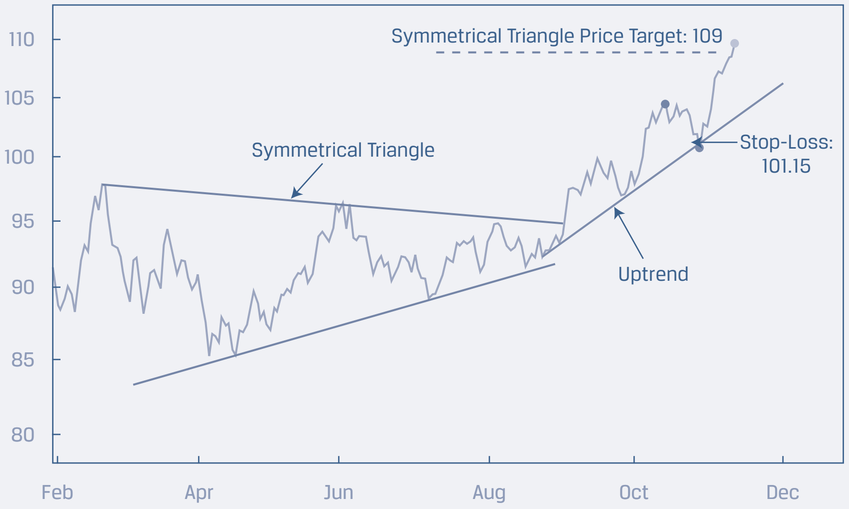
There is a knock on the door. The Head of Alternatives comes into your office and says, ***"Hey, gold has rallied and is well above our purchase price. I'm thinking it might be time to sell and take some profits on the investment."***

You respond by saying, ***"Remember, this investment in gold is a long-term asset allocation decision. We should not fall into the behavioral finance trap of letting greed guide our decisions."***

You decide to ask the technical analyst to again share her perspective.

The technical analyst prepares the following chart:

SPDR Gold Shares ETF – GLD (daily)



She also prepares the following comments:

- The symmetrical triangle price target has been met at the 109 level. The price target for the symmetrical triangle is calculated by taking the width of the chart pattern at the widest point and adding it to the breakout level.
- However, it is important to note that price can exceed calculated chart pattern price targets. The best way to capture long-term trends is either to use long-term averages to trail the price or to use trendlines to monitor the uptrend for a possible breakdown and a change in trend.

Then, she summarizes:

"So, even though the symmetrical triangle chart pattern price objective is met, the uptrend is still strong and is not violated. A new trendline (a steeper one) can be drawn, and the stop-loss can be raised to the most recent minor low of 101.15. In addition, the price continues to record higher lows, which is an indication of a steady uptrend."

You appreciate the additional perspective.

6.1 Technical Indicators

The technical analyst typically uses a variety of other technical indicators to supplement the basic information gleaned from chart patterns. A technical indicator is any measure based on price, market sentiment, or funds flow that can be used to predict changes in price. These mathematically calculated indicators often have a supply and demand underpinning. The moving average is the simplest of these techniques and has been used by statisticians to smooth data since the early 1900s.

6.1.1 *Price-Based Indicators*

Price-based indicators incorporate information contained in current and historical market prices. Indicators of this type range from simple (e.g., a moving average) to complex (e.g., a stochastic oscillator).

6.1.1.1 Moving Average A **moving average** is the average of the closing prices of a security over a specified number of periods. Moving averages smooth out short-term price fluctuations, giving the technician a clearer image of market trends. Technicians commonly use a simple moving average, which weights each price equally in the calculation of the average price. Some technicians prefer to use an exponential moving average (also called an exponentially smoothed moving average), which gives the greatest weight to recent prices while giving exponentially less weight to older prices.

The number of data points included in the moving average depends on the intended use of the moving average. A 20-day moving average is commonly used because a month contains roughly 20 trading days. Also, a 60-day average is commonly used because it represents a quarter year (three months) of trading activity.

Moving averages can be used in conjunction with a price trend or in conjunction with one another. Moving averages are also used to determine support and resistance levels.

Because a moving average is less volatile than price, this tool can be used in several ways. First, whether a price is above or below its moving average is important. As a general guideline, if the price is above a moving average, the trend is up. If the price is below a moving average, the trend is down. Second, the distance between the moving-average line and the price is also significant. Once price begins to move back up toward its moving-average line (in a reversion to the mean), this line can serve as a resistance level. Similarly, when price begins to move down after an upward trend, the moving average can serve as a support level. Even though moving averages for certain time periods (e.g., 200 days) are discussed and widely followed in the financial media, there is no ideal moving-average periodicity. It is important to note that the analyst, trader, or fund manager should pick the time period that is the best fit for the trading/investment horizon.

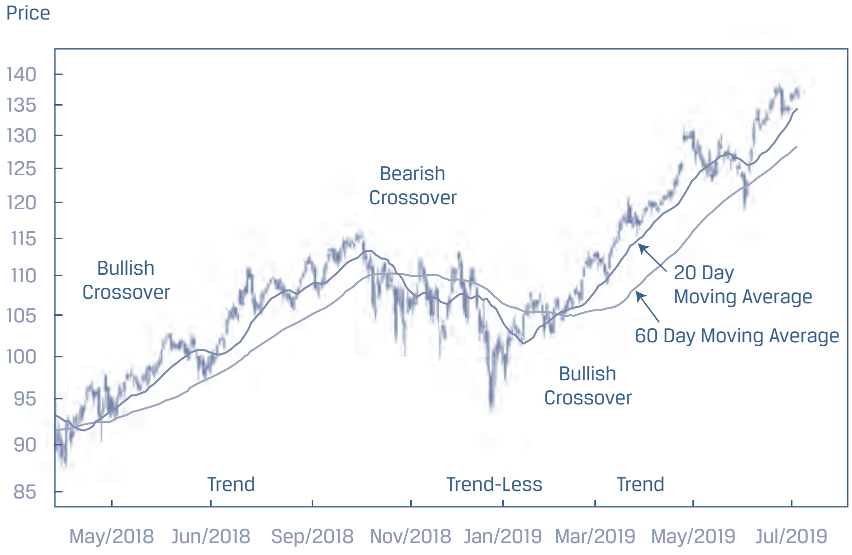
Two or more moving averages can be used in conjunction. A 5-day moving average is often used as a proxy for short-term momentum, while a 20-day or 60-day moving average can be an indicator of the intermediate trend. The 5-day moving average breaking up through the longer-term moving average can be used as a buy signal, indicating that momentum and trend are changing to the upside, and conversely, the 5-day breaking down through the longer-term moving average can be used as a sell signal.

Longer-horizon investors may want to simply focus on two intermediate-term moving averages. Exhibit 31 shows the price chart of Microsoft Corporation on the NASDAQ stock exchange overlaid with 20-day and 60-day moving averages during 2018–2019. Note that the longer the time frame used in the creation of a moving average, the smoother and less volatile the line will become. Investors often use moving-average crossovers as buy or sell signals. When a short-term moving average (e.g., one month) crosses a longer-term average (e.g., three months) from underneath, this movement is considered bullish. This crossover is called a **bullish crossover**. Conversely, when a short-term moving average crosses a longer-term moving average from above, this movement is considered bearish. This crossover is called a **bearish crossover**. In the case shown in Exhibit 31, a trading strategy of buying on bullish crosses and selling on bearish crosses would have been profitable.

A widely followed moving-average crossover signal is the one that takes place between the 50-day moving average and the 200-day moving average. Financial networks (such as Bloomberg and Reuters) refer this crossover signal as either a **golden cross** or a **death cross**. When the short-term (50-day) moving average crosses above the long-term (200-day) moving average, this movement is called a golden cross and is a bullish signal. Conversely, when the short-term average crosses below the long-term average, this movement is called a death cross and is a bearish signal.

Moving averages, while reducing noise by smoothing out day-to-day fluctuations, are useful for keeping the trader or investor on the right side of the price action and are key to trend-following strategies often used by commodity trading advisers (CTAs). When a short-term moving average is above the long-term average or the price is clearly above the average, traders and investors know that the trend in the given period of analysis is up, and therefore they should remain invested in the overall direction of the trend.

Exhibit 31 Daily Price Chart with 20-Day and 60-Day Moving Averages: Microsoft Corporation, April 2018–April 2019 (US$)

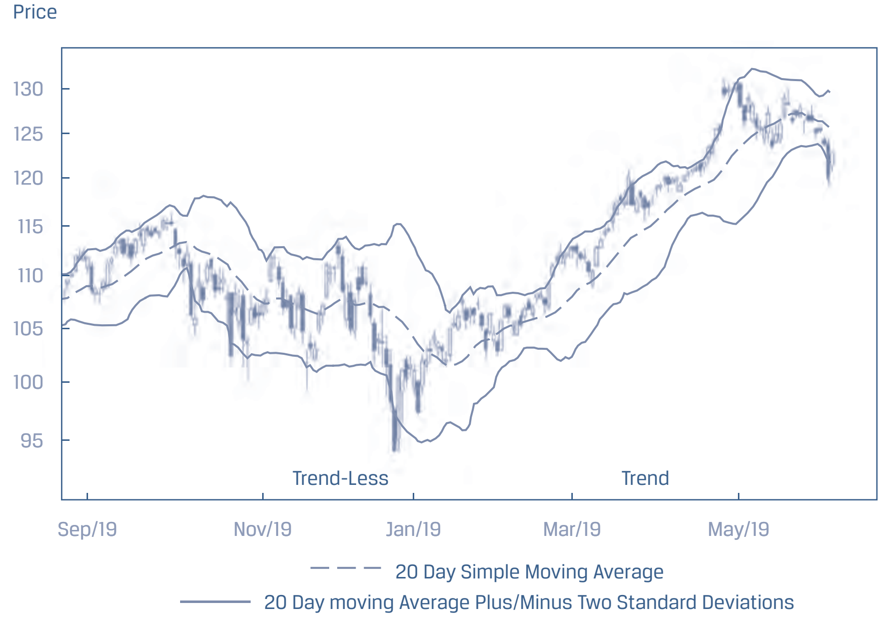


Moving averages are easy to construct, and simple trading rules can be created for using them. Computers can optimize the time lengths to set when using two moving averages. This optimization may take the form of changing the number of days included in each moving average or adding filter rules, such as waiting several days after a trade signal is given, or waiting until a minimum penetration distance is achieved, to make a trade. Reasons for optimization include the desire to minimize false signals and thereby manage capital drawdowns (maximize gains and minimize losses). Once the moving average is optimized, even if a profitable trading system is devised for that security, the strategy is unlikely to work for other securities, especially if they are dissimilar. Also, as market conditions change, a previously optimized trading system may no longer work. For that reason, one should not focus too much on the optimization of moving averages because this practice risks "overfitting" past data that may not necessarily repeat in the future. Moving averages should be used as trend filters.

When the price is above the moving average, this scenario suggests that the trend is upward. When the price is below the average, this scenario should warn the analyst of further weakness. When the short-term moving average is above the long-term moving average, the analyst should be alert for the continuation of that uptrend. When the short-term moving average is below the long-term average, this scenario should signal caution. It may suggest that any initial long exposure should be closed and that the investor should step to the sidelines in expectation of further weakness.

6.1.1.2 Bollinger Bands Market veteran John Bollinger combined his knowledge of technical analysis with his knowledge of statistics to create an indicator called **Bollinger Bands**. Bollinger Bands consist of a moving average (middle band), a higher line (upper band) representing the moving average plus a set number of standard deviations from the average price (for the same number of periods used to calculate the moving average), and a lower line (lower band) that is a moving average minus the same number of standard deviations. This indicator can help define a statistically reasonable range that the market is expected to trade within. Exhibit 32 depicts Bollinger Bands for Microsoft Corporation.

Exhibit 32 Bollinger Bands Using 20-Day Moving Average and Two Standard Deviations: Microsoft Corporation Daily Price Chart, September 2018–April 2019 (US$)



The more volatile the security being analyzed becomes, the wider the range becomes between the two outer lines or bands. Like moving averages, Bollinger Bands can be used to create trading strategies that can be easily computerized and tested. A common use is as a contrarian strategy, in which the investor sells when a security price reaches the upper band and buys the security back when the price reaches either the middle band (if within a perceived uptrend) or the lower band (when within a perceived range). This strategy assumes that the security price will stay within the bands. During trendless periods, this strategy works well. This type of strategy is likely to lead to a large number of trades, but it also limits risk because the trader can quickly exit unprofitable trades. In the event of a sharp price move, investors might instead buy on a significant breakout above the upper band because a major breakout would imply a change in trend that is likely to persist for some time.

The long-term investor would sell on a significant breakdown below the lower band in an attempt to limit downside risk. In this strategy, statistical significance would be defined as breaking above or below the band by a certain percentage (say, 5% or 10%) and/or for a certain period of time (say, a week for a daily price chart). Again, such rules can easily be computerized and tested. In the example of Microsoft, October 2018–February 2019 was a trendless period. A death cross took place in October 2018, and the stock remained choppy within a wide trading range. During this period, a trader could have applied the Bollinger Band strategy to buy at the touch of the lower band and sell at the touch of the upper band. This strategy would have worked well until the stock experienced a change in trend.

In February 2019, Microsoft experienced a bullish moving-average crossover resulting in a golden cross. Following the bullish signal, the price touched the upper Bollinger Band, suggesting a continuation of the uptrend. Selling at the touch of the upper band and waiting for prices to return to the lower band was not the right strategy; buying at the touch of the middle band would have worked better.

6.1.1.2.1 ***Bollinger Band Width Indicator*** Bollinger Band width is an indicator derived from Bollinger Bands. Band width measures the percentage difference between the upper band and the lower band. Band width decreases as Bollinger Bands narrow and increases as Bollinger Bands widen. Because Bollinger Bands are based on standard deviation, falling band width reflects decreasing volatility and rising band width reflects increasing volatility.

$$\text{Bollinger Band width} = [(\text{Upper band} - \text{Lower band})/\text{Middle band}] \times 100$$

When calculating band width, the first step is to subtract the value of the lower band from the value of the upper band. The result is the absolute difference. This difference is then divided by the value of the middle band to normalize the value. This normalized band width can then be compared across different time frames or with the band width values for other securities.

Band width is relative. Band width values should be gauged relative to prior band width values over a period of time. It is important to use an adequate look-back period to define the band width range for a particular ETF, index, or stock.

Bollinger Band width is best known for identifying a “squeeze.” A squeeze occurs when volatility falls to a very low level, as evidenced by the narrowing bands.

The theory is that periods of low volatility are followed by periods of high volatility. Relatively narrow band width can foreshadow a significant advance or decline. After a squeeze, a price surge and a subsequent band break signal the start of a new move. A new advance starts with a squeeze and a subsequent break above the upper band. A new decline starts with a squeeze and a subsequent break below the lower band.

Exhibit 33 Bollinger Band Width and Squeeze: Diös Fastigheter AB, March 2017–March 2019 Weekly Scale Price Chart (Swedish krona)

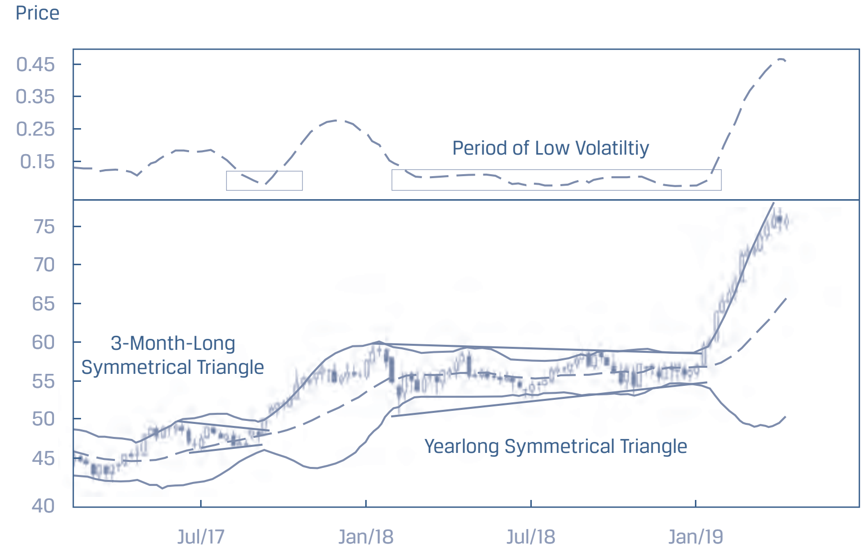


Exhibit 33 shows the Diös Fastigheter AB price chart that was featured earlier in this reading (see Exhibit 22). Every time the stock entered into a consolidation period, the volatility dropped to low levels. Low volatility was an indication of a possible high-volatility period to come. Low volatility begets high volatility, and vice versa. A breakout from low-volatility conditions can be seen on the Bollinger Band width indicator: Bollinger Bands start expanding, and Bollinger Band width starts rising from extremely low levels.

When combined with classical charting tools (such as a breakout above a predefined chart pattern boundary), Bollinger Bands have significant value. A breakout from a consolidation period at a time of historical low volatility suggests that the trend period that follows will have the strength to resume toward much higher levels. The two breakouts shown in Exhibit 33, one from a three-month-long symmetrical triangle and the other from a yearlong symmetrical triangle, both started from extreme low-volatility conditions.

TECHNICAL INDICATORS: OSCILLATORS, RELATIVE STRENGTH, AND SENTIMENT

7

g explain common technical indicators;

One of the key challenges in using indicators overlaid on a price chart is the difficulty of discerning changes in market movements that are out of the ordinary. **Momentum oscillators** are one tool intended to alleviate this problem. They are constructed from the rate of change in price data, but they are calculated so that they either oscillate between a low and a high (typically 0 and 100) or oscillate around a number (such as 0 or 100). Because of this construction, extreme highs and lows are easily discernible. These extremes can be viewed as graphic representations of market sentiment when selling or buying activity is more aggressive than has been historically typical. Because they are price based, momentum oscillators can be analyzed by using the same tools technicians use to analyze price, such as the concepts of trend, support, and resistance. It is important to note that indicators that are derived from price will lag price. An analyst will be able to calculate the latest value for the indicator only after the relevant day's, week's, or month's price data are recorded.

Technicians also look for **convergence** or **divergence** between oscillators and price. Convergence occurs when the oscillator moves in the same manner as the security being analyzed, and divergence occurs when the oscillator moves differently from the security. For example, when price reaches a new high, this sign is considered bullish, but if the momentum oscillator being used does not also reach a new high at the same time, the result is divergence. Divergence is considered to be an early warning of weakness, an indication that (in this case) the uptrend may soon end because momentum (or the rate of change in prices) is actually waning despite the new highs.

Momentum oscillators should be used in conjunction with an understanding of the existing market (price) trend. Oscillators alert a trader to **overbought** or **oversold** conditions. In overbought conditions, market sentiment is considered unsustainably bullish. In oversold conditions, market sentiment is considered unsustainably bearish. In other words, the oscillator *range* must be judged separately for every security. Some securities may experience wide variations, and others may experience only minor variations.

Oscillators have three main uses:

1. Oscillators can be used to determine the strength of a trend. Extreme overbought levels are warning signals for uptrends, and extreme oversold levels are warning signals for downtrends. We should note, however, that occasionally an indicator can stay overbought or oversold for an extended period of time. As a result, just because the indicator reached an overbought level doesn't mean we should anticipate a sudden change in trend. Uptrends and downtrends can sometimes resume despite overbought/oversold readings on oscillators.
2. When oscillators reach historically high or low levels, they may be signaling a pending trend reversal. For oscillators that move above and below zero, crossing the zero level signals a change in the direction of the trend. For oscillators that move above and below 100, crossing the 100 level signals a change in the direction of the trend.
3. In a trendless market, oscillators can be used for short-term trading decisions—that is, to sell at overbought levels and to buy at oversold levels.

7.1 Rate of Change Oscillator

The rate of change (ROC) oscillator is a pure momentum oscillator and is sometimes referred to as simply the momentum oscillator. The ROC oscillator is calculated by taking the most recent closing price, subtracting the closing price from a date that is a set number of days in the past, and multiplying the result by 100:

$$M = (V - Vx) \times 100,$$

where

M = momentum oscillator value
V = most recent closing price
Vx = closing price x days ago, typically 10 days

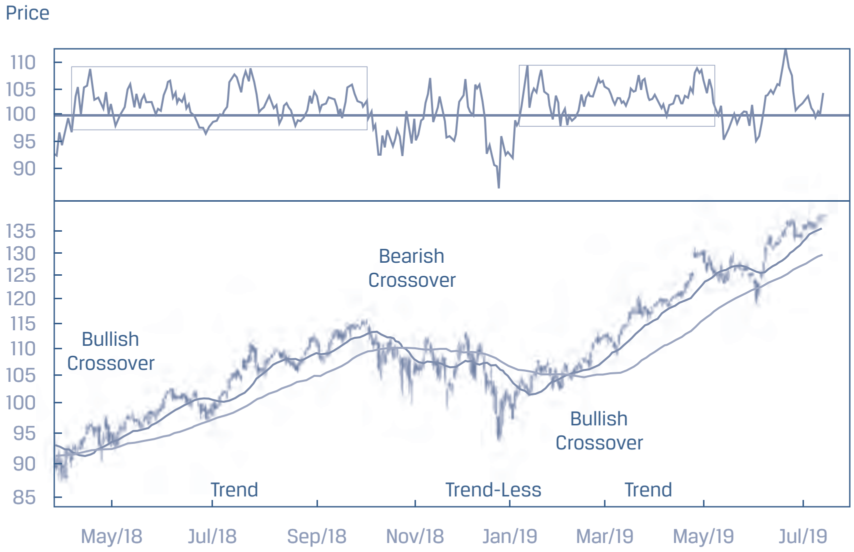
When the ROC oscillator crosses zero in the same direction as the direction of the trend, this movement is considered a buy or sell signal. For example, if the ROC oscillator crosses into positive territory during an uptrend, this movement is a buy signal. If it enters into negative territory during a downtrend, the crossover is considered a sell signal. The technician will ignore crossovers in opposition to the trend because when using oscillators, the technician must *always first* take into account the general trend.

An alternative method of constructing this oscillator is to set it to oscillate above and below 100, instead of 0, as follows:

$$M = \frac{V}{Vx} \times 100$$

This approach is shown in Exhibit 34 for Microsoft Corporation.

Exhibit 34 ROC Oscillator with 100 as Midpoint: Microsoft Corporation, April 2018–April 2019 (US$)



In Exhibit 34, the ROC oscillator for Microsoft stock, traded on the NASDAQ stock exchange, is set to move around 100, and x is 12 days. An extreme high means that the stock has posted its highest gain in any 12-day period, and an extreme low reading means it has posted its greatest loss over any 12-day period. When investors bid up the price of a security too rapidly, the indication is that sentiment may be unduly bullish and the market may be overbought. Exhibit 34 shows that overbought levels of the ROC oscillator coincide with temporary highs in the stock price. So, those levels would have been signals to sell the stock. We analyzed the moving-average crossovers and the trend and consolidation periods in preceding sections of this reading. It is important to incorporate that information in analyzing the price chart. During trend periods, we can see that the ROC oscillator moves approximately between 100 and 110. So once the uptrend is confirmed with a bullish crossover, we can use the ROC oscillator to time the minor peaks and troughs in the uptrend. A good strategy would be to take profits as the ROC approaches the 110 level and initiate new long positions as it falls back to the 100 level. During a trendless period (consolidation), we can see that the ROC oscillator moves in a wide range between 80 and 110. Following the bearish crossover, we should be aware that the trend period might be over and that we might see choppy price action. The timing of entry and exit might be more difficult in such a market environment.

However, with the help of Bollinger Bands and ROC oscillators, the analyst can capture overbought/oversold levels in the consolidation range. In the beginning of 2019, the ROC oscillator once again climbed above the 100 level, and the trend was confirmed with a bullish crossover, suggesting a new uptrend. In February 2019, the ROC oscillator found support around 100, the low level for the indicator during an uptrend. It was a good time to initiate a long position as the stock was entering a new uptrend period. In previous sections, we have seen that the Microsoft price also touchedthe upper Bollinger Band following the bullish crossover signal and generated a buy signal based on the Bollinger Band strategy. The other notable aspect of Exhibit 34 is the divergence when the share price hit a new high in October 2018 but the ROC oscillator did not reach its July high. This divergence would have been a bearish signal and would have been interpreted to mean that, although the share price hit a new high, market momentum was actually lower than it had been previously. In itself, this information would not have been enough to warrant selling the shares because an uptrend in price was still in place, but it alerted the technician to the fact that the trend might end soon. The technician could then look for further indications of the trend's end—such as a moving-average crossover to the downside—and with such confirmation, decide to sell the stock.

7.2 Relative Strength Index

Another tool often used to measure momentum is a **relative strength index (RSI)**, which is computed over a rolling time period. The RSI graphically compares a security's gains with its losses over the set period. The creator of the RSI, Welles Wilder, proposed a 14-day time period, and this is the period used by most technical analysis software. The technician should understand that this variable can be changed and that the optimal time range should be determined by how the technician intends to use the RSI information. Factors that influence the selection of a time period for the RSI are similar to those that influence the selection of a time period for moving averages.

The RSI is a momentum oscillator and provides information on whether or not an asset is overbought—in other words, whether an asset has advanced too quickly relative to a chance to digest recent advances. The formula is as follows:

$$RSI = 100 - \frac{100}{1 + RS},$$

where

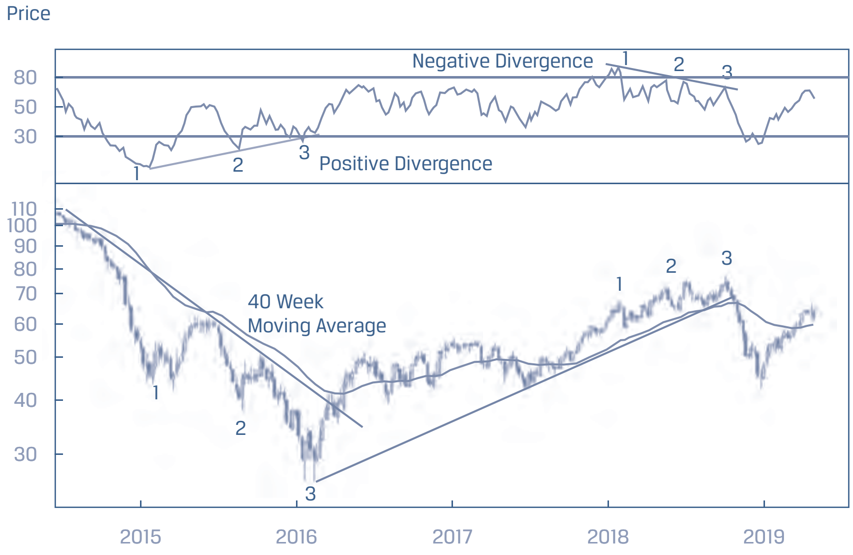
$$RS = \frac{\sum(\text{Up changes for the period under consideration})}{\sum(\text{Down changes for the period under consideration})}.$$

The RSI indicator is primarily used to analyze the inner momentum of a security. The analysis is performed by identifying divergences. A positive divergence takes place when the security price records a new low but the RSI records a higher low. This scenario suggests that while the price is reaching a new low, the inner momentum is improving; in other words, the selling pressure is decreasing.

A negative divergence takes place when the security price reaches a new high but the RSI fails to reach a new high. This scenario suggests that while the price is reaching a new high, the inner momentum is deteriorating; in other words, buying pressure is decreasing. Divergences are early signs of reversal that may foreshadow and anticipate a subsequent moving-average crossover reversal.

Note: The RSI is not to be confused with the "relative strength analysis" charting method. The RSI measures the internal strength of a single asset in terms of a ratio of the magnitude of "up days" to the magnitude of "down days" of that asset across a given period, whereas relative strength analysis compares the ratio of two security prices plotted over time or the relative performance of two assets.

Exhibit 35 Candlestick Chart with RSI: WTI Light Crude Oil First-Month Continuation Futures Price, June 2014–April 2019 (US$)



The candlestick chart of the WTI Light Crude Oil futures price in Exhibit 35 illustrates several aspects of the use of an RSI. For example, because the RSI oscillator was lower than 30 in the beginning of 2015, so the commodity was oversold at that time, a simple reading of the chart might have led to the conclusion that the trader should buy the commodity because an "oversold" indicator reading suggests the price is likely to rebound in the near term. However, this decision would have resulted in further losses as the price slid to new lows. A more careful technical interpretation that took the general trend into account would have indicated that the commodity was in a downtrend, as can be seen by the trendline and the fact that the price was trading below its long-term average, so further RSI readings below 30 could be expected. In other words, using the RSI alone may identify short-term reversal levels but risk early bottom-picking or top-picking that can still prove problematic. Buying at the time of an oversold condition might not be the best timing tool for trend reversals. Likewise, an overbought reading above the 80 level alone is not sufficient to conclude that price will reverse. It is best to combine the momentum indicator with elements of trend analysis, such as a trendline break.

Even if a stock, commodity, or currency is showing an oversold reading on the RSI, there is often not enough information to conclude that it is a good time to buy that security. In the second half of 2015, WTI reached a new low, but this time the RSI didn't reach a new low. Again, just because we have seen the first positive divergence (RSI lows labeled 1 and 2 in Exhibit 35) doesn't mean that a trend reversal is taking place. We need to see confirmation of the change in trend in the form of a breach of a trendline or a long-term average (confirmation by price action). In the beginning of 2016, the WTI price reached a new low but the RSI generated its second positive divergence. It was after this low that the price managed to breach the downtrend line and the long-term moving average on the upside, confirming a change in trend. A technician should be patient enough to wait for such periods of clear divergence late in the trend of a given asset.

Similarly, the top that formed during 2018 on the WTI price chart followed the same sequence of divergences. This time, the price continued to reach new highs but the RSI warned of a possible eventual change in trend direction. Only after the second negative divergence (RSI peaks labeled 2 and 3 in Exhibit 35) did the price of WTI break down the trendline (confirmation by price action) and the long-term average, resulting in a change in trend.

7.3 Stochastic Oscillator

The **stochastic oscillator** is a momentum indicator that compares a particular closing price of a security to a range of its prices over a certain period of time. The stochastic oscillator is based on the observation that in uptrends, prices tend to close at or near the high end of their recent range and in downtrends, they tend to close near the low end. The logic behind these patterns is that if the shares of a stock are constantly being bid up during the day but then lose value by the close, continuation of the rally is doubtful. If sellers have enough supply to overwhelm buyers, the rally is suspect. If a stock rallies during the day and is able to hold on to some or most of those gains by the close, that sign is bullish.

The stochastic oscillator oscillates between 0 and 100 and has a default 14-day period, which may be adjusted for the situation, as we discussed for the RSI. The oscillator is composed of two lines, called %K and %D, that are calculated as follows:

$$\%\mathrm{K} = 100\left(\frac{C - L14}{H14 - L14}\right),$$

where

C = the latest closing price
$L14$ = the lowest price in the past 14 days
$H14$ = the highest price in the past 14 daysand where
%D = the average of the last three %K values calculated daily

Analysts should think about %D in the same way they would a long-term moving-average line in conjunction with a short-term line. That is, %D, because it is the average of three %K values, is the slower-moving, smoother line and is called the signal line. And %K is the faster-moving line. The %K value means that the latest closing price (C) was in the %K percentile of the low–high range ($L14$ to $H14$).

The default oversold–overbought range for the stochastic oscillator is based on reading the signal line relative to readings of 20 and 80. The shorter the time frame is, the more volatile the oscillator becomes and the more false signals it generates.

The stochastic oscillator should be used with other technical tools, such as trend analysis, pattern analysis, and RSI analysis. If several methods suggest the same conclusion, the trader has convergence (or confirmation), but if they give conflicting signals, the trader has divergence, which is a warning signal suggesting that further analysis is necessary.

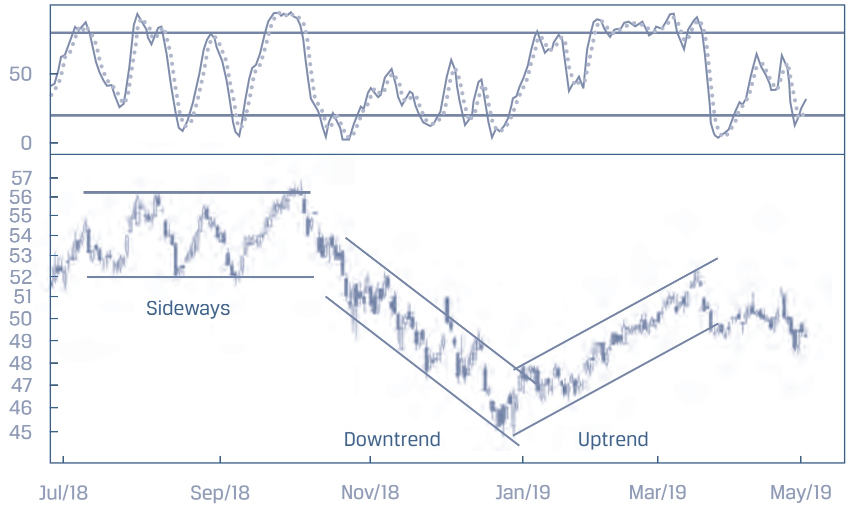
The absolute level of the two lines should be considered in light of their normal range. Movements or changes in trend in the middle of an 80–20 indexed range can most often be deemed short-term noise. To a technician, movements above this range indicate an overbought security and are considered bearish; movements below this range indicate an oversold security and are considered bullish. But it is important to note that an overbought indicator doesn’t warrant selling the security. Likewise, an oversold reading is not an indication to buy. One should generally look for changes in trend with the help of the %K line changing direction or a trendline break or a breach of moving average.

Crossovers of the two lines can also give trading signals, the same way crossovers of two moving averages give signals. When %K moves from below the %D line to above it, this move is considered a bullish short-term trading signal; conversely, when %K moves from above the %D line to below it, this movement is considered bearish. In practice, a trader can use technical analysis software to adjust trading rules and optimize the calculation of the stochastic oscillator for a particular security and investment purpose (e.g., short-term trading or long-term investing). While optimization can help to achieve better results for a certain period, as conditions change, the parameters might need to be revised.

The reason technicians use historical data to test their trading rules and find the optimal parameters for each security is that each security is different and may have its own natural rhythm. The group of market participants actively trading also differs from security to security. Just as each person has a different personality, so do groups of people. In effect, the group of active market participants trading each security is imparting its personality on the trading activity for that security. As this group changes over time, the ideal parameters for a particular security may change. The choice of either optimizing the indicator or following the default set of parameters is a continuous trade-off.

There are benefits as well as disadvantages to optimization. Exhibit 36 provides a good example of how the stochastic oscillator can be used together with trend analysis. The exhibit provides the daily price chart and stochastic oscillator for Total SA, which is traded on the Paris stock exchange, from July 2018 through April 2019. Note that during the sideways movement between July and October, the stochastic oscillator moved between 20 and 80 from one boundary to the other. It was one of the best times to trade the stock with the help of overbought/oversold conditions identified by the oscillator. Each time the oscillator reversed from the 80 level, it provided a valid sell signal, and each time it reversed from the 20 level, it provided a valid buy signal. These reversals coincided with reversals from horizontal support/resistance levels on the price chart. When the downtrend started in October, the stochastic oscillator reached oversold levels and hardly recovered above 50. Technicians should be able to identify the trend period (sideways, upward, or downward) and then adjust their expectations as to the levels the indicator can reach. During the downtrend, the oversold level remained around 20 but the overbought level shifted down to 50. Selling the rally around the 50 level in the established downtrend proved to be a good strategy. When the uptrend started at the beginning of January, the stochastic oscillator reached overbought levels. However, during the uptrend, the oscillator didn't reach oversold levels. While the overbought level remained around 80, the oversold level during the uptrend shifted up to 50. Buying the pullback when the indicator reached the 50 level proved to be a good strategy.

Exhibit 36 Daily Price Chart and Stochastic Oscillator: Total SA, July 2018–April 2019 (€)



7.4 Moving-Average Convergence/Divergence Oscillator (MACD)

The **moving-average convergence/divergence oscillator** is commonly referred to as the MACD. The MACD is the difference between short-term and long-term moving averages of the security's price. The MACD is constructed by calculating two lines, the MACD line and the signal line:

- The MACD line is the difference between two exponentially smoothed moving averages, generally over periods of 12 and 26 days.
- The signal line is the exponentially smoothed average of the MACD line, generally over 9 days.

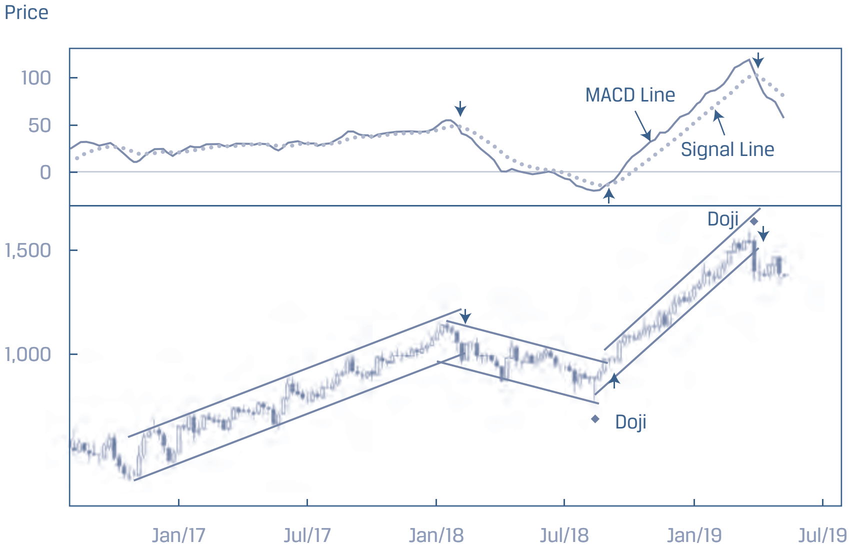
The indicator oscillates around zero and has no upper or lower limit. Rather than using a set overbought–oversold range for the MACD, the analyst compares the current level with the historical performance of the oscillator for a particular security to determine when a security is out of its normal sentiment range.

The MACD is used in technical analysis in three ways:

1. to note crossovers of the MACD line and the signal line, as discussed for moving averages and the stochastic oscillator. Crossovers of the two lines may indicate a change in trend.
2. to look for times when the MACD is outside its normal range for a given security.
3. to use trendlines on the MACD itself. When the MACD is trending in the same direction as price, this pattern is in convergence, and when the two are trending in opposite directions, the pattern is in divergence.

Exhibit 37 shows a weekly price chart for Palladium (at the top) with the MACD oscillator for August 2016 through April of 2019. The chart shows the effectiveness of MACD crossover signals when used in conjunction with trendline breakouts. This example highlights once again the importance of confirming any indicator signal with price action. During the uptrend between October 2016 and December 2018, the MACD generated several crossover signals. The early 2018 crossover signal was confirmed by a breakdown of the uptrend channel. Similarly, the trend reversal that took place in the last quarter of 2018 was confirmed by both the MACD crossover signal and a trend channel breakout. In August 2018, we can also identify the long-legged doji candlestick that marked the bottom. The uptrend that started in the last quarter of 2018 pushed the MACD to historical high levels when compared with previous readings. The reversal not only generated a MACD crossover but also broke down the upward trend channel.

Exhibit 37 MACD and Weekly Price Chart: Palladium, August 2016–April 2019 (US$)



7.5 Sentiment Indicators

Sentiment indicators attempt to gauge investor activity for signs of increasing bullishness or bearishness. Sentiment indicators come in two forms—investor polls and calculated statistical indexes.

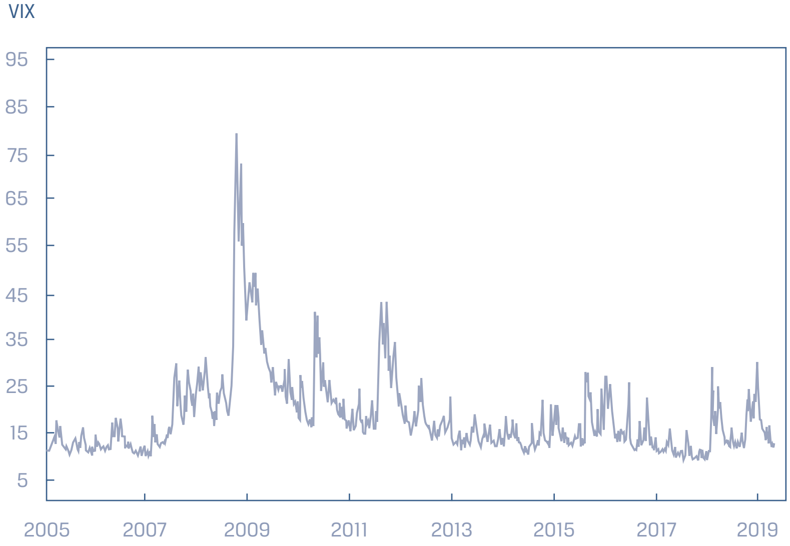
7.5.1 *Calculated Statistical Indexes*

Statistical indexes are sentiment indicators that are calculated from market data, such as security prices. The two most commonly used are derived from the options market; they are the put/call ratio and the volatility index. Additionally, many analysts look at margin debt and short interest.

The **put/call ratio** is the volume of put options traded divided by the volume of call options traded for a particular financial instrument. Investors who buy put options on a security are presumably bearish, and investors who buy call options are presumably bullish. The volume of call options traded is greater than the volume of put options over time, so the put/call ratio is normally below 1.0. The ratio is considered to be a contrarian indicator, meaning that higher values of put volume relative to call volume are considered bullish and lower values are considered bearish. But the ratio's usefulness as a contrarian indicator is limited, except at extreme low or high levels (relative to the historical trading level) of the put/call ratio for a particular financial instrument. The actual value of the put/call ratio and its normal range will be different for each security or market, so no standard definitions of overbought and oversold levels exist.

The **CBOE Volatility Index (VIX)** is a measure of near-term market volatility calculated by the Chicago Board Options Exchange. Since 2003, the VIX has been calculated from option prices on the stocks in the S&P 500. The VIX rises when market participants become fearful of an impending market decline. These participants bid up the price of puts, and the result is an increase in the VIX level. Technicians use the VIX in conjunction with trend, pattern, and oscillator tools, and it is interpreted from a contrarian perspective. When other indicators suggest that the market is oversold and the VIX is at an extreme high, this combination is considered bullish. Exhibit 38 shows the VIX from March 2005 to April 2019.

Exhibit 38 VIX, March 2005–April 2019

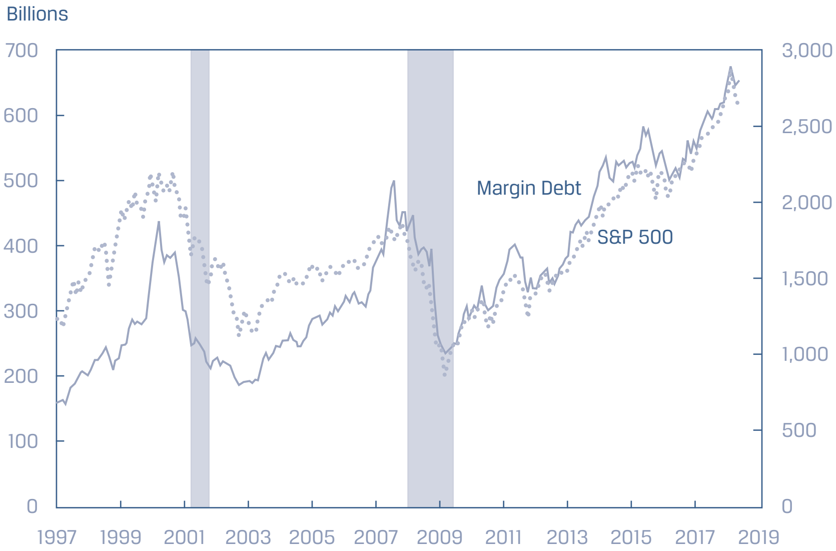


Margin debt is also often used as an intermediate- to long-term indication of sentiment. As a group, investors have a history of buying near market tops and selling at the bottom. When the market is rising and indexes reach new highs, investors are motivated to buy more equities in the hope of participating in the market rally. A margin account permits an investor to borrow part of the investment cost from the brokerage firm. This debt magnifies the gains or losses resulting from the investment.

Investor psychology plays an important role in the concept of margin debt as an indicator. When stock margin debt is increasing, investors are aggressively buying, and stock prices will move higher because of increased demand. Eventually, the margin traders use all of their available credit, so their buying power (and, therefore, demand) decreases, which fuels a price decline. Falling prices may trigger margin calls and forced selling, which drive prices even lower.

Brokerage firms must report activity in their customers' margin accounts, so keeping track of borrowing behavior is relatively easy. Exhibit 39 provides a 22-year comparison of margin debt with the S&P 500. The correlation is striking: Rising margin debt is generally associated with a rising index level, and falling margin debt is associated with a falling index level. In fact, for the two decades shown in Exhibit 39, the correlation coefficient between the level of margin debt and the S&P 500 is 80.2%. When margin debt peaked in the summer of 2007, the market also topped out. Margin debt dropped sharply during the latter part of 2008 as the subprime crisis took the market down. Investors began to use borrowed funds again in the first half of 2009, when heavily discounted shares became increasingly attractive. Margin debt was still well below the average of the last decade, but in this instance, the upturn would be viewed as a bullish sign by advocates of this indicator. Over the past decade, both margin debt and the S&P 500 have climbed to new highs. This outcome has been bullish for the market in the short term but is concerning in the longer term.

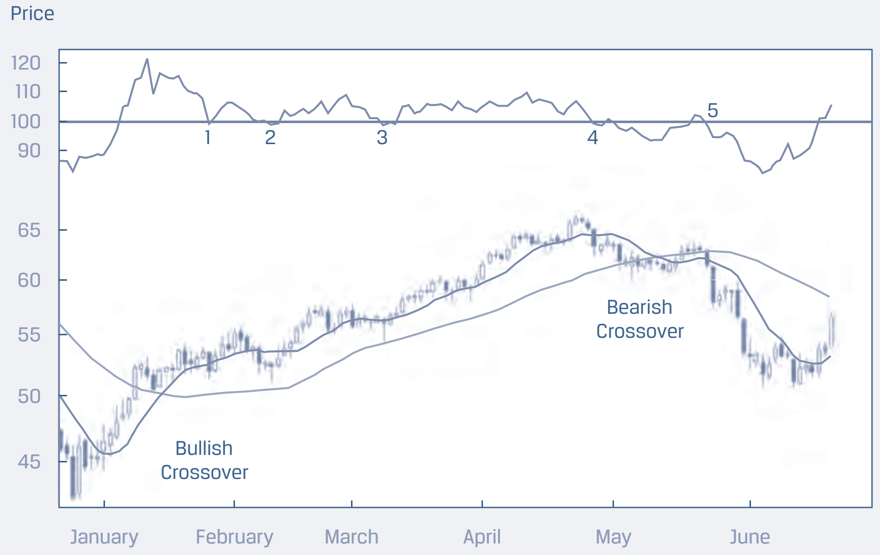
Exhibit 39 Margin Debt in US Market vs. S&P 500, 1997–2019



Source: FINRA.

Question 8

A trader wants to buy WTI Crude Oil futures. She wants to build a sizable position and is willing to enter a long position after the uptrend is underway. In other words, the trader aims to invest once the uptrend is established. For this reason, the trader wants to utilize moving-average crossovers as a trend filter and wants to buy after the bullish crossover. In addition to utilizing moving averages, she wants to utilize momentum as a timing tool to enter long positions. The trader knows that in uptrends, a pullback to the 100 level on the momentum indicator offers a buying opportunity. Find three points on the chart that meet those two criteria for a buying opportunity: an uptrend and a pullback to the 100 level.



a 3, 4, and 5

b 1, 4, and 5

c 1, 3, and 4

Answer: C. The bullish crossover took place in the middle of January 2019. At the time, the momentum indicator was above 110—an overbought level. Later, the indicator pulled back to the 100 level. The points labeled 1, 2, 3, and 4 were all buying opportunities in a steady uptrend. Even though point 4 met both criteria, the move failed to materialize into a new uptrend, and the indicator fell below the 100 level, indicating a change in trend. Weakness in momentum was followed by a bearish crossover on the moving averages.

Neither A nor B is the correct answer because at point 5, the momentum indicator was pulling back to the 100 level but the moving averages had already had a bearish crossover, suggesting a downtrend. All points except 5 were buying opportunities in an uptrend.

Question 9

When a stock price moves to a new high but the momentum indicator records a lower high, the analyst labels the divergence as a:

a positive divergence.

b negative divergence.

c positive reversal divergence.

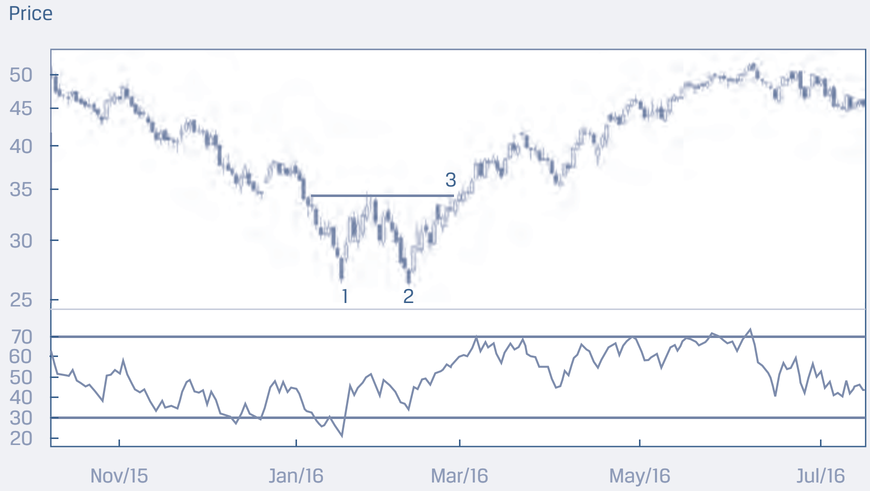
Answer: B. A negative divergence takes place when a stock price moves to a new high but the momentum indicator fails to keep up with the price and records a lower high.

A is not the correct answer because a positive divergence takes place when a stock moves to a new low but the indicator records a higher low.

C is not the correct answer because a positive reversal divergence takes place when the indicator reaches a new low but the price fails to record a new low.

Question 10

A trader wants to take a long position in WTI Crude Oil. He wants to utilize the RSI indicator to check for positive divergence as an early trend reversal signal but also wants to time his entry with a chart pattern breakout signal. Which number on the chart below is sufficient to make a trade decision based on the above requirements?



a 1

b 2

c 3

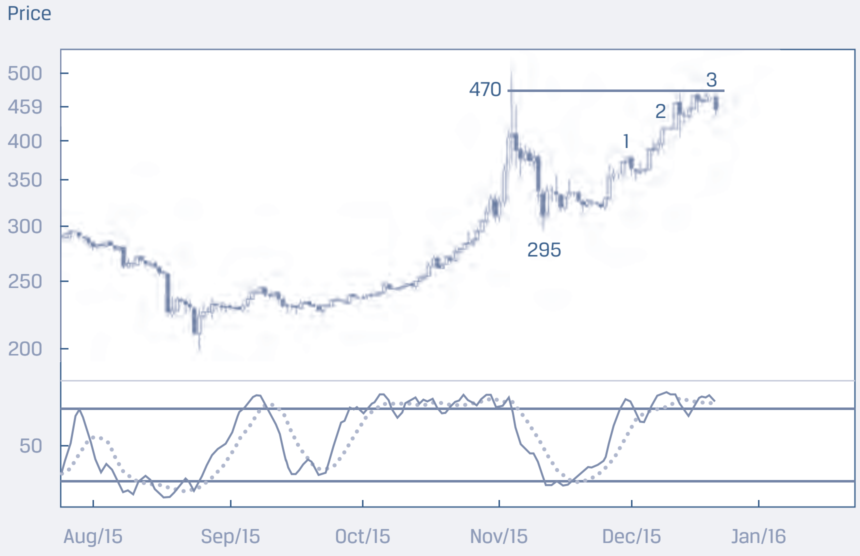
Answer: C. At point 3, the RSI has already reached an oversold level, rebounded, and formed a positive divergence. Meanwhile, the price holds at its previous low, forming a double bottom. Double bottom breakout confirmation takes place when the price breaches the neckline at point 3.

A is not the correct answer because at point 1, the price is in a steady downtrend and the indicator is at an oversold level. The price can continue to fall, and the indicator can remain in oversold territory.

B is not the correct answer because at point 2, the price is testing a support (previous low) and the indicator is showing a positive divergence. However, we don't have confirmation on the price chart in the form of a reversal chart pattern completion or a downtrend breakout. The trend reversal confirmation takes place at point 3.

Question 11

The chart below shows the price of bitcoin versus the US dollar between August 2015 and December 2015. In November 2015, bitcoin had a sharp advance, which turned into a buying frenzy and reached the 500 level. After the sharp advance, the price of bitcoin sold off to test 295. In late 2015, the bitcoin/US dollar price chart tested the November 2015 highs. During the uptrend between November and December, the stochastic oscillator reached overbought levels. An overbought indicator can remain overbought while price continues to move higher. Based on support/resistance levels and the stochastic oscillator, what would be your strategy at point 3?



a Buy more, because the price is in an uptrend and the indicator is overbought.

b Sell, because the price has reached an important resistance-level overbought reading on the stochastic oscillator.

c Sell, because the indicator has formed its fifth negative divergence.

Answer: B. The stochastic oscillator can reach overbought levels and stay in the overbought range (above 80). This scenario doesn't mean one should sell because of the overbought reading. Point 1 is an overbought reading on the stochastic oscillator, but the price is still far from the strong resistance area. Point 2 is similar. The stochastic oscillator is at an overbought level once again, but the price is below the strong resistance. At point 3, price meets both conditions: an overbought reading on the oscillator and a test of the strong resistance level. From a risk/reward perspective, it is a better decision to sell at the resistance with overbought conditions.

A is not the correct answer because the price of bitcoin is reaching the resistance level formed by the minor high at the 470 level. The test of the resistance takes place after the stochastic oscillator reaches an overbought level. The probability of a reversal is higher at point 3 than at 1 and 2. As a result, buying more bitcoin at the resistance is not the best strategy.

C is not the correct answer because the indicator hasn't formed its fifth negative divergence. Between the minor highs of points 2 and 3, the indicator forms its first negative divergence.

Question 12

A trader is looking to trade breakout opportunities as she thinks the cyclical nature of the markets suggests a strong trend period approaching. She knows that volatility is cyclical and that low volatility begets high volatility and vice versa. She opens charts with three different indicators: Bollinger Band width, the MACD, and the RSI. Which indicator and condition will guide her in identifying breakout opportunities by capturing low-volatility conditions?

a The MACD line crossing over the signal line when the indicator is below the zero level

b The RSI crossing above the 30 level after generating a positive divergence

c The Bollinger Band width indicator reaching its lowest level over a three-year look-back period

Answer: C. Bollinger Band width measures volatility with the help of Bollinger Bands. Band width is narrow and the indicator reaches low levels when price goes through a low-volatility condition. Low-volatility periods are usually followed by high-volatility periods, and these are usually trend periods that emerge from lengthy consolidations.

Neither A nor B is the correct answer because both the MACD and the RSI are price-based trend indicators. They don't have a volatility component in their calculation.

INTERMARKET ANALYSIS

h describe principles of intermarket analysis;

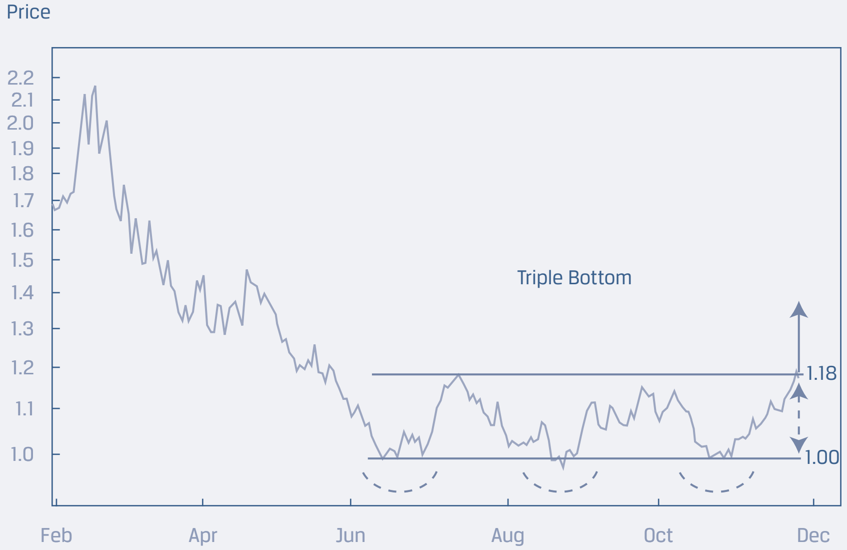
i explain technical analysis applications to portfolio management.

VIGNETTE

Scene 4

A few days pass, and the Deputy CIO is back from her Antarctic vacation. She thanks you for your efforts while she was away. She mentions that while viewing the penguins in Antarctica, she was thinking that although the gold investment was put in place for long-term asset allocation, she should find out about any tactical asset allocation opportunities. She asks you to look at gold versus oil. Before doing an in-depth analysis, you ask the technical analyst to provide perspective using intermarket analysis.

The technical analyst prepares the following chart:



She also prepares the following comments:

- For ease of understanding, the GLD/WTI ratio is indexed to 1.0 at the low in June 2009.
- Between February and June, GLD underperformed WTI. However, since June 2009, the GLD/WTI ratio has started forming a bottom. The GLD/WTI ratio has fluctuated between the 1.0 level and the 1.18 level.
- The sideways consolidation of the GLD/WTI ratio formed a triple bottom. A triple bottom can act as a major reversal and could result in a strong outperformance period for GLD.

Then, she summarizes:

"So, now is not the time to allocate from GLD to WTI because GLD could start outperforming WTI. In addition, a breakout on the relative performance chart of GLD/WTI would put the gold price on the radar of many investors."

You thank the technical analyst for her support. You are also glad that you had a chance to work with the technical analyst and feel that you better understand the role technical analysis can play in investment management. You are also happy that the Deputy CIO is back in the office! Well, mostly…

8.1 Principles of Intermarket Analysis

Intermarket analysis is a field within technical analysis that combines analyses of major categories of securities—namely, equities, bonds, currencies, and commodities—to identify market trends and possible inflections in a trend. Intermarket analysis also looks at industry subsectors and their relationship to sectors and industries. In addition, it measures the relative performance of major equity benchmarks around the globe.

In intermarket analysis, technicians often look for inflection points in one market as a sign or clue to start looking for a change in trend in a related market. To identify these intermarket relationships, a commonly used tool is relative strength analysis, which charts the price of one security divided by the price of another or the ratio of the values of two assets. A subtle shift may show up first in this kind of ratio analysis, where it may be clearer than in either individual chart pattern. Compared to traditional global macro analysis, using technical analysis to conduct intermarket analysis is considerably less time-consuming.

Exhibit 40 shows the price of 10-year US Treasury bonds compared with the S&P 500. Major trends can be clearly seen in periods of both outperformance and underperformance of the T-bond price relative to the S&P 500. The inflection points in this chart occur in 2000, 2003, 2007, and 2009. At each of these points, the relative performance ratio signaled that the time had come to move investments between these asset classes.

Exhibit 40 Relative Strength of 10-Year T-Bonds vs. S&P 500, 1999–2018

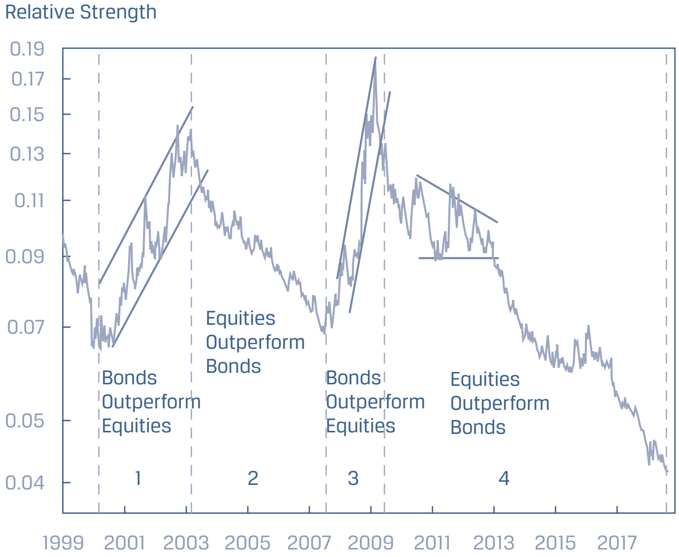
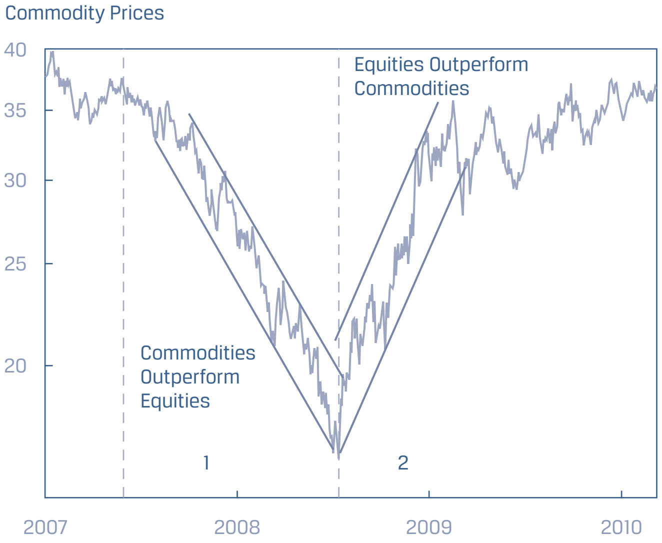


Exhibit 41 is a relative strength chart depicting the ratio of the S&P 500 to commodity prices. The chart shows a clear reversal of trend in mid-2008 from outperformance by commodities to outperformance by equities. This inflection point shows US stocks strengthening relative to commodities and indicates that moving funds away from commodities and into US equities might be appropriate. Of course, promptly interpreting such a sudden "V-reversal" in trend is still difficult to do, but a ratio chart may still help in the identification of a macro inflection point.

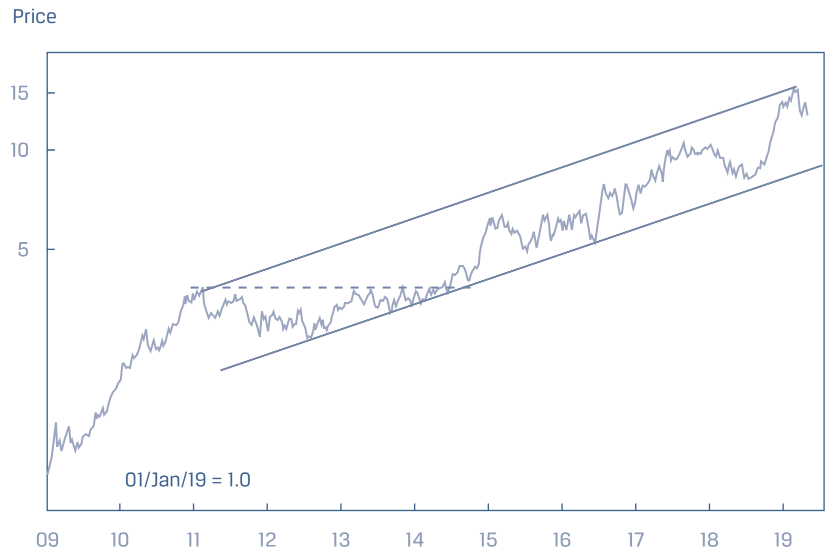
Exhibit 41 Historical Example: S&P 500 Index vs. Commodity Prices, 2007–2010



In addition to the preceding comparisons, once an asset category has been identified, relative strength analysis can be used to identify the best-performing securities in a sector. For example, if commodities look promising, an investor can analyze each of the major commodities relative to a broad commodity index in order to find the strongest commodities.

Exhibit 41 shows the relative price performance of Palladium versus commodities from 2009 to 2019. The ratio is constructed and indexed to 1.0 at the beginning of 2019. As of May 2019, the ratio is at the 13 level, meaning that Palladium outperformed commodities by a factor of 13. The year 2015 marked the acceleration of Palladium's outperformance versus the commodity index.

Exhibit 42 Palladium vs. Commodity Prices, 2009–2019



Intermarket analysis can also be used to identify sectors of the equity market to invest in—often in connection with technical observations of the business cycle at any given time. The equities of certain industry sectors tend to perform best at the beginning of an economic cycle. These sectors include financials, consumer non-durables, and transportation. As an economic recovery gets underway, retailers, manufacturers, health care, and consumer durables tend to outperform. Lagging sectors sometimes include those tied to commodity prices, such as energy and basic industrial commodities, which may catch up only in late-cycle, increasingly inflationary periods. In the final stage of a waning expansion, utility and consumer staple stocks may outperform because of their perceived safety, while transportation stocks often lead early market weakness in anticipation of a softening economy.

Observations based on intermarket analysis can also help in allocating funds across national markets. Certain countries' economies—for example, those of Australia, Canada, and South Africa—are closely tied to commodities. As economies evolve, these relationships change. So, the relationships must be monitored closely. For instance, a strong performance in industrial and precious metals can benefit mining stocks. The economies of Australia, Canada, and South Africa depend on the mining industry's performance, and as a result, strong metal prices can have a positive impact on the performance of these countries' equity markets.

8.2 Technical Analysis Applications to Portfolio Management

A technician acts very much like a doctor, who analyzes a patient for signs of wellness or sickness and then uses the clues discovered to formulate a composite opinion about fundamental changes in that patient's health. In this way, technical analysis complements fundamental analysis, whether the approach is top-down or bottom-up.

The top-down approach focuses on how the overall economy is affecting different sectors or industries. Analysts who apply a top-down approach believe that if a sector is doing well, stocks in that sector should also perform well.

Intermarket analysis is a key technical analysis tool that helps put the "big picture" into perspective. Intermarket analysis and relative strength analysis help the technician identify trends in countries, sectors, and industries and even among stocks within the same industry.

The bottom-up approach to identifying investment opportunities depends on rules and conditions. The analyst, trader, or investor starts his or her research with an *opportunity set* and tries to pick stocks that meet the *predefined criteria* irrespective of country, sector, or industry trends, which are sometimes considered as extraneous unknowable factors most easily considered as "All else being equal." Sometimes several opportunities in the same sector or industry will point to a developing overall *theme*. For example, finding several investment opportunities in mining stocks will alert the analyst, trader, or investor that metals (precious or industrial) might have strong performance. Or breakdowns in small-cap stocks might alert the analyst, trader, or investor that a possible "risk-off" environment is around the corner for the major equity benchmarks.

8.2.1 *Technical Analysis Applications to Portfolio Management: Top-Down Approach*

Analysts, traders, or investors whose focus is global equity markets often start their top-down approach by analyzing global benchmarks, such as MSCI and FTSE. The relative performance of major indexes will reveal important investment themes for investors with a long-term focus.

The same principles of charting and technical analysis discussed in the earlier sections of this reading can be applied to relative strength analysis. The aim is to identify consolidation periods and then participate in major trends once these consolidations are completed, whether on the upside or on the downside. In other words, the technician attempts to capture trend periods by focusing on major inflection points on price charts.

Exhibit 43 shows the relative strength of the MSCI Emerging Markets Index and the MSCI All Country World Index between 1997 and 2019. The highest point reached in 2010 is indexed to 1.0 for ease of analysis. Starting in July 1997, the Asian financial crisis was a period of turmoil in currency and equity markets that gripped much of East Asia and Southeast Asia and raised fears of a worldwide economic meltdown through financial contagion.

The relative strength chart for the Emerging Markets Index and the All Country World Index shows that 1997–1999 was a period of strong underperformance for the Emerging Markets Index. Indonesia, South Korea, and Thailand were the countries most affected by the crisis. Hong Kong SAR, Malaysia, and the Philippines were also hurt by the slump. These were all countries classified as emerging markets.

However, in 2001–2002, the performance ratio tested the 1998–99 lows and formed a double bottom chart pattern. Retesting of the previous lows and a reversal from the support area were the first signs of a change in trend. The actual change took place in 2003, when the ratio cleared its long-term average and the horizontal boundary (resistance) of the double bottom chart pattern. This movement signaled the beginning of a period of outperformance for the Emerging Markets Index against the All Country World Index. For investors, it was a time to focus on emerging market opportunities.

Strong outperformance for the Emerging Markets Index continued until 2012, when the ratio started challenging its long-term average and the upward-sloping trendline. A new trend was established: Since 2012, the developed markets have been outperforming the emerging markets. Global asset allocators would have captured much better performance by investing in developed market equities over the past six years.

In Exhibit 43, the two major inflection points are the breakout from the double bottom chart pattern with the breach of the long-term average in 2003 and the breakdown of the trendline and long-term average in 2012. These inflection points are labeled as "trend change" on the price chart.

Exhibit 43 MSCI Emerging Markets Index vs. MSCI ACWI, 1997–2019

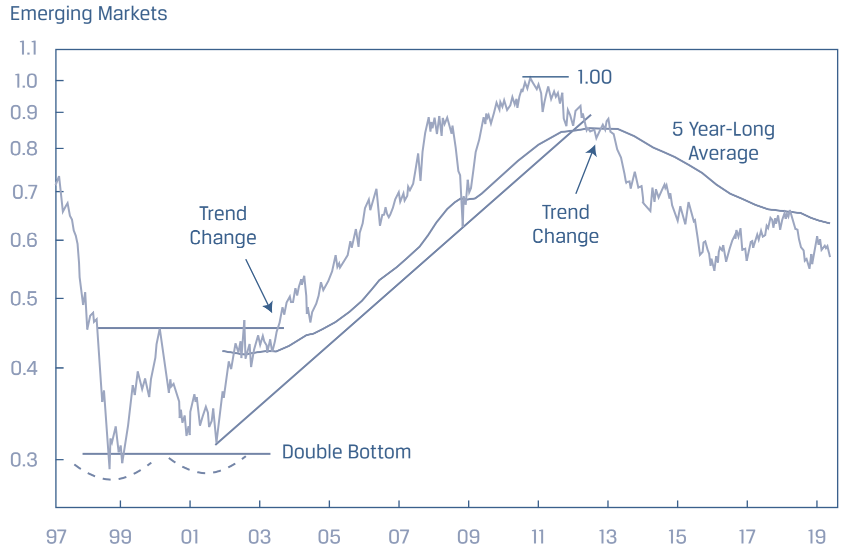


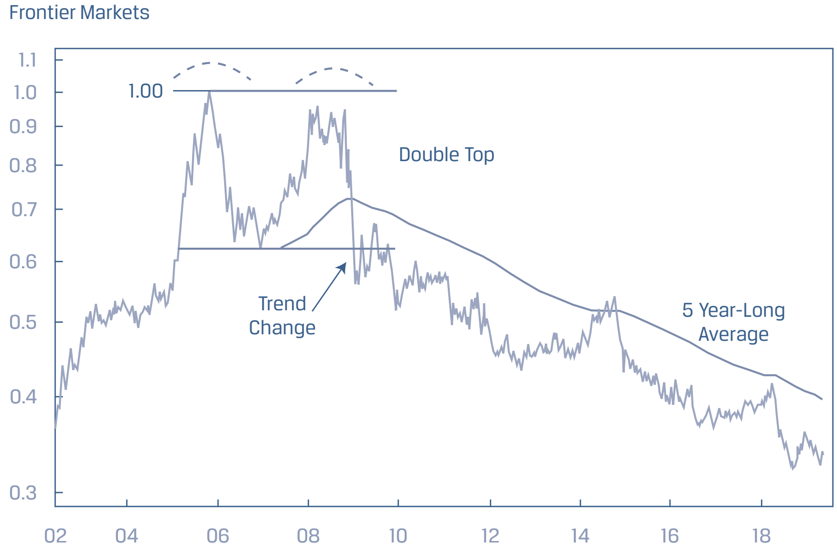
Exhibit 44 shows the relative strength of the MSCI Frontier Markets Index and the MSCI All Country World Index between 2002 and 2019. The highest point reached in 2005 is indexed to 1.0 for ease of analysis.

The charted ratio shows a strong outperformance for frontier market equities between 2002 and 2005. The ratio moved from the 0.35 level to as high as 1.00. This is almost 3 times outperformance during the three-year period. However, 2005 marked the highest point for the ratio over the analyzed period. In 2008, the ratio reached the previous highs and formed a double top. This chart pattern is the exact opposite of the one identified on the MSCI Emerging Markets versus MSCI ACWI chart during the 1998–2002 period.

A double top warns of a change in trend from up to down. At the beginning of 2009, the ratio not only breached the long-term average but also broke down the horizontal support formed by the neckline of the double top chart pattern.

Investors who had benefited from the outperformance of frontier market equities during 2002–2008 started allocating their funds out of frontier markets due to a change in trend. The following decade was a period of underperformance for the Frontier Markets Index. The ratio dropped from 0.7 to 0.35.

Exhibit 44 MSCI Frontier Markets Index vs. MSCI ACWI, 2002–2019



Out of the three major MSCI benchmarks—the developed, emerging, and frontier market indexes—both the Emerging Markets Index and the Frontier Markets Index started underperforming the All Country World Index after 2009–2010. At the same time, developed markets started outperforming the All Country World Index. The intermarket analysis suggested an allocation to developed markets.

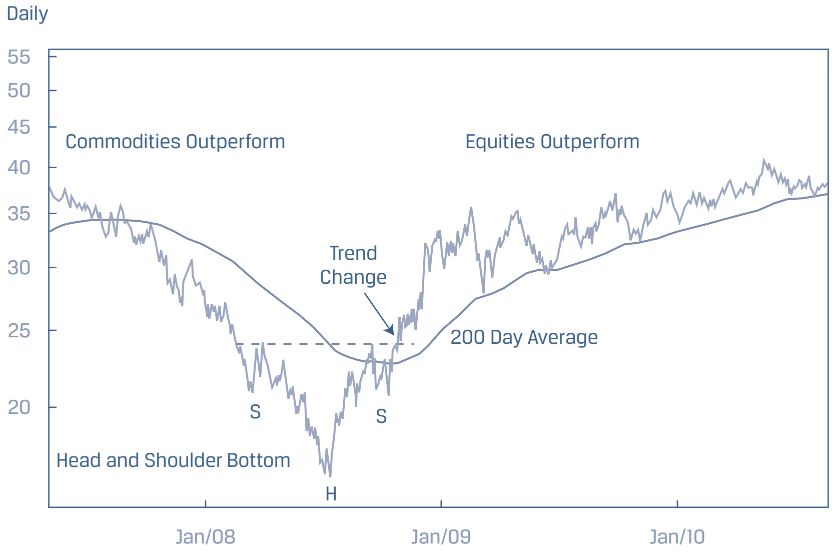
However, not all developed markets performed well after the 2009–10 period. The analyst or investor should also have been selective in picking the outperforming countries.

Intermarket analysis can be used to advantage in asset allocation decisions. **Tactical asset allocation (TAA)** is a portfolio strategy that shifts the percentages of assets held in various asset classes (or categories) to take advantage of market opportunities. Allocation shifts can occur within an asset class or across asset classes. This strategy allows portfolio managers to add value by taking advantage of certain situations in the marketplace.

Technical analysis and intermarket analysis help the portfolio manager make discretionary TAA decisions. In preparing to make these decisions, the portfolio manager studies trends and possible changes in the relationships on relative strength charts.

Exhibit 45 shows the relative strength of the S&P 500 Index and the Goldman Sachs Commodity Index on a daily scale. The ratio of these indexes was trending down between May 2007 and July 2008. The downtrend in the ratio signaled an outperformance for commodities. The ratio formed a head and shoulders bottom reversal during 2008. Completion of the head and shoulders bottom took place at the same time as a breach above the 200-day average. These two signals confirmed the change in trend, which favored equities over commodities. From a TAA perspective, intermarket analysis suggested shifting assets from commodities to equities.

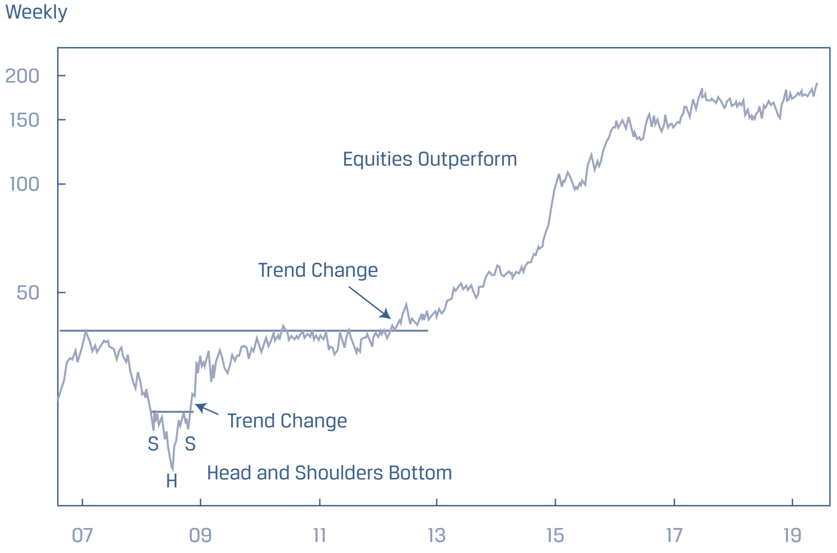
Exhibit 45 **S&P 500 Index vs. GSCI (Equities vs. Commodities), May 2007–August 2010**



Short-term charts such as the daily scale help us identify changes in trends at the early stages. Longer time frames will alert us to major shifts in demand and supply.

Exhibit 46 shows the relative performance of the S&P 500 Index and the Goldman Sachs Commodity Index on a weekly scale. Following the initial trend change in the last quarter of 2008, equities outperformed commodities until 2010. The ratio moved sideways between 2010 and 2012. In the first half of 2012, the S&P 500 to GSCI ratio cleared the horizontal resistance, suggesting a new phase of outperformance. Asset allocators would have higher conviction for increasing equity allocations following the renewed strength and breakout in the second half of 2012.

Exhibit 46 **S&P 500 Index vs. GSCI (Equities vs. Commodities), 2007–2019**



A similar relative strength chart can be drawn to compare bonds and commodities. Exhibit 47 shows the ratio of the 30-year US Treasury bond to the Goldman Sachs Commodity Index. Bonds started outperforming commodities in the second half of 2008. From the first quarter of 2009 till the last quarter of 2014, the ratio remained sideways and formed a symmetrical triangle. Consolidation periods are usually followed by trends. The breakout and trend change took place in the last quarter of 2014. Renewed strength in the ratio suggested outperformance for bonds versus commodities. From a TAA perspective, the intermarket analysis suggested shifting allocations from commodities to bonds.

Exhibit 47 30-Year US Treasury Bond vs. GSCI (Bonds vs. Commodities), 2006–2019

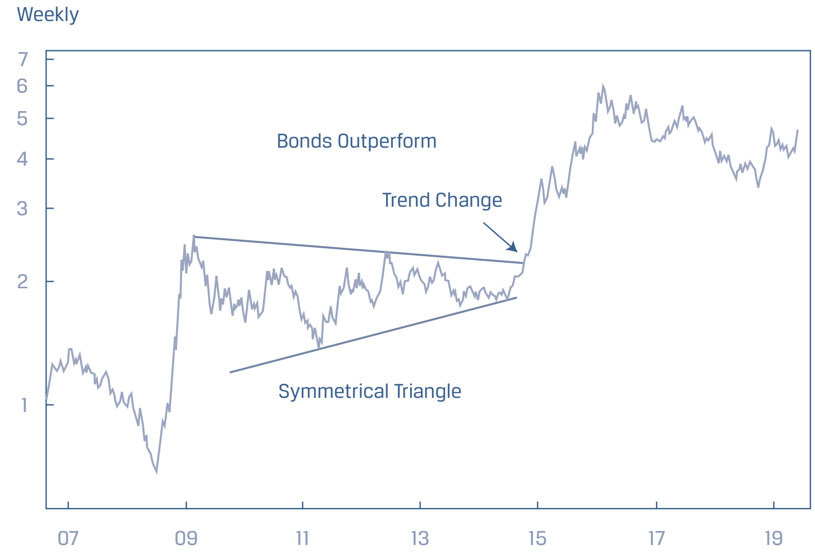
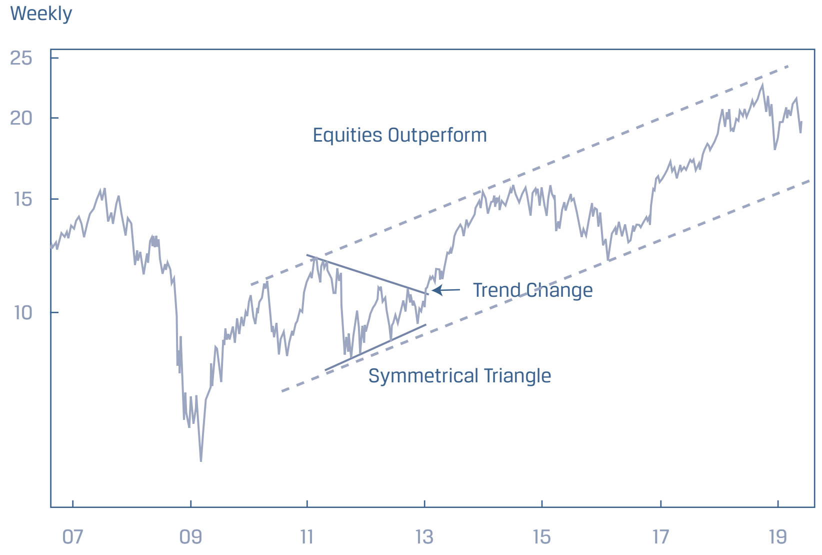


Exhibit 48 shows the relative strength of the S&P 500 Index and the 30-year US Treasury bond. Starting at the beginning of 2009, intermarket analysis suggested outperformance for equities versus bonds. Classical charting principles would have helped the asset allocator to identify the change in trend at the beginning of 2013, when the ratio completed a symmetrical triangle.

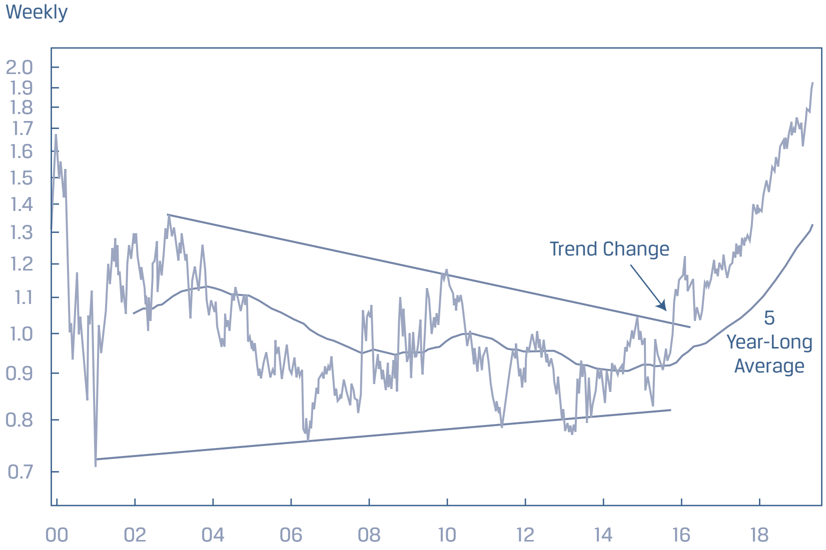
Exhibit 48 S&P 500 Index vs. 30-Year US Treasury Bond (Equities vs. Bonds), 2006–2019



We started our intermarket analysis with global equity markets. With the help of relative strength analysis and classical charting principles, we have identified trends favoring emerging and frontier markets and then developed markets.

As another example, Exhibit 49 shows the relative strength of Microsoft Corporation and the MSCI USA Index. Microsoft is a large holding in this benchmark. It is important for fund managers and portfolio managers to have a view on large holdings in the index. The only significant technical analysis signal was triggered when the ratio completed a multi-year-long symmetrical triangle consolidation at the beginning of 2016. The inflection point is labeled "trend change" on the chart and is indexed to 1.0. After the breakout, the ratio pulled back to test the previous support/resistance area formed by the apex of the symmetrical triangle. Microsoft entered a period of strong outperformance between 2016 and 2019. The ratio moved from 1.00 to 1.90 over the three-year period. The intermarket analysis suggested an overweight position for Microsoft.

Exhibit 49 Microsoft Corporation vs. MSCI USA Index, 2000–2019

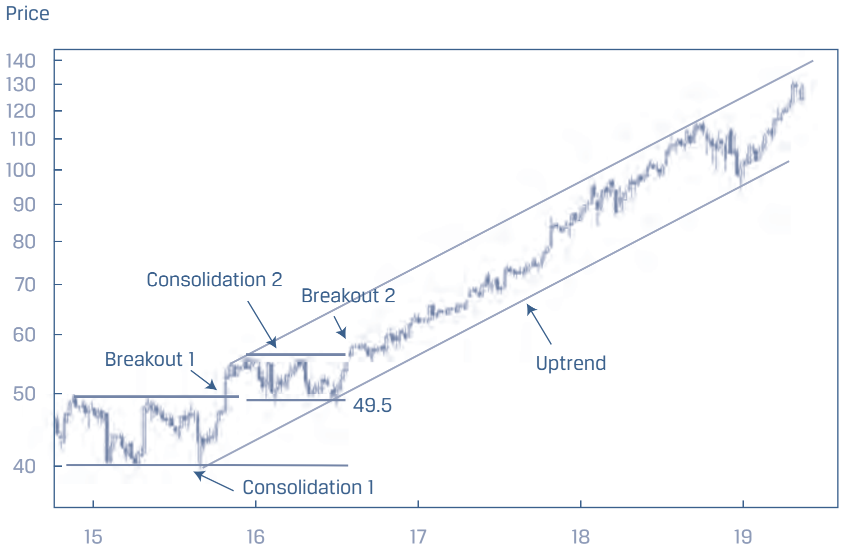


While the price ratio was generating a positive technical signal in favor of Microsoft, one would have questioned how to build positions in the stock. At this point, the analyst or investor would refer to the company's price chart to see trend and consolidation periods and the possibility to capture one of those well-defined breakout opportunities.

Exhibit 50 shows the price chart for Microsoft between October 2014 and May 2019. At the time when Microsoft started outperforming the MSCI USA Index, the price chart was trying to complete the yearlong rectangle labeled "Consolidation 1." The upper boundary of this yearlong consolidation acted as resistance at the 49.5 level. In the last quarter of 2015, Microsoft broke out of its rectangle chart pattern with a strong weekly candlestick. At the same time, the ratio of Microsoft to the MSCI USA (Exhibit 49) broke out of its multi-year-long symmetrical triangle consolidation.

The breakout on the Microsoft price chart was followed by a pullback and another multi-month-long consolidation. During this new consolidation, labeled "Consolidation 2" in Exhibit 50, previous resistance acted as support at the 49.5 level. Consolidation periods are price ranges in which investors with a longer-term focus assess the strength of the company and accumulate shares. If demand overcomes the supply, price eventually breaks out of the consolidation, and the trend resumes. Microsoft completed the second rectangle chart pattern in the second half of 2016 and resumed its uptrend. "Breakout 1" and "Breakout 2" are two inflection points on the price chart where Microsoft generated technical buy signals.

Exhibit 50 Microsoft Corporation, October 2014–May 2019 (US$)



8.2.2 *Technical Analysis Applications to Portfolio Management: Bottom-Up Approach*

A bottom-up investing approach focuses on the analysis of individual stocks. In bottom-up investing, the investor focuses his or her attention on a specific company's technicals rather than on the industry or index in which that company is classified. A bottom-up approach focuses on stock picking irrespective of market, industry, or more macro trends. It is possible that a thorough analysis of individual stocks will reveal investment themes in different sectors and industries. For example, several bottom reversals on individual health care stocks might suggest a turnaround in that sector due to anticipated news or a change in fundamentals.

To use the bottom-up approach, first, the investor or analyst needs an investment universe—a group of instruments to make a selection from. Second, the investor or analyst needs a method for selecting instruments using predefined criteria.

We will focus on momentum and breakout strategies. A breakout trader enters a long position after the stock price breaks above resistance or enters a short position after the stock breaks below support.

In our analysis, we will follow the following criteria, or trading rules, for stock selection:

1. The breakout should take place above the 200-day exponential moving average (long-only strategy).
2. Price should be in a low-volatility condition prior to the breakout.
3. The breakout should take place from a well-defined classical chart pattern between 3 and 24 months in duration.
4. Breakout confirmation should be a weekly close above the chart pattern boundary.

Our opportunity set will be developed market equities.

The examples below highlight breakout opportunities that were possible to identify during 2016. These breakouts not only captured directional movements on price charts but also pointed analysts and investors to a specific theme.

Our initial requirement for a stock is that it be going through a low-volatility period. We would like to capture those stocks that are in consolidation periods. If the opportunity set is large, the analyst can use filtering criteria to identify stocks with low Bollinger Band readings, as discussed in preceding sections. Following the initial screening, the analyst will use the **visual technique** to select one of the well-defined classical chart patterns.

Before we proceed to the examples below, we should clarify one important point. To use a bottom-up approach to stock selection, an analyst should have a method for identifying potential trades. Focusing on breakout and momentum is one of many different approaches to stock selection and investment decision making. Each analyst, investor, or trader can implement his or her own approach to identifying opportunities.

In our simplified scenario, these are the commonsense technical rules that we choose to apply:

- The breakout should take place above the 200-day exponential moving average (long-only strategy).

It is important to identify stocks that are already in a steady uptrend or are entering into an uptrend. Breakouts from consolidation periods in or preceding an uptrend are usually followed by a continuation of the existing trend. Investing in the overall direction of the trend is the *path of least resistance.*

- Price should be in a low-volatility condition.

Low-volatility conditions are usually followed by high volatility, and vice versa. Stocks that break out from low-volatility conditions are likely to experience high volatility, and high volatility in the direction of the existing trend will usually help the trend to pick up momentum.

- The breakout should take place from a well-defined classical chart pattern between 3 and 24 months in duration.

The longer the consolidation, the stronger the breakout and the subsequent trend period will be.

- Breakout confirmation should be a weekly close above the chart pattern boundary.

In an attempt to avoid premature breakout decisions, we will wait for a decisive weekly close above the chart pattern boundary.

Exhibit 51 shows the price chart for Belo Sun Mining Company from July 2013 to February 2017. The stock is listed on the Toronto Stock Exchange. The price chart formed a multi-month-long ascending triangle between 2014 and 2015. An ascending triangle is a bullish chart pattern. Its upward-sloping boundary gives this triangle its bullish bias. This pattern suggests that buyers are able to bid the price to higher levels after each correction.

Prior to the formation of an ascending triangle, the stock was in a steady downtrend. The ascending triangle acted as a consolidation and represented a period of low volatility. The breakout above the chart pattern boundary at 0.28 took place at the beginning of 2016. At this point, the stock was above its long-term average and ready for a new trend period. The breakout took place with a strong weekly candlestick. The bullish ascending triangle acted as a bottom reversal for the silver mining company, and a strong trend period followed. According to the bottom-up approach, in the last quarter of 2015, Belo Sun Mining should have been on our watch list for a possible breakout opportunity.

Exhibit 51 Belo Sun Mining Company, July 2013–February 2017 (Price in Canadian Dollars)

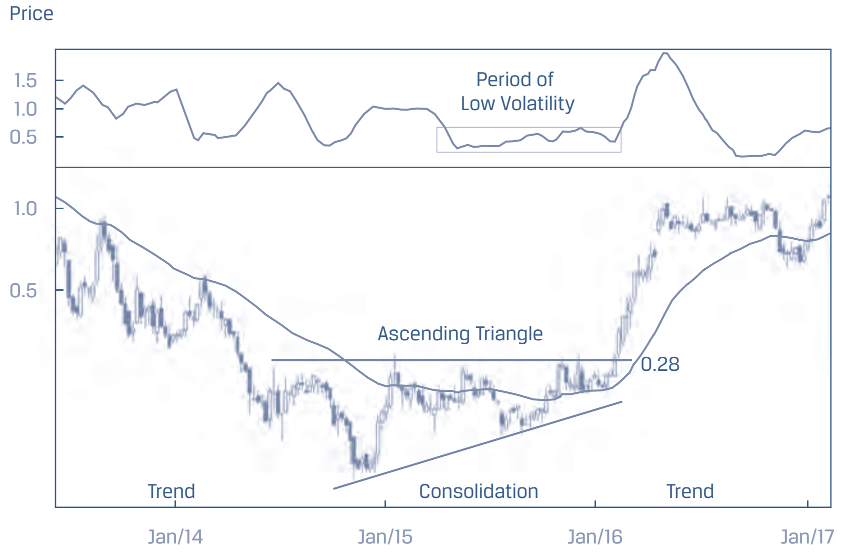
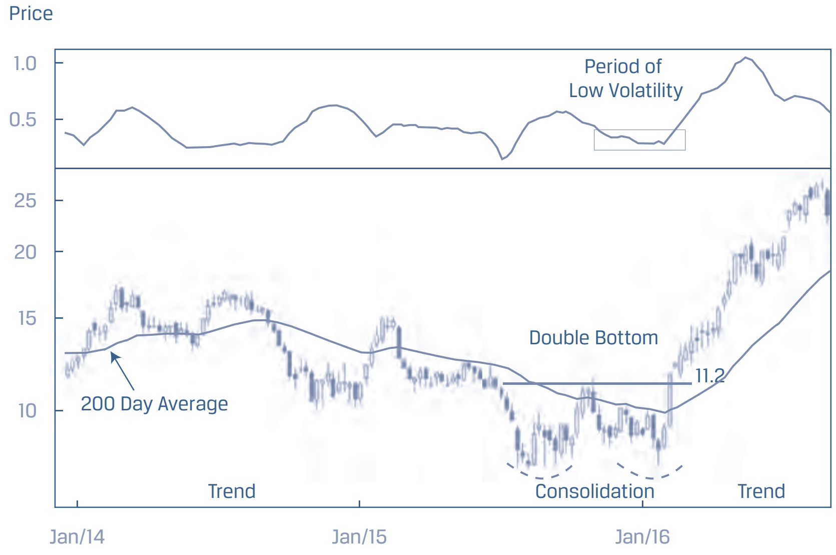


Exhibit 52 shows the price chart for Pan American Silver Corporation between January 2014 and September 2016. The stock is listed on the Toronto Stock Exchange. Pan American Silver formed a double bottom chart pattern, a bullish reversal, around the same time some of the silver mining companies formed consolidations. The double bottom formed in the second half of 2015. At the beginning of 2016, a breakout from the bullish reversal chart pattern above the 11.2 levels took place with a strong weekly candlestick.

Because these two stocks were going through low-volatility conditions and forming one of the well-defined chart patterns, they would have been on our watch list. Bullish bias in all the chart pattern developments would have alerted us to a possible reversal on silver prices and possibly other precious/industrial metal prices. This example shows how one can start with individual stocks and end up finding an investment theme.

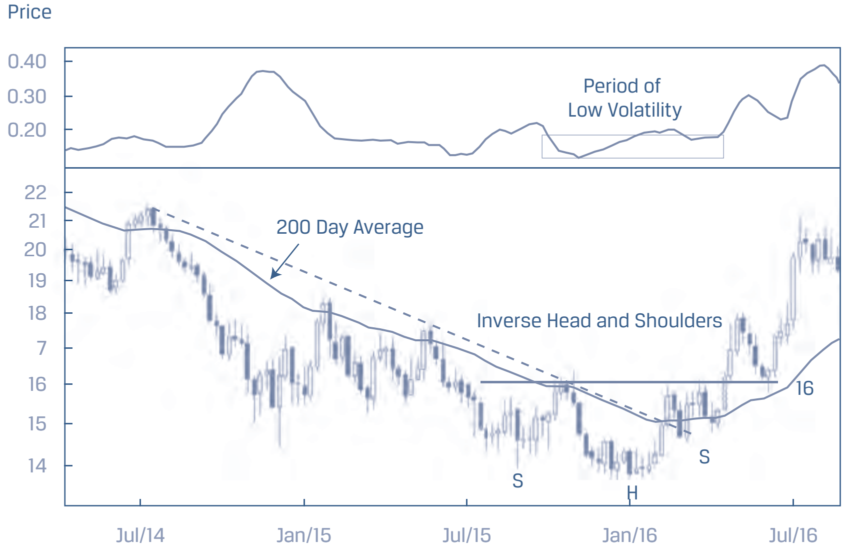
Exhibit 52 Pan American Silver Corporation, January 2014–September 2016 (Price in Canadian Dollars)



In this instance, strength in silver mining equities happened to lead the breakout in commodity prices.

Exhibit 53 shows the silver (per ounce) price between April 2014 and August 2016. Silver mining equities bottomed or completed bullish continuation chart patterns in the beginning of 2016. Silver prices formed an inverse head and shoulders chart pattern from the second half of 2015 to the first quarter of 2016. A breakout from the inverse head and shoulders chart pattern took place in April 2016, when price cleared the horizontal resistance at the 16 level. Following the breakout, price pulled back to the chart pattern boundary and resumed the uptrend.

Exhibit 53 Silver (per Ounce) Price, April 2014–August 2016 (US$)



8.3 The Role of the Technical Analyst in Fundamental Portfolio Management

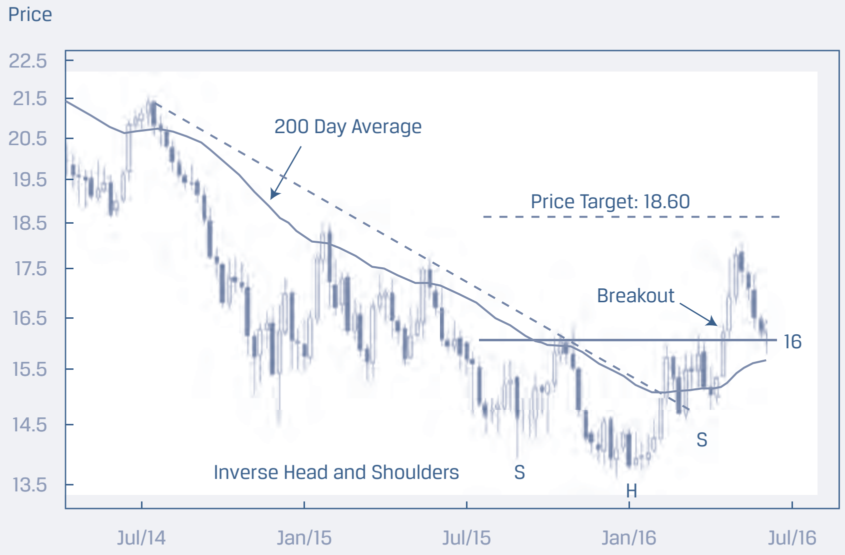
A technical analyst can serve a supporting role in a team of investors. The technician can conduct research and find investment opportunities by applying different strategies. As mentioned in the examples above, the technical analyst goes through the opportunity set and provides trade and investment ideas to the portfolio/fund managers.

The key value-added input would be in the form of timing of the purchase or sale of that security. The technical analyst should provide the rationale as well as potential price targets for the expected move and the price level at which the analysis would be invalidated. The technical analyst should follow a purist approach and follow the results of the research without being influenced by other inputs, such as news and fundamental data on the specific company—though news confirming the development of a technical trend is always welcome.

The technical analyst will typically not be directly involved in position sizing decisions, such as how much of the portfolio to allocate to a certain idea or by what margin the fund should overweight or underweight the specific stock. However, the portfolio manager may well use high-conviction inputs from the technical analyst in weighting decisions.

Question 13

A trader identifies an inverse head and shoulders chart pattern on the commodity XYZ. He plans to buy 100 units of the commodity at the time of the breakout. However, he only manages to get filled for 20 units because the price moves quickly. Over the next few weeks, the price moves close to the chart pattern price target at the 18.6 level but then reverses back to the breakout level. What should the trader do, given the price movement, assuming the chart pattern will hold?



a Sell the existing 20 units because the breakout failed.

b Buy the remaining 80 units (of the initial 100-unit purchase plan) because the price is giving another opportunity with a pullback.

c Wait for further correction toward the next support between 14.5 and 15.0.

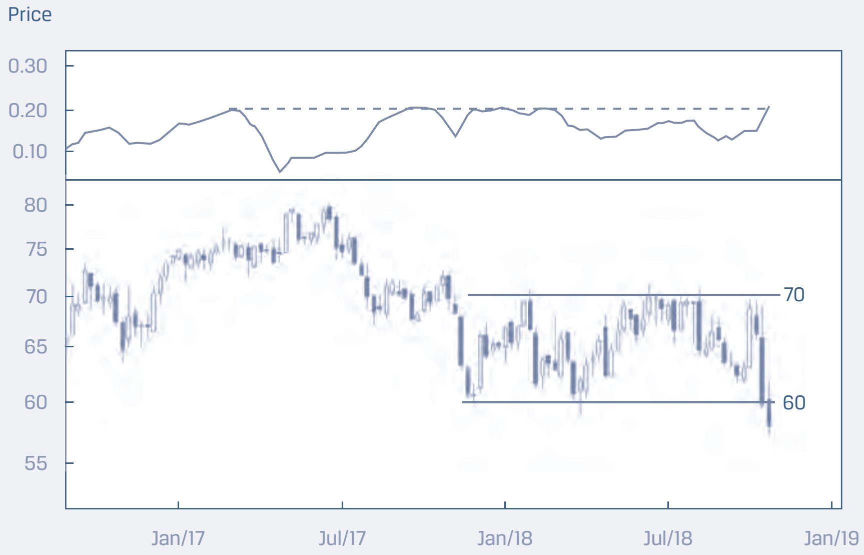
Answer: B. The price pullback to the chart pattern boundary provides an opportunity to complete the purchase of the planned quantity.

A is not the correct answer because the breakout didn't fail: The price held above the chart pattern boundary.

C is not the correct answer because a correction toward the 14.5–15.0 area would put the bullish interpretation in question and the chart pattern might lose its validity.

Question 14

A fund manager has been trading ABC stock between the horizontal chart pattern boundaries throughout 2018. Every time the price reached the 60 level (support), she was overweighting the stock, and every time it reached the 70 level (resistance), she was underweighting it. This strategy worked well until October 2018, when the stock price slipped below the horizontal support at the 60 level. The fund manager's mandate is to beat a benchmark, and the stock is one of the blue chips in the index with a high weighting. What should the fund manager do?



a Buy more of the ABC stock and remain overweight because the stock is now cheaper.

b Sell part of the exposure and remain underweight because the stock might be breaking down a lengthy consolidation.

c Expect the stock to remain sideways and maintain a market-weight position.

Answer: B.

The stock is completing a yearlong rectangle pattern, and the volatility measured by the Bollinger Band width is about to break out from a two-year-long consolidation. Therefore, there is a high probability that a trend period will follow and the stock will underperform. Selling part of the exposure and going underweight the benchmark at a time when the stock is likely to trend downward would result in outperformance versus the benchmark, adding value to the portfolio.

A is not the correct answer because buying more of the stock and overweighting the position at a time when the stock is likely to trend downward would result in underperformance versus the benchmark.

C is not the correct answer because maintaining a market-weight position at a time when the stock is likely to trend downward would result in a lost opportunity to outperform the benchmark by underweighting the position.

SUMMARY

- Technical analysis is a form of security analysis that uses price data and volume data, typically displayed graphically in charts. The charts are analyzed using various indicators in order to make investment recommendations.
- Technical analysis has three main principles and assumptions: (1) The market discounts everything, (2) prices move in trends and countertrends, and (3) price action is repetitive, with certain patterns reoccurring.
- Increasingly, analysts, fund managers, and individual investors are studying the basic principles of technical analysis to support their decision making in financial markets. Behavioral finance, which is the study of the influence of psychology on the behavior of investors, focuses on the fact that investors are not always rational, have limits to their self-control, and are influenced by their own biases. This relatively new field of finance is motivating more practitioners to consider technical analysis as a tool for understanding and explaining irrationalities in financial markets.
- Technical analysis can be used on any freely traded security in the global market and is used on a wide range of financial instruments, such as equities, bonds, commodities, currencies, and futures. However, in general, technical analysis is most effectively applied to liquid markets. Therefore, technical analysis has limited usefulness for illiquid securities, where a small trade can have a large impact on prices.
- The primary tools used in technical analysis are charts and indicators. Charts are graphical displays of price and volume data. Indicators are approaches to analyzing the charts. While the tools can be used on a standalone basis, many analysts, fund managers, and investors will find added value in combining the techniques of chart analysis with their own research and investment approach.
- Charts provide information about past price behavior and provide a basis for inferences about likely future price behavior. Basic charts include line charts, bar charts, and candlestick charts.
- Charts can be drawn either to a linear scale or to a logarithmic scale. A logarithmic scale is appropriate when the data move through a range of values representing several orders of magnitude (e.g., from 10 to 10,000), whereas a linear scale is better suited to narrower ranges (e.g., \$35 to \$50).
- Volume is an important element of technical analysis and is often included on charts. Volume can be viewed as a confirmation in that it indicates the strength or conviction of buyers and sellers in determining a security's price.
- One of the most important steps in successfully applying technical analysis is to define the time period being analyzed. Technical analysis and charting become more reliable as the time scale increases from intraday to daily, weekly, and even monthly. Analysts and investors whose primary research method is fundamental analysis will find more value in charting instruments on a weekly and/or a monthly scale. Longer time frames will allow analysts and investors to better identify the consolidation and trend periods and time their purchases or sales of securities.
- Several basic concepts can be applied to charts. These include relative strength analysis, trend, consolidation, support, resistance, and change in polarity.
- Relative strength analysis is based on the ratio of the prices of a security and a benchmark and is used to compare the performance of one asset with the performance of another asset.

- The concept of trend is an important aspect of technical analysis. An uptrend is defined as a sequence of higher highs and higher lows. To draw an uptrend line, a technician draws a line connecting the lows on the price chart. A downtrend is defined as a sequence of lower highs and lower lows. To draw a downtrend line, a technician draws a line connecting the highs on the price chart.
- Support is defined as a low price range in which the price stops declining because of buying activity. It is the opposite of resistance, which is a price range in which price stops rising because of selling activity.
- Chart patterns are formations appearing on price charts that create some type of recognizable shape. There are two major types of chart patterns: reversal patterns and continuation patterns.
- Reversal patterns signal the end of a trend. Common reversal patterns are head and shoulders (H&S), inverse H&S, double top, double bottom, triple top, and triple bottom.
- Continuation patterns indicate that a market trend that was in place prior to the pattern formation will continue once the pattern is completed. Common continuation patterns are triangles (symmetrical, ascending, and descending), rectangles (bullish and bearish), flags, and pennants.
- Technical indicators are used to derive additional information from basic chart patterns. An indicator is any measure based on price, market sentiment, or fund flows that can be used to predict changes in price. Mathematically calculated indicators usually have a supply and demand underpinning. Basic types of indicators include price-based indicators, momentum oscillators, and sentiment indicators.
- Price-based indicators incorporate information contained in market prices. Common price-based indicators include the moving average and Bollinger Bands.
- The moving average is the average of the closing prices of a security over a specified number of periods. Moving averages are a smoothing technique that gives the technical analyst a view of market trends. So, a moving average can be viewed as a trend filter. Long-term moving averages can provide important signals. A price move above the long-term moving average is a sign of an uptrend. A price move below the long-term moving average is a sign of a downtrend.
- When a short-term moving average crosses over a longer-term moving average from underneath, this movement is considered a bullish indicator and is called a “bullish crossover.” When a short-term moving-average crosses over a longer-term moving average from above, this movement is a bearish indicator and is called a “bearish crossover.”
- Bollinger Bands combine the concept of a moving average with standard deviations around the moving average. This tool is useful in defining a trading range for the security being analyzed. The Bollinger Band width indicator provides an indication of volatility. The idea is that periods of low volatility are followed by periods of high volatility, so that relatively narrow band width can foreshadow an advance or decline in the security under analysis.
- Momentum oscillators are constructed from price data, but they are calculated so that they fluctuate between a low and a high, typically between 0 and 100. Some examples of momentum oscillators include rate of change (ROC) oscillators, the relative strength index (RSI), stochastic oscillators, and the MACD (moving-average convergence/divergence oscillator).

- Momentum oscillators can be viewed as graphical representations of market sentiment that show when selling or buying activity is more aggressive than usual. Technical analysts also look for convergence or divergence between oscillators and price. For example, when the price reaches a new high, this outcome is usually considered "bullish." But if the momentum oscillator does not also reach a new high, this scenario is considered divergence and an early warning sign of weakness.
- Momentum oscillators also alert the technical analyst to overbought or oversold conditions. For example, in an oversold condition, market sentiment is considered unsustainably bearish.
- Sentiment indicators attempt to gauge investor activity for signs of increasing bullishness or bearishness. Commonly used calculated statistical indexes are the put/call ratio, the VIX, and margin debt.
- Intermarket analysis combines technical analysis of the major categories of securities—namely, equities, bonds, currencies, and commodities—to identify market trends and possible inflections in trends. Intermarket analysis also looks at industry subsectors and their relationship to sectors and industries. In addition, it measures the relative performance of major equity benchmarks around the globe.
- Technical analysis can use either a top-down approach or a bottom-up approach to analyze securities. The top-down method is useful for identifying outperforming asset classes, countries, or sectors. This approach can add value to asset allocation decisions. Allocation shifts can occur within an asset class or across asset classes. The bottom-up method is useful for identifying individual stocks, commodities, or currencies that are outperforming, irrespective of market, industry, or macro trends.
- The technical analyst can add value to an investment team by providing trading/investment ideas through either top-down or bottom-up analysis, depending on the nature of the investment firm or fund. In addition, technical analysis can add value to a fundamental portfolio approach by providing input on the timing of the purchase or sale of a security.

PRACTICE PROBLEMS

1 Technical analysis relies most importantly on:

- **A** price and volume data.
- **B** accurate financial statements.
- **C** fundamental analysis to confirm conclusions.

2 Which of the following is *not* an assumption of technical analysis?

- **A** Security markets are efficient.
- **B** The security under analysis is freely traded.
- **C** Market trends and patterns tend to repeat themselves.

3 Drawbacks of technical analysis include which of the following?

- **A** It identifies changes in trends only after the fact.
- **B** Deviations from intrinsic value can persist for long periods.
- **C** It usually requires detailed knowledge of the financial instrument under analysis.

4 Why is technical analysis especially useful in the analysis of commodities and currencies?

- **A** Valuation models cannot be used to determine fundamental intrinsic value for these securities.
- **B** Government regulators are more likely to intervene in these markets.
- **C** These types of securities display clearer trends than equities and bonds do.

5 Technical analysis is a form of security analysis that:

- **A** assesses past price action to project future prices.
- **B** requires in-depth knowledge of financial instruments.
- **C** is ineffective when evaluating long-term price movements.

6 One principle of technical analysis is that a security's price:

- **A** tends to move in a random fashion.
- **B** moves in patterns that tend to reoccur.
- **C** does not reflect all known factor information relating to the security.

7 A daily bar chart provides:

- **A** a logarithmically scaled horizontal axis.
- **B** a horizontal axis that represents changes in price.
- **C** high and low prices during the day and the day's opening and closing prices.

8 A candlestick chart is similar to a bar chart *except* that the candlestick chart:

- **A** represents upward movements in price with X's.
- **B** also graphically shows the range of the period's highs and lows.
- **C** has a body that is light or dark depending on whether the security closed higher or lower than its open.

9 In a candlestick chart, a shaded candlestick body indicates that the opening price was:

- **A** equal to the closing price.
- **B** lower than the closing price.

© 2020 CFA Institute. All rights reserved.

C higher than the closing price.

10 A chart constructed with a single data point per time interval is a:

A bar chart.

B line chart.

C candlestick chart.

11 In analyzing a price chart, high or increasing volume *most likely* indicates which of the following?

A Predicts a reversal in the price trend.

B Predicts that a trendless period will follow.

C Confirms a rising or declining trend in prices.

12 In constructing a chart, using a logarithmic scale on the vertical axis is likely to be *most useful* for which of the following applications?

A The price of gold for the past 100 years.

B The share price of a company over the past month.

C Yields on 10-year US Treasuries for the past 5 years.

13 A linear price scale is:

A inappropriate for a candlestick chart.

B better suited for analysis of short-term price movements.

C constructed with equal vertical distances corresponding to an equal percentage price change.

14 A downtrend line is constructed by drawing a line connecting:

A the lows of the price chart.

B the highs of the price chart.

C the highest high to the lowest low of the price chart.

15 Relative strength analysis typically compares the performance of an asset with that of a benchmark or other security using a:

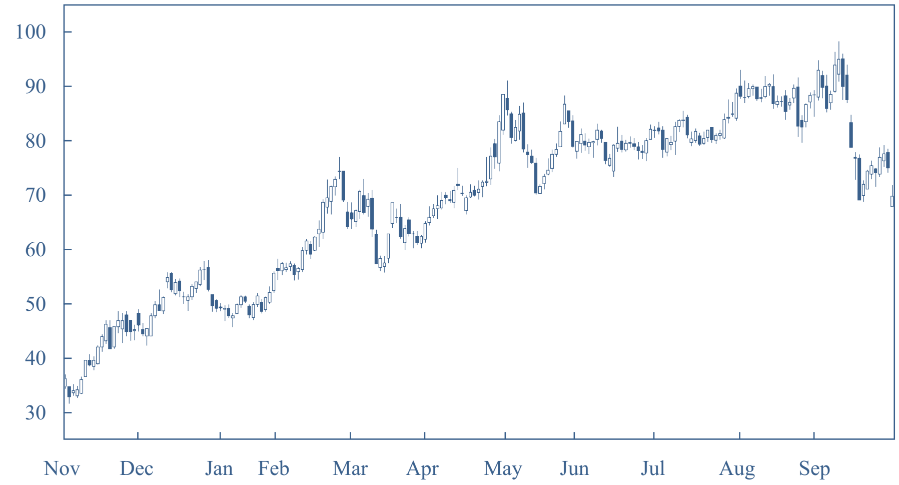
A bar chart that reflects the two assets' price history.

B line chart that reflects the ratio of the two assets' prices.

C candlestick chart that reflects ratios measuring the magnitude of each asset's up days versus down days.

16 Exhibit 1 depicts ABC Co., Ltd., ordinary shares, traded on the Shenzhen Stock Exchange, for the months of November through September in renminbi (RMB).

Exhibit 1 Candlestick Chart: ABC Co., Ltd. Price Data, November–September (Price Measured in RMB × 10)



Based on Exhibit 1, the uptrend was *most likely* broken at a price level nearest to:

A 7 RMB.

B 8.5 RMB.

C 10 RMB.

17 The "change in polarity" principle states which of the following?

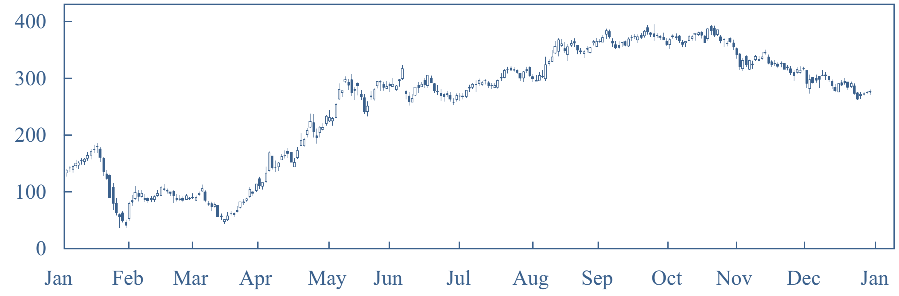
A Once an uptrend is broken, it becomes a downtrend.

B Once a resistance level is breached, it becomes a support level.

C The short-term moving average has crossed over the longer-term moving average.

18 Exhibit 2 depicts XYZ Co. ordinary shares, traded on the London Stock Exchange, in British pence.

Exhibit 2 Candlestick Chart: XYZ Co. Price Data, January–January (Price Measured in British Pence)



Based on Exhibit 2, Barclays appears to show resistance at a level nearest to:

A 50p.

B 275p.

C 390p.

19 When a security is not trending, it is considered to be in a:

A breakout.

B retracement.

C consolidation.

20 A technical analyst who observes a downtrending security break out of a consolidation range on the downside will *most likely* predict that the downtrend will:

A resume.

B reverse trend with an upside breakout.

C retrace back to the consolidation range.

21 Which of the following statements regarding technical support and resistance is correct?

A A breached support level becomes a new level of resistance.

B Support is a price range where selling activity is sufficient to stop a rise in price.

C Resistance is a price range where buying activity is sufficient to stop a decline in price.

22 Exhibit 3 depicts DGF Company common shares, traded on the New York Stock Exchange, for five years in US dollars.

Exhibit 3 Candlestick Chart: DGF Company, five years, February–February

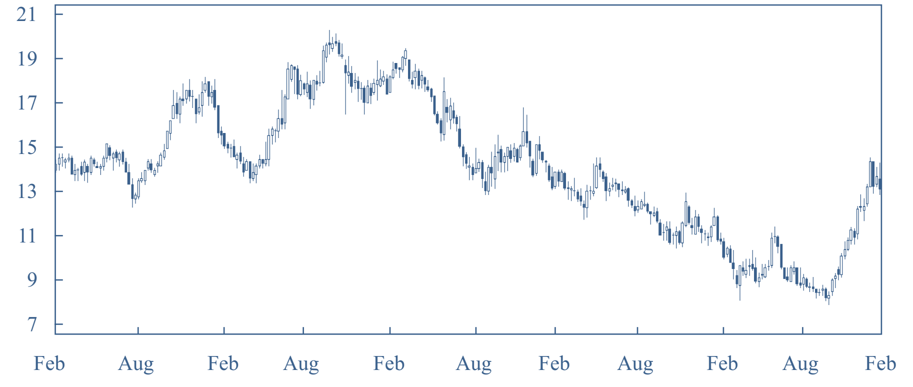


Exhibit 3 illustrates *most* clearly which type of pattern?

A Triangle.

B Triple top.

C Head and shoulders.

23 A triangle chart pattern that indicates a consolidation period and has bullish trading implications would *most likely* be classified as a(n):

A ascending triangle.

B descending triangle.

C symmetrical triangle.

24 In an inverted head and shoulders pattern, if the neckline is at €100, the shoulders at €90, and the head at €75, the price target is *closest* to which of the following?

A €50.

- **B** €110.
- **C** €125.

25 Which of the following chart patterns signals the end of an uptrend in price?

- **A** Bearish rectangle
- **B** Head and shoulders
- **C** Symmetrical triangle

26 An inverse head and shoulders acts as a reversal pattern for a preceding:

- **A** uptrend.
- **B** downtrend.
- **C** consolidation.

27 A fully formed head and shoulders pattern is *most likely* an indicator to:

- **A** buy.
- **B** sell.
- **C** hold.

28 To profit from a head and shoulders formation, a technician often sets a price target below the neckline price by an increment equal to the:

- **A** head minus neckline.
- **B** head minus top of right shoulder.
- **C** top of right shoulder minus neckline.

29 A "healthy correction" chart pattern:

- **A** is classified as a type of reversal pattern.
- **B** does not change long-term price trends since supply and demand remain in balance.
- **C** is formed when the price reaches a low, rebounds, and then sells off back to the first low level.

30 If the 5-day moving average for AZB Company crossed over its 60-day moving average from underneath, it would be considered a:

- **A** bullish indicator.
- **B** bearish indicator.
- **C** new level of resistance.

31 A trader observes that the 50-day moving average for the S&P 500 Index recently crossed below its long-term 200-day moving average. This situation is referred to as a:

- **A** death cross.
- **B** golden cross.
- **C** Bollinger Band.

32 Bollinger Bands are constructed by plotting:

- **A** a MACD line and a signal line.
- **B** a moving-average line with an uptrend line above and downtrend line below.
- **C** a moving-average line with upper and lower lines that are at a set number of standard deviations apart.

33 A Bollinger Band "squeeze" occurs when volatility:

- **A** falls to low levels as the Bollinger Band widens.
- **B** falls to low levels as the Bollinger Band narrows.
- **C** rises to high levels as the Bollinger Band narrows.

34 Which of the following is *not* a momentum oscillator?

- **A** MACD.
- **B** Stochastic oscillator.
- **C** Bollinger Bands.

35 Which of the following is a continuation pattern?

- **A** Triangle.
- **B** Triple top.
- **C** Head and shoulders.

36 Which of the following is a reversal pattern?

- **A** Pennant.
- **B** Rectangle.
- **C** Double bottom.

37 Which of the following is generally true of the head and shoulders pattern?

- **A** Volume is important in interpreting the data.
- **B** The neckline, once breached, becomes a support level.
- **C** Head and shoulders patterns are generally followed by an uptrend in the security's price.

38 Intermarket analysis focuses on the:

- **A** valuation drivers of intermarket asset price relationships.
- **B** bottom-up economic fundamentals of intermarket relationships.
- **C** identification of inflection points in intermarket relationships using relative strength indicators.

39 A technical analyst following a bottom-up investing approach focusing on momentum and breakout strategies should favor long positions in stocks with:

- **A** shorter consolidation periods.
- **B** high Bollinger Band readings.
- **C** low volatility prior to an upside breakout.

SOLUTIONS

1 A is correct. Almost all technical analysis relies on these data inputs.

2 A is correct. Technical analysis works because markets are *not* efficient and rational and because human beings tend to behave similarly in similar circumstances. The result is market trends and patterns that repeat themselves and are somewhat predictable.

3 A is correct. Trends generally must be in place for some time before they are recognizable. Thus, some time may be needed for a change in trend to be identified.

4 A is correct. Commodities and currencies do not have underlying financial statements or an income stream; thus, fundamental analysis is useless in determining theoretical values for them or whether they are over- or undervalued.

5 A is correct. The underlying logic of technical analysis is that past price action can be useful to anticipate and project potential future prices with charts and other technical tools

6 B is correct. According to technical analysts, market psychology leads to repetition of price movements. Because investor behavior repeats itself, price movement patterns can be charted out, allowing technicians to recognize patterns.

7 C is correct. The top and bottom of the bars indicate the highs and lows for the day; the line on the left indicates the opening price and the line on the right indicates the closing price.

8 C is correct. Dark and light shading is a unique feature of candlestick charts.

9 C is correct. A shaded candlestick body indicates that the price of the security closed down from its opening price, whereas a clear body indicates that the price closed up from its opening price. Thus, a shaded candlestick body indicates that the opening price was higher than the closing price.

10 B is correct. A line chart has one data point per time interval, with price on the vertical axis and unit of time on the horizontal axis.

11 C is correct. Rising volume shows conviction by many market participants, which is likely to lead to a continuation of the trend.

12 A is correct. The price of gold in nominal dollars was several orders of magnitude cheaper 100 years ago than it is today (roughly US$20 then versus US$1,100 today). Such a wide range of prices lends itself well to being graphically displayed on a logarithmic scale.

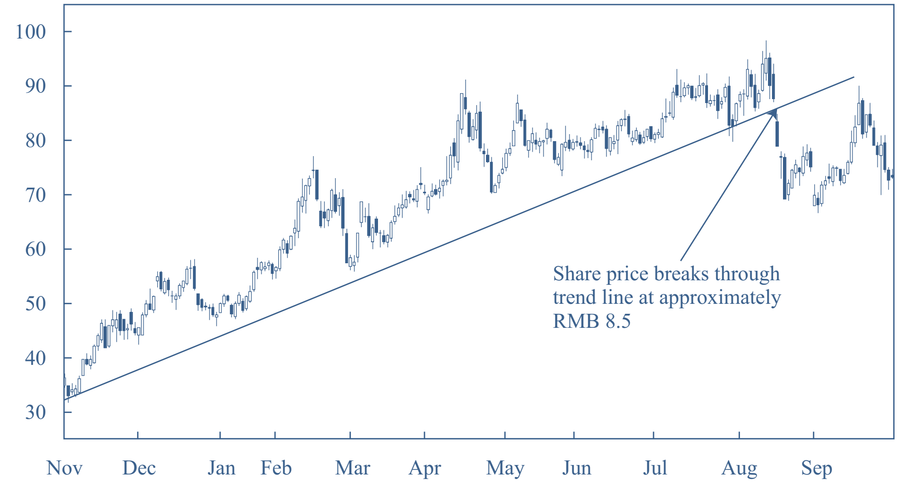
13 B is correct. A linearly scaled (rather than a logarithmically scaled) chart is better suited for analysis of short-term price movements. A linear scale plots price against a vertical axis with an equal distance between prices, whereas with a logarithmic scale, equal vertical distances correspond to an equal percentage change. The difference between a logarithmic price chart and an arithmetic (linear) price chart can be small when analyzing a chart in the short term. However, major differences are apparent when analyzing longer-term charts (more than two years of price data).

14 B is correct. A downtrend line is constructed by drawing a line connecting the highs of the price chart.

15 B is correct. In relative strength analysis, the analyst typically prepares a line chart of the ratio of the two assets' prices, with the asset under analysis as the numerator and with the benchmark or other security as the denominator. With this single line chart, the analyst can readily visualize relative performance by the positive or negative slope of the line. A rising line shows the asset is performing better than the index or other stock; a declining line shows the opposite.

16 B is correct. It is demonstrated in the following chart:

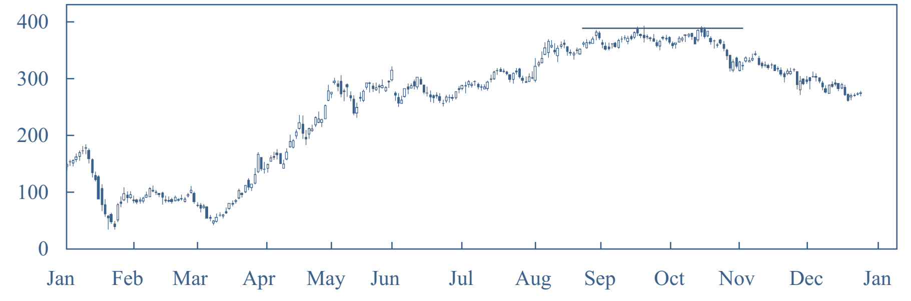
Exhibit 1 Candlestick Chart: ABC Co., Ltd. Price Data, November–September (Price Measured in RMB × 10)



17 B is correct.

18 C is correct. As shown in the following chart, XYZ Co. shares traded up to 390p on three occasions, each several weeks apart, and declined thereafter each time.

Exhibit 2 Candlestick Chart: XYZ Co. Price Data, January–January (Price Measured in British Pence)



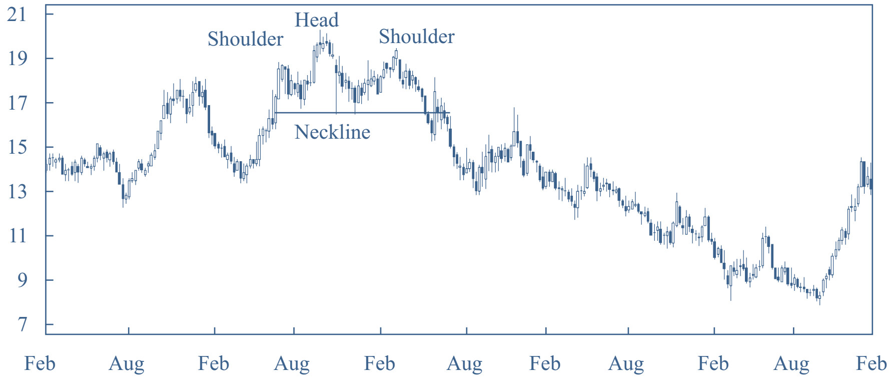
19 C is correct. In technical analysis, when a security is not trending, it is considered to be in consolidation. A consolidation phase is characterized by a fairly narrow trading range in which the price moves sideways (indicative of relative supply/demand balance) without much upward or downward movement. The key to technical analysis is the ability to differentiate between a consolidation and a trend phase.

20 A is correct. When a previously downtrending price breaks out of its consolidation range on the downside, it suggests that the price will enter a new downtrend phase.

21 A is correct. According to the change in polarity principle, once a support level is breached, it becomes a level of resistance. Likewise, when a resistance level is breached, it becomes a level of support.

22 C is correct. The left shoulder formed at around US$18.50, the head formed at around US$20.50, and the second shoulder formed at around US$19.

Exhibit 3 Candlestick Chart: DGF Company, five years, February–February



23 A is correct. In the ascending triangle pattern, irrespective of where the ascending triangle is identified, it should be considered as a consolidation with bullish implications.

24 C is correct. Target = Neckline + (Neckline − Head): €100 + (€100 − €75) = €125

25 B is correct. A head and shoulders pattern is a reversal pattern that signals the end of an uptrend in price. Once a head and shoulders pattern has been formed, the expectation is that the price will decline through the neckline price of the formation.

26 B is correct. An inverse head and shoulders is a reversal pattern for a downtrend that preceded the formation of the pattern. An inverse head and shoulders is also referred to as a head and shoulders bottom and signals a potential reversal from the preceding downtrend to an uptrend. An inverse head and shoulders is the opposite of a head and shoulders pattern that signals a reversal from a preceding uptrend to a downtrend.

27 B is correct. A head and shoulders formation is a sell indicator that signals a reversal of a preceding uptrend. Once the head and shoulders pattern forms, the expectation is that the price will decline through the neckline price, setting the stage for a downtrend phase.

28 A is correct. When attempting to profit from a head and shoulders formation, a technician will often use the price difference between the head and neckline to set a price target below the neckline. The price target is therefore set as follows: Price target = Neckline – (Head – Neckline).

29 B is correct. With a "healthy correction," a type of continuation pattern, the long-term price trend does not change as supply and demand remains in balance while ownership transitions from one investor group to another.

30 A is correct. When the short-term moving average crosses above the long-term moving average, it can be viewed as a bullish buy signal. For example, a 5-day short-term moving average (a proxy for short-term momentum) breaking up through a 60-day longer-term moving average (an indicator of intermediate trend) can be used as a buy signal.

31 A is correct. A widely followed moving average crossover signal is the one that takes place between the 50-day moving average and the 200-day moving average. When the short-term (50-day) moving average crosses below the long-term (200-day) moving average, it is called a death cross and is a bearish signal.

32 C is correct. Bollinger Bands consist of a moving average and a higher line representing the moving average plus a set number of standard deviations from average price (for the same number of periods as used to calculate the moving average) and a lower line that is a moving average minus the same number of standard deviations.

33 B is correct. A Bollinger Band "squeeze" occurs when volatility falls to a very low level, as evidenced by the narrowing bands.

34 C is correct. Bollinger Bands are price-based indicators, *not* momentum oscillators, which are constructed so that they oscillate between a high and a low or around 0 or 100.

35 A is correct. Triangles are one of several continuation patterns.

36 C is correct. It is one of several reversal patterns.

37 A is correct. Volume is necessary to confirm the various market rallies and reversals during the formation of the head and shoulders pattern.

38 C is correct. Intermarket analysis is a field within technical analysis that combines analysis of major categories of securities, industries, benchmarks, and geographies to identify intermarket relationships, trends, and possible inflections in a trend. To identify these intermarket relationships and trends, a commonly used tool is relative strength analysis, which charts the ratio of the price of two assets.

39 C is correct. Securities that break out from low-volatility conditions are likely to experience high volatility, and high volatility in the direction of an existing trend will usually help the trend pick up momentum. Breakouts from consolidation periods in an uptrend or entering an uptrend are usually followed by a continuation of the existing trend.

READING

55

Fintech in Investment Management

by Robert Kissell, PhD, and Barbara J. Mack

Robert Kissell, PhD, is at Molloy College and Kissell Research Group (USA). Barbara J. Mack is at Pingry Hill Enterprises, Inc. (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe "fintech;"
☐	**b.** describe Big Data, artificial intelligence, and machine learning;
☐	**c.** describe fintech applications to investment management;
☐	**d.** describe financial applications of distributed ledger technology.

INTRODUCTION AND WHAT IS FINTECH

1

a describe "fintech;"

The meeting of finance and technology, commonly known as *fintech*, is changing the landscape of investment management. Advancements include the use of Big Data, artificial intelligence, and machine learning to evaluate investment opportunities, optimize portfolios, and mitigate risks. These developments are affecting not only quantitative asset managers but also fundamental asset managers who make use of these tools and technologies to engage in hybrid forms of investment decision making.

Investment advisory services are undergoing changes with the growth of automated wealth advisers or "robo-advisers." Robo-advisers may assist investors without the intervention of a human adviser, or they may be used in combination with a human adviser. The desired outcome is the ability to provide tailored, actionable advice to investors with greater ease of access and at lower cost.

In the area of financial record keeping, blockchain and distributed ledger technology (DLT) are creating new ways to record, track, and store transactions for financial assets. An early example of this trend is the cryptocurrency bitcoin, but the technology is being considered in a broader set of applications.

This reading is divided into seven main sections, which together define fintech and outline some of its key areas of impact in the field of investment management. Section 1 explains the concept of and areas of fintech. Sections 2 and 3 discuss Big

© 2019 CFA Institute. All rights reserved.

Data, artificial intelligence, and machine learning. Section 4 discusses data science, and Sections 5–8 provide applications of fintech to investment management. Sections 9 and 10 examine DLT. A summary of key points completes the reading.

1.1 What Is Fintech?

In its broadest sense, the term “fintech” generally refers to technology-driven innovation occurring in the financial services industry. For the purposes of this reading, **fintech** refers to technological innovation in the design and delivery of financial services and products. Note, however, that in common usage, fintech can also refer to companies (often new, startup companies) involved in developing the new technologies and their applications, as well as the business sector that comprises such companies. Many of these innovations are challenging the traditional business models of incumbent financial services providers.

Early forms of fintech included data processing and the automation of routine tasks. Then followed systems that provided execution of decisions according to specified rules and instructions. Fintech has since advanced into decision-making applications based on complex machine-learning logic, where computer programs are able to “learn” how to complete tasks over time. In some applications, advanced computer systems are performing tasks at levels far surpassing human capabilities. Fintech has changed the financial services industry in many ways, giving rise to new systems for investment advice, financial planning, business lending, and payments.

Whereas fintech covers a broad range of services and applications, areas of fintech development that are more directly relevant to the investment industry include the following:

- **Analysis of large datasets.** In addition to growing amounts of traditional data, such as security prices, corporate financial statements, and economic indicators, massive amounts of alternative data generated from non-traditional data sources, such as social media and sensor networks, can now be integrated into a portfolio manager’s investment decision-making process and used to help generate alpha and reduce losses.
- **Analytical tools.** For extremely large datasets, techniques involving **artificial intelligence** (AI)—computer systems capable of performing tasks that previously required human intelligence—may be better suited to identify complex, non-linear relationships than traditional quantitative methods and statistical analysis. Advances in AI-based techniques are enabling different data analysis approaches. For example, analysts are turning to artificial intelligence to sort through the enormous amounts of data from company filings, annual reports, and earnings calls to determine which data are most important and to help uncover trends and generate insights relating to human sentiment and behavior.
- **Automated trading.** Executing investment decisions through computer algorithms or automated trading applications may provide a number of benefits to investors, including more efficient trading, lower transaction costs, anonymity, and greater access to market liquidity.
- **Automated advice. Robo-advisers** or automated personal wealth management services provide investment services to a larger number of retail investors at lower cost than traditional adviser models can provide.
- **Financial record keeping.** New technology, such as DLT, may provide secure ways to track ownership of financial assets on a peer-to-peer (P2P) basis. By allowing P2P interactions—in which individuals or firms transact directly with each other without mediation by a third party—DLT reduces the need for financial intermediaries.

Drivers underlying fintech development in these areas include extremely rapid growth in data—including their quantity, types, sources, and quality—and technological advances that enable the capture and extraction of information from them. The data explosion is addressed in Section 2, and selected technological advances and data science are addressed in Sections 3 and 4, respectively.

BIG DATA

2

b describe Big Data, artificial intelligence, and machine learning;

As noted, datasets are growing rapidly in terms of the size and diversity of data types that are available for analysis. The term **Big Data** has been in use since the late 1990s and refers to the vast amount of data being generated by industry, governments, individuals, and electronic devices. Big Data includes data generated from traditional sources—such as stock exchanges, companies, and governments—as well as non-traditional data types, also known as **alternative data**, arising from the use of electronic devices, social media, sensor networks, and company exhaust (data generated in the normal course of doing business).

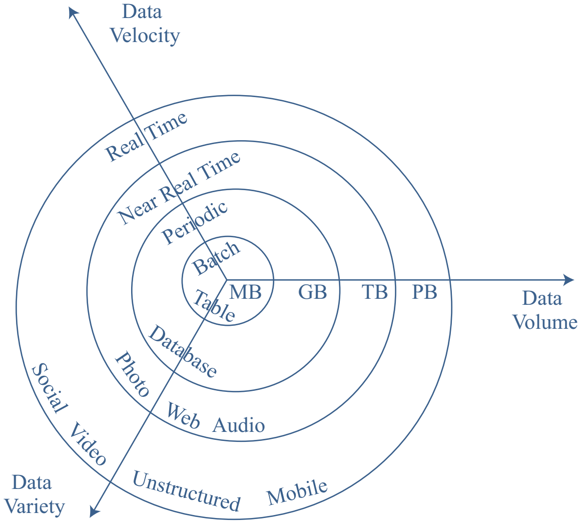
Traditional data sources include corporate data in the form of annual reports, regulatory filings, sales and earnings figures, and conference calls with analysts. Traditional data also include data that are generated in the financial markets, including trade prices and volumes. Because the world has become increasingly connected, we can now obtain data from a wide range of devices, including smart phones, cameras, microphones, radio-frequency identification (RFID) readers, wireless sensors, and satellites that are now in use all over the world. As the internet and the presence of such networked devices have grown, the use of non-traditional data sources, or alternative data sources—including social media (posts, tweets, and blogs), email and text communications, web traffic, online news sites, and other electronic information sources—has risen.

The term *Big Data* typically refers to datasets having the following characteristics:

- **Volume**: The amount of data collected in files, records, and tables is very large, representing many millions, or even billions, of data points.
- **Velocity**: The speed with which the data are communicated is extremely great. Real-time or near-real-time data have become the norm in many areas.
- **Variety**: The data are collected from many different sources and in a variety of formats, including structured data (e.g., SQL tables or CSV files), semi-structured data (e.g., HTML code), and unstructured data (e.g., video messages).

Features relating to Big Data's volume, velocity, and variety are shown in Exhibit 1.

Exhibit 1 Big Data Characteristics: Volume, Velocity, and Variety



Data	Volume Key	Bytes of Information
MB	Megabyte	One Million
GB	Gigabyte	One Billion
TB	Terabyte	One Trillion
PB	Petabyte	One Quadrillion

Source: http://whatis.techtarget.com/definition/3Vs.

Exhibit 1 shows that data volumes are growing from megabytes (MB) and gigabytes (GB) to far larger sizes, such as terabytes (TB) and petabytes (PB), as more data are being generated, captured, and stored. At the same time, more data, traditional and non-traditional, are available on a real-time or near-real-time basis with far greater variety in data types than ever before.

When using big data for inference or prediction, there is a "fourth V": Veracity relates to the credibility and reliability of different data sources. Determining the credibility and reliability of data sources is an important part of any empirical investigation. The issue of veracity becomes critically important for big data, however, because of the varied sources of these large datasets. Big data amplifies the age-old challenge of disentangling quality from quantity.

Big Data may be structured, semi-structured, or unstructured data. Structured data items can be organized in tables and are commonly stored in a database where each field represents the same type of information. Unstructured data may be disparate, unorganized data that cannot be represented in tabular form. Unstructured data, such as those generated by social media, email, text messages, voice recordings, pictures, blogs, scanners, and sensors, often require different, specialized applications or custom programs before they can be useful to investment professionals. For example, in order to analyze data contained in emails or texts, specially developed or customized computer code may be required to first process these files. Semi-structured data may have attributes of both structured and unstructured data.

2.1 Sources of Big Data

Big Data, therefore, encompasses data generated by:

- financial markets (e.g., equity, fixed income, futures, options, and other derivatives),
- businesses (e.g., corporate financials, commercial transactions, and credit card purchases),
- governments (e.g., trade, economic, employment, and payroll data),
- individuals (e.g., credit card purchases, product reviews, internet search logs, and social media posts),
- sensors (e.g., satellite imagery, shipping cargo information, and traffic patterns), and, in particular,
- the Internet of Things, or IoT (e.g., data generated by "smart" buildings, where the building is providing a steady stream of information about climate control, energy consumption, security, and other operational details).

In gathering business intelligence, historically, analysts have tended to draw on traditional data sources, employing statistical methods to measure performance, predict future growth, and analyze sector and market trends. In contrast, the analysis of Big Data incorporates the use of alternative data sources.

From retail sales data to social media sentiment to satellite imagery that may reveal information about agriculture, shipping, and oil rigs, alternative datasets may provide additional insights about consumer behavior, firm performance, trends, and other factors important for investment-related activities. Such information is having a significant effect on the way that professional investors, particularly quantitative investors, approach financial analysis and decision-making processes.

There are three main sources of alternative data:

- data generated by individuals,
- data generated by business processes, and
- data generated by sensors.

Data generated by individuals are often produced in text, video, photo, and audio formats and may also be generated through such means as website clicks or time spent on a webpage. This type of data tends to be unstructured. The volume of this type of data is growing dramatically as people participate in greater numbers and more frequently in online activities, such as social media and e-commerce, including online reviews of products, services, and entire companies, and as they make personal data available through web searches, email, and other electronic trails.

Business process data include information flows from corporations and other public entities. These data tend to be structured data and include direct sales information, such as credit card data, as well as corporate exhaust. Corporate exhaust includes corporate supply chain information, banking records, and retail point-of-sale scanner data. Business process data can be leading or real-time indicators of business performance, whereas traditional corporate metrics may be reported only on a quarterly or even yearly basis and are typically lagging indicators of performance.

Sensor data are collected from such devices as smart phones, cameras, RFID chips, and satellites that are usually connected to computers via wireless networks. Sensor data can be unstructured, and the volume of data is many orders of magnitude greater than that of individual or business process datastreams. This form of data is growing exponentially because microprocessors and networking technology are increasingly present in a wide array of personal and commercial electronic devices. Extended to office buildings, homes, vehicles, and many other physical forms, this culminates in a network arrangement, known as the **Internet of Things**, that is formed by the vast array of physical devices, home appliances, smart buildings, vehicles, and other items that are embedded with electronics, sensors, software, and network connections that enable the objects in the system to interact and share information.

Exhibit 2 shows a classification of alternative data sources and includes examples for each.

Exhibit 2 Classification of Alternative Data Sources

Individuals	Business Processes	Sensors
Social media	Transaction data	Satellites
News, reviews	Corporate data	Geolocation
Web searches, personal data		Internet of Things
		Other sensors

In the search to identify new factors that may affect security prices, enhance asset selection, improve trade execution, and uncover trends, alternative data are being used to support data-driven investment models and decisions. As interest in alternative data has risen, there has been a growth in the number of specialized firms that collect, aggregate, and sell alternative datasets.

While the marketplace for alternative data is expanding, investment professionals should understand potential legal and ethical issues related to information that is not in the public domain. For example, the scraping of web data could potentially capture personal information that is protected by regulations or that may have been published or provided without the explicit knowledge and consent of the individuals involved. Best practices are still in development in many jurisdictions, and because of varying approaches taken by national regulators, there may be conflicting forms of guidance.

2.2 Big Data Challenges

Big Data poses several challenges when it is used in investment analysis, including the quality, volume, and appropriateness of the data. Key issues revolve around the following questions, among others: Does the dataset have selection bias, missing data, or data outliers? Is the volume of collected data sufficient? Is the dataset well suited for the type of analysis? In most instances, the data must be sourced, cleansed, and organized before analysis can occur. This process can be extremely difficult with alternative data owing to the unstructured characteristics of the data involved, which are more often qualitative (e.g., texts, photos, and videos) than quantitative in nature.

Given the size and complexity of alternative datasets, traditional analytical methods cannot always be used to interpret and evaluate these datasets. To address this challenge, artificial intelligence and machine learning techniques have emerged that support work on such large and complex sources of information.

3

ADVANCED ANALYTICAL TOOLS: ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

b describe Big Data, artificial intelligence, and machine learning;

Artificial intelligence computer systems are capable of performing tasks that have traditionally required human intelligence. AI technology has enabled the development of computer systems that exhibit cognitive and decision-making ability comparable or superior to that of human beings.

An early example of AI was the “expert system,” a type of computer programming that attempted to simulate the knowledge base and analytical abilities of human experts in specific problem-solving contexts. This was often accomplished through the use of “if-then” rules. By the late 1990s, faster networks and more powerful processors enabled AI to be deployed in logistics, data mining, financial analysis, medical diagnosis, and other areas. Since the 1980s, financial institutions have made use of AI—particularly, **neural networks**, programming based on how our brain learns and processes information—to detect abnormal charges or claims in credit card fraud detection systems.

Machine learning (ML) involves computer-based techniques that seek to extract knowledge from large amounts of data without making any assumptions on the data’s underlying probability distribution. The goal of ML algorithms is to automate decision-making processes by generalizing, or “learning,” from known examples to determine an underlying structure in the data. The emphasis is on the ability of the algorithm to generate structure or predictions without any help from a human. Simply put, ML algorithms aim to “find the pattern, apply the pattern.”

As it is currently used in the investing context, ML requires massive amounts of data for “training,” so although some ML techniques have existed for years, insufficient data have historically limited broader application. Previously, these algorithms lacked access to the large amounts of data needed to model relationships successfully. The growth in Big Data has provided ML algorithms, such as neural networks, with sufficient data to improve modeling and predictive accuracy, and greater use of ML techniques is now possible.

In ML, the computer algorithm is given “inputs” (a set of variables or datasets) and may be given “outputs” (the target data). The algorithm “learns” from the data provided how best to model inputs to outputs (if provided) or how to identify or describe underlying data structure if no outputs are given. Training occurs as the algorithm identifies relationships in the data and uses that information to refine its learning process.

ML involves splitting the dataset into three distinct subsets: a training dataset, a validation dataset, and a test dataset. The training dataset allows the algorithm to identify relationships between inputs and outputs based on historical patterns in the data. These relationships are then validated, and the model tuned, using the validation dataset. The test dataset is used to test the model’s ability to predict well on new data. Once an algorithm has been trained, validated, and tested, the ML model can be used to predict outcomes based on other datasets.

ML still requires human judgement in understanding the underlying data and selecting the appropriate techniques for data analysis. Before they can be used, the data must be clean and free of biases and spurious data. As noted, ML models also require sufficiently large amounts of data and may not perform well where there may not be enough available data to train and validate the model.

Analysts must also be cognizant of errors that may arise from **overfitting** the data, because models that overfit the data may discover “false” relationships or “unsubstantiated” patterns that will lead to prediction errors and incorrect output forecasts. Overfitting occurs when the ML model learns the input and target dataset too precisely. In such cases, the model has been “over-trained” on the data and treats noise in the data as true parameters. An ML model that has been overfitted is not able to accurately predict outcomes using a different dataset and may be too complex. When a model has been underfitted, the ML model treats true parameters as if they are noise and is not able to recognize relationships within the training data. In such cases, the model may be too simplistic. Underfitted models will typically fail to fully discover patterns that underlie the data.

In addition, since they are not explicitly programmed, ML techniques can appear to be opaque or "black box" approaches, which arrive at outcomes that may not be entirely understood or explainable.

3.1 Types of Machine Learning

ML approaches can help identify relationships between variables, detect patterns or trends, and create structure from data, including data classification. Machine learning can be broadly divided into three distinct classes of techniques: supervised learning, unsupervised learning, and deep learning.

In **supervised learning**, computers learn to model relationships based on labeled training data. In supervised learning, inputs and outputs are labeled, or identified, for the algorithm. After learning how best to model relationships for the labeled data, the trained algorithms are used to model or predict outcomes for new datasets. Trying to identify the best signal, or variable, to forecast future returns on a stock or trying to predict whether local stock market performance will be up, down, or flat during the next business day are problems that may be approached using supervised learning techniques.

In **unsupervised learning**, computers are not given labeled data but instead are given only data from which the algorithm seeks to describe the data and their structure. Trying to group companies into peer groups based on their characteristics rather than using standard sector or country groupings is a problem that may be approached using unsupervised learning techniques.

Underlying AI advances have been key developments relating to neural networks. In **deep learning**, (or **deep learning nets**), computers use neural networks, often with many hidden layers, to perform multistage, non-linear data processing to identify patterns. Deep learning may use supervised or unsupervised machine learning approaches. By taking a layered or multistage approach to data analysis, deep learning develops an understanding of simple concepts that informs analysis of more complex concepts. Neural networks have existed since 1958 and have been used for many applications, such as forecasting and pattern recognition, since the early 1990s. Improvements in the algorithms underlying neural networks are providing more accurate models that better incorporate and learn from data. As a result, these algorithms are now far better at such activities as image, pattern, and speech recognition. In many cases, the advanced algorithms require less computing power than the earlier neural networks, and their improved solution enables analysts to discover insights and identify relationships that were previously too difficult or too time consuming to uncover.

Advances in Artificial Intelligence outside Finance

Non-finance-related AI breakthroughs include victories in the general knowledge game-show Jeopardy (by IBM's Watson in 2011) and in the ancient Chinese board game Go (by Google's DeepMind in 2016). Not only is AI providing solutions where there is perfect information (all players have equal access to the same information), such as checkers, chess, and Go, but AI is also providing insight in cases where information may be imperfect and players have hidden information; AI successes at the game of poker (by DeepStack) are an example. AI has also been behind the rise of virtual assistants, such as Siri (from Apple), Google's Translate app, and Amazon's product recommendation engine.

The ability to analyze Big Data using ML techniques, alongside more traditional statistical methods, represents a significant development in investment research, supported by the presence of greater data availability and advances in the algorithms themselves. Improvements in computing power and software processing speeds and falling storage costs have further supported this evolution.

ML techniques are being used for Big Data analysis to help predict trends or market events, such as the likelihood of a successful merger or an outcome to a political election. Image recognition algorithms can now analyze data from satellite-imaging systems to provide intelligence on the number of consumers in retail store parking lots, shipping activity and manufacturing facilities, and yields on agricultural crops, to name just a few examples.

Such information may provide insights into individual firms or at national or global levels and may be used as inputs into valuation or economic models.

DATA SCIENCE: EXTRACTING INFORMATION FROM BIG DATA

4

c describe fintech applications to investment management;

Data science can be defined as an interdisciplinary field that harnesses advances in computer science (including machine learning), statistics, and other disciplines for the purpose of extracting information from Big Data (or data in general). Companies rely on the expertise of data scientists/analysts to extract information and insights from Big Data for a wide variety of business and investment purposes.

An important consideration for the data scientist is the structure of the data. As noted in the discussion on Big Data, because of their unstructured nature, alternative data often require specialized treatment before they can be used for analysis.

4.1 Data Processing Methods

To help determine the best data management technique needed for Big Data analysis, data scientists use various data processing methods, including capture, curation, storage, search, and transfer.

- Capture—Data capture refers to how the data are collected and transformed into a format that can be used by the analytical process. Low-latency systems—systems that operate on networks that communicate high volumes of data with minimal delay (latency)—are essential for automated trading applications that make decisions based on real-time prices and market events. In contrast, high-latency systems do not require access to real-time data and calculations.
- Curation—Data curation refers to the process of ensuring data quality and accuracy through a data cleaning exercise. This process consists of reviewing all data to detect and uncover data errors—bad or inaccurate data—and making adjustments for missing data when appropriate.
- Storage—Data storage refers to how the data will be recorded, archived, and accessed and the underlying database design. An important consideration for data storage is whether the data are structured or unstructured and whether analytical needs require low-latency solutions.

- Search—Search refers to how to query data. Big Data has created the need for advanced applications capable of examining and reviewing large quantities of data to locate requested data content.
- Transfer—Transfer refers to how the data will move from the underlying data source or storage location to the underlying analytical tool. This could be through a direct data feed, such as a stock exchange's price feed.

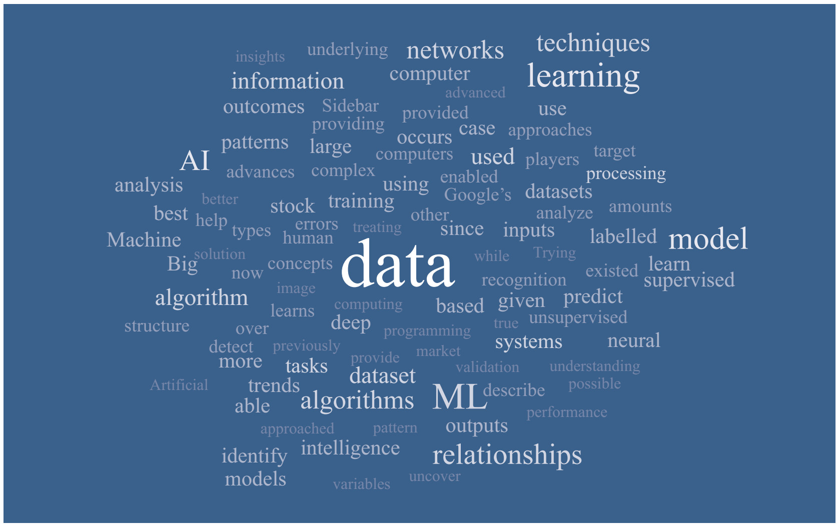
4.2 Data Visualization

Data visualization is an important tool for understanding Big Data. Visualization refers to how the data will be formatted, displayed, and summarized in graphical form. Traditional structured data can be visualized using tables, charts, and trends, whereas non-traditional unstructured data require new techniques of data visualization. These visualization tools include, for example, interactive three-dimensional (3D) graphics, where users can focus in on specified data ranges and rotate the data across 3D axes to help identify trends and uncover relationships. Multidimensional data analysis consisting of more than three variables requires additional data visualization techniques—for example, adding color, shapes, and sizes to the 3D charts. Further, a wide variety of solutions exists to reflect the structure of the data through the geometry of the visualization, with interactive graphics allowing for especially rich possibilities. Examples include heat maps, tree diagrams, and network graphs.

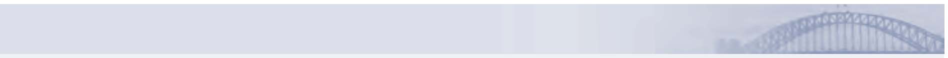
Another valuable Big Data visualization technique that is applicable to textual data is a "tag cloud," where words are sized and displayed on the basis of the frequency of the word in the data file. For example, words that appear more often are shown with a larger font, and words that appear less often are shown with a smaller font. A "mind map" is another data visualization technique; it is a variation of the tag cloud, but rather than displaying the frequency of words, a mind map shows how different concepts are related to each other.

Exhibit 3 shows an example of a "tag cloud" based on a section of this reading. The more frequently a word is found within the text, the larger it becomes in the tag cloud. As shown in the tag cloud, the words appearing most frequently in the section include "data," "ML," "learning," "AI," "techniques," "model," and "relationships."

Exhibit 3 Data Visualization Tag Cloud: Section 3, Advanced Analytical Tools


insights
underlying
networks
techniques
information
computer
learning
advanced
outcomes
Sidebar
provided
use
providing
occurs
case
approaches
patterns
large
computers
used
players
target
AI
advances
complex
processing
enabled
analysis
using
Google's
datasets
better
stock
training
analyze
amounts
best
help
other
errors
since
inputs
labelled
model
types
treating
Machine
human
data
while
Trying
solution
Big
now
concepts
existed
learn
recognition
supervised
image
algorithm
computing
based
given
predict
learns
deep
true
unsupervised
structure
over
programming
systems
neural
detect
previously
provide
market
more
tasks
validation
understanding
Artificial
dataset
possible
trends
algorithms
ML
describe
able
performance
approached
pattern
outputs
identify
intelligence
relationships
models
uncover
variables

Source: https://worditout.com/word-cloud/create.



Programming Languages and Databases

Some of the more common programming languages used in data science include the following:

- **Python:** Python is an open source, free programming language that does not require an in-depth understanding of computer programming. Python allows individuals with little or no programming experience to develop computer applications for advanced analytical use and is the basis for many fintech applications.
- **R:** R is an open source, free programming language traditionally used for statistical analysis. R has mathematical packages for statistical analysis, machine learning, optimization, econometrics, and financial analysis.
- **Java:** Java is a programming language that can run on different computers, servers, and operating systems. Java is the underlying program language used in many internet applications.
- **C/C++:** C/C++ is a specialized programming language that provides the ability to optimize source code to achieve superior calculation speed and processing performance. C/C++ is used in applications for algorithmic and high-frequency trading.
- **Excel VBA:** Excel VBA helps bridge the gap between programming and manual data processing by allowing users to run macros to automate tasks, such as updating data tables and formulas, running data queries and collecting data from different web locations, and performing calculations. Excel VBA allows users to develop customized reports and analyses that rely on data that are updated from different applications and databases.

Some of the more common databases in use include the following:

- **SQL:** SQL is a database for structured data where the data can be stored in tables with rows and columns. SQL databases need to be run on a server that is accessed by users.

- **SQLite:** SQLite is a database for structured data. SQLite databases are embedded into the program and do not need to be run on a server. It is the most common database for mobile apps that require access to data.
- **NoSQL:** NoSQL is a database used for unstructured data where the data cannot be summarized in traditional tables with rows and columns.

5 SELECTED APPLICATIONS OF FINTECH TO INVESTMENT MANAGEMENT; TEXT ANALYTICS & NATURAL LANGUAGE PROCESSING

c describe fintech applications to investment management;

Fintech is being used in numerous areas of investment management. Applications for investment management include text analytics and natural language processing, robo-advisory services, risk analysis, and algorithmic trading.

5.1 Text Analytics and Natural Language Processing

Text analytics involves the use of computer programs to analyze and derive meaning typically from large, unstructured text- or voice-based datasets, such as company filings, written reports, quarterly earnings calls, social media, email, internet postings, and surveys. Text analytics includes using computer programs to perform automated information retrieval from different, unrelated sources in order to aid the decision-making process. More analytical usage includes lexical analysis, or the analysis of word frequency in a document and pattern recognition based on key words and phrases. Text analytics may be used in predictive analysis to help identify indicators of future performance, such as consumer sentiment.

Natural language processing (NLP) is a field of research at the intersection of computer science, artificial intelligence, and linguistics that focuses on developing computer programs to analyze and interpret human language. Within the larger field of text analytics, NLP is an important application. Automated tasks using NLP include translation, speech recognition, text mining, sentiment analysis, and topic analysis. NLP may also be employed in compliance functions to review employee voice and electronic communications for adherence to company or regulatory policy, inappropriate conduct, or fraud or for ensuring private or customer information is kept confidential.

Consider that all the public corporations worldwide generate millions of pages of annual reports and tens of thousands of hours of earnings calls each year. This is more information than any individual analyst or team of researchers can assess. NLP, especially when aided by ML algorithms, can analyze annual reports, call transcripts, news articles, social media posts, and other text- and audio-based data to identify trends in shorter timespans and with greater scale and accuracy than is humanly possible.

For example, NLP may be used to monitor analyst commentary to aid investment decision making. Financial analysts may generate earnings-per-share (EPS) forecasts reflecting their views on a company's near-term prospects. Focusing on forecasted EPS numbers could mean investors miss subtleties contained in an analyst's written research report. Since analysts tend not to change their buy, hold, and sell recommendations for a company frequently, they may instead offer nuanced commentary without making a change in their investment recommendation. After analyzing analyst commentary, NLP can assign sentiment ratings ranging from very negative to very positive for each. NLP can, therefore, be used to detect, monitor, and tag shifts in sentiment, potentially ahead of an analyst's recommendation change. Machine capabilities enable this analysis to scale across thousands of companies worldwide, performing work previously done by humans.

Similarly, communications and transcripts from policymakers, such as the European Central Bank or the US Federal Reserve, offer an opportunity for NLP-based analysis, because officials at these institutions may send subtle messages through their choice of topics, words, and inferred tone. NLP can help analyze nuances within text to provide insights around trending or waning topics of interest, such as interest rate policy, aggregate output, or inflation expectations.

Models using NLP analysis may incorporate non-traditional information to evaluate what people are saying—via their preferences, opinions, likes, or dislikes—in an attempt to identify trends and short-term indicators about a company, a stock, or an economic event that might have a bearing on future performance. Past research has evaluated the predictive power of Twitter sentiment regarding IPO performance, for example.[1] The effect of positive and negative news sentiment on stock returns has also been researched.[2]

ROBO-ADVISORY SERVICES

6

c describe fintech applications to investment management;

Since their emergence in 2008, a number of startup firms, as well as large asset managers, have introduced robo-advisory services, which provide investment solutions through online platforms, reducing the need for direct interaction with financial advisers.

As robo-advisers have been incorporated into the investment landscape, they have drawn the attention of regulatory authorities. In the United States, robo-advisers must be established as registered investment advisers, and they are regulated by the Securities and Exchange Commission. In the United Kingdom, they are regulated by the Financial Conduct Authority. In Australia, all financial advisers must obtain an Australian Financial Services license, with guidance on digital advisers coming from the Australian Securities and Investments Commission. Robo-advisers are also on the rise in parts of Asia and the rest of the world. Although regulatory conditions vary, robo-advisers are likely to be held to a similar level of scrutiny and code of conduct as other investment professionals in the given region.

Robo-advice tends to start with an investor questionnaire, which may include many of the categories and subcategories shown in Exhibit 4. Exhibit 4 is a synthesis of questionnaires created by the researchers attributed in the source below. Once assets, liabilities, risk preferences, and target investment returns have been digitally entered by a client, the robo-adviser software produces recommendations, based on algorithmic rules and historical market data, that incorporate the client's stated investment parameters. According to research by Michael Tertilt and Peter Scholz, robo-advisers do not seem to incorporate the full range of available information into their recommendations;[3] further research will be necessary over time to see how this may affect performance and the evolution of digital advisory services. Nevertheless, current robo-advisory services include automated asset allocation, trade execution, portfolio optimization, tax-loss harvesting, and rebalancing for investor portfolios.

---

[1] Jim Kyung-Soo Liew and Garrett Zhengyuan Wang, "Twitter Sentiment and IPO Performance: A Cross-Sectional Examination," *Journal of Portfolio Management*, vol. 42, no. 4 (Summer 2016): 129–135.

[2] Steven L. Heston and Nitish Ranjan Sinha, "News vs. Sentiment: Predicting Stock Returns from News Stories," *Financial Analysts Journal*, vol. 73, no. 3 (Third Quarter 2017): 67–83. (https://www.tandfonline.com/doi/abs/10.2469/faj.v73.n3.3).

Exhibit 4 Categories and Subcategories for Investor Questionnaires

General Information	Risk Tolerance
Income	Age
Investment Amount	Association with Investing
Job Description	Association with Risk
Other	Choose Portfolio Risk Level
Source of Income	Comfort Investing in Stock
Spendings	Credit Based Investments
Time to Retirement	Dealing with Financial Decisions
Type of Account	Degree of Financial Risk Taken
Working Status	Education
Risk Capacity	Ever Interested in Risky Asset for Thrill
Dependence on Withdrawal of Investment Amount	Experience of Drop/Reaction on Drop/Max Drop before Selling
Income Prediction	Family and Household Status
Investment Amount/Savings Rate Ratio	Financial Knowledge
Investment Amount/Total Capital Ratio	Gender
Investment Horizon	Investment Experience
Liabilities	Investment Goal
Savings Rate	Investor Type/Self-Assessment Risk Tolerance
Total Capital	Preference Return vs. Risk

Source: Michael Tertilt and Peter Scholz, 2017 “To Advise, or Not to Advise—How Robo-Advisors Evaluate the Risk Preferences of Private Investors,” working paper (13 June): Table 1: Categories and Subcategories for Questionnaires.

Although their analyses and recommendations can cover both active and passive management styles, most robo-advisers follow a passive investment approach. These robo-advisers typically have low fees and low account minimums, implementing their recommendations with low-cost, diversified index mutual funds or exchange-traded funds (ETFs). A diverse range of asset classes can be managed in this manner, including stocks, bonds, commodities, futures, and real estate. Because of their low-cost structure, robo-advisers can reach underserved populations, such as the mass affluent or mass market segments, which are less able to afford a traditional financial adviser.

Two types of wealth management services dominate the robo-advice sector: fully automated digital wealth managers and adviser-assisted digital wealth managers.

■ Fully Automated Digital Wealth Managers

---

[3] Michael Tertilt and Peter Scholz, To Advise, or Not to Advise — How Robo-Advisors Evaluate the Risk Preferences of Private Investors (June 12, 2017). Available at SSRN: https://ssrn.com/abstract=2913178 or http://dx.doi.org/10.2139/ssrn.2913178

The fully automated model does not rely on assistance from a human financial adviser. These services seek to offer a low-cost solution to investing and recommend an investment portfolio, which is often composed of ETFs. The service package may include direct deposits, periodic rebalancing, and dividend reinvestment options.

- **Adviser-Assisted Digital Wealth Managers**

  Adviser-assisted digital wealth managers provide automated investment services along with a virtual financial adviser, who is available to offer basic financial planning advice and periodic reviews by phone. Adviser-assisted digital wealth managers are capable of providing additional services that may involve a more holistic analysis of a client's assets and liabilities.

Wealthy and ultra-wealthy individuals typically have had access to human advisory teams, but there has been a gap in the availability and quality of advisers to serve investors with less wealth. The advent of robo-advisers offers a cost-effective and easily accessible form of financial guidance. In following a typically passive investment approach, research suggests that robo-advisers tend to offer fairly conservative advice.

However, critics of robo-advisers have wondered what would happen in a time of crisis, when people most often look to human expertise for guidance. It may not always be completely transparent why a robo-adviser chooses to make a recommendation or take a trading action that it did, unlike a human adviser who can provide his or her rationale. And finally, there may be trust issues in allowing computers to make these decisions, including worries of instances where robo-advisers might recommend inappropriate investments.

As the complexity and size of an investor's portfolio grows, robo-advisers may not be able to sufficiently address the particular preferences and needs of the investor. In the case of extremely affluent investors who may own a greater number of asset types—including alternative investments (e.g., venture capital, private equity, hedge funds, and real estate)—in addition to global stocks and bonds and have greater demands for customization, the need for a team of human advisers, each with particular areas of investment or wealth-management expertise, is likely to endure.

RISK ANALYSIS

7

c describe fintech applications to investment management;

As mandated by regulators worldwide, the global investment industry has undertaken major steps in stress testing and risk assessment that involve the analysis of vast amounts of quantitative and qualitative risk data. Required data include information on the liquidity of the firm and its trading partners, balance sheet positions, credit exposures, risk-weighted assets, and risk parameters. Stress tests may also take qualitative information into consideration, such as capital planning procedures, expected business plan changes, business model sustainability, and operational risk.

There is increasing interest in monitoring risk in real time. To do so, relevant data must be taken by a firm, mapped to known risks, and identified as it moves within the firm. Data may be aggregated for reporting purposes or used as inputs to risk models. Big Data may provide insights into real-time and changing market circumstances to help identify weakening market conditions and adverse trends in advance, allowing managers to employ risk management techniques and hedging practices sooner to help preserve asset value. For example, evaluation of alternative data using ML techniques may help foreshadow declining company earnings and future stock performance. Furthermore, analysis of real-time market data and trading patterns may help analysts detect buying or selling pressure in the stock.

ML techniques may be used to help assess data quality. To help ensure accurate and reliable data that may originate from numerous alternative data sources, ML techniques can help validate data quality by identifying questionable data, potential errors, and data outliers before integration with traditional data for use in risk models and in risk management applications.

Portfolio risk management often makes use of scenario analysis—analyzing the likely performance of the portfolio and liquidation costs under a hypothetical stress scenario or the repeat of a historical stress event. For example, to understand the implications of holding or liquidating positions during adverse or extreme market periods, such as the financial crisis, fund managers may perform "what-if" scenario analysis and portfolio backtesting using point-in-time data to understand liquidation costs and portfolio consequences under differing market conditions. These backtesting simulations are often computationally intense and may be facilitated through the use of advanced AI-based techniques.

8

ALGORITHMIC TRADING

c describe fintech applications to investment management;

Algorithmic trading is the computerized buying and selling of financial instruments, in accordance with pre-specified rules and guidelines. Algorithmic trading is often used to execute large institutional orders, slicing orders into smaller pieces and executing across different exchanges and trading venues. Algorithmic trading provides investors with many benefits, including speed of execution, anonymity, and lower transaction costs. Over the course of a day, algorithms may continuously update and revise their execution strategy on the basis of changing prices, volumes, and market volatility. Algorithms may also determine the best way to price the order (e.g., limit or market order) and the most appropriate trading venue (e.g., exchange or dark pool) to route for execution.

High-frequency trading (HFT) is a form of algorithmic trading that makes use of vast quantities of granular financial data (tick data, for example) to automatically place trades when certain conditions are met. Trades are executed on ultra-high-speed, low-latency networks in fractions of a second. HFT algorithms decide what to buy or sell and where to execute on the basis of real-time prices and market conditions, seeking to earn a profit from intraday market mispricings.

Global financial markets have undergone substantial change as markets have fragmented into multiple trading destinations consisting of electronic exchanges, alternative trading systems, and so-called dark pools, and average trade sizes have fallen. In this environment, and with markets continuously reflecting real-time information, algorithmic trading has been viewed as an important tool.

9

DISTRIBUTED LEDGER TECHNOLOGY, AND PERMISSIONED AND PERMISSIONLESS NETWORKS

d describe financial applications of distributed ledger technology.

Distributed ledger technology—technology based on a distributed ledger (defined below)—represents a fintech development that offers potential improvements in the area of financial record keeping. DLT networks are being considered as an efficient means to create, exchange, and track ownership of financial assets on a peer-to-peer basis. Potential benefits include greater accuracy, transparency, and security in record keeping; faster transfer of ownership; and peer-to-peer interactions. However, the technology is not fully secure, and breaches in privacy and data protection are possible. In addition, the processes underlying DLT generally require massive amounts of energy to verify transaction activity.

A **distributed ledger** is a type of database that may be shared among entities in a network. In a distributed ledger, entries are recorded, stored, and distributed across a network of participants so that each participant has a matching copy of the digital database. Basic elements of a DLT network include a digital ledger, a consensus mechanism used to confirm new entries, and a participant network.

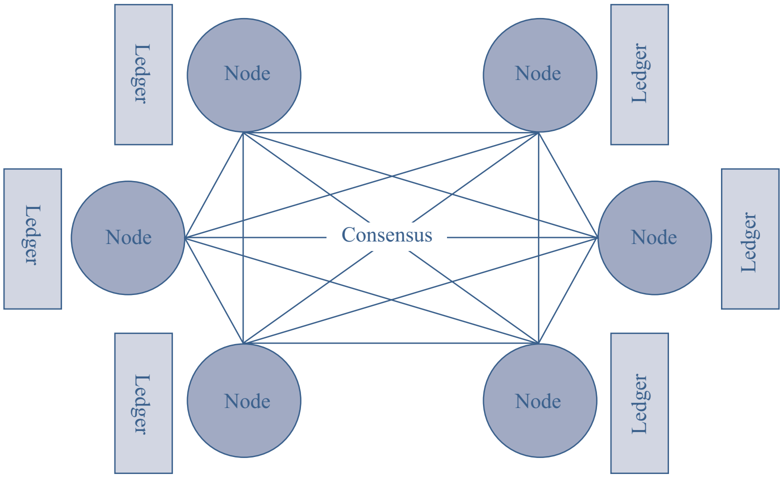
The consensus mechanism is the process by which the computer entities (or nodes) in a network agree on a common state of the ledger. Consensus generally involves two steps: transaction validation and agreement on ledger update by network parties. These features enable the creation of records that are, for the most part, considered immutable, or unchangeable, yet they are transparent and accessible to network participants on a near-real-time basis.

Features of DLT include the use of **cryptography**—an algorithmic process to encrypt data, making the data unusable if received by unauthorized parties—which enables a high level of network security and database integrity. For example, DLT uses cryptographic methods of proof to verify network participant identity and for data encryption.

DLT has the potential to accommodate "**smart contracts**," which are computer programs that self-execute on the basis of pre-specified terms and conditions agreed to by the parties to a contract. Examples of smart contract use are the automatic execution of contingent claims for derivatives and the instantaneous transfer of collateral in the event of default.

Exhibit 5 illustrates a distributed ledger network in which all nodes are connected to one another, each having a copy of the distributed ledger. The term "Consensus" is shown in the center of the network and represents the consensus mechanism in which the nodes agree on new transactions and ledger updates.

Exhibit 5 Distributed Ledger Network Setup



Source: https://blockgeeks.com/guides/what-is-hyperledger/.

Blockchain is a type of digital ledger in which information, such as changes in ownership, is recorded sequentially within blocks that are then linked or "chained" together and secured using cryptographic methods. Each block contains a grouping of transactions (or entries) and a secure link (known as a hash) to the previous block. New transactions are inserted into the chain only after validation via a consensus mechanism in which authorized members agree on the transaction and the preceding order, or history, in which previous transactions have occurred.

The consensus mechanism used to verify a transaction includes a cryptographic problem that must be solved by some computers on the network (known as miners) each time a transaction takes place. The process to update the blockchain can require substantial amounts of computing power, making it very difficult and extremely expensive for an individual third party to manipulate historical data. To manipulate historical data, an individual or entity would have to control the majority of nodes in the network. The success of the network, therefore, relies on broad network participation.

Blockchain (Distributed Ledger) Network—How Do Transactions Get Added?

Outlined below are the steps involved in adding a transaction to a blockchain distributed ledger.

1. Transaction takes place between buyer and seller
2. Transaction is broadcast to the network of computers (nodes)
3. Nodes validate the transaction details and parties to the transaction
4. Once verified, the transaction is combined with other transactions to form a new block (of predetermined size) of data for the ledger
5. This block of data is then added or linked (using a cryptographic process) to the previous block(s) containing data
6. Transaction is considered complete and ledger has been updated

9.1 Permissioned and Permissionless Networks

DLT can take the form of permissionless or permissioned networks. **Permissionless networks** are open to any user who wishes to make a transaction, and all users within the network can see all transactions that exist on the blockchain. In a permissionless, or open, DLT system, network participants can perform all network functions.

The main benefit of a permissionless network is that it does not depend on a centralized authority to confirm or deny the validity of transactions, because this takes place through the consensus mechanism. This means no single point of failure exists, since all transactions are recorded on a single distributed database and every node stores a copy of the database. Once a transaction has been added to the blockchain, it cannot be changed, barring manipulation; the distributed ledger becomes a permanent and immutable record of all previous transactions. In a permissionless network, trust is not a requirement between transacting parties.

Bitcoin is a well-known use of an open permissionless network. Using blockchain technology, Bitcoin was created in 2009 to serve as the public ledger for all transactions occurring on its virtual currency. Since the introduction of bitcoin, many more cryptocurrencies, or digital currencies, which use permissionless DLT networks, have been created.

In **permissioned networks**, network members may be restricted from participating in certain network activities. Controls, or permissions, may be used to allow varying levels of access to the ledger, from adding transactions (e.g., a participant) to viewing transactions only (e.g., a regulator) to viewing selective details of the transactions but not the full record.

APPLICATIONS OF DISTRIBUTED LEDGER TECHNOLOGY TO INVESTMENT MANAGEMENT

10

d describe financial applications of distributed ledger technology.

Potential applications of DLT to investment management include cryptocurrencies, tokenization, post-trade clearing and settlement, and compliance.

10.1 Cryptocurrencies

A **cryptocurrency**, also known as a digital currency, operates as electronic currency and allows near-real-time transactions between parties without the need for an intermediary, such as a bank. As electronic mediums of exchange, cryptocurrencies lack physical form and are issued privately by individuals, companies, and other organizations. Most issued cryptocurrencies utilize open DLT systems in which a decentralized distributed ledger is used to record and verify all digital currency transactions. Cryptocurrencies have not traditionally been government backed or regulated. Central banks around the world, however, are recognizing potential benefits and examining use cases for their own cryptocurrency versions.

Many cryptocurrencies have a self-imposed limit on the total amount of currency they may issue. Although such limits could help maintain their store of value, it is important to note that many cryptocurrencies have experienced high levels of price volatility. A lack of clear fundamentals underlying these currencies has contributed to their volatility.

Cryptocurrencies have proven to be an attractive means for companies looking to raise capital. An **initial coin offering** (ICO) is an unregulated process whereby companies sell their crypto tokens to investors in exchange for fiat money or for another agreed upon cryptocurrency. An ICO is typically structured to issue digital tokens to investors that can be used to purchase future products or services being developed by the issuer. ICOs provide an alternative to traditional, regulated capital-raising processes, such as initial public offerings (IPOs). Compared to the regulated IPO market, ICOs may have lower associated issuance costs and shorter capital raising time frames. However, most ICOs do not typically have attached voting rights. Regulation for ICOs is under consideration in a number of jurisdictions, and there have been numerous instances of investor loss resulting from fraudulent schemes.

10.2 Tokenization

Transactions involving physical assets, such as real estate, luxury goods, and commodities, often require substantial efforts in ownership verification and examination each time a transfer in ownership takes place. Through **tokenization**, the process of representing ownership rights to physical assets on a blockchain or distributed ledger, DLT has the potential to streamline this process by creating a single, digital record of ownership with which to verify ownership title and authenticity, including all historical activity. Real estate transactions that require ownership and identify verification may be one area to benefit from tokenization, because these transactions are typically labor intensive and costly, involving decentralized, paper-based records and multiple parties.

10.3 Post-Trade Clearing and Settlement

In the financial securities markets, post-trade processes to confirm, clear, and settle transactions are often complex and labor intensive, requiring multiple interactions between counterparties and financial intermediaries. DLT has the ability to streamline existing post-trade processes by providing near-real-time trade verification, reconciliation, and settlement, thereby reducing the complexity, time, and costs associated with processing transactions. A single distributed record of ownership between network peers would eliminate the need for independent and duplicative reconciliation efforts between parties and reduce the need for third-party facilitation. A shortened settlement time frame could lessen the time exposed to counterparty risk and associated collateral requirements while increasing the potential liquidity of assets and funds. Additionally, the use of automated contracts may also help to reduce post-trade time frames, lowering exposure to counterparty credit risk and trade fails.

10.4 Compliance

Regulators worldwide have imposed more stringent reporting requirements and demand greater transparency and access to data. To meet these requirements, many firms have added staff to their post-trade and compliance groups. But these functions remain predominantly manual. To comply with regulations, firms need to maintain and process large amounts of risk-related data. DLT may allow regulators and firms to maintain near-real-time review over transactions and other compliance-related processes. Improved post-trade reconciliation and automation through DLT could lead to more accurate record keeping and create operational efficiencies for a firm's compliance and regulatory reporting processes, while providing greater transparency and auditability for external authorities and regulators.

DLT-based compliance may better support shared information, communications, and transparency within and between firms, exchanges, custodians, and regulators. Closed or permissioned networks could offer advantages in security and privacy. These platforms could store highly sensitive information in a way that is secure but easily accessible to internal and external authorities. DLT could help uncover fraudulent activity and reduce compliance costs associated with know-your-customer and anti-money-laundering regulations, which entail verifying the identity of clients and business partners.

DLT Challenges

A number of challenges exist before DLT may be successfully adopted by the investment industry. These include the following:

- There is a lack of DLT network standardization, as well as difficulty integrating with legacy systems.
- DLT processing capabilities may not be financially competitive with existing solutions.
- Increasing the scale of DLT systems requires substantial (storage) resources.
- Immutability of transactions means accidental or “canceled” trades can be undone only by submitting an equal and offsetting trade.
- DLT requires huge amounts of computer power normally associated with high electricity usage.
- Regulatory approaches may differ by jurisdiction.

SUMMARY

- The term “fintech” refers to technological innovation in the design and delivery of financial services and products.
- Areas of fintech development include the analysis of large datasets, analytical techniques, automated trading, automated advice, and financial record keeping.
- Big Data is characterized by the three Vs—volume, velocity, and variety—and includes both traditional and non-traditional (or alternative) datasets.
- Among the main sources of alternative data are data generated by individuals, business processes, and sensors.
- Artificial intelligence computer systems are capable of performing tasks that traditionally required human intelligence at levels comparable (or superior) to those of human beings.
- Machine learning seeks to extract knowledge from large amounts of data by “learning” from known examples and then generating structure or predictions. Simply put, ML algorithms aim to “find the pattern, apply the pattern.” Main types of ML include supervised learning, unsupervised learning, and deep learning.

- Natural language processing is an application of text analytics that uses insight into the structure of human language to analyze and interpret text- and voice-based data.
- Robo-advisory services are providing automated advisory services to increasing numbers of retail investors. Services include asset allocation, portfolio optimization, trade execution, rebalancing, and tax strategies.
- Big Data and ML techniques may provide insights into real-time and changing market circumstances to help identify weakening or adverse trends in advance, allowing for improved risk management and investment decision making.
- Algorithmic traders use automated trading programs to determine when, where, and how to trade an order on the basis of pre-specified rules and market conditions. Benefits include speed of executions, lower trading costs, and anonymity.
- Blockchain and distributed ledger technology (DLT) may offer a new way to store, record, and track financial assets on a secure, distributed basis. Applications include cryptocurrencies and tokenization. Additionally, DLT may bring efficiencies to post-trade and compliance processes through automation, smart contracts, and identity verification.

PRACTICE PROBLEMS

1 A correct description of fintech is that it:

A is driven by rapid growth in data and related technological advances.

B increases the need for intermediaries.

C is at its most advanced state using systems that follow specified rules and instructions.

2 A characteristic of Big Data is that:

A one of its traditional sources is business processes.

B it involves formats with diverse types of structures.

C real-time communication of it is uncommon due to vast content.

3 In the use of machine learning (ML):

A some techniques are termed “black box” due to data biases.

B human judgment is not needed because algorithms continuously learn from data.

C training data can be learned too precisely, resulting in inaccurate predictions when used with different datasets.

4 Text Analytics is appropriate for application to:

A economic trend analysis.

B large, structured datasets.

C public but not private information.

5 In providing investment services, robo-advisers are *most likely* to:

A rely on their cost effectiveness to pursue active strategies.

B offer fairly conservative advice as easily accessible guidance.

C be free from regulation when acting as fully-automated wealth managers.

6 Which of the following statements on fintech’s use of data as part of risk analysis is correct?

A Stress testing requires precise inputs and excludes qualitative data.

B Machine learning ensures that traditional and alternative data are fully segregated.

C For real-time risk monitoring, data may be aggregated for reporting and used as model inputs.

7 A factor associated with the widespread adoption of algorithmic trading is increased:

A market efficiency.

B average trade sizes.

C trading destinations.

8 A benefit of distributed ledger technology (DLT) favoring its use by the investment industry is its:

A scalability of underlying systems.

B ease of integration with existing systems.

C streamlining of current post-trade processes.

© 2018 CFA Institute. All rights reserved.

9 What is a distributed ledger technology (DLT) application suited for physical assets?

A Tokenization

B Cryptocurrencies

C Permissioned networks

SOLUTIONS

1. A is correct. Drivers of fintech include extremely rapid growth in data (including their quantity, types, sources, and quality) and technological advances enabling the capture and extraction of information from it.
2. B is correct. Big Data is collected from many different sources and is in a variety of formats, including structured data (e.g., SQL tables or CSV files), semi-structured data (e.g., HTML code), and unstructured data (e.g., video messages).
3. C is correct. Overfitting occurs when the ML model learns the input and target dataset too precisely. In this case, the model has been "over trained" on the data and is treating noise in the data as true parameters. An ML model that has been overfitted is not able to accurately predict outcomes using a different dataset and may be too complex.
4. A is correct. Through the Text Analytics application of natural language processing (NLP), models using NLP analysis may incorporate non-traditional information to evaluate what people are saying—via their preferences, opinions, likes, or dislikes—in the attempt to identify trends and short-term indicators about a company, a stock, or an economic event that might have a bearing on future performance.
5. B is correct. Research suggests that robo-advisers tend to offer fairly conservative advice, providing a cost-effective and easily accessible form of financial guidance to underserved populations, such as the mass affluent and mass market segments.
6. C is correct. There is increasing interest in monitoring risk in real-time. To do so, relevant data must be taken by a firm, mapped to known risks, and identified while moving within the firm. Data may be aggregated for reporting purposes or used as inputs to risk models.
7. C is correct. Global financial markets have undergone substantial change as markets have fragmented into multiple trading destinations consisting of electronic exchanges, alternative trading systems, and so-called dark pools. In such an environment, when markets are continuously reflecting real-time information and continuously changing conditions, algorithmic trading has been viewed as an important tool.
8. C is correct. DLT has the potential to streamline the existing, often complex and labor intensive post-trade processes in securities markets by providing close to real-time trade verification, reconciliation, and settlement, thereby reducing related complexity, time, and costs.
9. A is correct. Through tokenization—the process of representing ownership rights to physical assets on a blockchain or distributed ledger—DLT has the potential to streamline this rights process by creating a single, digital record of ownership with which to verify ownership title and authenticity, including all historical activity.

Ethical and Professional Standards

STUDY SESSION

Study Session 19	Ethical and Professional Standards

TOPIC LEVEL LEARNING OUTCOME

The candidate should be able to explain the need for high ethical standards in the investment industry and the ethical responsibilities required by the CFA Institute Code of Ethics and Standards of Professional Conduct as well as to demonstrate an understanding of the application of the Code and Standards. The candidate should also be able to demonstrate an understanding of the Global Investment Performance Standards.

Trust in the investment profession is achieved only if those practicing within the industry adhere to the highest levels of ethical conduct and behavior. The CFA Institute Code of Ethics and Standards of Professional Conduct (Code and Standards) serve as the ethical foundation for the CFA Institute self-regulatory program.

The Standards of Practice Handbook provides practical application of the Code and Standards by explaining the purpose and scope of each standard, presenting recommended procedures for compliance, and providing examples of each standard in practice.

The Global Investment Performance Standards (GIPS®) establish global standards for performance reporting by investment managers. By providing a consistent set of standards and methodology, GIPS facilitate the fair and accurate comparison of managers around the world, while minimizing the potential for ambiguous or misleading performance reporting practices.

© 2021 CFA Institute. All rights reserved.

ETHICAL AND PROFESSIONAL STANDARDS

STUDY SESSION

19

Ethical and Professional Standards

This study session introduces ethics, related challenges to ethical behavior, and the role played by ethics and professionalism in the investment industry. A framework to support ethical decision-making is provided to help guide behavior. The CFA Institute Code of Ethics and Standards of Professional Conduct (Code and Standards) are examined, with attention given to each standard and its application. The session concludes with coverage of the Global Investment Performance Standards.

READING ASSIGNMENTS

Reading 56	Ethics and Trust in the Investment Profession by Bidhan L. Parmar, PhD, Dorothy C. Kelly, CFA, Colin McLean, MBA, FIA, FSIP, Nitin Mehta, CFA, FSIP, and David B. Stevens, CIMC, CFA
Reading 57	Code of Ethics and Standards of Professional Conduct *Standards of Practice Handbook*, Eleventh Edition
Reading 58	Guidance for Standards I–VII *Standards of Practice Handbook*, Eleventh Edition
Reading 59	Introduction to the Global Investment Performance Standards (GIPS)
Reading 60	Ethics Application

© 2021 CFA Institute. All rights reserved.

READING

56

Ethics and Trust in the Investment Profession

by Bidhan L. Parmar, PhD, Dorothy C. Kelly, CFA, Colin McLean, MBA, FIA, FSIP, Nitin Mehta, CFA, FSIP, and David B. Stevens, CIMC, CFA

Bidhan L. Parmar, PhD, is at the University of Virginia (USA). Dorothy C. Kelly, CFA, is at McIntire School of Commerce, University of Virginia (USA). Colin McLean, MBA, FIA, FSIP, is at SVM Asset Management (United Kingdom). Nitin Mehta, CFA, FSIP (United Kingdom). David B. Stevens, CIMC, CFA, is at Wells Fargo Private Bank (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** explain ethics;
☐	**b.** describe the role of a code of ethics in defining a profession;
☐	**c.** describe professions and how they establish trust;
☐	**d.** describe the need for high ethical standards in investment management;
☐	**e.** explain professionalism in investment management;
☐	**f.** identify challenges to ethical behavior;
☐	**g.** compare and contrast ethical standards with legal standards;
☐	**h.** describe a framework for ethical decision making.

INTRODUCTION

1

As a candidate in the CFA Program, you are both expected and required to meet high ethical standards. This reading introduces ideas and concepts that will help you understand the importance of ethical behavior in the investment industry. You will be introduced to various types of ethical issues within the investment profession and learn about the CFA Institute Code of Ethics.

The readings covering ethics and professional standards demonstrate that ethical behavior is central to creating trust. Professional behavior is equally important. Professions help maintain trust in an industry by establishing codes and setting

© 2016 CFA Institute. All rights reserved.

standards that put a framework around ethical behavior and technical competence. Professions also set the wider goal of gaining and maintaining the trust of society as a whole. In this regard, professions have a sense of purpose that society values.

Imagine that you are employed in the research department of a large financial services firm. You and your colleagues spend your days researching, analyzing, and valuing the shares of publicly traded companies and sharing your investment recommendations with clients. You love your work and take great satisfaction in knowing that your recommendations can help the firm's investing clients make informed investment decisions that will help them meet their financial goals and improve their lives.

Several months after starting at the firm, you learn that an analyst at the firm has been terminated for writing and publishing research reports that misrepresented the fundamental risks of some companies to investors. You learn that the analyst wrote the reports with the goal of pleasing the management of the companies that were the subjects of the research reports. He hoped that these companies would hire your firm's investment banking division for its services and he would be rewarded with large bonuses for helping the firm increase its investment banking fees. Some clients bought shares based on the analyst's reports and suffered losses. They posted stories on the internet about their losses and the misleading nature of the reports. When the media investigated and published the story, the firm's reputation for investment research suffered. Investors began to question the firm's motives and the objectivity of its research recommendations. The firm's investment clients started to look elsewhere for investment advice, and company clients begin to transfer their business to firms with untarnished reputations. With business declining, management is forced to trim staff. Along with many other hard-working colleagues, you lose your job—through no fault of your own.

Imagine how you would feel in this situation. Most people would feel upset and resentful that their hard and honest work was derailed by someone else's unethical behavior. Yet, this type of scenario is not uncommon. Around the world, unsuspecting employees at such companies as SAC Capital, Stanford Financial Group, Everbright Securities, Enron, Satyam Computer Services, Arthur Andersen, and other large companies have experienced such career setbacks when someone else's actions destroyed trust in their companies and industries.

Businesses and financial markets thrive on trust—defined as a strong belief in the reliability of a person or institution. In a 2016 study on trust, investors indicated that to earn their trust, the top two attributes of an investment manager should be that it (1) has transparent and open business practices, and (2) has ethical business practices.[1] Although these attributes are valued by customers and clients in any industry, this reading will explore why they are of particular importance to the investment industry.

People may think that ethical behavior is simply about following laws, regulations, and other rules, but throughout our lives and careers we will encounter situations in which there is no definitive rule that specifies how to act, or the rules that exist may be unclear or even in conflict with each other. Responsible people, including investment professionals, must be willing and able to identify potential ethical issues and create solutions to them even in the absence of clearly stated rules.

---

[1] CFA Institute From Trust to Loyalty: A Global Survey of What Investors Want (2013): http://www.cfapubs.org/doi/pdf/10.2469/ccb.v2013.n14.1.(2016): https://www.cfainstitute.org/research/survey-reports/from-trust-to-loyalty

ETHICS

2

a explain ethics;

Through our individual actions, each of us can affect the lives of others. Our decisions and behavior can harm or benefit a variety of **stakeholders**—individuals or groups of individuals who could be affected either directly or indirectly by a decision and thus have an interest, or stake, in the decision. Examples of stakeholders in decisions made by investment industry professionals include our colleagues, our clients, our employers, the communities in which we live and work, the investment profession, trade associations, regulators, and other financial market participants. In some cases, our actions may benefit all of these stakeholder groups; in other cases, our actions may benefit only some stakeholder groups; and in still other cases, our actions may benefit some stakeholder groups and harm others. For example, recall the research analyst in the introduction who wrote misleading research reports with the aim of increasing the financial benefit to himself and his employer. In the very short term, his conduct seemed to directly benefit some stakeholders (certain clients, himself, and his employer) and to harm other stakeholders (clients who invested based on his reports). Over a longer time period, his conduct resulted in harm to himself and many other stakeholders—his employer, his employer's clients, his colleagues, investors, and through loss of trust when the story was published, the larger financial market.

Ethics encompasses a set of moral principles and rules of conduct that provide guidance for our behavior. The word "ethics" comes from the Greek word "ethos," meaning character, used to describe the guiding beliefs or ideals characterizing a society or societal group. Beliefs are assumptions or thoughts we hold to be true. A principle is defined as a belief or fundamental truth that serves as the foundation for a system of belief or behavior or a chain of reasoning. Our beliefs form our values—those things we deem to have worth or merit.

Moral principles or **ethical principles** are beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior. Ethical principles may refer to beliefs regarding behavior that an individual expects of himself or herself, as well as shared beliefs regarding standards of behavior expected or required by a community or societal group.

The study of ethics examines the role of consequences and personal character in defining what is considered good, or ethical, conduct.

Ethical conduct is behavior that follows moral principles and balances self-interest with both the direct and the indirect consequences of the behavior on others. Ethical actions are those actions that are perceived as beneficial and conforming to the ethical expectations of society. An action may be considered beneficial if it improves the outcomes or consequences for stakeholders affected by the action. Telling the truth about the risks or costs associated with a recommended investment, for example, is an ethical action—that is, one that conforms to the ethical expectations of society in general and clients in particular. Telling the truth is also beneficial; telling the truth builds trust with customers and clients and enables them to make more informed decisions, which should lead to better outcomes for them and higher levels of client/customer satisfaction for you and your employer.

Widely acknowledged ethical principles include honesty, transparency, fairness or justice, diligence, and respect for the rights of others. Most societal groups share these fundamental ethical principles and build on them, establishing a shared set of rules regarding how members should behave in certain situations. The principles or rules may take different forms depending on the community establishing them.

Governments and related entities, for example, may establish laws and/or regulations to reflect widely shared beliefs about obligatory and forbidden conduct. Laws and regulations are rules of conduct specified by a governing body, such as a legislature or a regulator, identifying how individuals and entities under its jurisdiction should behave in certain situations. Most countries have laws and regulations governing the investment industry and the conduct of its participants. Differences in laws may reflect differences in beliefs and values.

In some countries, for example, the law requires that an investment adviser act in the best interests of his or her clients. Other countries require that investment professionals recommend investments that are suitable for their clients. These differing requirements can also hold true within one country where some advisers are held to a suitability standard and others to the fiduciary standard of the client's best interests. Investment advisers and portfolio managers who are required by law to act in their clients' best interests must always put their clients' interests ahead of their own or their employers' interests. An investment adviser who is required by law to act in a client's best interest must understand the client's financial objectives and risk tolerance, research and investigate multiple investment opportunities, and recommend the investment or investment portfolio that is *most* suitable for the client in terms of meeting his or her long-term financial objectives. In addition, the investment adviser would be expected to monitor the client's financial situation and investments to ensure that the investments recommended remain the *best* overall option for meeting the client's long-term financial objectives. In countries with only a suitability requirement, it is legal for investment professionals to recommend a suitable investment to a client even if other, similar suitable investments with lower fees are available. These differences in laws reflect differences in beliefs and values.

Specific communities or societal groups in which we live and work sometimes codify their beliefs about obligatory and forbidden conduct in a written set of principles, often called a **code of ethics**. Universities, employers, and professional associations often adopt a code of ethics to communicate the organization's values and overall expectations regarding member behavior. The code of ethics serves as a general guide for how community members should act. Some communities will also expand on their codes of ethics and adopt explicit rules or standards that identify specific behaviors required of community members. These **standards of conduct** serve as benchmarks for the minimally acceptable behavior of community members and can help clarify the code of ethics. Members can choose behaviors that demonstrate even higher standards. By joining the community, members are agreeing to adhere to the community's code of ethics and standards of conduct. To promote their code of ethics and reduce the incidence of violations, communities frequently display their codes in prominent locations and in written materials. In addition, most communities require that members commit to their codes in writing on an annual or more frequent basis.

Violations of a community's established code of ethics and/or standards of conduct can harm the community in a variety of ways. Violations have the potential to damage the community's reputation among external stakeholders and the general public. Violations can also damage the community's reputation internally and lead to reduced trust among community members and can cause the organization to fracture or splinter from within. To protect the reputation of its membership and limit potential harm to innocent members, the community may take corrective actions to investigate possible violations, repair any damages, and attempt to discipline the violator or, in severe cases, revoke the violator's membership in the community.

EXAMPLE 1

Ethics

1 Which of the following statements is *most* accurate? Ethics can be described as:

- A a commitment to upholding the law.
- B an individual's personal opinion about right and wrong.
- C a set of moral principles that provide guidance for our behavior.

2 Which of the following statements is *most* accurate? Standards of conduct:

- A are a necessary component of any code of ethics.
- B serve as a general guide regarding proper conduct by members of a group.
- C serve as benchmarks for the minimally acceptable behavior required of members of a group.

Solution to 1:

C is correct. Ethics can be described as a set of moral principles that provide guidance for our behavior; these may be moral principles shared by a community or societal group.

Solution to 2:

C is correct. Standards of conduct serve as benchmarks for the minimally acceptable behavior required of members of a group. Some organizations will adopt only a code of ethics, which communicates the organization's values and overall expectations regarding member behavior. Others may adopt both a code of ethics and standards of conduct. Standards of conduct identify specific behavior required of community members and serve as benchmarks for the minimally acceptable behavior of community members.

ETHICS AND PROFESSIONALISM

3

- b describe the role of a code of ethics in defining a profession;
- c describe professions and how they establish trust;
- d describe the need for high ethical standards in investment management;

A **profession** is an occupational community that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition. Most professions emphasize an ethical approach, the importance of good service, and empathy with the client.

Professions have grown in size and number over the last century: the rise of new specialist areas of expertise has created new professions. Driving forces of a new profession include governments and regulators, which encourage the formation of an ethical relationship between professionals and society at large. There is also demand for professions from individuals who see an advantage in working as a professional and from clients who desire to work with professionals.

Professions have not developed in every country. But in most countries, those who work in specialized areas—such as doctors, lawyers, actuaries, accountants, architects, and engineers—are subject to some combination of licensed status and technical standards. These standards distinguish professions from the craft guilds and trade bodies that were established in many countries. In particular, the requirement for members of professions to uphold high ethical standards is one clear difference. Another difference is that trade bodies do not normally have a mission to serve society or to set and enforce professional conduct rules for practitioners.

3.1 How Professions Establish Trust

For a profession to be credible, a primary goal is to establish trust among clients and among society in general. In doing so, professions have a number of common characteristics that, when combined, greatly increase confidence and credibility in professionals and their organizations.

Professions normalize practitioner behavior. Professionalism is underpinned by codes and standards developed by professional bodies. Regulators typically support professional ethics and recognize the framework for ethics that professions can provide. Many regulators around the world have engaged closely with professional bodies to understand their codes and standards, as well as how they are enforced. Codes and standards developed by practitioners can be complementary to regulations, codifying many more individual practices than the high-level principles set by regulation.

Many governments have recognized that a profession can develop a more sophisticated system of standards than a regulator can, via continuous practitioner input and a strong mutual interest within the profession to maintain good standards and adopt best practices. Government support of professions is attributable to the role of professions in helping the public and ensuring expert and principled performance of complex services.

Professions provide a service to society. There is an obligation for professionals to go beyond codes and standards. Professionals should advocate for higher educational and ethical standards in the industry, individually and through their companies. Professions can widen access to services and support economic activity by encouraging trust in the industries they serve. Professions have realized that earning community trust not only creates professional pride and acceptance but also delivers commercial benefits. A profession that earns trust may ultimately have greater flexibility and independence from government regulators to manage its own affairs, which allows members of the profession to develop service models that are both useful to clients and beneficial to members.

Professions are client focused. An integral part of a profession's mission is to develop and administer codes, best practice guidelines, and standards that guide an industry. These codes, standards, and guidelines help ensure that all professionals place the integrity of their profession and the interests of clients above their own personal interests. At a minimum, professionals must act in the best interest of the client, exercising a reasonable level of care, skill, and diligence. The obligation to deliver a high standard of care when acting for the benefit of another party is called *fiduciary duty*. Other entities, including employers, regulators, trade associations, and not-for-profit organizations, may also support an industry but are not the same as professional bodies. Unlike professions, these other entities generally do not exist to set and maintain professional standards. Most employers encourage employees to be members of relevant professions, and many give financial support for this membership to ideally improve the quality of client service and reinforce ethical awareness.

Professions have high entry standards. Membership in a profession is a signal to the market that the professional will deliver high-quality service of a promised standard, going beyond simply academic credentials. Professions develop curricula that equip future professionals with competence, including technical skills, knowledge, and ethics.

Professions possess a body of expert knowledge. A repository of knowledge, developed by experienced and skilled practitioners, is made available to all members of a profession. This knowledge helps members work effectively and ethically and is based on best practice.

Professions encourage and facilitate continuing education. Entry into a profession does not, on its own, guarantee that an individual will maintain competency and continue to uphold professional standards. After qualification and throughout the working life of a professional, there will be changes in knowledge and technical skills to perform certain jobs, in technology and standards of ethical behavior, in services that can be offered, and in the legal and business environment in which professional services are delivered. These all require the development of competence and ethical awareness. Most professional bodies make it a condition of membership that a specific amount of new learning is undertaken each year. Typically, such conditions specify a time commitment, which may be separated into different competencies and types of learning activity. This is often referred to as *continuing professional development* and is seen as an important part of maintaining professional standards. The training and education that professionals undertake increase the value of human capital, which can contribute to economic growth and social mobility.

Professions monitor professional conduct. Members of a profession must be held accountable for their conduct to maintain the integrity and reputation of an industry. Doing so often involves self-regulation by professional bodies through monitoring and imposition of sanctions on members.

Professions are collegial. Professionals should be respectful to each other, even when they are competing. At the very least, they must respect the rights, dignity, and autonomy of others.

Professions are recognized oversight bodies. Many professional bodies are not-for-profit organizations with a mission emphasizing excellence, integrity, and public service. Although it is the responsibility of individual professionals to remain competent, an oversight body typically monitors this responsibility. Such bodies provide individuals with ongoing educational resources and access to information about changes in standards and imposes a framework of discipline. Continuing membership indicates sustained competence in (and updating of) practical skills while maintaining ongoing compliance with an ethical code of conduct.

Professions encourage the engagement of members. Participation by members as volunteers is part of the essence of a profession. Professionals are more likely to refer to, use, and adhere to values that they have helped develop, and they typically have the power as members to revise these values. A good professional will want to mentor and inspire others who recently entered or wish to enter the profession. Professionals should be willing to volunteer to advance the profession and engage with peers to develop expertise and ethics. Professionals should volunteer to help educate new generations in ethical knowledge and ethical decision making and to foster a productive debate about new areas of ethics. Most professionals find that the experience of volunteering within the profession enhances their skills and widens their contacts within the industry. Membership in a professional body allows the necessary engagement with other professionals.

3.2 Professions Are Evolving

No profession stands still. Such trends as greater transparency and public accountability force professions to adapt to change. Meanwhile, technology opens up possibilities for new services and different ways of working. In addition, key processes of a profession's responsibilities may need to be reviewed by a government agency or independent public body. In general, professions often engage with non-member individuals. This can help a profession evaluate the viewpoints of the public, clients, or other stakeholders when determining policy and practice and can encourage public trust for a profession's conduct and disciplinary process.

Effective professions continue to develop their role to account for changing best practices. Some medical professional bodies, for instance, have been established for more than 500 years but may now have the same need to adapt as the much younger investment management profession. This means that at any point in time, society may recognize an area of work as a profession even if it has not fully or universally implemented all the expectations. As the requirements for a profession evolve, gaps open up that may take time to remedy. Effective professions also actively learn from other professions, particularly in the area of ethics. New standards of conduct in the accounting profession might be an influence on standards considered in investment management, for example.

3.3 Professionalism in Investment Management

> Successful investing professionals are disciplined and consistent and they think a great deal about what they do and how they do it.
>
> —Benjamin Graham, *The Intelligent Investor* (1949)

Investment management is a relatively young profession, which means that public understanding of its practice and codes is still developing. Recognition by regulators and employers also lags established professions. Not everyone engaged in investment management is a professional; some practitioners have not undertaken specific investment training or are not members of a professional body. That creates a challenge for the investment management profession to gain trust, because not all practitioners need to be committed to high ethical standards. However, key elements of the profession have been steadily established over several decades. For example, the publication of Graham and Dodd's *Security Analysis* in 1934 was an important step in establishing a body of knowledge for investment.

The investment management profession meets most, but currently not all, of the expectations of a profession. In most countries, some form of certification or licensing is needed to practice, but there may not be a requirement to join a professional body. Globally, the trend is to require examined entry to practice investment management and to maintain competence. But few professions have perfect implementation of all the expected attributes. The investment management profession, similar to other professions, is on a journey to improve implementation and keep up with changing demands.

The investment management profession has become increasingly global as capital markets have opened up around the world. Investment management professionals may seek cross-border opportunities or may need to relocate between offices within multinational asset management firms. Regulatory coordination across borders and the emergence of technology are contributing factors to this globalization of investment management. Various investment management professional bodies have developed in individual countries, and several of these bodies have expanded internationally. In addition, several other professional bodies, including those focused on actuarial and accountancy services, have investment management professionals as members.

3.4 Trust in Investment Management

The investment management professional today has similarities with professionals in longer-established professions, such as medicine and law. Like doctors and lawyers, investment management professionals are trusted to draw on a body of formal knowledge and apply that knowledge with care and judgement. In comparison to clients, investment professionals are also expected to have superior financial expertise, technical knowledge, and knowledge of the applicable laws and regulations. There is a risk that clients may not be fully aware of the conflicts, risks, and fees involved, so investment management professionals must always handle and fully disclose these issues in a way that serves the best interests of clients. Compliance with codes of ethics and professional standards is essential, and practice must be guided by care, transparency, and integrity.

The investment management profession and investment firms must be interdependent to maintain trust. Employers and regulators have their own standards and practices that may differ from regulations and standards set by professional bodies. The investment management professional bodies typically direct professionals in how to resolve these differences.

In many developed economies, the investment management profession affects many key aspects of the economy, including savings, retirement planning, and the pricing and allocation of capital. In most countries, skilled evaluation of securities leads to more efficient capital allocation and, combined with ethical corporate governance, can assist in attracting investment from international investors. The investment management profession can deliver more value to society when higher levels of trust and better capital allocation reduce transaction costs and help meet client objectives. These reasons explain why practitioners, clients, regulators, and governments have supported the development of an investment management profession.

3.5 CFA Institute as an Investment Management Professional Body

CFA Institute is the largest body for investment management professionals.[2] Reflecting the globalization of investment management, CFA Institute moved beyond North America in the 1980s. CFA Institute initiated a number of other changes in line with the growth of investment management. One significant change occurred in 2015, when CFA Institute decided to implement the highest standards of governance in the US not-for-profit sector. The Board of Governors resolved “to implement US Public Company Standards and US not-for-profit leading practices, unless the Board determines that it is not in the best interest of the membership or organization to do so.”

The mission of CFA Institute is “to lead the investment profession globally, by promoting the highest standards of ethics, education, and professional excellence for the ultimate benefit of society.” The CFA Institute Code of Ethics and Standards of Professional Conduct (Code and Standards) promote the integrity of charterholders and establish a model for ethical behavior. CFA Institute candidates and charterholders must meet the highest standards among those established by CFA Institute, regulators, or the employer. If candidates and charterholders do not meet these standards, there are negative consequences. Where client interests and market interests conflict, the Code and Standards set an investment management professional's duty to market integrity as the overriding obligation. The advocacy efforts of CFA Institute aim to build market integrity by calling for regulations that align the interests of firms and clients.

---

[2] Eligibility and requirements for becoming a member of CFA Institute vary by jurisdiction. Please consult www.cfainstitute.org for further details.

As a professional body, CFA Institute gathers knowledge from practicing investment professionals, conducts rigorous examinations, and ensures practitioner involvement in developing its codes and values. The CFA Institute Global Body of Investment Knowledge (GBIK) and Candidate Body of Knowledge (CBOK) are updated on an ongoing basis through a process known as *practice analysis*. Through interactions with practicing investment management professionals, practice analysis helps ensure that the body of knowledge for the investment management profession remains current and globally relevant. The CFA Program ensures that candidates have sufficiently mastered the core knowledge, skills, and abilities (competencies) necessary that are generally accepted and applied by investment professionals. CFA Institute also contributes to the dissemination of new research and ideas in finance with the publication of the *Financial Analysts Journal*; CFA Institute Research Foundation books, research briefs, and reviews; and *CFA Institute Magazine*.

CFA Institute encourages charterholders to engage in their professional communities and involves charterholders in its initiatives. CFA Institute local societies keep charterholders connected and engaged in their communities. CFA Institute assists local societies with providing continuing education programs and events that facilitate charterholders engagement. For CFA charterholders, a local CFA society is an important route to maintaining professionalism, particularly for continuing professional development.

CFA charterholders and CFA Program candidates are required to adhere to the Code and Standards and to sign annually a statement attesting to that continued adherence. Charterholders and candidates must maintain and improve their professional competence and strive to maintain and improve the competence of other investment professionals.

EXAMPLE 2

Ethics and Professionalism

1. Which of the following statements is *most* accurate? Investment professionals have a special responsibility to act ethically because:
   - **A** the industry is heavily regulated.
   - **B** they are entrusted to protect clients' assets.
   - **C** the profession requires compliance with its code of ethics.
2. Which of the following statements *best* completes the following sentence? Professionals use their specialized knowledge and skills:
   - **A** in service to others.
   - **B** to advance their career.
   - **C** for the exclusive benefit of their employers.
3. Which of the following statements is *most* accurate? A profession's code of ethics:
   - **A** includes standards of conduct or specific benchmarks for behavior.
   - **B** ensures that all members of a profession will act ethically at all times.
   - **C** publicly communicates the shared principles and expected behaviors of a profession's members.

Solution to 1:

B is correct. Investment professionals have a special responsibility because clients entrust them to protect the clients' assets.

Solution to 2:

A is correct. Professionals use specialized knowledge and skills in service to others. Their career and employer may benefit, but those results are not the primary focus of a professional's use of his or her specialized knowledge and skills.

Solution to 3:

C is correct. A profession's code of ethics publicly communicates the shared principles and expected behaviors of a profession's members. The existence of a code of ethics does not ensure that all members will behave in a manner consistent with the code and act ethically at all times. A profession will often establish a disciplinary process to address alleged violations of the code of ethics. A profession may adopt standards of conduct to enhance and clarify the code of ethics.

CHALLENGES TO ETHICAL CONDUCT

4

f identify challenges to ethical behavior;

Professionals generally aim to be responsible and to adhere to high moral standards, so what is the benefit of studying ethics? Throughout our careers, we may find ourselves in difficult or at least unfamiliar situations in which an appropriate course of action is not immediately clear and/or there may be more than one seemingly acceptable choice; studying ethics helps us prepare for such situations. This section addresses challenges to engaging in ethical conduct. Failure to acknowledge, understand, or consider these challenges can lead to poor decision making, resulting in unintentional consequences, such as unethical conduct and potential violations of the Code and Standards.

Several challenges can make adherence to ethical conduct difficult. First, people tend to believe that they are ethical people and that their ethical standards are higher than average. Of course, everyone cannot be above average. However, surveys show this belief in above averageness remains.

These survey results illustrate overconfidence, a common behavioral bias that can lead to faulty decision making. Studies have shown that our beliefs and emotions frequently interfere with our cognitive reasoning and result in behavioral bias, a tendency to behave in a way that is not strictly rational. As a result of the overconfidence bias, we are more likely to overestimate the morality of our own behavior, particularly in situations that we have not faced before. The overconfidence bias can result in a failure to consider, explicitly or implicitly, important inputs and variables needed to form the best decision from an ethical perspective. In general, the overconfidence bias leads us to place too much importance on internal traits and intrinsic motivations, such as "I'm honest and would not lie," even though studies have shown that internal traits are generally not the main determinant of whether or not someone will behave ethically in a given situation.

A second challenge is that decision makers often fail to recognize and/or significantly underestimate the effect of situational influences, such as what other people around them are doing. **Situational influences** are external factors, such as environmental or cultural elements, that shape our thinking, decision making, and behavior. Social psychologists have studied how much situational influences affect our behavior and have found that even good people with honorable motives can and often will be influenced to do unethical things when put into difficult situations. Experiments have shown that even people who consider themselves strong, independent, free thinkers will conform to social pressures in many situations. The bystander effect, for example, demonstrates that people are less likely to intervene in an emergency when others are present. Fortunately, experiments have also shown that situational influences can induce people to act more ethically. For example, people tend to behave more ethically when they think someone else is watching or when there is a mirror placed close to them. The important concept to understand is that situational influences have a very powerful and often unrecognized effect on our thinking and behavior. Thus, learning to recognize situational influences is critical to making good decisions.

Common situational influences in the investment industry that can shape thinking and behavior include money and prestige. One experiment found that simply mentioning money can reduce ethical behavior. In the experiment, participants were less likely to cooperate when playing a game if the game was called the Wall Street Game, rather than the Community Game. In the investment industry, large financial rewards—including individual salaries, bonuses, and/or investment gains—can induce honest and well-intentioned individuals to act in ways that others might not consider ethical. Large financial rewards and/or prestige can motivate individuals to act in their own short-term self-interests, ignoring possible short-term risks or consequences to themselves and others as well as long-term risks or consequences for both themselves and others. Another extremely powerful situational influence is loyalty. Loyalty to supervisors or organizations, fellow employees, and other colleagues can tempt individuals to make compromises and take actions that they would reject under different situational influences or judge harshly when taken by others.

Situational influences often blind people to other important considerations. Bonuses, promotions, prestige, and loyalty to employer and colleagues are examples of situational influences that frequently have a disproportionate weight in our decision making. Our brains more easily and quickly identify, recognize, and consider these short-term situational influences than longer-term considerations, such as a commitment to maintaining our integrity and contributing to the integrity of the financial markets. Although absolutely important, these long-term considerations often have less immediate consequences than situational influences, making them less obvious as factors to consider in a decision and, therefore, less likely to influence our overall decision making. Situational influences shift our brain's focus from the long term to the short or immediate term. When our decision making is too narrowly focused on short-term factors and/or self-interest, we tend to ignore and/or minimize the longer-term risks and/or costs and consequences to ourselves and others, and the likelihood of suffering ethical lapses and making poor decisions increases.

Loyalty to employer and/or colleagues is an extremely powerful situational influence. Our colleagues can influence our thinking and behavior in both positive and negative ways. For example, colleagues may have encouraged you to signal your commitment to your career and high ethical standards by enrolling in the CFA Program. If you work for or with people who are not bound by the Code and Standards, they might encourage you to take actions that are consistent with local law, unaware that the recommended conduct falls short of the Code and Standards.

Well-intentioned firms may adopt or develop strong compliance programs to encourage adherence to rules, regulations, and policies. A strong compliance policy is a good start to developing an ethical culture, but a focus on adherence to rules may not be sufficient. A compliance approach may not encourage decision makers to consider the larger picture and can oversimplify decision making. Taken to the extreme, a strong compliance culture can become another situational influence that blinds employees to other important considerations. In a firm focused primarily on compliance, employees may adopt a "check the box" mentality rather than an ethical decision-making approach. Employees may ask the question "What *can* I do?" rather than "What *should* I do?"

EXAMPLE 3

Challenges to Ethical Conduct

1 Which of the following will *most likely* determine whether an individual will behave unethically?

- **A** The person's character
- **B** The person's internal traits and intrinsic motivation
- **C** External factors, such as environmental or cultural elements

2 Which of the following statements is *most* accurate?

- **A** Large financial rewards, such as bonuses, are the most powerful situational influences.
- **B** When decision making focuses on short-term factors, the likelihood of ethical conduct increases.
- **C** Situational influences can motivate individuals to act in their short-term self-interests without recognizing the long-term risks or consequences for themselves and others.

Solution to 1:

C is correct. Social psychologists have shown that even good people may behave unethically in difficult situations. Situational influences, which are external factors (e.g., environmental or cultural elements), can shape our thinking, decision making, and behavior and are more likely to lead to unethical behavior than internal traits or character.

Solution to 2:

C is correct. Situational influences can motivate individuals to act in their short-term self-interests without recognizing the long-term risks or consequences for themselves and others. Large financial rewards are powerful situational influences, but in some situations, other situational influences, such as loyalty to colleagues, may be even more powerful.

ETHICAL VS. LEGAL STANDARDS

5

g compare and contrast ethical standards with legal standards;

Many times, stakeholders have common ethical expectations. Other times, different stakeholders will have different perceptions and perspectives and use different criteria to decide whether something is beneficial and/or ethical.

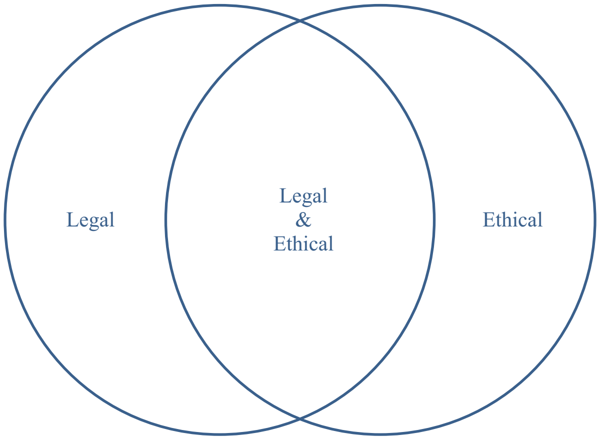
Laws and regulations often codify ethical actions that lead to better outcomes for society or specific groups of stakeholders. For example, some laws and regulations require businesses and their representatives to tell the truth. They require specific written disclosures in marketing and other materials. Complying with such rules is considered an ethical action; it creates a more satisfactory outcome that conforms to stakeholders' ethical expectations. As an example, consider disclosure requirements mandated by securities regulators regarding the risks of investing. Complying with such rules creates better outcomes for you, your clients, and your employer. First, compliance with the rule reduces the risk that clients will invest in securities without understanding the risks involved, which, in turn, reduces the risk that clients will file complaints and/or take legal action if their investments decline in value. Complying with the rules also reduces the risk that regulators will initiate an investigation, file charges, or/and discipline or sanction you and/or your employer. Any of these actions could jeopardize the reputation and future prospects of you and your employer. Conduct that reduces these risks (e.g., following disclosure rules) would be considered ethical; it leads to better outcomes for you, your clients, and your employer and conforms to the ethical expectations of various stakeholders.

Although laws frequently codify ethical actions, legal and ethical conduct are not always the same. Think about the diagram in Exhibit 1. Many types of conduct are both legal and ethical, but some conduct may be one and not the other. Some legal behaviors or activities may be considered unethical, and some behaviors or activities considered ethical may be deemed illegal in certain jurisdictions. Acts of civil disobedience, such as peaceful protests, may be in response to laws that individuals consider unethical. The act of civil disobedience may itself be considered ethical, and yet it violates existing local laws.

The investment industry has examples of conduct that may be legal but considered by some to be unethical. Some countries, for example, do not have laws prohibiting trading while in possession of material nonpublic information, but many investment professionals and CFA Institute consider such trading unethical.

Another area in which ethics and laws may conflict is the area of "whistleblowing." Whistleblowing refers to the disclosure by an individual of dishonest, corrupt, or illegal activity by an organization or government. Depending on the circumstances, a whistleblower may violate organizational policies and even local laws with the disclosure; thus, a whistleblower's actions may be deemed illegal and yet considered by some to be ethical.

Exhibit 1 Types of Conduct



Some people advocate that increased regulation and monitoring of the behavior of participants in the investment industry will increase trust in the financial markets. Although this approach may work in some circumstances, the law is not always the best mechanism to reduce unethical behavior for several reasons. First, laws typically follow market practices; regulators may proactively design laws and regulations to address existing or anticipated practices that may adversely affect the fairness and efficiency of markets or reactively design laws and regulations in response to a crisis or an event that resulted in significant monetary losses and loss of confidence/trust in the financial system. Regulators' responses typically take significant time, during which the problematic practice may continue or even grow. Once enacted, a new law may be vague, conflicting, and/or too narrow in scope. A new law may reduce or even eliminate the existing activity while simultaneously creating an opportunity for a different, but similarly problematic, activity. Additionally, laws vary across countries or jurisdictions, allowing questionable practices to move to places that lack laws relevant to the questionable practice. Laws are also subject to interpretation and compliance by market participants, who may choose to interpret the law in the most advantageous way possible or delay compliance until a later date. For these reasons, laws and regulations are insufficient to ensure the ethical behavior of investment professionals and market participants.

Ethical conduct goes beyond what is legally required and encompasses what different societal groups or communities, including professional associations, consider to be ethically correct behavior. To act ethically, individuals need to be able to think through the facts of the situation and make good choices even in the absence of clear laws or rules. In many cases, there is no simple algorithm or formula that will always lead to an ethical course of action. Ethics requires judgment—the ability to make considered decisions and reach sensible conclusions. Good ethical judgment requires actively considering the interests of stakeholders and trying to benefit multiple stakeholders—clients, family, colleagues, employers, market participants, and so forth—and minimize risks, including reputational risk.

EXAMPLE 4

Ethical vs. Legal Standards

1. Which of the following statements is *most* accurate?
   - **A** All legal behavior is ethical behavior.
   - **B** Some ethical behavior may be illegal.
   - **C** Legal standards represent the highest standard.
2. Which of the following statements is *most* accurate?
   - **A** Increased regulations are the most useful means to reduce unethical behavior by market participants.
   - **B** Regulators quickly design and implement laws and regulations to address practices that adversely affect the fairness and efficiency of markets.
   - **C** New laws designed to reduce or eliminate conduct that adversely affects the markets can create opportunities for different, but similarly problematic, conduct.

Solution to 1:

B is correct. Some ethical behavior may be illegal. Civil disobedience is an example of what may be illegal behavior that some consider to be ethical. Legal and ethical behavior often coincide but not always. Standards of conduct based on ethical principles may represent a higher standard of behavior than the behavior required by law.

Solution to 2:

C is correct. New laws designed to reduce or eliminate conduct that adversely affects the markets can create opportunities for different, but similarly problematic, conduct.

6

ETHICAL DECISION-MAKING FRAMEWORKS

h describe a framework for ethical decision making.

Laws, regulations, professional standards, and codes of ethics can guide ethical behavior, but individual judgment is a critical ingredient in making principled choices and engaging in appropriate conduct. One strategy to increase trust in the investment industry is to increase the ability and motivation of market participants to act ethically and help them minimize the likelihood of unethical actions. By integrating ethics into the decision-making activities of employees, firms can enhance the ability and the motivation of employees to act ethically, thereby reducing the likelihood of unethical actions. The ability to relate an ethical decision-making framework to a firm's or profession's code of ethics allows investment professionals to bring the principles of the code of ethics to life. An investment professional's natural desire to "do the right thing" can be reinforced by building a culture of integrity and accountability in the workplace. Development, maintenance, and demonstration of a strong culture of integrity within the firm by senior management may be the single most important factor in promoting ethical behavior among the firm's employees.

Adopting a code that clearly lays out the ethical principles that guide the thought processes and conduct the firm expects from its employees is a critical first step. But a code of ethics, although necessary, is insufficient. Simply nurturing an inclination to do right is no match for the multitude of daily decisions that investment professionals make. We need to exercise ethical decision-making skills to develop the muscle memory necessary for fundamentally ethical people to make good decisions despite the reality of conflicts and our natural instinct for self-preservation. Just as coaching and practice transform our natural ability to run across a field into the technique and endurance required to run a race, teaching, reinforcing, and practicing ethical decision-making skills prepare us to confront the hard issues effectively. It is good for business, individuals, firms, the industry, and the markets, as well as society as a whole, to engage in the investment management profession in a highly ethical manner. A strong ethical culture, built on a defined set of principles, that helps honest, ethical people engage in ethical behavior will foster the trust of investors, lead to robust global financial markets, and ultimately benefit society. That is why ethics matter.

6.1 The Framework for Ethical Decision-Making

When faced with decisions that can affect multiple stakeholders, investment professionals must have a well-developed set of principles; otherwise, their thought processes can lead to, at best, indecision and, at worst, fraudulent conduct and destruction of the public trust. Establishing an ethical framework to guide your internal thought process regarding how to act is a crucial step to engaging in ethical conduct. Investment professionals are generally comfortable analyzing and making decisions from an economic (profit/loss) perspective. Given the importance of ethical behavior in carrying out professional responsibilities, it is also important to analyze decisions and their potential consequences from an ethical perspective. Using a framework for ethical decision making will help investment professionals to effectively examine their choices in the context of conflicting interests common to their professional obligations (e.g., researching and gathering information, developing investment recommendations, and managing money for others). Such a framework will allow investment professionals to analyze and choose options in a way that allows them to meet high standards of ethical behavior. An ethical decision-making framework provides investment professionals with a tool to help them adhere to a code of ethics. By applying the framework and analyzing the particular circumstances of each available alternative, investment professionals are able to determine the best course of action to fulfill their responsibilities in an ethical manner.

An ethical decision-making framework will help a decision maker see the situation from multiple perspectives and pay attention to aspects of the situation that may be less evident with a short-term, self-focused perspective. The goal of getting a broader picture of a situation is to be able to create a plan of action that is less likely to harm stakeholders and more likely to benefit them. If a decision maker does not know or understand the effects of his or her actions on stakeholders, the likelihood of making a decision and taking action that harms stakeholders is more likely to occur, even if unintentionally. Finally, an ethical decision-making framework helps decision makers explain and justify their actions to a broader audience of stakeholders.

Ethical decision-making frameworks are designed to facilitate the decision-making process for all decisions. They help people look at and evaluate a decision from multiple perspectives, enabling them to identify important issues they might not otherwise consider. Using an ethical decision-making framework consistently will help you develop sound judgment and decision-making skills and avoid making decisions that have unanticipated ethical consequences. Ethical decision-making frameworks come in many forms with varying degrees of detail. A general ethical decision-making framework is shown in Exhibit 2.

Exhibit 2 Ethical Decision-Making Framework

- Identify: Relevant facts, stakeholders and duties owed, ethical principles, conflicts of interest
- Consider: Situational influences, additional guidance, alternative actions
- Decide and act
- Reflect: Was the outcome as anticipated? Why or why not?

The ethical decision-making process includes multiple phases, each of which has multiple components. The process is often iterative, and you, the decision maker, may move between phases in an order different from what is presented. For simplicity, we will discuss the phases sequentially. In the initial phase, you will want to identify the important facts that you have available to you, as well as information that you may not have but would like to have to give yourself a more complete understanding of the situation. It is helpful to distinguish between facts and personal opinion, judgements, and biases. You will also want to identify the stakeholders—clients, family, colleagues, your employer, market participants, and so forth—and the duties you have to each of them. You will then want to identify relevant ethical principles and/or legal requirements that might apply to the situation. You should also identify any potential conflicts of interest inherent in the situation or conflicts in the duties you hold to others. For example, your duty to your client may conflict with your duty to your employer.

In the second phase of ethical decision making, you will take time to consider the situational influences as well as personal behavioral biases that could affect your thinking and thus decision making. These situational influences and biases could include a desire to please your boss, to be seen as successful by your peers and family, to gain acceptance, to earn a large bonus, and so on. During this phase, you may seek additional guidance from trusted sources—perhaps a family member, colleague, or mentor who can help you think through the situation and help you identify and evaluate alternative actions. You may turn to your compliance department for assistance or you may even consult outside legal counsel. Seeking additional guidance is a critical step in viewing the situation from different perspectives. You should seek guidance from someone, possibly external to the firm, who is not affected by the same situational influences and behavioral biases as you are and can, therefore, provide a fresh perspective. You should also seek guidance from your firm's policies and procedures and the CFA Institute Code and Standards. A helpful technique might be to imagine how an ethical peer or role model might act in the situation.

The next phase of the framework is to make a decision and act. After you have acted on your decision, you should take the time to reflect on and assess your decision and its outcome. Was the outcome what you anticipated? Why or why not? Had you properly identified all the important facts, stakeholders, duties to stakeholders, conflicts of interest, and relevant ethical principles? Had you considered the situational influences? Did you identify personal behavioral biases that might affect your thinking? Had you sought sufficient guidance? Had you considered and properly evaluated a variety of alternative actions? You may want to reflect on the decision multiple times as the immediate and longer-term consequences of your decision and actions become apparent.

The process is often iterative. After identifying the relevant facts and considering situational influences, you may, for example, decide that you cannot make a decision without more information. You may seek additional guidance on how to obtain the information you need. You may also begin considering alternative actions regarding how to proceed based on expectations of what the additional information will reveal, or you may wait until you have more information, reflect on what you have done and learned so far, and start the process over again. Sometimes cases can be complicated and multiple iterations may reveal that no totally acceptable solution can be created. Applying an ethical decision-making framework can help you evaluate the situation so you can make the best possible decision. The next section shows applications of the framework shown in Exhibit 2.

6.2 Applying the Framework

To illustrate how the framework could be applied in your career, consider the scenario in Example 5.

EXAMPLE 5

Applying an Ethical Decision-Making Framework I

You have been hired as a junior analyst with a major investment bank. When you join the bank, you receive a copy of the firm's policies as well as training on the policies. Your supervisor is the senior technology analyst for the investment bank. As part of your duties, you gather information, draft documents, conduct analysis, and perform other support functions for the senior analyst.

Your employer is one of several investment banks working on the initial public offering (IPO) of a well-known technology company. The IPO is expected to generate significant revenues for the investment banks participating in the offering. The IPO has been highly anticipated and is in the news every day.

You are thrilled when your supervisor asks you to work on several research projects related to analyzing and valuing the upcoming IPO for investors. You eagerly compile information and draft a one-page outline. You stop to consider what other information you could add to improve the report before proceeding. You realize that you have two excellent contacts in the technology industry who could review your work and provide some additional and potentially valuable perspectives. You draft an email to your contacts reading:

> I am working on an analysis and valuation of Big Tech Company for investors. My employer is one of the banks participating in the IPO, and I want to make sure I have considered everything. I was hoping you could give me feedback on the prospects and risks facing Big Tech. Please treat all the attached material as confidential.

Before hitting the send button, you stop and think about the ethical decision-making framework you have studied. You decide to apply the framework and jot down some notes as you work through the process: On the first page, you work through the identification phase and make a list of the relevant facts, stakeholders to whom you owe a duty, potential conflicts of interest, and ethical principles. This list is shown is Exhibit 3.

Exhibit 3 Identification Phase

1 Relevant facts:
- *Working on the deal/IPO of the decade*
- *Employer is one of several investment banks working on IPO*
- *The IPO is highly anticipated*
- *A successful IPO could lead to additional investment banking deals and revenues for the firm*
- *Supervisor is relying on me*
- *Employer has documented policies and procedures*
- *Industry is regulated, with many rules and regulations in place*

2 Stakeholders and duties owed. I have a duty to the following:
- *Supervisor*
- *Employer*
- *Employer's corporate client, the technology company*
- *Employer's asset management and other investing clients*
- *Employer's partners in the IPO*
- *Investors and market participants interested in the IPO*
- *All capital market participants*
- *Profession*
- *Society as a whole*

3 Conflicts or potential conflicts of interest include the following:
- *Gathering additional research from external sources versus maintaining confidentiality*

(continued)

Exhibit 3 (Continued)

- *Duty to supervisor versus desire to impress*
- *Duty to corporate client versus duty to other clients of the firm*
- *The firm's corporate client benefits from a high IPO price whereas the firm's asset management clients would benefit from a low IPO price*
- *Making the IPO look attractive to the market (sell-side marketing) versus objective analysis of the investment potential of this deal (buy-side analysis)*
- *My bonus, compensation, and career prospects are tied to my supervisor's and the IPO's success; duty to employer, employer's investing clients, profession*

4 Ethical principles that are relevant to this situation include the following:

- *Duty of loyalty to employer*
- *Client interests come first*
- *Maintain confidences and confidentiality of information*
- *Objectivity of analysis*
- *Fairness to market participants*

On the next page, you write notes relating to the second phase of the framework, considering the various situational factors and the guidance available to you before considering alternative actions. These notes are shown in Exhibit 4.

Exhibit 4 Consideration Phase

1 Situational influences:

- *The firm's written policies*
- *The bank will earn big fees from the IPO*
- *I want to impress my boss—and potential future bosses*
- *My bonus, compensation, and career prospects will be influenced by my contribution to the success of this deal and other deals*
- *I am one of very few people working on this deal; it is a real honor, and others would be impressed that I am working on this deal*
- *My employer is filled with successful and wealthy people who are go-getters; I want to be successful and wealthy like them*

2 Additional guidance. I could seek guidance from the following:

- *The firm's code of ethics*
- *The firm's written policies*
- *A peer in my firm*
- *My supervisor, the senior analyst*
- *The compliance department*
- *A mentor either at the firm or perhaps from university or industry*
- *The CFA Institute Code and Standards*

Exhibit 4 (Continued)

- *Outside legal counsel*

3 Alternative actions. I could consider the following:

- *Asking contacts what they have heard*
- *Submitting the report as a draft and suggesting that contacts in the industry might be able to provide more perspective*
- *Sending a survey to various technology industry veterans soliciting their viewpoints on developments*

After completing these steps, you decide to check the firm's policies. Under a section entitled "Research Analyst Role in Securities Offerings," the manual states, "You may not distribute any written (which include email, fax, electronic, and other means) material related to companies and/or their offerings . . . during the course of any offering and the related quiet period."

You read further and note a section entitled "Wall Crossing Policy and Procedures" that states that "employees with confidential information may not communicate the information to anyone who does not have a valid need to know" without first obtaining clearance from the legal and compliance department.

You decide that your contacts do not have a "valid need to know" and that it is unlikely the firm's legal and compliance professionals would approve sharing the information. You then decide to mention your contacts to the senior research analyst. He suggests that they may have some useful perspective and that you might talk to them to hear their perspective and cautions you not to disclose any information about any of the firm's clients, pending deals, or research. You return to your desk, delete the email, and following the senior research analyst's advice, call your contacts on the telephone to discuss the technology sector, its prospects, and its challenges. During the calls, you take care not to reveal any details about Big Tech Company or its offering.

Whatever action you take, you should take time afterward to reflect on the decision and the outcome. Was the outcome as anticipated? Why or why not?

The initial facts presented in the example are based on the real-life experience of a young junior analyst working on a highly anticipated IPO. The junior analyst may or may not have used an ethical decision-making framework to evaluate his situation. Without seeking additional guidance, the junior analyst sent an email similar to the one in the example with an attachment that included confidential, proprietary information, including the senior analyst's analysis and forecasts. Months later, long after the IPO offering, the junior analyst's email was discovered by his employer. When questioned, he admitted that he had received training regarding the firm's policies and that he did not discuss or seek approval from anyone before sending the email. Two days later, the firm terminated the junior analyst's employment and reported to regulatory authorities that he had been terminated for distributing written materials, by email, during a securities offering in violation of firm policies that prohibit the dissemination of any written materials during the course of a securities offering and related periods. The junior analyst's supervisor also lost his job for failing to properly supervise the analyst. Multiple regulators investigated the matter, and the firm was fined millions of dollars for failing to supervise its employees properly. The information regarding the junior analyst's termination was posted and remains available on the regulator's website for all to see. Future employers conducting routine background checks will know that the analyst was terminated for violating firm policies relating to a securities offering.

The example presented is similar to situations faced by many analysts. Using an ethical decision-making framework will help you evaluate situations from multiple perspectives, avoid poor decision making, and avoid the consequences that can result from taking an ill-conceived course of action. Using an ethical decision-making framework is no guarantee of a positive outcome but may help you avoid making unethical decisions.

EXAMPLE 6

Applying an Ethical Decision-Making Framework II

A financial adviser has been saving a portion of his salary to purchase a new vehicle. He is on track to have enough saved within the next three months. His employer has offered a special bonus for this quarter, which will go to the team that attracts the most new investors into the firm's investment funds. In addition to the potential bonus, the firm pays a 5% commission to employees who sell shares in the firm's investment funds. Several of the funds are highly rated, including one designed to provide steady income to investors.

The financial adviser has added only a few new investors to the firm's funds, but his teammates have been very successful in their efforts. The end of the quarter is one week away, and his team is competing closely with another team for the bonus. One of his teammates informs the financial adviser that he really needs the bonus so his elderly mother can receive medical treatment.

Later that day, the financial adviser meets with an elderly client on a limited income who is seeking more income from his investment portfolio. The client is 89 years old and in poor health. According to the client's will, the client's investment portfolio will go to his favorite charity upon his death.

1. Which of the following situational influences is likely to have the *most* effect on the financial adviser's efforts to get new clients to invest in the funds? His relationship with his:
   - **A** client.
   - **B** employer.
   - **C** teammates.
2. Which of the following statements is *most* accurate? An ethical decision-making framework:
   - **A** is only beneficial when a firm lacks a code of ethics.
   - **B** is used to improve compliance with laws and regulations.
   - **C** is a tool for analyzing the potential alternative actions and consequences of a decision.
3. Which of the following is *most* accurate? Ethical decision-making frameworks:
   - **A** raise awareness of different perspectives.
   - **B** focus attention on short-term consequences.
   - **C** allocate more weight to those who will directly benefit from the decision.
4. Which of the following is *most* accurate? Ethical decision-making frameworks:

- **A** are not needed if behavior is legal.
- **B** identify who gains the most from a decision.
- **C** can help reduce unanticipated ethical lapses and unexpected consequences.

5 Using an ethical decision-making framework, which of the following duties would *most likely* take precedence in the scenario described? The financial adviser's duty to his:

- **A** client.
- **B** employer.
- **C** colleagues.

6 Using an ethical decision-making framework, the financial adviser would *most likely*:

- **A** recommend that the elderly client invest at least some of his assets in the highly rated fund.
- **B** research other investments that can provide steady income before making a recommendation to his elderly client.
- **C** disclose the commission he would earn before recommending that the elderly client invest at least some of his assets in the highly rated fund.

Solution to 1:

C is correct. The financial adviser's relationship with his teammates is likely to have the most effect on the financial adviser's efforts. The teammates share in earning the bonus from the employer. In addition, the team works with each other on a regular basis where there is a likelihood that social influences of working together play a role in decision making and effort.

Solution to 2:

C is correct. An ethical decision-making framework is a tool for analyzing the potential alternative actions and consequences of a decision.

Solution to 3:

A is correct. Ethical decision-making frameworks raise awareness of different perspectives. The framework should consider short-term consequences, but they are not the focus of the framework. Similarly, the framework may allocate more weight to those who directly benefit from decisions, but this is not the primary goal of an ethical decision-making framework.

Solution to 4:

C is correct. Ethical decision-making frameworks can help avoid unanticipated ethical consequences. As it relates to A, ethics standards are often higher than legal standards so an ethical decision-making framework would be needed. Although B is accurate, identifying who gains most from the decision is a small component of the framework.

Solution to 5:

A is correct. Using an ethical decision-making framework, the financial adviser's relationship with his client would most likely take precedence in this scenario. The adviser should put his client's interests first. The exception to client interests taking precedence occurs when market integrity effects take precedence.

Solution to 6:

B is correct. Using an ethical decision-making framework, the financial adviser would identify the relevant facts, stakeholders, duties owed, and potential conflicts. In this scenario, the financial adviser owes a duty to his client as well as his employer. His client's interests take precedence over all other interests. The bonus and his colleague's desire to help his mother are situational influences. To navigate this situation, the financial adviser should seek additional information; he should research the risk and return parameters and fee structures of other investments that can provide steady income before making a recommendation to his client.

7 CONCLUSION

This reading introduced ideas and concepts that will help you understand the importance of ethical behavior in the investment industry as well as the challenges to adhering to high ethical standards. A code of ethics will communicate an organization's values and the expected behavior of its members as well as provide guidance for decision making. A code of ethics may be further enhanced and clarified by the adoption of standards of conduct. Professions and professional organizations, such as the CFA Institute, can help to establish codes and standards that provide a framework and technical competence for practitioners. An ethical decision-making framework combined with a code of ethics may help investment professionals analyze their decisions in a way that identifies potential conflicts and negative consequences.

Knowing the rules to apply in a particular situation, although important, may not be sufficient to ensure ethical conduct if used alone. Responsible professionals in the investment industry must be able both to recognize areas that are prone to ethical pitfalls and to identify and process those circumstances and influences that can impair judgment and lead to ethical lapses.

SUMMARY

- Ethics refers to the study of making good choices. Ethics encompasses a set of moral principles and rules of conduct that provide guidance for our behavior.
- Situational influences are external factors that may shape our behavior.
- Challenges to ethical behavior include being overconfident in our own morality, underestimating the effect of situational influences, and focusing on the immediate rather than long-term outcomes or consequences of a decision.
- In any given profession, the code of ethics publicly communicates the established principles and expected behavior of its members.
- Members of a profession use specialized knowledge and skills to serve others; they share and agree to adhere to a common code of ethics to serve others and advance the profession.
- A code of ethics helps foster public confidence that members of the profession will use their specialized skills and knowledge to serve their clients and others.

- A profession is an occupational group that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition.
- The requirement to uphold high ethical standards is one clear difference between professions and craft guilds or trade bodies.
- A primary goal of professions is to establish trust among clients and among society in general.
- Common characteristics of professions include normalization of practitioner behavior, service to society, client focus, high entry standards, a body of expert knowledge, encouragement and facilitation of continuing education, monitoring of professional conduct, collegiality, recognized overseeing bodies, and encouragement of member engagement.
- The investment management profession has become increasingly global, driven by the opening of capital markets, coordination of regulation across borders, and the emergence of technology.
- Investment management professionals are trusted to draw on a body of formal knowledge and apply that knowledge with care and judgement. In comparison to clients, investment professionals are also expected to have superior financial expertise, technical knowledge, and knowledge of the applicable laws and regulations.
- As a professional body, CFA Institute gathers knowledge from practicing investment professionals, conducts rigorous examinations, and ensures practitioner involvement in developing its codes and values.
- Investment management professionals are likely to encounter dilemmas, including those with ethical implications. Professionals should consider carefully how to determine the facts of the issue and assess the implications.
- High ethical standards always matter and are of particular importance in the investment management profession, which is based almost entirely on trust. Clients trust investment professionals to use their specialized skills and knowledge to serve clients and protect client assets. All stakeholders gain long-term benefits when investment professionals adhere to high ethical standards.
- Legal standards are often rule based. Ethical conduct goes beyond legal standards, balancing self-interest with the direct and indirect consequences of behavior on others.
- A framework for ethical decision making can help people look at and evaluate a decision from different perspectives, enabling them to identify important issues, make wise decisions, and limit unintended consequences.

REFERENCES

Beaton, George Ramsay. 2010. *Why Professionalism is Still Relevant.* University of Melbourne Legal Studies Research Paper No. 445. Available at SSRN: https://ssrn.com/abstract=1545509.

Bellis, C. S. 2000. "Professions in Society." *British Actuarial Journal* 6 (2): 317–64.

Bogle, John C. 2017. "Balancing Professional Values and Business Values." *Financial Analysts Journal* 73 (2): 14–23.

CFA Institute. 2017. *Future State of the Investment Profession.* www.cfainstitute.org/learning/future/Documents/future_state_of_investment_profession.pdf

CFA Society UK.April 2016. The Value of the Investment Profession: A Report on Stakeholders' Views. www.cfauk.org/-/media/files/pdf/professionalism/value-of-the-investment-profession-report.pdf

Graham, Benjamin. 1949. *The Intelligent Investor*. New York: HarperCollins.

Graham, Benjamin, and David Dodd. 1934. *Security Analysis*. New York: McGraw-Hill.

Intelligence, Edelman. 2017. "*2017 Edelman Trust Barometer*" (17 January). https://www.edelman.com/trust2017/.

Sanders, Deen, and Alex Roberts. 2015. *Professionalisation of Financial Services.* White Paper. https://www.psc.gov.au/sites/default/files/NEW-PSC%20Whitepaper_final.pdf.

PRACTICE PROBLEMS

1 Benchmarks for minimally acceptable behaviors of community members are:

A a code of ethics.

B laws and regulations.

C standards of conduct.

2 Specialized knowledge and skills, a commitment to serve others, and a shared code of ethics *best* characterize a(n):

A vocation.

B profession.

C occupation.

3 Which of the following *best* identifies an internal trait that may lead to poor ethical decision making?

A Overconfidence

B Loyalty to employer

C Promise of money or prestige

4 Situational influences in decision making will *most likely* be minimized if:

A strong compliance programs are in place.

B longer-term consequences are considered.

C individuals believe they are truthful and honest.

5 Decision makers who use a compliance approach are *most likely* to:

A avoid situational influences.

B oversimplify decision making.

C consider more factors than when using an ethical decision-making approach.

6 When unethical behavior erodes trust in an investment firm, that firm is *more likely* to experience:

A lower revenues only.

B higher expenses only.

C lower revenues and higher expenses.

7 Which is an example of an activity that may be legal but that CFA Institute considers unethical?

A Making legally required disclosures in marketing materials

B Trading while in possession of material nonpublic information

C Disclosure by an employee of his or her own company's dishonest activity

8 An ethical decision-making framework will *most likely*:

A include a pre-determined, uniform sequence.

B focus exclusively on confirmable facts and relationships.

C help avoid a decision that has unanticipated ethical consequences.

9 High ethical standards are distinguishing features of which of the following bodies?

A Craft guilds

© 2017 CFA Institute. All rights reserved.

B Trade bodies

C Professional bodies

10 Fiduciary duty is a standard *most likely* to be upheld by members of a(n):

A employer.

B profession.

C not-for-profit body.

11 To maintain trust, the investment management profession must be interdependent with:

A regulators.

B employers.

C investment firms.

12 When an ethical dilemma occurs, an investment professional should *most likely* first raise the issue with a:

A mentor outside the firm.

B professional body's hotline.

C senior individual in the firm.

SOLUTIONS

1 C is correct. Standards of conduct are applied to specific communities or societal groups and identify specific behaviors required of community members. These standards of conduct serve as benchmarks for the minimally acceptable behavior of community members. Codes of ethics serve as a general guide for how community members should act; they communicate the organization's values and overall expectations regarding member behavior, but they do not identify specific behaviors required of community members. Laws and regulations are rules of conduct defined by governments and related entities about obligatory and forbidden conduct broadly applicable for individuals and entities under their jurisdiction.

2 B is correct. A profession has several characteristics that distinguish it from an occupation or vocation, such as specialized knowledge and skills, service to others, and a code of ethics shared by its members. A profession is the ultimate evolution of an occupation, resulting from excellence in practice, a mastery mindest, and expected adherence to a code of ethics and standards of practice.

3 A is correct. An overconfidence bias can lead individuals to put too much importance on internal traits and intrinsic motivations, such as their own perceptions of personal honesty, that can lead to faulty decision making. Loyalty to an employer and promise of money or prestige are situational influences that can lead to faulty decision making.

4 B is correct. Consciously considering long-term consequences will help offset situational influences. We more easily recognize and consider short-term situational influences than longer-term considerations because longer-term considerations have fewer immediate consequences than situational influences do. When decision making is too narrowly focused on short-term factors, we tend to ignore longer-term risks and consequences, and the likelihood of poor ethical decision making increases. A strong compliance policy is a good first step toward developing an ethical culture; a focus on rules adherence may not be sufficient. Emphasis on compliance may not encourage decision makers to consider the larger picture and can oversimplify decision making. Taken to the extreme, a strong compliance culture can become another situational influence that blinds employees to other important considerations. An overconfidence bias can place too much importance on internal traits and intrinsic motivations, such as "I'm honest and would not lie," even though studies have shown that internal traits are generally not the main determinant of whether or not someone will behave ethically in a given situation.

5 B is correct. A compliance approach can oversimplify decision making and may not encourage decision makers to consider the larger picture. A strong compliance culture may be a good start in developing an ethical culture but can become another situational influence that may result in employees failing to consider other important factors.

6 C is correct. Unethical behavior ultimately harms investment firms. Clients are not attracted if they suspect unethical behavior, leading to less business and lower revenues. Investment firms may also experience higher relative costs because regulators are more likely to have cause to initiate costly investigations.

7 B is correct. The investment industry has examples of conduct that may be legal but that CFA Institute considers unethical. Trading while in possession of material nonpublic information is not prohibited by law worldwide and can, therefore, be legal, but CFA Institute considers such trading unethical.

8 C is correct. Using an ethical decision-making framework consistently will help you develop sound judgment and decision-making skills and avoid making decisions that have unanticipated ethical consequences. The decision-making process is often iterative, and the decision maker may move between phases of the framework. A decision maker should consider more than confirmable facts and relationships; for example, the decision maker should consider situational influences and personal biases.

9 C is correct. High ethical standards distinguish professions from the craft guilds or trade bodies. Unlike trade bodies, professional bodies also typically have a mission to serve society and enforce professional conduct rules for practitioners.

10 B is correct. Fiduciary duty is an obligation to deliver a high standard of care when acting for the benefit of another party. Professionals must act in the best interest of the client, exercising a reasonable level of care, skill, and diligence. Other entities—including employers, regulators, trade associations, and not-for-profit bodies—may also support an industry but are not the same as professional bodies. Unlike professions, these other entities generally do not exist to set and maintain professional standards.

11 C is correct. The investment management profession and investment firms must be interdependent to maintain trust. Employers and regulators have their own standards and practices, which may differ from regulations and standards set by professional bodies.

12 C is correct. When a dilemma occurs, raising an issue internally with a senior employee is often a good starting place and creates an opportunity for an independent internal review. Protecting the client and the firm may take priority over the position of an individual professional raising a concern.

READING

57

Code of Ethics and Standards of Professional Conduct

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** describe the structure of the CFA Institute Professional Conduct Program and the process for the enforcement of the Code and Standards;
☐	**b.** identify the six components of the Code of Ethics and the seven Standards of Professional Conduct;
☐	**c.** explain the ethical responsibilities required by the Code and Standards, including the sub-sections of each Standard.

PREFACE

The *Standards of Practice Handbook* (*Handbook*) provides guidance to the people who grapple with real ethical dilemmas in the investment profession on a daily basis; the *Handbook* addresses the professional intersection where theory meets practice and where the concept of ethical behavior crosses from the abstract to the concrete. The *Handbook* is intended for a diverse and global audience: CFA Institute members navigating ambiguous ethical situations; supervisors and direct/indirect reports determining the nature of their responsibilities to each other, to existing and potential clients, and to the broader financial markets; and candidates preparing for the Chartered Financial Analyst (CFA) examinations.

Recent events in the global financial markets have tested the ethical mettle of financial market participants, including CFA Institute members. The standards taught in the CFA Program and by which CFA Institute members and candidates must abide represent timeless ethical principles and professional conduct for all market conditions. Through adherence to these standards, which continue to serve as the model for ethical behavior in the investment profession globally, each market participant does his or her part to improve the integrity and efficient operations of the financial markets.

The *Handbook* provides guidance in understanding the interconnectedness of the aspirational and practical principles and provisions of the Code of Ethics and Standards of Professional Conduct (Code and Standards). The Code contains high-level

This reading is a verbatim reprint of *Standards of Practice Handbook*, Eleventh Edition, updated March 2018, p. i to 12 included (www.cfainstitute.org/en/ethics/codes/standards-practice-handbook).

© 2021 CFA Institute. All rights reserved.

aspirational ethical principles that drive members and candidates to create a positive and reputable investment profession. The Standards contain practical ethical principles of conduct that members and candidates must follow to achieve the broader industry expectations. However, applying the principles individually may not capture the complexity of ethical requirements related to the investment industry. The Code and Standards should be viewed and interpreted as an interwoven tapestry of ethical requirements. Through members' and candidates' adherence to these principles as a whole, the integrity of and trust in the capital markets are improved.

Evolution of the CFA Institute Code of Ethics and Standards of Professional Conduct

Generally, changes to the Code and Standards over the years have been minor. CFA Institute has revised the language of the Code and Standards and occasionally added a new standard to address a prominent issue of the day. For instance, in 1992, CFA Institute added the standard addressing performance presentation to the existing list of standards.

Major changes came in 2005 with the ninth edition of the *Handbook*. CFA Institute adopted new standards, revised some existing standards, and reorganized the standards. The revisions were intended to clarify the requirements of the Code and Standards and effectively convey to its global membership what constitutes "best practice" in a number of areas relating to the investment profession.

The Code and Standards must be regularly reviewed and updated if they are to remain effective and continue to represent the highest ethical standards in the global investment industry. CFA Institute strongly believes that revisions of the Code and Standards are not undertaken for cosmetic purposes but to add value by addressing legitimate concerns and improving comprehension.

Changes to the Code and Standards have far-reaching implications for the CFA Institute membership, the CFA Program, and the investment industry as a whole. CFA Institute members and candidates are *required* to adhere to the Code and Standards. In addition, the Code and Standards are increasingly being adopted, in whole or in part, by firms and regulatory authorities. Their relevance goes well beyond CFA Institute members and candidates.

Standards of Practice Handbook

The periodic revisions of the Code and Standards have come in conjunction with updates of the *Standards of Practice Handbook*. The *Handbook* is the fundamental element of the ethics education effort of CFA Institute and the primary resource for guidance in interpreting and implementing the Code and Standards. The *Handbook* seeks to educate members and candidates on how to apply the Code and Standards to their professional lives and thereby benefit their clients, employers, and the investing public in general. The *Handbook* explains the purpose of the Code and Standards and how they apply in a variety of situations. The sections discuss and amplify each standard and suggest procedures to prevent violations.

Examples in the "Application of the Standard" sections are meant to illustrate how the standard applies to hypothetical but factual situations. The names contained in the examples are fictional and are not meant to refer to any actual person or entity. Unless otherwise stated (e.g., one or more people specifically identified), individuals in each example are CFA Institute members and holders of the CFA designation. Because factual circumstances vary so widely and often involve gray areas, the explanatory material and examples are not intended to be all inclusive. Many examples set forth in the application sections involve standards that have legal counterparts; ***members***

are strongly urged to discuss with their supervisors and legal and compliance departments the content of the Code and Standards and the members' general obligations under the Code and Standards.

CFA Institute recognizes that the presence of any set of ethical standards may create a false sense of security unless the documents are fully understood, enforced, and made a meaningful part of everyday professional activities. The *Handbook* is intended to provide a useful frame of reference that suggests ethical professional behavior in the investment decision-making process. This book cannot cover every contingency or circumstance, however, and it does not attempt to do so. The development and interpretation of the Code and Standards are evolving processes; the Code and Standards will be subject to continuing refinement.

Summary of Changes in the Eleventh Edition

The comprehensive review of the Code and Standards in 2005 resulted in principle requirements that remain applicable today. The review carried out for the eleventh edition focused on market practices that have evolved since the tenth edition. Along with updates to the guidance and examples within the *Handbook*, the eleventh edition includes an update to the Code of Ethics that embraces the members' role of maintaining the social contract between the industry and investors. Additionally, there are three changes to the Standards of Professional Conduct, which recognize the importance of proper supervision, clear communications with clients, and the expanding educational programs of CFA Institute.

Inclusion of Updated CFA Institute Mission

The CFA Institute Board of Governors approved an updated mission for the organization that is included in the Preamble to the Code and Standards. The new mission conveys the organization's conviction in the investment industry's role in the betterment of society at large.

> ***Mission:***
>
> To lead the investment profession globally by promoting the highest standards of ethics, education, and professional excellence for the ultimate benefit of society.

Updated Code of Ethics Principle

One of the bullets in the Code of Ethics was updated to reflect the role that the capital markets have in the greater society. As members work to promote and maintain the integrity of the markets, their actions should also help maintain the social contract with investors.

> ***Old:***
>
> Promote the integrity of and uphold the rules governing capital markets.
>
> ***New:***
>
> Promote the integrity and viability of the global capital markets for the ultimate benefit of society.

New Standard Regarding Responsibilities of Supervisors [IV(C)]

The standard for members and candidates with supervision or authority over others within their firms was updated to bring about improvements in preventing illegal and unethical actions from occurring. The prior version of Standard IV(C) focused on the detection and prevention of violations. The updated version stresses broader compliance expectations, which include the detection and prevention aspects of the original version.

> ***Old:***
>
> Members and Candidates must make reasonable efforts to detect and prevent violations of applicable laws, rules, regulations, and the Code and Standards by anyone subject to their supervision or authority.
>
> ***New:***
>
> Members and Candidates must make reasonable efforts to ensure that anyone subject to their supervision or authority complies with applicable laws, rules, regulations, and the Code and Standards.

Additional Requirement under the Standard for Communication with Clients and Prospective Clients [V(B)]

Given the constant development of new and exotic financial instruments and strategies, the standard regarding communicating with clients now includes an implicit requirement to discuss the risks and limitations of recommendations being made to clients. The new principle and related guidance take into account the fact that levels of disclosure will differ between products and services. Members and candidates, along with their firms, must determine the specific disclosures their clients should receive while ensuring appropriate transparency of the individual firms' investment processes.

> ***Addition:***
>
> Disclose to clients and prospective clients significant limitations and risks associated with the investment process.

Modification to Standard VII(A)

Since this standard was developed, CFA Institute has launched additional educational programs. The updated standard not only maintains the integrity of the CFA Program but also expands the same ethical considerations when members or candidates participate in such programs as the CIPM Program and the CFA Institute Investment Foundations certificate program. Whether participating as a member assisting with the curriculum or an examination or as a sitting candidate within a program, we expect them to engage in these programs as they would participate in the CFA Program.

> ***Old:***
>
> Conduct as Members and Candidates in the CFA Program
>
> Members and Candidates must not engage in any conduct that compromises the reputation or integrity of CFA Institute or the CFA designation or the integrity, validity, or security of the CFA examinations.

New:

> Conduct as Participants in CFA Institute Programs
>
> Members and Candidates must not engage in any conduct that compromises the reputation or integrity of CFA Institute or the CFA designation or the integrity, validity, or security of CFA Institute programs.

General Guidance and Example Revision

The guidance and examples were updated to reflect practices and scenarios applicable to today's investment industry. Two concepts that appear frequently in the updates in this edition relate to the increased use of social media for business communications and the use of and reliance on the output of quantitative models. The use of social media platforms has increased significantly since the publication of the tenth edition. And although financial modeling is not new to the industry, this update reflects upon actions that are viewed as possible contributing factors to the financial crises of the past decade.

CFA Institute Professional Conduct Program

All CFA Institute members and candidates enrolled in the CFA Program are required to comply with the Code and Standards. The CFA Institute Board of Governors maintains oversight and responsibility for the Professional Conduct Program (PCP), which, in conjunction with the Disciplinary Review Committee (DRC), is responsible for enforcement of the Code and Standards. The DRC is a volunteer committee of CFA charterholders who serve on panels to review conduct and partner with Professional Conduct staff to establish and review professional conduct policies. The CFA Institute Bylaws and Rules of Procedure for Professional Conduct (Rules of Procedure) form the basic structure for enforcing the Code and Standards. The Professional Conduct division is also responsible for enforcing testing policies of other CFA Institute education programs as well as the professional conduct of Certificate in Investment Performance Measurement (CIPM) certificants.

Professional Conduct inquiries come from a number of sources. First, members and candidates must self-disclose on the annual Professional Conduct Statement all matters that question their professional conduct, such as involvement in civil litigation or a criminal investigation or being the subject of a written complaint. Second, written complaints received by Professional Conduct staff can bring about an investigation. Third, CFA Institute staff may become aware of questionable conduct by a member or candidate through the media, regulatory notices, or another public source. Fourth, candidate conduct is monitored by proctors who complete reports on candidates suspected to have violated testing rules on exam day. Lastly, CFA Institute may also conduct analyses of scores and exam materials after the exam, as well as monitor online and social media to detect disclosure of confidential exam information.

When an inquiry is initiated, the Professional Conduct staff conducts an investigation that may include requesting a written explanation from the member or candidate; interviewing the member or candidate, complaining parties, and third parties; and collecting documents and records relevant to the investigation. Upon reviewing the material obtained during the investigation, the Professional Conduct staff may conclude the inquiry with no disciplinary sanction, issue a cautionary letter, or continue proceedings to discipline the member or candidate. If the Professional Conduct staff believes a violation of the Code and Standards or testing policies has occurred, the member or candidate has the opportunity to reject or accept any charges and the proposed sanctions.

If the member or candidate does not accept the charges and proposed sanction, the matter is referred to a panel composed of DRC members. Panels review materials and presentations from Professional Conduct staff and from the member or candidate. The panel's task is to determine whether a violation of the Code and Standards or testing policies occurred and, if so, what sanction should be imposed.

Sanctions imposed by CFA Institute may have significant consequences; they include public censure, suspension of membership and use of the CFA designation, and revocation of the CFA charter. Candidates enrolled in the CFA Program who have violated the Code and Standards or testing policies may be suspended or prohibited from further participation in the CFA Program.

Adoption of the Code and Standards

The Code and Standards apply to individual members of CFA Institute and candidates in the CFA Program. CFA Institute does encourage firms to adopt the Code and Standards, however, as part of their code of ethics. Those who claim compliance should fully understand the requirements of each of the principles of the Code and Standards.

Once a party—nonmember or firm—ensures its code of ethics meets the principles of the Code and Standards, that party should make the following statement whenever claiming compliance:

> "[Insert name of party] claims compliance with the CFA Institute Code of Ethics and Standards of Professional Conduct. This claim has not been verified by CFA Institute."

CFA Institute welcomes public acknowledgement, when appropriate, that firms are complying with the CFA Institute Code of Ethics and Standards of Professional Conduct and encourages firms to notify us of the adoption plans. For firms that would like to distribute the Code and Standards to clients and potential clients, attractive one-page copies of the Code and Standards, including translations, are available on the CFA Institute website (www.cfainstitute.org).

CFA Institute has also published the Asset Manager Code of Professional Conduct, which is designed, in part, to help asset managers comply with the regulations mandating codes of ethics for investment advisers. Whereas the Code and Standards are aimed at individual investment professionals who are members of CFA Institute or candidates in the CFA Program, the Asset Manager Code was drafted specifically for firms. The Asset Manager Code provides specific, practical guidelines for asset managers in six areas: loyalty to clients, the investment process, trading, compliance, performance evaluation, and disclosure. The Asset Manager Code and the appropriate steps to acknowledge adoption or compliance can be found on the CFA Institute website (www.cfainstitute.org).

Acknowledgments

CFA Institute is a not-for-profit organization that is heavily dependent on the expertise and intellectual contributions of member volunteers. Members devote their time because they share a mutual interest in the organization's mission to promote and achieve ethical practice in the investment profession. CFA Institute owes much to the volunteers' abundant generosity and energy in extending ethical integrity.

The CFA Institute Standards of Practice Council (SPC), a group consisting of CFA charterholder volunteers from many different countries, is charged with maintaining and interpreting the Code and Standards and ensuring that they are effective. The SPC draws its membership from a broad spectrum of organizations in the securities field, including brokers, investment advisers, banks, and insurance companies. In most instances, the SPC members have important supervisory responsibilities within their firms.

The SPC continually evaluates the Code and Standards, as well as the guidance in the *Handbook*, to ensure that they are

- representative of high standards of professional conduct,
- relevant to the changing nature of the investment profession,
- globally applicable,
- sufficiently comprehensive, practical, and specific,
- enforceable, and
- testable for the CFA Program.

The SPC has spent countless hours reviewing and discussing revisions to the Code and Standards and updates to the guidance that make up the eleventh edition of the *Handbook*. Following is a list of the current and former members of the SPC who generously donated their time and energy to this effort.

James E. Hollis III, CFA, Chair
Rik Albrecht, CFA
Terence E. Burns, CFA
Laura Dagan, CFA
Samuel B. Jones, Jr., CFA
Ulrike Kaiser-Boeing, CFA
Jinliang (Jack) Li, CFA
Christopher C. Loop, CFA,
James M. Meeth, CFA
Guy G. Rutherfurd, Jr., CFA
Edouard Senechal, CFA
Wenliang (Richard) Wang, CFA
Peng Lian Wee, CFA

ETHICS AND THE INVESTMENT INDUSTRY

Society ultimately benefits from efficient markets where capital can freely flow to the most productive or innovative destination. Well-functioning capital markets efficiently match those needing capital with those seeking to invest their assets in revenue-generating ventures. In order for capital markets to be efficient, investors must be able to trust that the markets are fair and transparent and offer them the opportunity to be rewarded for the risk they choose to take. Laws, regulations, and enforcement play a vital role but are insufficient alone to guarantee fair and transparent markets. The markets depend on an ethical foundation to guide participants' judgment and behavior. CFA Institute maintains and promotes the Code of Ethics and Standards of Professional Conduct in order to create a culture of ethics for the ultimate benefit of society.

Why Ethics Matters

Ethics can be defined as a set of moral principles or rules of conduct that provide guidance for our behavior when it affects others. Widely acknowledged fundamental ethical principles include honesty, fairness, diligence, and care and respect for others. Ethical conduct follows those principles and balances self-interest with both the direct and the indirect consequences of that behavior for other people.

Not only does unethical behavior by individuals have serious personal consequences—ranging from job loss and reputational damage to fines and even jail—but unethical conduct from market participants, investment professionals, and those who service investors can damage investor trust and thereby impair the sustainability of the global capital markets as a whole. Unfortunately, there seems to be an unending parade of stories bringing to light accounting frauds and manipulations, Ponzi schemes, insider-trading scandals, and other misdeeds. Not surprisingly, this has led to erosion in public confidence in investment professionals. Empirical evidence from numerous surveys documents the low standing in the eyes of the investing public of banks and financial services firms—the very institutions that are entrusted with the economic well-being and retirement security of society.

Governments and regulators have historically tried to combat misconduct in the industry through regulatory reform, with various levels of success. Global capital markets are highly regulated to protect investors and other market participants. However, compliance with regulation alone is insufficient to fully earn investor trust. Individuals and firms must develop a "culture of integrity" that permeates all levels of operations and promotes the ethical principles of stewardship of investor assets and working in the best interests of clients, above and beyond strict compliance with the law. A strong ethical culture that helps honest, ethical people engage in ethical behavior will foster the trust of investors, lead to robust global capital markets, and ultimately benefit society. That is why ethics matters.

Ethics, Society, and the Capital Markets

CFA Institute recently added the concept "for the ultimate benefit of society" to its mission. The premise is that we want to live in a socially, politically, and financially stable society that fosters individual well-being and welfare of the public. A key ingredient for this goal is global capital markets that facilitate the efficient allocation of resources so that the available capital finds its way to places where it most benefits that society. These investments are then used to produce goods and services, to fund innovation and jobs, and to promote improvements in standards of living. Indeed, such a function serves the interests of the society. Efficient capital markets, in turn, provide a host of benefits to those providing the investment capital. Investors are provided the opportunity to transfer and transform risk because the capital markets serve as an information exchange, create investment products, provide liquidity, and limit transaction costs.

However, a well-functioning and efficient capital market system is dependent on trust of the participants. If investors believe that capital market participants—investment professionals and firms—cannot be trusted with their financial assets or that the capital markets are unfair such that only insiders can be successful, they will be unlikely to invest or, at the very least, will require a higher risk premium. Decreased investment capital can reduce innovation and job creation and hurt the economy and society as a whole. Reduced trust in capital markets can also result in a less vibrant, if not smaller, investment industry.

Ethics for a global investment industry should be universal and ultimately support trust and integrity above acceptable local or regional customs and culture. Universal ethics for a global industry strongly supports the efficiency, values, and mission of the industry as a whole. Different countries may be at different stages of development in establishing standards of practice, but the end goal must be to achieve rules, regulations, and standards that support and promote fundamental ethical principles on a global basis.

Capital Market Sustainability and the Actions of One

Individuals and firms also have to look at the indirect impacts of their actions on the broader investment community. The increasingly interconnected nature of global finance brings to the fore an added consideration of market sustainability that was, perhaps, less appreciated in years past. In addition to committing to the highest levels of ethical behavior, today's investment professionals and their employers should consider the long-term health of the market as a whole.

As recent events have demonstrated, apparently isolated and unrelated decisions, however innocuous when considered on an individual basis, in aggregate can precipitate a market crisis. In an interconnected global economy and marketplace, each participant should strive to be aware of how his or her actions or the products he or she distributes may have an impact on capital market participants in other regions or countries.

Investment professionals should consider how their investment decision-making processes affect the global financial markets in the broader context of how they apply their ethical and professional obligations. Those in positions of authority have a special responsibility to consider the broader context of market sustainability in their development and approval of corporate policies, particularly those involving risk management and product development. In addition, corporate compensation strategies should not encourage otherwise ethically sound individuals to engage in unethical or questionable conduct for financial gain. Ethics, sustainability, and properly functioning capital markets are components of the same concept of protecting the best interests of all. To always place the interests of clients ahead of both investment professionals' own interests and those of their employer remains a key ethos.

The Relationship between Ethics and Regulations

Some equate ethical behavior with legal behavior: If you are following the law, you must be acting appropriately. Ethical principles, like laws and regulations, prescribe appropriate constraints on our natural tendency to pursue self-interest that could harm the interests of others. Laws and regulations often attempt to guide people toward ethical behavior, but they do not cover all unethical behavior. Ethical behavior is often distinguished from legal conduct by describing legal behavior as what is required and ethical behavior as conduct that is morally correct. Ethical principles go beyond that which is legally sufficient and encompass what is the right thing to do.

Given many regulators' lack of sufficient resources to enforce well-conceived rules and regulations, relying on a regulatory framework to lead the charge in establishing ethical behavior has its challenges. Therefore, reliance on compliance with laws and regulation alone is insufficient to ensure ethical behavior of investment professionals or to create a truly ethical culture in the industry.

The recent past has shown us that some individuals will succeed at circumventing the regulatory rules for their personal gain. Only the application of strong ethical principles, at both the individual level and the firm level, will limit abuses. Knowing the rules or regulations to apply in a particular situation, although important, may not be sufficient to ensure ethical conduct. Individuals must be able both to recognize areas that are prone to ethical pitfalls and to identify and process those circumstances and influences that can impair ethical judgment.

Applying an Ethical Framework

Laws, regulations, professional standards, and codes of ethics can guide ethical behavior, but individual judgment is a critical ingredient in making principled choices and engaging in appropriate conduct. When faced with an ethical dilemma, individuals must have a well-developed set of principles; otherwise, their thought processes can lead to, at best, equivocation and indecision and, at worst, fraudulent conduct and destruction of the public trust. Establishing an ethical framework for an internal thought process prior to deciding to act is a crucial step in engaging in ethical conduct.

Most investment professionals are used to making decisions from a business (profit/loss) outlook. But given the importance of ethical behavior in carrying out professional responsibilities, it is critical to also analyze decisions and potential conduct from an ethical perspective. Utilizing a framework for ethical decision making will help investment professionals effectively examine their conduct in the context of conflicting interests common to their professional obligations (e.g., researching and gathering information, developing investment recommendations, and managing money for others). Such a framework will allow investment professionals to analyze their conduct in a way that meets high standards of ethical behavior.

An ethical decision-making framework can come in many forms but should provide investment professionals with a tool for following the principles of the firm's code of ethics. Through analyzing the particular circumstances of each decision, investment professionals are able to determine the best course of action to fulfill their responsibilities in an ethical manner.

Commitment to Ethics by Firms

A firm's code of ethics risks becoming a largely ignored, dusty compilation if it is not truly integrated into the fabric of the business. The ability to relate an ethical decision-making framework to a firm's code of ethics allows investment professionals to bring the aspirations and principles of the code of ethics to life—transforming it from a compliance exercise to something that is at the heart of a firm's culture.

An investment professional's natural desire to "do the right thing" must be reinforced by building a culture of integrity in the workplace. Development, maintenance, and demonstration of a strong culture of integrity within the firm by senior management may be the single most important factor in promoting ethical behavior among the firm's employees. Adopting a code that clearly lays out the ethical principles that guide the thought processes and conduct the firm expects from its employees is a critical first step. But a code of ethics, while necessary, is insufficient.

Simply nurturing an inclination to do right is no match for the multitude of daily decisions that investment managers make. We need to exercise ethical decision-making skills to develop the muscle memory necessary for fundamentally ethical people to make good decisions despite the reality of agent conflicts. Just as coaching and practice transform our natural ability to run across a field into the technique and endurance required to run a race, teaching, reinforcing, and practicing ethical decision-making skills prepare us to confront the hard issues effectively. It is good for business, individuals, firms, the industry, and the markets, as well as society as a whole, to engage in the investment management profession in a highly ethical manner.

Ethical Commitment of CFA Institute

An important goal of CFA Institute is to ensure that the organization and its members and candidates develop, promote, and follow the highest ethical standards in the investment industry. The CFA Institute Code of Ethics (Code) and Standards of Professional Conduct (Standards) are the foundation supporting the organization's quest to uphold the industry's highest standards of individual and corporate practice and to help serve the greater good. The Code is a set of principles that define the overarching conduct CFA Institute expects from its members and CFA Program candidates. The Code works in tandem with the Standards, which outline professional conduct that constitutes fair and ethical business practices.

For more than 50 years, CFA Institute members and candidates have been required to abide by the organization's Code and Standards. Periodically, CFA Institute has revised and updated its Code and Standards to ensure that they remain relevant to the changing nature of the investment profession and representative of the highest standard of professional conduct. Within this *Handbook*, CFA Institute addresses ethical principles for the profession, including individual professionalism; responsibilities to capital markets, clients, and employers; ethics involved in investment analysis, recommendations, and actions; and possible conflicts of interest. Although the investment world has become a far more complex place since the first publication of the *Standard of Practice Handbook*, distinguishing right from wrong remains the paramount principle of the Code and Standards.

New challenges will continually arise for members and candidates in applying the Code and Standards because many decisions are not unambiguously right or wrong. The dilemma exists because the choice between right and wrong is not always clear. Even well-intentioned investment professionals can find themselves in circumstances that may tempt them to cut corners. Situational influences can overpower the best of intentions.

CFA Institute has made a significant commitment to providing members and candidates with the resources to extend and deepen their understanding of how to appropriately apply the principles of the Code and Standards. The product offerings from CFA Institute offer a wealth of material. Through publications, conferences, webcasts, and podcasts, the ethical challenges of investment professionals are brought to light. Archived issues of these items are available on the CFA Institute website (www.cfainstitute.org).

By reviewing these resources and discussing with their peers, market participants can further enhance their abilities to apply an effective ethical decision-making framework. In time, this should help restore some of the trust recently lost by investors.

Markets function to an important extent on trust. Recent events have shown the fragility of this foundation and the devastating consequences that can ensue when it is fundamentally questioned. Investment professionals should remain mindful of the long-term health of financial markets and incorporate this concern for the market's sustainability in their investment decision making. CFA Institute and the Standards of Practice Council hope this edition of the *Handbook* will assist and guide investment professionals in meeting the ethical demands of the highly interconnected global capital markets for the ultimate benefit of society.

CFA INSTITUTE CODE OF ETHICS AND STANDARDS OF PROFESSIONAL CONDUCT

Preamble

The CFA Institute Code of Ethics and Standards of Professional Conduct are fundamental to the values of CFA Institute and essential to achieving its mission to lead the investment profession globally by promoting the highest standards of ethics, education, and professional excellence for the ultimate benefit of society. High ethical standards are critical to maintaining the public's trust in financial markets and in the investment profession. Since their creation in the 1960s, the Code and Standards have promoted the integrity of CFA Institute members and served as a model for measuring the ethics of investment professionals globally, regardless of job function, cultural differences, or local laws and regulations. All CFA Institute members (including holders of the Chartered Financial Analyst [CFA] designation) and CFA candidates have the personal responsibility to embrace and uphold the provisions of the Code and Standards and are encouraged to notify their employer of this responsibility. Violations may result in disciplinary sanctions by CFA Institute. Sanctions can include revocation of membership, revocation of candidacy in the CFA Program, and revocation of the right to use the CFA designation.

The Code of Ethics

Members of CFA Institute (including CFA charterholders) and candidates for the CFA designation ("Members and Candidates") must:

- Act with integrity, competence, diligence, and respect and in an ethical manner with the public, clients, prospective clients, employers, employees, colleagues in the investment profession, and other participants in the global capital markets.
- Place the integrity of the investment profession and the interests of clients above their own personal interests.
- Use reasonable care and exercise independent professional judgment when conducting investment analysis, making investment recommendations, taking investment actions, and engaging in other professional activities.
- Practice and encourage others to practice in a professional and ethical manner that will reflect credit on themselves and the profession.
- Promote the integrity and viability of the global capital markets for the ultimate benefit of society.
- Maintain and improve their professional competence and strive to maintain and improve the competence of other investment professionals.

Standards of Professional Conduct

I. PROFESSIONALISM

A Knowledge of the Law

Members and Candidates must understand and comply with all applicable laws, rules, and regulations (including the CFA Institute Code of Ethics and Standards of Professional Conduct) of any government, regulatory organization, licensing agency, or professional association governing their professional activities. In the event of conflict, Members and Candidates must comply with the more strict law, rule, or regulation. Members and Candidates must not knowingly participate or assist in and must dissociate from any violation of such laws, rules, or regulations.

B Independence and Objectivity

Members and Candidates must use reasonable care and judgment to achieve and maintain independence and objectivity in their professional activities. Members and Candidates must not offer, solicit, or accept any gift, benefit, compensation, or consideration that reasonably could be expected to compromise their own or another's independence and objectivity.

C Misrepresentation

Members and Candidates must not knowingly make any misrepresentations relating to investment analysis, recommendations, actions, or other professional activities.

D Misconduct

Members and Candidates must not engage in any professional conduct involving dishonesty, fraud, or deceit or commit any act that reflects adversely on their professional reputation, integrity, or competence.

II. INTEGRITY OF CAPITAL MARKETS

A Material Nonpublic Information

Members and Candidates who possess material nonpublic information that could affect the value of an investment must not act or cause others to act on the information.

B Market Manipulation

Members and Candidates must not engage in practices that distort prices or artificially inflate trading volume with the intent to mislead market participants.

III. DUTIES TO CLIENTS

A Loyalty, Prudence, and Care

Members and Candidates have a duty of loyalty to their clients and must act with reasonable care and exercise prudent judgment. Members and Candidates must act for the benefit of their clients and place their clients' interests before their employer's or their own interests.

B Fair Dealing

Members and Candidates must deal fairly and objectively with all clients when providing investment analysis, making investment recommendations, taking investment action, or engaging in other professional activities.

C Suitability

1 When Members and Candidates are in an advisory relationship with a client, they must:

- **a** Make a reasonable inquiry into a client's or prospective client's investment experience, risk and return objectives, and financial constraints prior to making any investment recommendation or taking investment action and must reassess and update this information regularly.
- **b** Determine that an investment is suitable to the client's financial situation and consistent with the client's written objectives, mandates, and constraints before making an investment recommendation or taking investment action.
- **c** Judge the suitability of investments in the context of the client's total portfolio.

2 When Members and Candidates are responsible for managing a portfolio to a specific mandate, strategy, or style, they must make only investment recommendations or take only investment actions that are consistent with the stated objectives and constraints of the portfolio.

D Performance Presentation

When communicating investment performance information, Members and Candidates must make reasonable efforts to ensure that it is fair, accurate, and complete.

E Preservation of Confidentiality

Members and Candidates must keep information about current, former, and prospective clients confidential unless:

1 The information concerns illegal activities on the part of the client or prospective client,

2 Disclosure is required by law, or

3 The client or prospective client permits disclosure of the information.

IV. DUTIES TO EMPLOYERS

A Loyalty

In matters related to their employment, Members and Candidates must act for the benefit of their employer and not deprive their employer of the advantage of their skills and abilities, divulge confidential information, or otherwise cause harm to their employer.

B Additional Compensation Arrangements

Members and Candidates must not accept gifts, benefits, compensation, or consideration that competes with or might reasonably be expected to create a conflict of interest with their employer's interest unless they obtain written consent from all parties involved.

C Responsibilities of Supervisors

Members and Candidates must make reasonable efforts to ensure that anyone subject to their supervision or authority complies with applicable laws, rules, regulations, and the Code and Standards.

V. INVESTMENT ANALYSIS, RECOMMENDATIONS, AND ACTIONS

A Diligence and Reasonable Basis

Members and Candidates must:

1 Exercise diligence, independence, and thoroughness in analyzing investments, making investment recommendations, and taking investment actions.

2 Have a reasonable and adequate basis, supported by appropriate research and investigation, for any investment analysis, recommendation, or action.

B Communication with Clients and Prospective Clients

Members and Candidates must:

1 Disclose to clients and prospective clients the basic format and general principles of the investment processes they use to analyze investments, select securities, and construct portfolios and must promptly disclose any changes that might materially affect those processes.

2 Disclose to clients and prospective clients significant limitations and risks associated with the investment process.

3 Use reasonable judgment in identifying which factors are important to their investment analyses, recommendations, or actions and include those factors in communications with clients and prospective clients.

4 Distinguish between fact and opinion in the presentation of investment analysis and recommendations.

C Record Retention

Members and Candidates must develop and maintain appropriate records to support their investment analyses, recommendations, actions, and other investment-related communications with clients and prospective clients.

VI. CONFLICTS OF INTEREST

A Disclosure of Conflicts

Members and Candidates must make full and fair disclosure of all matters that could reasonably be expected to impair their independence and objectivity or interfere with respective duties to their clients, prospective clients, and employer. Members and Candidates must ensure that such disclosures are prominent, are delivered in plain language, and communicate the relevant information effectively.

B Priority of Transactions

Investment transactions for clients and employers must have priority over investment transactions in which a Member or Candidate is the beneficial owner.

C Referral Fees

Members and Candidates must disclose to their employer, clients, and prospective clients, as appropriate, any compensation, consideration, or benefit received from or paid to others for the recommendation of products or services.

VII. RESPONSIBILITIES AS A CFA INSTITUTE MEMBER OR CFA CANDIDATE

A Conduct as Participants in CFA Institute Programs

Members and Candidates must not engage in any conduct that compromises the reputation or integrity of CFA Institute or the CFA designation or the integrity, validity, or security of CFA Institute programs.

B Reference to CFA Institute, the CFA Designation, and the CFA Program

When referring to CFA Institute, CFA Institute membership, the CFA designation, or candidacy in the CFA Program, Members and Candidates must not misrepresent or exaggerate the meaning or implications of membership in CFA Institute, holding the CFA designation, or candidacy in the CFA Program.

PRACTICE PROBLEMS

1 The Standards of Practice Handbook provides guidance:

- **A** regarding the penalties incurred as a result of ethical violations.
- **B** to which all CFA Institute members and candidates must adhere.
- **C** through explanatory material and examples intended to be all inclusive.

2 Which of the following statements *best* describes an aspect of the Professional Conduct Program process?

- **A** Inquiries are not initiated in response to information provided by the media.
- **B** Investigations result in Disciplinary Review Committee panels for each case.
- **C** Investigations may include requesting a written explanation from the member or candidate.

3 A current Code of Ethics principle reads in full, "Promote the integrity:

- **A** and viability of the global capital markets."
- **B** of and uphold the rules governing capital markets."
- **C** and viability of the global capital markets for the ultimate benefit of society."

4 As stated in the revised 11th edition, the Standards of Professional Conduct:

- **A** require supervisors to focus on the detection and prevention of violations.
- **B** adopt separate ethical considerations for programs such as CIPM and Investment Foundations.
- **C** address the risks and limitations of recommendations being made to clients.

5 According to the Code of Ethics, members of CFA Institute and candidates for the CFA designation must:

- **A** maintain their professional competence to exercise independent professional judgment.
- **B** place the integrity of the investment profession and the interests of clients above their own personal interests.
- **C** practice in a professional and ethical manner with the public, clients, and others in the global capital markets.

6 Which of the following statements *best* describes an aspect of the Standards of Professional Conduct? Members and candidates are required to:

- **A** ensure any portfolio mandate followed is fair, accurate, and complete.
- **B** promptly disclose changes that might materially affect investment processes.
- **C** have a reasonable and adequate basis for decisions about client confidentiality.

7 Which of the following responses most completely represents an ethical principle of CFA Institute as outlined in the *Standards of Practice Handbook*?

- **A** Individual professionalism
- **B** Responsibilities to clients and employers
- **C** Ethics involved in investment analysis and recommendations

8 A CFA Institute member would violate the standard for material nonpublic information by:

- **A** conducting price distortion practices.

© 2019 CFA Institute. All rights reserved.

B inappropriately causing others to act.

C inadequately maintaining investment records.

9 According to the Duties to Clients standard, suitability requires members and candidates in an advisory relationship with a client to:

A place their clients' interests before their own interests.

B consider investments in the context of the client's total portfolio.

C not knowingly make misrepresentations relating to recommendations.

10 As part of the Duties to Clients standard, members and candidates must:

A document client financial constraints after an initial investment action.

B maintain an equal balance of interests owed to their clients and employers.

C deal fairly and objectively with all clients when engaging in professional activities.

11 The Duties to Employers standard states that members and candidates must not:

A accept any gifts that might compromise their independence and objectivity.

B deprive their employer of their skills and abilities as related to their employment.

C accept compensation competing with their employer's interest and with the written consent of all parties involved.

12 The Investment Analysis, Recommendations, and Actions standard states that members and candidates must:

A find an investment suitable for their client before making a recommendation.

B make reasonable efforts to ensure that performance presentation is fair, accurate, and complete.

C distinguish between fact and opinion in the presentation of investment analysis and recommendations.

13 Based on the Conflicts of Interest standard, members and candidates must:

A disclose, as required by law, those conflicts interfering with their professional duties.

B disclose, as appropriate, any benefit paid to others for the recommendation of products.

C seek employer approval before prioritizing their investment transactions over those clients.

14 The Responsibilities as a CFA Institute Member or CFA Candidate Standard explicitly states a requirement regarding:

A loyalty.

B responsibility of supervisors.

C reference to the CFA Program.

SOLUTIONS

1 B is correct. The *Standards of Practice Handbook* provides guidance to which CFA Institute members and candidates are required to adhere.

A is incorrect because the *Handbook* provides guidance in understanding the interconnectedness of the aspirational and practical principles (not regarding penalties for violations) of the Code of Ethics and Standards of Conduct.

C is incorrect because although the *Standards of Practice Handbook* provides hypothetical but factual situations, the explanatory material and examples are not intended to be all inclusive.

2 C is correct. When an inquiry is initiated, the Professional Conduct staff conducts an investigation that may include requesting a written explanation from the member or candidate.

A is incorrect because Professional Conduct inquiries can be initiated in response to information provided by the media. CFA Institute staff may become aware of questionable conduct by a member or candidate through the media, regulatory notices, or another public source.

B is incorrect because although the Disciplinary Review Committee (DRC) is responsible for enforcement of the Code and Standards in conjunction with the Professional Conduct Program (PCP), only in the event that a member or candidate does not accept the charges and proposed sanction is the matter referred to a panel composed of DRC members.

3 C is correct. One of the principles in the Code of Ethics was updated to reflect the role that the capital markets have in society as a whole.

A is incorrect because it is incomplete, missing the additional language to reflect the role that the capital markets have in society as a whole.

B is incorrect because this is the old principle as written in the Code of Ethics, which was recently updated to reflect the role of the capital markets in society as a whole.

4 C is correct. Given the constant development of new and exotic financial instruments and strategies, the standard regarding communicating with clients now includes an implicit requirement to discuss the risks and limitations of recommendations being made to clients.

A is incorrect because the updated standard for members and candidates with supervision or authority over others within their firms stresses broader compliance expectations, which include the detection and prevention aspects of the original version that was the prior focus.

B is incorrect because the updated standard not only maintains the integrity of the CFA Program but also expands the same (not separate) ethical considerations when members or candidates participate in such programs as the CIPM Program and the Investment Foundations Certificate.

5 B is correct. Members of CFA Institute and candidates for the CFA designation must place the integrity of the investment profession and the interests of clients above their own personal interests.

A is incorrect because members of CFA Institute and candidates for the CFA designation must maintain and improve their professional competence and strive to maintain and improve the competence of other investment professionals. The exercise of independent professional judgment is associated with using reasonable care.

C is incorrect because members of CFA Institute and candidates for the CFA designation must practice and encourage others to practice in a professional and ethical manner that will reflect credibly on themselves and the profession. Members are supposed to act with integrity, competence, diligence, and respect and in an ethical manner with the public, clients, and other market participants.

6 B is correct. The current Standards of Professional Conduct requires members and candidates to promptly disclose any changes that might materially affect investment processes.

A is incorrect because under Standard III.C.2 Suitability, when members and candidates are responsible for managing a portfolio according to a specific mandate, they must take only investment actions that are consistent with the stated objectives of the portfolio. The "fair, accurate, and complete" criterion relates to the Standard III D Performance Presentation.

C is incorrect because under Standard III.E.1, 2, 3 Preservation of Confidentiality, members and candidates must keep information about current clients confidential unless the information concerns illegal activities on the part of the client, disclosure is required by law, or the client permits disclosure. No decisions on confidentiality are required, with the "reasonable and adequate basis" criterion related to Standard V.A.2 Diligence and Reasonable Basis.

7 A is correct. Within the Standards of Practice Handbook, CFA Institute addresses ethical principles for the profession in the following Standards: individual professionalism; integrity in capital markets; responsibilities to clients, responsibilities to employers; ethics involved in investment analysis, recommendations, and actions; and possible conflicts of interest. B is incorrect because it represents, and combines, two ethical principles, those relating to the Standards "Duties to Clients" and "Duties to Employers." C is incorrect because the ethical principle (and Standard) relating to ethics in investment analysis and recommendations also includes actions.

8 B is correct. Under Standard II.A Material Nonpublic Information, members having material nonpublic information that could affect the value of an investment must not cause others to act on the information.

A is incorrect because price distortion is mentioned in the Standard II.B Market Manipulation, not Standard II.A Material Nonpublic Information.

C is incorrect because the maintenance of appropriate records to support investment analyses is noted in Standard V.C Record Retention, not Standard II.A Material Nonpublic Information.

9 B is correct. Standard III.C.1c Suitability states that when members and candidates are in an advisory relationship with a client, they must judge the suitability of investments in the context of the client's total portfolio.

A is incorrect because this is a requirement addressed under Standard III.A Loyalty, Prudence, and Care, not Standard III.C.1c Suitability.

C is incorrect because this is a requirement addressed under Standard I.C Misrepresentation, not Standard III.C.1c Suitability.

10 C is correct. Under the III.B Fair Dealing section of the Duty to Clients standard, members and candidates must deal fairly and objectively with all clients when providing investment analysis, making investment recommendations, taking investment action, or engaging in other professional activities.

A is incorrect because under Standard III.C.1a Suitability, a section of Duties to Clients, members and candidates in an advisory relationship must make a reasonable inquiry into a client's financial constraints prior to (not after) taking investment action and must reassess and update this regularly.

B is incorrect because under Standard III.A Loyalty, Prudence, and Care, members and candidates must act for the benefit of their clients and place their clients' interests before (not maintain an equal balance with) their employer's or their own interests.

11 B is correct. The IV.A Loyalty section of the Duties to Employers standard states that members and candidates cannot deprive their employer of the advantage of their skills and abilities in matters related to their employment.

A is incorrect because accepting gifts that might compromise a member or candidate's independence and objectivity is addressed by Standard I.B Independence and Objectivity, a section of Professionalism, not under Standard IV Duties to Employers.

C is incorrect because IV.B Additional Compensation Arrangements, part of the Duties to Employers standard, permits members and candidates to accept compensation that competes with their employer's interest if they obtain written consent from all parties involved.

12 C is correct. The V.B.4 Communications with Clients and Prospective Clients section of the Investment Analysis, Recommendations, and Actions standard states that members and candidates must distinguish between fact and opinion in the presentation of investment analysis and recommendations.

A is incorrect because this standard is discussed in the III.C.1b Suitability section of the Duties to Clients standard.

B is incorrect because performance presentation is discussed in the III.D Performance Presentation section of the Duties to Clients standard.

13 B is correct. The VI.C Referral Fees section of the Conflicts of Interest standard requires members and candidates to disclose to their employer, clients, and prospective clients, as appropriate, any compensation, consideration, or benefit received from or paid to others for the recommendation of products or services.

A is incorrect because the VI.A Disclosure of Conflicts section of the Conflicts of Interest standard requires members and candidates to make full and fair disclosure of all matters (not limited to legal requirements) that could reasonably be expected to impair their independence and objectivity or interfere with respective duties to their clients, prospective clients, and employer.

C is incorrect because the VI.B Priority of Transactions section of the Conflicts of Interest standard requires members and candidates to give priority to investment transactions for clients and employers versus those in which a member or candidate is the beneficial owner. This requirement is not waived by an employer's approval.

14 C is correct. The VII.B Reference to CFA Institute, the CFA Designation, and the CFA Program section of the Responsibilities as a CFA Institute Member or CFA Candidate standard explicitly states the appropriate manner to make reference to CFA Institute, CFA Institute membership, the CFA designation, or candidacy in the CFA Program.

A is incorrect because Standard VII Responsibilities as a CFA Institute Member or CFA Candidate standard does not refer to loyalty. Loyalty is addressed in two other standards, Standard III.A Loyalty, Prudence, and Care and Standard IV.A Loyalty.

B is incorrect because Standard VII Responsibilities as a CFA Institute Member or CFA Candidate standard does not refer to the responsibility of supervisors. The responsibility of supervisors is addressed in Standard IV.C Responsibility of Supervisors.

READING

58

Guidance for Standards I–VII

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;
☐	**b.** identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;
☐	**c.** recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

STANDARD I(A): PROFESSIONALISM - KNOWLEDGE OF THE LAW

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Standard I(A) Knowledge of the Law

Members and Candidates must understand and comply with all applicable laws, rules, and regulations (including the CFA Institute Code of Ethics and Standards of Professional Conduct) of any government, regulatory organization, licensing agency, or professional association governing their professional activities. In the event of conflict, Members and Candidates must comply with the more strict law, rule, or regulation. Members and Candidates must not knowingly participate or assist in and must dissociate from any violation of such laws, rules, or regulations.

This reading is a verbatim reprint of *Standards of Practice Handbook*, Eleventh Edition, updated March 2018, p. 13 to 174 included (www.cfainstitute.org/en/ethics/codes/standards-practice-handbook).

© 2014 CFA Institute. All rights reserved.

Guidance

Highlights:

- *Relationship between the Code and Standards and Applicable Law*
- *Participation in or Association with Violations by Others*
- *Investment Products and Applicable Laws*

Members and candidates must understand the applicable laws and regulations of the countries and jurisdictions where they engage in professional activities. These activities may include, but are not limited to, trading of securities or other financial instruments, providing investment advice, conducting research, or performing other investment services. On the basis of their reasonable and good faith understanding, members and candidates must comply with the laws and regulations that directly govern their professional activities and resulting outcomes and that protect the interests of the clients.

When questions arise, members and candidates should know their firm's policies and procedures for accessing compliance guidance. This standard does not require members and candidates to become experts, however, in compliance. Additionally, members and candidates are not required to have detailed knowledge of or be experts on all the laws that could potentially govern their activities.

During times of changing regulations, members and candidates must remain vigilant in maintaining their knowledge of the requirements for their professional activities. New financial products and processes, along with uncovered ethical missteps, create an environment for recurring and potentially wide-ranging regulatory changes. Members and candidates are also continually provided improved and enhanced methods of communicating with both clients and potential clients, such as mobile applications and web-based social networking platforms. As new local, regional, and global requirements are updated to address these and other changes, members, candidates, and their firms must adjust their procedures and practices to remain in compliance.

Relationship between the Code and Standards and Applicable Law

Some members or candidates may live, work, or provide investment services to clients living in a country that has no law or regulation governing a particular action or that has laws or regulations that differ from the requirements of the Code and Standards. When applicable law and the Code and Standards require different conduct, members and candidates must follow the more strict of the applicable law or the Code and Standards.

"Applicable law" is the law that governs the member's or candidate's conduct. Which law applies will depend on the particular facts and circumstances of each case. The "more strict" law or regulation is the law or regulation that imposes greater restrictions on the action of the member or candidate or calls for the member or candidate to exert a greater degree of action that protects the interests of investors. For example, applicable law or regulation may not require members and candidates to disclose referral fees received from or paid to others for the recommendation of investment products or services. Because the Code and Standards impose this obligation, however, members and candidates must disclose the existence of such fees.

Members and candidates must adhere to the following principles:

- Members and candidates must comply with applicable laws or regulations related to their professional activities.

- Members and candidates must not engage in conduct that constitutes a violation of the Code and Standards, even though it may otherwise be legal.
- In the absence of any applicable law or regulation or when the Code and Standards impose a higher degree of responsibility than applicable laws and regulations, members and candidates must adhere to the Code and Standards. Applications of these principles are outlined in Exhibit 1.

The applicable laws governing the responsibilities of a member or candidate should be viewed as the minimal threshold of acceptable actions. When members and candidates take actions that exceed the minimal requirements, they further support the conduct required of Standard I(A).

CFA Institute members are obligated to abide by the CFA Institute Articles of Incorporation, Bylaws, Code of Ethics, Standards of Professional Conduct, Rules of Procedure, Membership Agreement, and other applicable rules promulgated by CFA Institute, all as amended periodically. CFA candidates who are not members must also abide by these documents (except for the Membership Agreement) as well as rules and regulations related to the administration of the CFA examination, the Candidate Responsibility Statement, and the Candidate Pledge.

Participation in or Association with Violations by Others

Members and candidates are responsible for violations in which they *knowingly* participate or assist. Although members and candidates are presumed to have knowledge of all applicable laws, rules, and regulations, CFA Institute acknowledges that members may not recognize violations if they are not aware of all the facts giving rise to the violations. Standard I(A) applies when members and candidates know or should know that their conduct may contribute to a violation of applicable laws, rules, or regulations or the Code and Standards.

If a member or candidate has reasonable grounds to believe that imminent or ongoing client or employer activities are illegal or unethical, the member or candidate must dissociate, or separate, from the activity. In extreme cases, dissociation may require a member or candidate to leave his or her employment. Members and candidates may take the following intermediate steps to dissociate from ethical violations of others when direct discussions with the person or persons committing the violation are unsuccessful. The first step should be to attempt to stop the behavior by bringing it to the attention of the employer through a supervisor or the firm's compliance department. If this attempt is unsuccessful, then members and candidates have a responsibility to step away and dissociate from the activity. Dissociation practices will differ on the basis of the member's or candidate's role in the investment industry. It may include removing one's name from written reports or recommendations, asking for a different assignment, or refusing to accept a new client or continue to advise a current client. Inaction combined with continuing association with those involved in illegal or unethical conduct may be construed as participation or assistance in the illegal or unethical conduct.

CFA Institute strongly encourages members and candidates to report potential violations of the Code and Standards committed by fellow members and candidates. Although a failure to report is less likely to be construed as a violation than a failure to dissociate from unethical conduct, the impact of inactivity on the integrity of capital markets can be significant. Although the Code and Standards do not compel members and candidates to report violations to their governmental or regulatory organizations unless such disclosure is mandatory under applicable law (voluntary reporting is often referred to as whistleblowing), such disclosure may be prudent under certain circumstances. Members and candidates should consult their legal and compliance advisers for guidance.

Additionally, CFA Institute encourages members, nonmembers, clients, and the investing public to report violations of the Code and Standards by CFA Institute members or CFA candidates by submitting a complaint in writing to the CFA Institute Professional Conduct Program via e-mail (pcprogram@cfainstitute.org) or the CFA Institute website (www.cfainstitute.org).

Investment Products and Applicable Laws

Members and candidates involved in creating or maintaining investment services or investment products or packages of securities and/or derivatives should be mindful of where these products or packages will be sold as well as their places of origination. The applicable laws and regulations of the countries or regions of origination and expected sale should be understood by those responsible for the supervision of the services or creation and maintenance of the products or packages. Members or candidates should make reasonable efforts to review whether associated firms that are distributing products or services developed by their employing firm also abide by the laws and regulations of the countries and regions of distribution. Members and candidates should undertake the necessary due diligence when transacting cross-border business to understand the multiple applicable laws and regulations in order to protect the reputation of their firm and themselves.

Given the complexity that can arise with business transactions in today's market, there may be some uncertainty surrounding which laws or regulations are considered applicable when activities are being conducted in multiple jurisdictions. Members and candidates should seek the appropriate guidance, potentially including the firm's compliance or legal departments and legal counsel outside the organization, to gain a reasonable understanding of their responsibilities and how to implement them appropriately.

Exhibit 1 Global Application of the Code and Standards

Members and candidates who practice in multiple jurisdictions may be subject to varied securities laws and regulations. If applicable law is stricter than the requirements of the Code and Standards, members and candidates must adhere to applicable law; otherwise, they must adhere to the Code and Standards. The following chart provides illustrations involving a member who may be subject to the securities laws and regulations of three different types of countries:

NS: country with no securities laws or regulations

LS: country with ***less*** strict securities laws and regulations than the Code and Standards

MS: country with ***more*** strict securities laws and regulations than the Code and Standards

Applicable Law	Duties	Explanation
Member resides in NS country, does business in LS country; LS law applies.	Member must adhere to the Code and Standards.	Because applicable law is less strict than the Code and Standards, the member must adhere to the Code and Standards.
Member resides in NS country, does business in MS country; MS law applies.	Member must adhere to the law of MS country.	Because applicable law is stricter than the Code and Standards, member must adhere to the more strict applicable law.

Exhibit 1 (Continued)

Applicable Law	Duties	Explanation
Member resides in LS country, does business in NS country; LS law applies.	Member must adhere to the Code and Standards.	Because applicable law is less strict than the Code and Standards, member must adhere to the Code and Standards.
Member resides in LS country, does business in MS country; MS law applies.	Member must adhere to the law of MS country.	Because applicable law is stricter than the Code and Standards, member must adhere to the more strict applicable law.
Member resides in LS country, does business in NS country; LS law applies, but it states that law of locality where business is conducted governs.	Member must adhere to the Code and Standards.	Because applicable law states that the law of the locality where the business is conducted governs and there is no local law, the member must adhere to the Code and Standards.
Member resides in LS country, does business in MS country; LS law applies, but it states that law of locality where business is conducted governs.	Member must adhere to the law of MS country.	Because applicable law of the locality where the business is conducted governs and local law is stricter than the Code and Standards, member must adhere to the more strict applicable law.
Member resides in MS country, does business in LS country; MS law applies.	Member must adhere to the law of MS country.	Because applicable law is stricter than the Code and Standards, member must adhere to the more strict applicable law.
Member resides in MS country, does business in LS country; MS law applies, but it states that law of locality where business is conducted governs.	Member must adhere to the Code and Standards.	Because applicable law states that the law of the locality where the business is conducted governs and local law is less strict than the Code and Standards, member must adhere to the Code and Standards.

(continued)

Exhibit 1 (Continued)

Applicable Law	Duties	Explanation
Member resides in MS country, does business in LS country with a client who is a citizen of LS country; MS law applies, but it states that the law of the client's home country governs.	Member must adhere to the Code and Standards.	Because applicable law states that the law of the client's home country governs (which is less strict than the Code and Standards), member must adhere to the Code and Standards.
Member resides in MS country, does business in LS country with a client who is a citizen of MS country; MS law applies, but it states that the law of the client's home country governs.	Member must adhere to the law of MS country.	Because applicable law states that the law of the client's home country governs and the law of the client's home country is stricter than the Code and Standards, the member must adhere to the more strict applicable law.

STANDARD I(A): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Members and Candidates

Suggested methods by which members and candidates can acquire and maintain understanding of applicable laws, rules, and regulations include the following:

- *Stay informed*: Members and candidates should establish or encourage their employers to establish a procedure by which employees are regularly informed about changes in applicable laws, rules, regulations, and case law. In many instances, the employer's compliance department or legal counsel can provide such information in the form of memorandums distributed to employees in the organization. Also, participation in an internal or external continuing education program is a practical method of staying current.
- *Review procedures*: Members and candidates should review, or encourage their employers to review, the firm's written compliance procedures on a regular basis to ensure that the procedures reflect current law and provide adequate guidance to employees about what is permissible conduct under the law and/or the Code and Standards. Recommended compliance procedures for specific items of the Code and Standards are discussed in this *Handbook* in the "Guidance" sections associated with each standard.
- *Maintain current files*: Members and candidates should maintain or encourage their employers to maintain readily accessible current reference copies of applicable statutes, rules, regulations, and important cases.

Distribution Area Laws

Members and candidates should make reasonable efforts to understand the applicable laws—both country and regional—for the countries and regions where their investment products are developed and are most likely to be distributed to clients.

Legal Counsel

When in doubt about the appropriate action to undertake, it is recommended that a member or candidate seek the advice of compliance personnel or legal counsel concerning legal requirements. If a potential violation is being committed by a fellow employee, it may also be prudent for the member or candidate to seek the advice of the firm's compliance department or legal counsel.

Dissociation

When dissociating from an activity that violates the Code and Standards, members and candidates should document the violation and urge their firms to attempt to persuade the perpetrator(s) to cease such conduct. To dissociate from the conduct, a member or candidate may have to resign his or her employment.

Firms

The formality and complexity of compliance procedures for firms depend on the nature and size of the organization and the nature of its investment operations. Members and candidates should encourage their firms to consider the following policies and procedures to support the principles of Standard I(A):

- *Develop and/or adopt a code of ethics*: The ethical culture of an organization starts at the top. Members and candidates should encourage their supervisors or managers to adopt a code of ethics. Adhering to a code of ethics facilitates solutions when people face ethical dilemmas and can prevent the need for employees to resort to a "whistleblowing" solution publicly alleging concealed misconduct. CFA Institute has published the *Asset Manager Code of Professional Conduct*, which firms may adopt or use as the basis for their codes (visit www.cfainstitute.org).
- *Provide information on applicable laws*: Pertinent information that highlights applicable laws and regulations might be distributed to employees or made available in a central location. Information sources might include primary information developed by the relevant government, governmental agencies, regulatory organizations, licensing agencies, and professional associations (e.g., from their websites); law firm memorandums or newsletters; and association memorandums or publications (e.g., *CFA Institute Magazine*).
- *Establish procedures for reporting violations*: Firms might provide written protocols for reporting suspected violations of laws, regulations, or company policies.

STANDARD I(A): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Notification of Known Violations):

Michael Allen works for a brokerage firm and is responsible for an underwriting of securities. A company official gives Allen information indicating that the financial statements Allen filed with the regulator overstate the issuer's earnings. Allen seeks the advice of the brokerage firm's general counsel, who states that it would be difficult for the regulator to prove that Allen has been involved in any wrongdoing.

> *Comment*: Although it is recommended that members and candidates seek the advice of legal counsel, the reliance on such advice does not absolve a member or candidate from the requirement to comply with the law or regulation. Allen should report this situation to his supervisor, seek an independent legal opinion, and determine whether the regulator should be notified of the error.

Example 2 (Dissociating from a Violation):

Lawrence Brown's employer, an investment banking firm, is the principal underwriter for an issue of convertible debentures by the Courtney Company. Brown discovers that the Courtney Company has concealed severe third-quarter losses in its foreign operations. The preliminary prospectus has already been distributed.

> *Comment*: Knowing that the preliminary prospectus is misleading, Brown should report his findings to the appropriate supervisory persons in his firm. If the matter is not remedied and Brown's employer does not dissociate from the underwriting, Brown should sever all his connections with the underwriting. Brown should also seek legal advice to determine whether additional reporting or other action should be taken.

Example 3 (Dissociating from a Violation):

Kamisha Washington's firm advertises its past performance record by showing the 10-year return of a composite of its client accounts. Washington discovers, however, that the composite omits the performance of accounts that have left the firm during the 10-year period, whereas the description of the composite indicates the inclusion of all firm accounts. This omission has led to an inflated performance figure. Washington is asked to use promotional material that includes the erroneous performance number when soliciting business for the firm.

> *Comment*: Misrepresenting performance is a violation of the Code and Standards. Although she did not calculate the performance herself, Washington would be assisting in violating Standard I(A) if she were to use the inflated performance number when soliciting clients. She must dissociate herself from the activity. If discussing the misleading number with the person responsible is not an option for correcting the problem, she

> can bring the situation to the attention of her supervisor or the compliance department at her firm. If her firm is unwilling to recalculate performance, she must refrain from using the misleading promotional material and should notify the firm of her reasons. If the firm insists that she use the material, she should consider whether her obligation to dissociate from the activity requires her to seek other employment.

Example 4 (Following the Highest Requirements):

James Collins is an investment analyst for a major Wall Street brokerage firm. He works in a developing country with a rapidly modernizing economy and a growing capital market. Local securities laws are minimal—in form and content—and include no punitive prohibitions against insider trading.

> *Comment*: Collins must abide by the requirements of the Code and Standards, which might be more strict than the rules of the developing country. He should be aware of the risks that a small market and the absence of a fairly regulated flow of information to the market represent to his ability to obtain information and make timely judgments. He should include this factor in formulating his advice to clients. In handling material nonpublic information that accidentally comes into his possession, he must follow Standard II(A)–Material Nonpublic Information.

Example 5 (Following the Highest Requirements):

Laura Jameson works for a multinational investment adviser based in the United States. Jameson lives and works as a registered investment adviser in the tiny, but wealthy, island nation of Karramba. Karramba's securities laws state that no investment adviser registered and working in that country can participate in initial public offerings (IPOs) for the adviser's personal account. Jameson, believing that, as a US citizen working for a US-based company, she should comply only with US law, has ignored this Karrambian law. In addition, Jameson believes that as a charterholder, as long as she adheres to the Code and Standards requirement that she disclose her participation in any IPO to her employer and clients when such ownership creates a conflict of interest, she is meeting the highest ethical requirements.

> *Comment*: Jameson is in violation of Standard I(A). As a registered investment adviser in Karramba, Jameson is prevented by Karrambian securities law from participating in IPOs regardless of the law of her home country. In addition, because the law of the country where she is working is stricter than the Code and Standards, she must follow the stricter requirements of the local law rather than the requirements of the Code and Standards.

Example 6 (Laws and Regulations Based on Religious Tenets):

Amanda Janney is employed as a fixed-income portfolio manager for a large international firm. She is on a team within her firm that is responsible for creating and managing a fixed-income hedge fund to be sold throughout the firm's distribution centers to high-net-worth clients. Her firm receives expressions of interest from potential clients from the Middle East who are seeking investments that comply with Islamic law. The marketing and promotional materials for the fixed-income hedge fund do not specify whether or not the fund is a suitable investment for an investor seeking compliance with Islamic law. Because the fund is being distributed globally, Janney is concerned about the reputation of the fund and the firm and believes disclosure of whether or not the fund complies with Islamic law could help minimize potential mistakes with placing this investment.

> *Comment*: As the financial market continues to become globalized, members and candidates will need to be aware of the differences between cultural and religious laws and requirements as well as the different governmental laws and regulations. Janney and the firm could be proactive in their efforts to acknowledge areas where the new fund may not be suitable for clients.

Example 7 (Reporting Potential Unethical Actions):

Krista Blume is a junior portfolio manager for high-net-worth portfolios at a large global investment manager. She observes a number of new portfolios and relationships coming from a country in Europe where the firm did not have previous business and is told that a broker in that country is responsible for this new business. At a meeting on allocation of research resources to third-party research firms, Blume notes that this broker has been added to the list and is allocated payments for research. However, she knows the portfolios do not invest in securities in the broker's country, and she has not seen any research come from this broker. Blume asks her supervisor about the name being on the list and is told that someone in marketing is receiving the research and that the name being on the list is OK. She believes that what may be going on is that the broker is being paid for new business through the inappropriate research payments, and she wishes to dissociate from the misconduct.

> *Comment*: Blume should follow the firm's policies and procedures for reporting potential unethical activity, which may include discussions with her supervisor or someone in a designated compliance department. She should communicate her concerns appropriately while advocating for disclosure between the new broker relationship and the research payments.

Example 8 (Failure to Maintain Knowledge of the Law):

Colleen White is excited to use new technology to communicate with clients and potential clients. She recently began posting investment information, including performance reports and investment opinions and recommendations, to her Facebook page. In addition, she sends out brief announcements, opinions, and thoughts via her Twitter account (for example, "Prospects for future growth of XYZ company look good! #makingmoney4U"). Prior to White's use of these social media platforms, the local regulator had issued new requirements and guidance governing online electronic communication. White's communications appear to conflict with the recent regulatory announcements.

> *Comment*: White is in violation of Standard I(A) because her communications do not comply with the existing guidance and regulation governing use of social media. White must be aware of the evolving legal requirements pertaining to new and dynamic areas of the financial services industry that are applicable to her. She should seek guidance from appropriate, knowledgeable, and reliable sources, such as her firm's compliance department, external service providers, or outside counsel, unless she diligently follows legal and regulatory trends affecting her professional responsibilities.

STANDARD I(B): PROFESSIONALISM - INDEPENDENCE AND OBJECTIVITY

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Members and Candidates must use reasonable care and judgment to achieve and maintain independence and objectivity in their professional activities. Members and Candidates must not offer, solicit, or accept any gift, benefit, compensation, or consideration that reasonably could be expected to compromise their own or another's independence and objectivity.

Guidance

Highlights:

- *Buy-Side Clients*
- *Fund Manager and Custodial Relationships*
- *Investment Banking Relationships*
- *Performance Measurement and Attribution*
- *Public Companies*
- *Credit Rating Agency Opinions*
- *Influence during the Manager Selection/Procurement Process*
- *Issuer-Paid Research*
- *Travel Funding*

Standard I(B) states the responsibility of CFA Institute members and candidates in the CFA Program to maintain independence and objectivity so that their clients will have the benefit of their work and opinions unaffected by any potential conflict of interest or other circumstance adversely affecting their judgment. Every member and candidate should endeavor to avoid situations that could cause or be perceived to cause a loss of independence or objectivity in recommending investments or taking investment action.

External sources may try to influence the investment process by offering analysts and portfolio managers a variety of benefits. Corporations may seek expanded research coverage, issuers and underwriters may wish to promote new securities offerings, brokers may want to increase commission business, and independent rating agencies may be influenced by the company requesting the rating. Benefits may include gifts, invitations to lavish functions, tickets, favors, or job referrals. One type of benefit is the allocation of shares in oversubscribed IPOs to investment managers for their personal accounts. This practice affords managers the opportunity to make quick profits that may not be available to their clients. Such a practice is prohibited under Standard I(B). Modest gifts and entertainment are acceptable, but special care must be taken by members and candidates to resist subtle and not-so-subtle pressures to act in conflict with the interests of their clients. Best practice dictates that members and candidates reject any offer of gift or entertainment that could be expected to threaten their independence and objectivity.

Receiving a gift, benefit, or consideration from a *client* can be distinguished from gifts given by entities seeking to influence a member or candidate to the detriment of other clients. In a client relationship, the client has already entered some type of compensation arrangement with the member, candidate, or his or her firm. A gift from a client could be considered supplementary compensation. The potential for obtaining influence to the detriment of other clients, although present, is not as great as in situations where no compensation arrangement exists. When possible, prior to accepting "bonuses" or gifts from clients, members and candidates should disclose to their employers such benefits offered by clients. If notification is not possible prior to acceptance, members and candidates must disclose to their employer benefits previously accepted from clients. Disclosure allows the employer of a member or candidate to make an independent determination about the extent to which the gift may affect the member's or candidate's independence and objectivity.

Members and candidates may also come under pressure from their own firms to, for example, issue favorable research reports or recommendations for certain companies with potential or continuing business relationships with the firm. The situation may be aggravated if an executive of the company sits on the bank or investment firm's board and attempts to interfere in investment decision making. Members and candidates acting in a sales or marketing capacity must be especially mindful of their objectivity in promoting appropriate investments for their clients.

Left unmanaged, pressures that threaten independence place research analysts in a difficult position and may jeopardize their ability to act independently and objectively. One of the ways that research analysts have coped with these pressures in the past is to use subtle and ambiguous language in their recommendations or to temper the tone of their research reports. Such subtleties are lost on some investors, however, who reasonably expect research reports and recommendations to be straightforward and transparent and to communicate clearly an analyst's views based on unbiased analysis and independent judgment.

Members and candidates are personally responsible for maintaining independence and objectivity when preparing research reports, making investment recommendations, and taking investment action on behalf of clients. Recommendations must convey the member's or candidate's true opinions, free of bias from internal or external pressures, and be stated in clear and unambiguous language.

Members and candidates also should be aware that some of their professional or social activities within CFA Institute or its member societies may subtly threaten their independence or objectivity. When seeking corporate financial support for conventions, seminars, or even weekly society luncheons, the members or candidates responsible for the activities should evaluate both the actual effect of such solicitations on their independence and whether their objectivity might be perceived to be compromised in the eyes of their clients.

Buy-Side Clients

One source of pressure on sell-side analysts is buy-side clients. Institutional clients are traditionally the primary users of sell-side research, either directly or with soft dollar brokerage. Portfolio managers may have significant positions in the security of a company under review. A rating downgrade may adversely affect the portfolio's performance, particularly in the short term, because the sensitivity of stock prices to ratings changes has increased in recent years. A downgrade may also affect the manager's compensation, which is usually tied to portfolio performance. Moreover, portfolio performance is subject to media and public scrutiny, which may affect the manager's professional reputation. Consequently, some portfolio managers implicitly or explicitly support sell-side ratings inflation.

Portfolio managers have a responsibility to respect and foster the intellectual honesty of sell-side research. Therefore, it is improper for portfolio managers to threaten or engage in retaliatory practices, such as reporting sell-side analysts to the covered company in order to instigate negative corporate reactions. Although most portfolio managers do not engage in such practices, the perception by the research analyst that a reprisal is possible may cause concern and make it difficult for the analyst to maintain independence and objectivity.

Fund Manager and Custodial Relationships

Research analysts are not the only people who must be concerned with maintaining their independence. Members and candidates who are responsible for hiring and retaining outside managers and third-party custodians should not accepts gifts, entertainment, or travel funding that may be perceived as impairing their decisions. The use of secondary fund managers has evolved into a common practice to manage specific asset allocations. The use of third-party custodians is common practice for independent investment advisory firms and helps them with trading capabilities and reporting requirements. Primary and secondary fund managers, as well as third-party custodians, often arrange educational and marketing events to inform others about their business strategies, investment process, or custodial services. Members and candidates must review the merits of each offer individually in determining whether they may attend yet maintain their independence.

Investment Banking Relationships

Some sell-side firms may exert pressure on their analysts to issue favorable research reports on current or prospective investment banking clients. For many of these firms, income from investment banking has become increasingly important to overall firm profitability because brokerage income has declined as a result of price competition. Consequently, firms offering investment banking services work hard to develop and maintain relationships with investment banking clients and prospects. These companies are often covered by the firm's research analysts because companies often select their investment banks on the basis of the reputation of their research analysts, the quality of their work, and their standing in the industry.

In some countries, research analysts frequently work closely with their investment banking colleagues to help evaluate prospective investment banking clients. In other countries, because of past abuses in managing the obvious conflicts of interest, regulators have established clear rules prohibiting the interaction of these groups. Although collaboration between research analysts and investment banking colleagues may benefit the firm and enhance market efficiency (e.g., by allowing firms to assess risks more accurately and make better pricing assumptions), it requires firms to carefully balance the conflicts of interest inherent in the collaboration. Having analysts work with investment bankers is appropriate only when the conflicts are adequately and effectively managed and disclosed. Firm managers have a responsibility to provide an environment in which analysts are neither coerced nor enticed into issuing research that does not reflect their true opinions. Firms should require public disclosure of actual conflicts of interest to investors.

Members, candidates, and their firms must adopt and follow perceived best practices in maintaining independence and objectivity in the corporate culture and protecting analysts from undue pressure by their investment banking colleagues. The "firewalls" traditionally built between these two functions must be managed to minimize conflicts of interest; indeed, enhanced firewall policies may go as far as prohibiting all communications between these groups. A key element of an enhanced firewall is separate reporting structures for personnel on the research side and personnel on the investment banking side. For example, investment banking personnel should not have any authority to approve, disapprove, or make changes to research reports or recommendations. Another element should be a compensation arrangement that minimizes the pressures on research analysts and rewards objectivity and accuracy. Compensation arrangements should not link analyst remuneration directly to investment banking assignments in which the analyst may participate as a team member. Firms should also regularly review their policies and procedures to determine whether analysts are adequately safeguarded and to improve the transparency of disclosures relating to conflicts of interest. The highest level of transparency is achieved when disclosures are prominent and specific rather than marginalized and generic.

Performance Measurement and Attribution

Members and candidates working within a firm's investment performance measurement department may also be presented with situations that challenge their independence and objectivity. As performance analysts, their analyses may reveal instances where managers may appear to have strayed from their mandate. Additionally, the performance analyst may receive requests to alter the construction of composite indexes owing to negative results for a selected account or fund. The member or candidate must not allow internal or external influences to affect their independence and objectivity as they faithfully complete their performance calculation and analysis-related responsibilities.

Public Companies

Analysts may be pressured to issue favorable reports and recommendations by the companies they follow. Not every stock is a "buy," and not every research report is favorable—for many reasons, including the cyclical nature of many business activities and market fluctuations. For instance, a "good company" does not always translate into a "good stock" rating if the current stock price is fully valued. In making an investment recommendation, the analyst is responsible for anticipating, interpreting, and assessing a company's prospects and stock price performance in a factual manner. Many company managers, however, believe that their company's stock is undervalued, and these managers may find it difficult to accept critical research reports or ratings downgrades. Company managers' compensation may also be dependent on stock performance.

Due diligence in financial research and analysis involves gathering information from a wide variety of sources, including public disclosure documents (such as proxy statements, annual reports, and other regulatory filings) and also company management and investor-relations personnel, suppliers, customers, competitors, and other relevant sources. Research analysts may justifiably fear that companies will limit their ability to conduct thorough research by denying analysts who have "negative" views direct access to company managers and/or barring them from conference calls and other communication venues. Retaliatory practices include companies bringing legal action against analysts personally and/or their firms to seek monetary damages for the economic effects of negative reports and recommendations. Although few companies engage in such behavior, the perception that a reprisal is possible is a reasonable concern for analysts. This concern may make it difficult for them to conduct the comprehensive research needed to make objective recommendations. For further information and guidance, members and candidates should refer to the CFA Institute publication *Best Practice Guidelines Governing Analyst/Corporate Issuer Relations* (www.cfainstitute.org).

Credit Rating Agency Opinions

Credit rating agencies provide a service by grading the fixed-income products offered by companies. Analysts face challenges related to incentives and compensation schemes that may be tied to the final rating and successful placement of the product.

Members and candidates employed at rating agencies should ensure that procedures and processes at the agencies prevent undue influences from a sponsoring company during the analysis. Members and candidates should abide by their agencies' and the industry's standards of conduct regarding the analytical process and the distribution of their reports.

The work of credit rating agencies also raises concerns similar to those inherent in investment banking relationships. Analysts may face pressure to issue ratings at a specific level because of other services the agency offers companies—namely, advising on the development of structured products. The rating agencies need to develop the necessary firewalls and protections to allow the independent operations of their different business lines.

When using information provided by credit rating agencies, members and candidates should be mindful of the potential conflicts of interest. And because of the potential conflicts, members and candidates may need to independently validate the rating granted.

Influence during the Manager Selection/Procurement Process

Members and candidates may find themselves on either side of the manager selection process. An individual may be on the hiring side as a representative of a pension organization or an investment committee member of an endowment or a charitable organization. Additionally, other members may be representing their organizations in attempts to earn new investment allocation mandates. The responsibility of members and candidates to maintain their independence and objectivity extends to the hiring or firing of those who provide business services beyond investment management.

When serving in a hiring capacity, members and candidates should not solicit gifts, contributions, or other compensation that may affect their independence and objectivity. Solicitations do not have to benefit members and candidates personally to conflict with Standard I(B). Requesting contributions to a favorite charity or political organization may also be perceived as an attempt to influence the decision-making process. Additionally, members and candidates serving in a hiring capacity should refuse gifts, donations, and other offered compensation that may be perceived to influence their decision-making process.

When working to earn a new investment allocation, members and candidates should not offer gifts, contributions, or other compensation to influence the decision of the hiring representative. The offering of these items with the intent to impair the independence and objectivity of another person would not comply with Standard I(B). Such prohibited actions may include offering donations to a charitable organization or political candidate referred by the hiring representative.

A clear example of improperly influencing hiring representatives was displayed in the "pay-to-play" scandal involving government-sponsored pension funds in the United States. Managers looking to gain lucrative allocations from the large funds made requested donations to the political campaigns of individuals directly responsible for the hiring decisions. This scandal and other similar events have led to new laws requiring additional reporting concerning political contributions and bans on hiring—or hiring delays for—managers that made campaign contributions to representatives associated with the decision-making process.

Issuer-Paid Research

In light of the recent reduction of sell-side research coverage, many companies, seeking to increase visibility both in the financial markets and with potential investors, have hired analysts to produce research reports analyzing their companies. These reports bridge the gap created by the lack of coverage and can be an effective method of communicating with investors.

Issuer-paid research conducted by independent analysts, however, is fraught with potential conflicts. Depending on how the research is written and distributed, investors may be misled into believing that the research is from an independent source when, in reality, it has been paid for by the subject company.

Members and candidates must adhere to strict standards of conduct that govern how the research is to be conducted and what disclosures must be made in the report. Analysts must engage in thorough, independent, and unbiased analysis and must fully disclose potential conflicts of interest, including the nature of their compensation. Otherwise, analysts risk misleading investors.

Investors need clear, credible, and thorough information about companies, and they need research based on independent thought. At a minimum, issuer-paid research should include a thorough analysis of the company's financial statements based on publicly disclosed information, benchmarking within a peer group, and industry analysis. Analysts must exercise diligence, independence, and thoroughness in conducting their research in an objective manner. Analysts must distinguish between fact and opinion in their reports. Conclusions must have a reasonable and adequate basis and must be supported by appropriate research.

Independent analysts must also strictly limit the type of compensation that they accept for conducting issuer-paid research. Otherwise, the content and conclusions of the reports could reasonably be expected to be determined or affected by compensation from the sponsoring companies. Compensation that might influence the research report could be direct, such as payment based on the conclusions of the report, or indirect, such as stock warrants or other equity instruments that could increase in value on the basis of positive coverage in the report. In such instances, the independent analyst has an incentive to avoid including negative information or making negative conclusions. Best practice is for independent analysts, prior to writing their reports, to negotiate only a flat fee for their work that is not linked to their conclusions or recommendations.

Travel Funding

The benefits related to accepting paid travel extend beyond the cost savings to the member or candidate and his firm, such as the chance to talk exclusively with the executives of a company or learning more about the investment options provided by an investment organization. Acceptance also comes with potential concerns; for example, members and candidates may be influenced by these discussions when flying on a corporate or chartered jet or attending sponsored conferences where many expenses, including airfare and lodging, are covered. To avoid the appearance of compromising their independence and objectivity, best practice dictates that members and candidates always use commercial transportation at their expense or at the expense of their firm rather than accept paid travel arrangements from an outside company. Should commercial transportation be unavailable, members and candidates may accept modestly arranged travel to participate in appropriate information-gathering events, such as a property tour.

STANDARD I(B): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Members and candidates should adhere to the following practices and should encourage their firms to establish procedures to avoid violations of Standard I(B):

- *Protect the integrity of opinions*: Members, candidates, and their firms should establish policies stating that every research report concerning the securities of a corporate client should reflect the unbiased opinion of the analyst. Firms should also design compensation systems that protect the integrity of the investment decision process by maintaining the independence and objectivity of analysts.
- *Create a restricted list*: If the firm is unwilling to permit dissemination of adverse opinions about a corporate client, members and candidates should encourage the firm to remove the controversial company from the research universe and put it on a restricted list so that the firm disseminates only factual information about the company.
- *Restrict special cost arrangements*: When attending meetings at an issuer's headquarters, members and candidates should pay for commercial transportation and hotel charges. No corporate issuer should reimburse members or candidates for air transportation. Members and candidates should encourage issuers to limit the use of corporate aircraft to situations in which commercial transportation is not available or in which efficient movement could not otherwise be arranged. Members and candidates should take particular care that when frequent meetings are held between an individual issuer and an individual member or candidate, the issuer should not always host the member or candidate.
- *Limit gifts*: Members and candidates must limit the acceptance of gratuities and/or gifts to token items. Standard I(B) does not preclude customary, ordinary business-related entertainment as long as its purpose is not to influence or reward members or candidates. Firms should consider a strict value limit for acceptable gifts that is based on the local or regional customs and should address whether the limit is per gift or an aggregate annual value.
- *Restrict investments*: Members and candidates should encourage their investment firms to develop formal policies related to employee purchases of equity or equity-related IPOs. Firms should require prior approval for employee participation in IPOs, with prompt disclosure of investment actions taken following the offering. Strict limits should be imposed on investment personnel acquiring securities in private placements.
- *Review procedures*: Members and candidates should encourage their firms to implement effective supervisory and review procedures to ensure that analysts and portfolio managers comply with policies relating to their personal investment activities.
- *Independence policy*: Members, candidates, and their firms should establish a formal written policy on the independence and objectivity of research and implement reporting structures and review procedures to ensure that research analysts do not report to and are not supervised or controlled by any department of the firm that could compromise the independence of the analyst. More detailed recommendations related to a firm's policies regarding research objectivity are set forth in the CFA Institute statement *Research Objectivity Standards* (www.cfainstitute.org).
- *Appointed officer*: Firms should appoint a senior officer with oversight responsibilities for compliance with the firm's code of ethics and all regulations concerning its business. Firms should provide every employee with the procedures and policies for reporting potentially unethical behavior, violations of regulations, or other activities that may harm the firm's reputation.

STANDARD I(B): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Travel Expenses):

Steven Taylor, a mining analyst with Bronson Brokers, is invited by Precision Metals to join a group of his peers in a tour of mining facilities in several western US states. The company arranges for chartered group flights from site to site and for accommodations in Spartan Motels, the only chain with accommodations near the mines, for three nights. Taylor allows Precision Metals to pick up his tab, as do the other analysts, with one exception—John Adams, an employee of a large trust company who insists on following his company's policy and paying for his hotel room himself.

> *Comment*: The policy of the company where Adams works complies closely with Standard I(B) by avoiding even the appearance of a conflict of interest, but Taylor and the other analysts were not necessarily violating Standard I(B). In general, when allowing companies to pay for travel and/or accommodations in these circumstances, members and candidates must use their judgment. They must be on guard that such arrangements not impinge on a member's or candidate's independence and objectivity. In this example, the trip was strictly for business and Taylor was not accepting irrelevant or lavish hospitality. The itinerary required chartered flights, for which analysts were not expected to pay. The accommodations were modest. These arrangements are not unusual and did not violate Standard I(B) as long as Taylor's independence and objectivity were not compromised. In the final analysis, members and candidates should consider both whether they can remain objective and whether their integrity might be perceived by their clients to have been compromised.

Example 2 (Research Independence):

Susan Dillon, an analyst in the corporate finance department of an investment services firm, is making a presentation to a potential new business client that includes the promise that her firm will provide full research coverage of the potential client.

> *Comment*: Dillon may agree to provide research coverage, but she must not commit her firm's research department to providing a favorable recommendation. The firm's recommendation (favorable, neutral, or unfavorable) must be based on an independent and objective investigation and analysis of the company and its securities.

Example 3 (Research Independence and Intrafirm Pressure):

Walter Fritz is an equity analyst with Hilton Brokerage who covers the mining industry. He has concluded that the stock of Metals & Mining is overpriced at its current level, but he is concerned that a negative research report will hurt the good relationship between Metals & Mining and the investment banking division of his firm. In fact, a senior manager of Hilton Brokerage has just sent him a copy of a proposal his firm has made to Metals & Mining to underwrite a debt offering. Fritz needs to produce a report right away and is concerned about issuing a less-than-favorable rating.

> *Comment*: Fritz's analysis of Metals & Mining must be objective and based solely on consideration of company fundamentals. Any pressure from other divisions of his firm is inappropriate. This conflict could have been eliminated if, in anticipation of the offering, Hilton Brokerage had placed Metals & Mining on a restricted list for its sales force.

Example 4 (Research Independence and Issuer Relationship Pressure):

As in Example 3, Walter Fritz has concluded that Metals & Mining stock is overvalued at its current level, but he is concerned that a negative research report might jeopardize a close rapport that he has nurtured over the years with Metals & Mining's CEO, chief finance officer, and investment relations officer. Fritz is concerned that a negative report might result also in management retaliation—for instance, cutting him off from participating in conference calls when a quarterly earnings release is made, denying him the ability to ask questions on such calls, and/or denying him access to top management for arranging group meetings between Hilton Brokerage clients and top Metals & Mining managers.

> *Comment*: As in Example 3, Fritz's analysis must be objective and based solely on consideration of company fundamentals. Any pressure from Metals & Mining is inappropriate. Fritz should reinforce the integrity of his conclusions by stressing that his investment recommendation is based on relative valuation, which may include qualitative issues with respect to Metals & Mining's management.

Example 5 (Research Independence and Sales Pressure):

As support for the sales effort of her corporate bond department, Lindsey Warner offers credit guidance to purchasers of fixed-income securities. Her compensation is closely linked to the performance of the corporate bond department. Near the quarter's end, Warner's firm has a large inventory position in the bonds of Milton, Ltd., and has been unable to sell the bonds because of Milton's recent announcement of an operating problem. Salespeople have asked her to contact large clients to push the bonds.

> *Comment*: Unethical sales practices create significant potential violations of the Code and Standards. Warner's opinion of the Milton bonds must not be affected by internal pressure or compensation. In this case, Warner must refuse to push the Milton bonds unless she is able to justify that the market price has already adjusted for the operating problem.

Example 6 (Research Independence and Prior Coverage):

Jill Jorund is a securities analyst following airline stocks and a rising star at her firm. Her boss has been carrying a "buy" recommendation on International Airlines and asks Jorund to take over coverage of that airline. He tells Jorund that under no circumstances should the prevailing buy recommendation be changed.

> *Comment*: Jorund must be independent and objective in her analysis of International Airlines. If she believes that her boss's instructions have compromised her, she has two options: She can tell her boss that she cannot cover the company under these constraints, or she can take over coverage of the company, reach her own independent conclusions, and if they conflict with her boss's opinion, share the conclusions with her boss or other supervisors in the firm so that they can make appropriate recommendations. Jorund must issue only recommendations that reflect her independent and objective opinion.

Example 7 (Gifts and Entertainment from Related Party):

Edward Grant directs a large amount of his commission business to a New York–based brokerage house. In appreciation for all the business, the brokerage house gives Grant two tickets to the World Cup in South Africa, two nights at a nearby resort, several meals, and transportation via limousine to the game. Grant fails to disclose receiving this package to his supervisor.

> *Comment*: Grant has violated Standard I(B) because accepting these substantial gifts may impede his independence and objectivity. Every member and candidate should endeavor to avoid situations that might cause or be perceived to cause a loss of independence or objectivity in recommending investments or taking investment action. By accepting the trip, Grant has opened himself up to the accusation that he may give the broker favored treatment in return.

Example 8 (Gifts and Entertainment from Client):

Theresa Green manages the portfolio of Ian Knowlden, a client of Tisbury Investments. Green achieves an annual return for Knowlden that is consistently better than that of the benchmark she and the client previously agreed to. As a reward, Knowlden offers Green two tickets to Wimbledon and the use of Knowlden's flat in London for a week. Green discloses this gift to her supervisor at Tisbury.

> *Comment*: Green is in compliance with Standard I(B) because she disclosed the gift from one of her clients in accordance with the firm's policies. Members and candidates may accept bonuses or gifts from clients as long as they disclose them to their employer because gifts in a client relationship are deemed less likely to affect a member's or candidate's objectivity and independence than gifts in other situations. Disclosure is required, however, so that supervisors can monitor such situations to guard against employees favoring a gift-giving client to the detriment of other fee-paying clients (such as by allocating a greater proportion of IPO stock to the gift-giving client's portfolio).
>
> Best practices for monitoring include comparing the transaction costs of the Knowlden account with the costs of other accounts managed by Green and other similar accounts within Tisbury. The supervisor could also compare the performance returns with the returns of other clients with the same mandate. This comparison will assist in determining whether a pattern of favoritism by Green is disadvantaging other Tisbury clients or the possibility that this favoritism could affect her future behavior.

Example 9 (Travel Expenses from External Manager):

Tom Wayne is the investment manager of the Franklin City Employees Pension Plan. He recently completed a successful search for a firm to manage the foreign equity allocation of the plan's diversified portfolio. He followed the plan's standard procedure of seeking presentations from a number of qualified firms and recommended that his board select Penguin Advisors because of its experience, well-defined investment strategy, and performance record. The firm claims compliance with the Global Investment Performance Standards (GIPS) and has been verified. Following the selection of Penguin, a reporter from the *Franklin City Record* calls to ask if there was any connection between this action and the fact that Penguin was one of the sponsors of an "investment fact-finding trip to Asia" that Wayne made earlier in the year. The trip was one of several conducted by the Pension Investment Academy, which had arranged the itinerary of meetings with economic, government, and corporate officials in major cities in several Asian countries. The Pension Investment Academy obtains support for the cost of these trips from a number of investment managers, including Penguin Advisors; the Academy then pays the travel expenses of the various pension plan managers on the trip and provides all meals and accommodations. The president of Penguin Advisors was also one of the travelers on the trip.

> *Comment*: Although Wayne can probably put to good use the knowledge he gained from the trip in selecting portfolio managers and in other areas of managing the pension plan, his recommendation of Penguin Advisors may be tainted by the possible conflict incurred when he participated in a trip partly paid for by Penguin Advisors and when he was in the daily company of the president of Penguin Advisors. To avoid violating Standard I(B), Wayne's basic expenses for travel and accommodations should have been paid by his employer or the pension plan; contact with the president of Penguin Advisors should have been limited to informational or educational events only; and the trip, the organizer, and the sponsor should have been made a matter of public record. Even if his actions were not in violation of Standard I(B), Wayne should have been sensitive to the public perception of the trip when reported in the newspaper and the extent to which the subjective elements of his decision might have been affected by the familiarity that the daily contact of such a trip would encourage. This advantage would probably not be shared by firms competing with Penguin Advisors.

Example 10 (Research Independence and Compensation Arrangements):

Javier Herrero recently left his job as a research analyst for a large investment adviser. While looking for a new position, he was hired by an investor-relations firm to write a research report on one of its clients, a small educational software company. The investor-relations firm hopes to generate investor interest in the technology company. The firm will pay Herrero a flat fee plus a bonus if any new investors buy stock in the company as a result of Herrero's report.

> *Comment*: If Herrero accepts this payment arrangement, he will be in violation of Standard I(B) because the compensation arrangement can reasonably be expected to compromise his independence and objectivity. Herrero will receive a bonus for attracting investors, which provides an incentive to draft a positive report regardless of the facts and to ignore or play down any negative information about the company. Herrero should accept only a flat fee that is not tied to the conclusions or recommendations

> of the report. Issuer-paid research that is objective and unbiased can be done under the right circumstances as long as the analyst takes steps to maintain his or her objectivity and includes in the report proper disclosures regarding potential conflicts of interest.

Example 11 (Recommendation Objectivity and Service Fees):

Two years ago, Bob Wade, trust manager for Central Midas Bank, was approached by Western Funds about promoting its family of funds, with special interest in the service-fee class of funds. To entice Central to promote this class, Western Funds offered to pay the bank a service fee of 0.25%. Without disclosing the fee being offered to the bank, Wade asked one of the investment managers to review Western's funds to determine whether they were suitable for clients of Central Midas Bank. The manager completed the normal due diligence review and determined that the new funds were fairly valued in the market with fee structures on a par with competitors. Wade decided to accept Western's offer and instructed the team of portfolio managers to exclusively promote these funds and the service-fee class to clients seeking to invest new funds or transfer from their current investments.

Now, two years later, the funds managed by Western begin to underperform their peers. Wade is counting on the fees to reach his profitability targets and continues to push these funds as acceptable investments for Central's clients.

> *Comment*: Wade is violating Standard I(B) because the fee arrangement has affected the objectivity of his recommendations. Wade is relying on the fee as a component of the department's profitability and is unwilling to offer other products that may affect the fees received.
>
> See also Standard VI(A)–Disclosure of Conflicts.

Example 12 (Recommendation Objectivity):

Bob Thompson has been doing research for the portfolio manager of the fixed-income department. His assignment is to do sensitivity analysis on securitized subprime mortgages. He has discussed with the manager possible scenarios to use to calculate expected returns. A key assumption in such calculations is housing price appreciation (HPA) because it drives "prepays" (prepayments of mortgages) and losses. Thompson is concerned with the significant appreciation experienced over the previous five years as a result of the increased availability of funds from subprime mortgages. Thompson insists that the analysis should include a scenario run with −10% for Year 1, −5% for Year 2, and then (to project a worst-case scenario) 0% for Years 3 through 5. The manager replies that these assumptions are too dire because there has never been a time in their available database when HPA was negative.

Thompson conducts his research to better understand the risks inherent in these securities and evaluates these securities in the worst-case scenario, an unlikely but possible environment. Based on the results of the enhanced scenarios, Thompson does not recommend the purchase of the securitization. Against the general market trends, the manager follows Thompson's recommendation and does not invest. The following year, the housing market collapses. In avoiding the subprime investments, the manager's portfolio outperforms its peer group that year.

> *Comment*: Thompson's actions in running the worst-case scenario against the protests of the portfolio manager are in alignment with the principles of Standard I(B). Thompson did not allow his research to be pressured by the general trends of the market or the manager's desire to limit the research to historical norms.

> See also Standard V(A)–Diligence and Reasonable Basis.

Example 13 (Influencing Manager Selection Decisions):

Adrian Mandel, CFA, is a senior portfolio manager for ZZYY Capital Management who oversees a team of investment professionals who manage labor union pension funds. A few years ago, ZZYY sought to win a competitive asset manager search to manage a significant allocation of the pension fund of the United Doughnut and Pretzel Bakers Union (UDPBU). UDPBU's investment board is chaired by a recognized key decision maker and long-time leader of the union, Ernesto Gomez. To improve ZZYY's chances of winning the competition, Mandel made significant monetary contributions to Gomez's union reelection campaign fund. Even after ZZYY was hired as a primary manager of the pension, Mandel believed that his firm's position was not secure. Mandel continued to contribute to Gomez's reelection campaign chest as well as to entertain lavishly the union leader and his family at top restaurants on a regular basis. All of Mandel's outlays were routinely handled as marketing expenses reimbursed by ZZYY's expense accounts and were disclosed to his senior management as being instrumental in maintaining a strong close relationship with an important client.

> *Comment*: Mandel not only offered but actually gave monetary gifts, benefits, and other considerations that reasonably could be expected to compromise Gomez's objectivity. Therefore, Mandel was in violation of Standard I(B).

Example 14 (Influencing Manager Selection Decisions):

Adrian Mandel, CFA, had heard about the manager search competition for the UDPBU Pension Fund through a broker/dealer contact. The contact told him that a well-known retired professional golfer, Bobby "The Bear" Finlay, who had become a licensed broker/dealer serving as a pension consultant, was orchestrating the UDPBU manager search. Finlay had gained celebrity status with several labor union pension fund boards by entertaining their respective board members and regaling them with colorful stories of fellow pro golfers' antics in clubhouses around the world. Mandel decided to improve ZZYY's chances of being invited to participate in the search competition by befriending Finlay to curry his favor. Knowing Finlay's love of entertainment, Mandel wined and dined Finlay at high-profile bistros where Finlay could glow in the fan recognition lavished on him by all the other patrons. Mandel's endeavors paid off handsomely when Finlay recommended to the UDPBU board that ZZYY be entered as one of three finalist asset management firms in its search.

> *Comment*: Similar to Example 13, Mandel lavished gifts, benefits, and other considerations in the form of expensive entertainment that could reasonably be expected to influence the consultant to recommend the hiring of his firm. Therefore, Mandel was in violation of Standard I(B).

Example 15 (Fund Manager Relationships):

Amie Scott is a performance analyst within her firm with responsibilities for analyzing the performance of external managers. While completing her quarterly analysis, Scott notices a change in one manager's reported composite construction. The change concealed the bad performance of a particularly large account by placing that account into a new residual composite. This change allowed the manager to remain at the top of the list of manager performance. Scott knows her firm has a large allocation to this manager, and the fund's manager is a close personal friend of the CEO. She needs to deliver her final report but is concerned with pointing out the composite change.

> *Comment*: Scott would be in violation of Standard I(B) if she did not disclose the change in her final report. The analysis of managers' performance should not be influenced by personal relationships or the size of the allocation to the outside managers. By not including the change, Scott would not be providing an independent analysis of the performance metrics for her firm.

Example 16 (Intrafirm Pressure):

Jill Stein is head of performance measurement for her firm. During the last quarter, many members of the organization's research department were removed because of the poor quality of their recommendations. The subpar research caused one larger account holder to experience significant underperformance, which resulted in the client withdrawing his money after the end of the quarter. The head of sales requests that Stein remove this account from the firm's performance composite because the performance decline can be attributed to the departed research team and not the client's adviser.

> *Comment*: Pressure from other internal departments can create situations that cause a member or candidate to violate the Code and Standards. Stein must maintain her independence and objectivity and refuse to exclude specific accounts from the firm's performance composites to which they belong. As long as the client invested under a strategy similar to that of the defined composite, it cannot be excluded because of the poor stock selections that led to the underperformance and asset withdrawal.

STANDARD I(C): PROFESSIONALISM – MISREPRESENTATION

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Members and Candidates must not knowingly make any misrepresentations relating to investment analysis, recommendations, actions, or other professional activities.

Guidance

Highlights:

- *Impact on Investment Practice*
- *Performance Reporting*
- *Social Media*
- *Omissions*

- *Plagiarism*
- *Work Completed for Employer*

Trust is the foundation of the investment profession. Investors must be able to rely on the statements and information provided to them by those with whom the investors have trusted their financial well-being. Investment professionals who make false or misleading statements not only harm investors but also reduce the level of investor confidence in the investment profession and threaten the integrity of capital markets as a whole.

A misrepresentation is any untrue statement or omission of a fact or any statement that is otherwise false or misleading. A member or candidate must not knowingly omit or misrepresent information or give a false impression of a firm, organization, or security in the member's or candidate's oral representations, advertising (whether in the press or through brochures), electronic communications, or written materials (whether publicly disseminated or not). In this context, "knowingly" means that the member or candidate either knows or should have known that the misrepresentation was being made or that omitted information could alter the investment decision-making process.

Written materials include, but are not limited to, research reports, underwriting documents, company financial reports, market letters, newspaper columns, and books. Electronic communications include, but are not limited to, internet communications, webpages, mobile applications, and e-mails. Members and candidates who use webpages should regularly monitor materials posted on these sites to ensure that they contain current information. Members and candidates should also ensure that all reasonable precautions have been taken to protect the site's integrity and security and that the site does not misrepresent any information and does provide full disclosure.

Standard I(C) prohibits members and candidates from guaranteeing clients any specific return on volatile investments. Most investments contain some element of risk that makes their return inherently unpredictable. For such investments, guaranteeing either a particular rate of return or a guaranteed preservation of investment capital (e.g., "I can guarantee that you will earn 8% on equities this year" or "I can guarantee that you will not lose money on this investment") is misleading to investors. Standard I(C) does not prohibit members and candidates from providing clients with information on investment products that have guarantees built into the structure of the products themselves or for which an institution has agreed to cover any losses.

Impact on Investment Practice

Members and candidates must not misrepresent any aspect of their practice, including (but not limited to) their qualifications or credentials, the qualifications or services provided by their firm, their performance record and the record of their firm, and the characteristics of an investment. Any misrepresentation made by a member or candidate relating to the member's or candidate's professional activities is a breach of this standard.

Members and candidates should exercise care and diligence when incorporating third-party information. Misrepresentations resulting from the use of the credit ratings, research, testimonials, or marketing materials of outside parties become the responsibility of the investment professional when it affects that professional's business practices.

Investing through outside managers continues to expand as an acceptable method of investing in areas outside a firm's core competencies. Members and candidates must disclose their intended use of external managers and must not represent those managers' investment practices as their own. Although the level of involvement of outside managers may change over time, appropriate disclosures by members and candidates are important in avoiding misrepresentations, especially if the primary activity is to invest directly with a single external manager. Standard V(B)–Communication with Clients and Prospective Clients discusses in further detail communicating the firm's investment practices.

Performance Reporting

The performance benchmark selection process is another area where misrepresentations may occur. Members and candidates may misrepresent the success of their performance record through presenting benchmarks that are not comparable to their strategies. Further, clients can be misled if the benchmark's results are not reported on a basis comparable to that of the fund's or client's results. Best practice is selecting the most appropriate available benchmark from a universe of available options. The transparent presentation of appropriate performance benchmarks is an important aspect in providing clients with information that is useful in making investment decisions.

However, Standard I(C) does not require that a benchmark always be provided in order to comply. Some investment strategies may not lend themselves to displaying an appropriate benchmark because of the complexity or diversity of the investments included. Furthermore, some investment strategies may use reference indexes that do not reflect the opportunity set of the invested assets—for example, a hedge fund comparing its performance with a "cash plus" basis. When such a benchmark is used, members and candidates should make reasonable efforts to ensure that they disclose the reasons behind the use of this reference index to avoid misrepresentations of their performance. Members and candidates should discuss with clients on a continuous basis the appropriate benchmark to be used for performance evaluations and related fee calculations.

Reporting misrepresentations may also occur when valuations for illiquid or non-traded securities are available from more than one source. When different options are available, members and candidates may be tempted to switch providers to obtain higher security valuations. The process of shopping for values may misrepresent a security's worth, lead to misinformed decisions to sell or hold an investment, and result in overcharging clients advisory fees.

Members and candidates should take reasonable steps to provide accurate and reliable security pricing information to clients on a consistent basis. Changing pricing providers should not be based solely on the justification that the new provider reports a higher current value of a security. Consistency in the reported information will improve the perception of the valuation process for illiquid securities. Clients will likely have additional confidence that they were able to make an informed decision about continuing to hold these securities in their portfolios.

Social Media

The advancement of online discussion forums and communication platforms, commonly referred to as "social media," is placing additional responsibilities on members and candidates. When communicating through social media channels, members and candidates should provide only the same information they are allowed to distribute to clients and potential clients through other traditional forms of communication. The online or interactive aspects of social media do not remove the need to be open and honest about the information being distributed.

Along with understanding and following existing and newly developing rules and regulations regarding the allowed use of social media, members and candidates should also ensure that all communications in this format adhere to the requirements of the Code and Standards. The perceived anonymity granted through these platforms may entice individuals to misrepresent their qualifications or abilities or those of their employer. Actions undertaken through social media that knowingly misrepresent investment recommendations or professional activities are considered a violation of Standard I(C).

Omissions

The omission of a fact or outcome can be misleading, especially given the growing use of models and technical analysis processes. Many members and candidates rely on such models and processes to scan for new investment opportunities, to develop investment vehicles, and to produce investment recommendations and ratings. When inputs are knowingly omitted, the resulting outcomes may provide misleading information to those who rely on it for making investment decisions. Additionally, the outcomes from models shall not be presented as fact because they represent the expected results based on the inputs and analysis process incorporated.

Omissions in the performance measurement and attribution process can also misrepresent a manager's performance and skill. Members and candidates should encourage their firms to develop strict policies for composite development to prevent cherry picking—situations in which selected accounts are presented as representative of the firm's abilities. The omission of any accounts appropriate for the defined composite may misrepresent to clients the success of the manager's implementation of its strategy.

Plagiarism

Standard I(C) also prohibits plagiarism in the preparation of material for distribution to employers, associates, clients, prospects, or the general public. Plagiarism is defined as copying or using in substantially the same form materials prepared by others without acknowledging the source of the material or identifying the author and publisher of such material. Members and candidates must not copy (or represent as their own) original ideas or material without permission and must acknowledge and identify the source of ideas or material that is not their own.

The investment profession uses a myriad of financial, economic, and statistical data in the investment decision-making process. Through various publications and presentations, the investment professional is constantly exposed to the work of others and to the temptation to use that work without proper acknowledgment.

Misrepresentation through plagiarism in investment management can take various forms. The simplest and most flagrant example is to take a research report or study done by another firm or person, change the names, and release the material as one's own original analysis. This action is a clear violation of Standard I(C). Other practices include (1) using excerpts from articles or reports prepared by others either verbatim or with only slight changes in wording without acknowledgment, (2) citing specific quotations as attributable to "leading analysts" and "investment experts" without naming the specific references, (3) presenting statistical estimates of forecasts prepared by others and identifying the sources but without including the qualifying statements or caveats that may have been used, (4) using charts and graphs without stating their sources, and (5) copying proprietary computerized spreadsheets or algorithms without seeking the cooperation or authorization of their creators.

In the case of distributing third-party, outsourced research, members and candidates may use and distribute such reports as long as they do not represent themselves as the report's authors. Indeed, the member or candidate may add value for the client by sifting through research and repackaging it for clients. In such cases, clients should be fully informed that they are paying for the ability of the member or candidate to find the best research from a wide variety of sources. Members and candidates must not misrepresent their abilities, the extent of their expertise, or the extent of their work in a way that would mislead their clients or prospective clients. Members and candidates should disclose whether the research being presented to clients comes from another source—from either within or outside the member's or candidate's firm. This allows clients to understand who has the expertise behind the report or whether the work is being done by the analyst, other members of the firm, or an outside party.

Standard I(C) also applies to plagiarism in oral communications, such as through group meetings; visits with associates, clients, and customers; use of audio/video media (which is rapidly increasing); and telecommunications, including electronic data transfer and the outright copying of electronic media.

One of the most egregious practices in violation of this standard is the preparation of research reports based on multiple sources of information without acknowledging the sources. Examples of information from such sources include ideas, statistical compilations, and forecasts combined to give the appearance of original work. Although there is no monopoly on ideas, members and candidates must give credit where it is clearly due. Analysts should not use undocumented forecasts, earnings projections, asset values, and so on. Sources must be revealed to bring the responsibility directly back to the author of the report or the firm involved.

Work Completed for Employer

The preceding paragraphs address actions that would constitute a violation of Standard I(C). In some situations, however, members or candidates may use research conducted or models developed by others within the same firm without committing a violation. The most common example relates to the situation in which one (or more) of the original analysts is no longer with the firm. Research and models developed while employed by a firm are the property of the firm. The firm retains the right to continue using the work completed after a member or candidate has left the organization. The firm may issue future reports without providing attribution to the prior analysts. A member or candidate cannot, however, reissue a previously released report solely under his or her name.

STANDARD I(C): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Factual Presentations

Members and candidates can prevent unintentional misrepresentations of their qualifications or the services they or their firms provide if each member and candidate understands the limit of the firm’s or individual’s capabilities and the need to be accurate and complete in presentations. Firms can provide guidance for employees who make written or oral presentations to clients or potential clients by providing a written list of the firm’s available services and a description of the firm’s qualifications. This list should suggest ways of describing the firm’s services, qualifications, and compensation that are both accurate and suitable for client or customer presentations. Firms can also help prevent misrepresentation by specifically designating which employees are authorized to speak on behalf of the firm. Regardless of whether the firm provides guidance, members and candidates should make certain that they understand the services the firm can perform and its qualifications.

Qualification Summary

In addition, to ensure accurate presentations to clients, each member and candidate should prepare a summary of his or her own qualifications and experience and a list of the services the member or candidate is capable of performing. Firms can assist member and candidate compliance by periodically reviewing employee correspondence and documents that contain representations of individual or firm qualifications.

Verify Outside Information

When providing information to clients from a third party, members and candidates share a responsibility for the accuracy of the marketing and distribution materials that pertain to the third party's capabilities, services, and products. Misrepresentation by third parties can damage the member's or candidate's reputation, the reputation of the firm, and the integrity of the capital markets. Members and candidates should encourage their employers to develop procedures for verifying information of third-party firms.

Maintain Webpages

Members and candidates who publish a webpage should regularly monitor materials posted on the site to ensure that the site contains current information. Members and candidates should also ensure that all reasonable precautions have been taken to protect the site's integrity, confidentiality, and security and that the site does not misrepresent any information and provides full disclosure.

Plagiarism Policy

To avoid plagiarism in preparing research reports or conclusions of analysis, members and candidates should take the following steps:

- *Maintain copies*: Keep copies of all research reports, articles containing research ideas, material with new statistical methodologies, and other materials that were relied on in preparing the research report.
- *Attribute quotations*: Attribute to their sources any direct quotations, including projections, tables, statistics, model/product ideas, and new methodologies prepared by persons other than recognized financial and statistical reporting services or similar sources.
- *Attribute summaries*: Attribute to their sources any paraphrases or summaries of material prepared by others. For example, to support his analysis of Brown Company's competitive position, the author of a research report on Brown might summarize another analyst's report on Brown's chief competitor, but the author of the Brown report must acknowledge in his own report the reliance on the other analyst's report.

STANDARD I(C): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Disclosure of Issuer-Paid Research):

Anthony McGuire is an issuer-paid analyst hired by publicly traded companies to electronically promote their stocks. McGuire creates a website that promotes his research efforts as a seemingly independent analyst. McGuire posts a profile and a strong buy recommendation for each company on the website indicating that the stock is expected to increase in value. He does not disclose the contractual relationships with the companies he covers on his website, in the research reports he issues, or in the statements he makes about the companies in internet chat rooms.

> *Comment*: McGuire has violated Standard I(C) because the website is misleading to potential investors. Even if the recommendations are valid and supported with thorough research, his omissions regarding the true relationship between himself and the companies he covers constitute a misrepresentation. McGuire has also violated Standard VI(A)–Disclosure of Conflicts by not disclosing the existence of an arrangement with the companies through which he receives compensation in exchange for his services.

Example 2 (Correction of Unintentional Errors):

Hijan Yao is responsible for the creation and distribution of the marketing materials for his firm, which claims compliance with the GIPS standards. Yao creates and distributes a presentation of performance by the firm's Asian equity composite that states the composite has ¥350 billion in assets. In fact, the composite has only ¥35 billion in assets, and the higher figure on the presentation is a result of a typographical error. Nevertheless, the erroneous material is distributed to a number of clients before Yao catches the mistake.

> *Comment*: Once the error is discovered, Yao must take steps to cease distribution of the incorrect material and correct the error by informing those who have received the erroneous information. Because Yao did not knowingly make the misrepresentation, however, he did not violate Standard I(C). Because his firm claims compliance with the GIPS standards, it must also comply with the GIPS Guidance Statement on Error Correction in relation to the error.

Example 3 (Noncorrection of Known Errors):

Syed Muhammad is the president of an investment management firm. The promotional material for the firm, created by the firm's marketing department, incorrectly claims that Muhammad has an advanced degree in finance from a prestigious business school in addition to the CFA designation. Although Muhammad attended the school for a short period of time, he did not receive a degree. Over the years, Muhammad and others in the firm have distributed this material to numerous prospective clients and consultants.

> *Comment*: Even though Muhammad may not have been directly responsible for the misrepresentation of his credentials in the firm's promotional material, he used this material numerous times over an extended period and should have known of the misrepresentation. Thus, Muhammad has violated Standard I(C).

Example 4 (Plagiarism):

Cindy Grant, a research analyst for a Canadian brokerage firm, has specialized in the Canadian mining industry for the past 10 years. She recently read an extensive research report on Jefferson Mining, Ltd., by Jeremy Barton, another analyst. Barton provided extensive statistics on the mineral reserves, production capacity, selling rates, and marketing factors affecting Jefferson's operations. He also noted that initial drilling results on a new ore body, which had not been made public, might show the existence of mineral zones that could increase the life of Jefferson's main mines, but Barton cited no specific data as to the initial drilling results. Grant called an officer of Jefferson, who gave her the initial drilling results over the telephone. The data indicated that the expected life of the main mines would be tripled. Grant added these statistics to Barton's report and circulated it within her firm as her own report.

> *Comment*: Grant plagiarized Barton's report by reproducing large parts of it in her own report without acknowledgment.

Example 5 (Misrepresentation of Information):

When Ricki Marks sells mortgage-backed derivatives called "interest-only strips" (IOs) to public pension plan clients, she describes them as "guaranteed by the US government." Purchasers of the IOs are entitled only to the interest stream generated by the mortgages, however, not the notional principal itself. One particular municipality's investment policies and local law require that securities purchased by its public pension plans be guaranteed by the US government. Although the underlying mortgages are guaranteed, neither the investor's investment nor the interest stream on the IOs is guaranteed. When interest rates decline, causing an increase in prepayment of mortgages, interest payments to the IOs' investors decline, and these investors lose a portion of their investment.

> *Comment*: Marks violated Standard I(C) by misrepresenting the terms and character of the investment.

Example 6 (Potential Information Misrepresentation):

Khalouck Abdrabbo manages the investments of several high-net-worth individuals in the United States who are approaching retirement. Abdrabbo advises these individuals that a portion of their investments be moved from equity to bank-sponsored certificates of deposit and money market accounts so that the principal will be "guaranteed" up to a certain amount. The interest is not guaranteed.

> *Comment*: Although there is risk that the institution offering the certificates of deposits and money market accounts could go bankrupt, in the United States, these accounts are insured by the US government through

> the Federal Deposit Insurance Corporation. Therefore, using the term "guaranteed" in this context is not inappropriate as long as the amount is within the government-insured limit. Abdrabbo should explain these facts to the clients.

Example 7 (Plagiarism):

Steve Swanson is a senior analyst in the investment research department of Ballard and Company. Apex Corporation has asked Ballard to assist in acquiring the majority ownership of stock in the Campbell Company, a financial consulting firm, and to prepare a report recommending that stockholders of Campbell agree to the acquisition. Another investment firm, Davis and Company, had already prepared a report for Apex analyzing both Apex and Campbell and recommending an exchange ratio. Apex has given the Davis report to Ballard officers, who have passed it on to Swanson. Swanson reviews the Davis report and other available material on Apex and Campbell. From his analysis, he concludes that the common stocks of Campbell and Apex represent good value at their current prices; he believes, however, that the Davis report does not consider all the factors a Campbell stockholder would need to know to make a decision. Swanson reports his conclusions to the partner in charge, who tells him to "use the Davis report, change a few words, sign your name, and get it out."

> *Comment*: If Swanson does as requested, he will violate Standard I(C). He could refer to those portions of the Davis report that he agrees with if he identifies Davis as the source; he could then add his own analysis and conclusions to the report before signing and distributing it.

Example 8 (Plagiarism):

Claude Browning, a quantitative analyst for Double Alpha, Inc., returns from a seminar in great excitement. At that seminar, Jack Jorrely, a well-known quantitative analyst at a national brokerage firm, discussed one of his new models in great detail, and Browning is intrigued by the new concepts. He proceeds to test the model, making some minor mechanical changes but retaining the concepts, until he produces some very positive results. Browning quickly announces to his supervisors at Double Alpha that he has discovered a new model and that clients and prospective clients should be informed of this positive finding as ongoing proof of Double Alpha's continuing innovation and ability to add value.

> *Comment*: Although Browning tested Jorrely's model on his own and even slightly modified it, he must still acknowledge the original source of the idea. Browning can certainly take credit for the final, practical results; he can also support his conclusions with his own test. The credit for the innovative thinking, however, must be awarded to Jorrely.

Example 9 (Plagiarism):

Fernando Zubia would like to include in his firm's marketing materials some "plain-language" descriptions of various concepts, such as the price-to-earnings (P/E) multiple and why standard deviation is used as a measure of risk. The descriptions come from other sources, but Zubia wishes to use them without reference to the original authors. Would this use of material be a violation of Standard I(C)?

> *Comment*: Copying verbatim any material without acknowledgement, including plain-language descriptions of the P/E multiple and standard deviation, violates Standard I(C). Even though these concepts are general, best practice would be for Zubia to describe them in his own words or cite the sources from which the descriptions are quoted. Members and candidates would be violating Standard I(C) if they either were responsible for creating marketing materials without attribution or knowingly use plagiarized materials.

Example 10 (Plagiarism):

Through a mainstream media outlet, Erika Schneider learns about a study that she would like to cite in her research. Should she cite both the mainstream intermediary source as well as the author of the study itself when using that information?

> *Comment*: In all instances, a member or candidate must cite the actual source of the information. Best practice for Schneider would be to obtain the information directly from the author and review it before citing it in a report. In that case, Schneider would not need to report how she found out about the information. For example, suppose Schneider read in the *Financial Times* about a study issued by CFA Institute; best practice for Schneider would be to obtain a copy of the study from CFA Institute, review it, and then cite it in her report. If she does not use any interpretation of the report from the *Financial Times* and the newspaper does not add value to the report itself, the newspaper is merely a conduit of the original information and does not need to be cited. If she does not obtain the report and review the information, Schneider runs the risk of relying on second-hand information that may misstate facts. If, for example, the *Financial Times* erroneously reported some information from the original CFA Institute study and Schneider copied that erroneous information without acknowledging CFA Institute, she could be the object of complaints. Best practice would be either to obtain the complete study from its original author and cite only that author or to use the information provided by the intermediary and cite both sources.

Example 11 (Misrepresentation of Information):

Paul Ostrowski runs a two-person investment management firm. Ostrowski's firm subscribes to a service from a large investment research firm that provides research reports that can be repackaged by smaller firms for those firms' clients. Ostrowski's firm distributes these reports to clients as its own work.

> *Comment*: Ostrowski can rely on third-party research that has a reasonable and adequate basis, but he cannot imply that he is the author of such research. If he does, Ostrowski is misrepresenting the extent of his work in a way that misleads the firm's clients or prospective clients.

Example 12 (Misrepresentation of Information):

Tom Stafford is part of a team within Appleton Investment Management responsible for managing a pool of assets for Open Air Bank, which distributes structured securities to offshore clients. He becomes aware that Open Air is promoting the structured securities as a much less risky investment than the investment management policy followed by him and the team to manage the original pool of assets. Also, Open Air has procured an independent rating for the pool that significantly overstates the quality of the investments. Stafford communicates his concerns to his supervisor, who responds that Open Air owns the product and is responsible for all marketing and distribution. Stafford's supervisor goes on to say that the product is outside of the US regulatory regime that Appleton follows and that all risks of the product are disclosed at the bottom of page 184 of the prospectus.

> *Comment*: As a member of the investment team, Stafford is qualified to recognize the degree of accuracy of the materials that characterize the portfolio, and he is correct to be worried about Appleton's responsibility for a misrepresentation of the risks. Thus, he should continue to pursue the issue of Open Air's inaccurate promotion of the portfolio according to the firm's policies and procedures.
>
> The Code and Standards stress protecting the reputation of the firm and the sustainability and integrity of the capital markets. Misrepresenting the quality and risks associated with the investment pool may lead to negative consequences for others well beyond the direct investors.

Example 13 (Avoiding a Misrepresentation):

Trina Smith is a fixed-income portfolio manager at a pension fund. She has observed that the market for highly structured mortgages is the focus of salespeople she meets and that these products represent a significant number of trading opportunities. In discussions about this topic with her team, Smith learns that calculating yields on changing cash flows within the deal structure requires very specialized vendor software. After more research, they find out that each deal is unique and that deals can have more than a dozen layers and changing cash flow priorities. Smith comes to the conclusion that, because of the complexity of these securities, the team cannot effectively distinguish between potentially good and bad investment options. To avoid misrepresenting their understanding, the team decides that the highly structured mortgage segment of the securitized market should not become part of the core of the fund's portfolio; they will allow some of the less complex securities to be part of the core.

> *Comment*: Smith is in compliance with Standard I(C) by not investing in securities that she and her team cannot effectively understand. Because she is not able to describe the risk and return profile of the securities to the pension fund beneficiaries and trustees, she appropriately limits the fund's exposure to this sector.

Example 14 (Misrepresenting Composite Construction):

Robert Palmer is head of performance for a fund manager. When asked to provide performance numbers to fund rating agencies, he avoids mentioning that the fund manager is quite liberal in composite construction. The reason accounts are included/excluded is not fully explained. The performance values reported to the rating agencies for the composites, although accurate for the accounts shown each period, may not present a true representation of the fund manager's ability.

> *Comment*: "Cherry picking" accounts to include in either published reports or information provided to rating agencies conflicts with Standard I(C). Moving accounts into or out of a composite to influence the overall performance results materially misrepresents the reported values over time.

> Palmer should work with his firm to strengthen its reporting practices concerning composite construction to avoid misrepresenting the firm's track record or the quality of the information being provided.

Example 15 (Presenting Out-of-Date Information):

David Finch is a sales director at a commercial bank, where he directs the bank's client advisers in the sale of third-party mutual funds. Each quarter, he holds a division-wide training session where he provides fact sheets on investment funds the bank is allowed to offer to clients. These fact sheets, which can be redistributed to potential clients, are created by the fund firms and contain information about the funds, including investment strategy and target distribution rates.

Finch knows that some of the fact sheets are out of date; for example, one long-only fund approved the use of significant leverage last quarter as a method to enhance returns. He continues to provide the sheets to the sales team without updates because the bank has no control over the marketing material released by the mutual fund firms.

> *Comment*: Finch is violating Standard I(C) by providing information that misrepresents aspects of the funds. By not providing the sales team and, ultimately, the clients with the updated information, he is misrepresenting the potential risks associated with the funds with outdated fact sheets. Finch can instruct the sales team to clarify the deficiencies in the fact sheets with clients and ensure they have the most recent fund prospectus document before accepting orders for investing in any fund.

Example 16 (Overemphasis of Firm Results):

Bob Anderson is chief compliance officer for Optima Asset Management Company, a firm currently offering eight funds to clients. Seven of the eight had 10-year returns below the median for their respective sectors. Anderson approves a recent advertisement, which includes this statement: "Optima Asset Management is achieving excellent returns for its investors. The Optima Emerging Markets Equity fund, for example, has 10-year returns that exceed the sector median by more than 10%."

> *Comment*: From the information provided it is difficult to determine whether a violation has occurred as long as the sector outperformance is correct. Anderson may be attempting to mislead potential clients by citing the performance of the sole fund that achieved such results. Past performance is often used to demonstrate a firm's skill and abilities in comparison to funds in the same sectors.
>
> However, if all the funds outperformed their respective benchmarks, then Anderson's assertion that the company "is achieving excellent returns" may be factual. Funds may exhibit positive returns for investors, exceed benchmarks, and yet have returns below the median in their sectors.
>
> Members and candidates need to ensure that their marketing efforts do not include statements that misrepresent their skills and abilities to remain compliant with Standard I(C). Unless the returns of a single fund reflect the performance of a firm as a whole, the use of a singular fund for performance comparisons should be avoided.

STANDARD I(D): PROFESSIONALISM – MISCONDUCT

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Members and Candidates must not engage in any professional conduct involving dishonesty, fraud, or deceit or commit any act that reflects adversely on their professional reputation, integrity, or competence.

Guidance

Whereas Standard I(A) addresses the obligation of members and candidates to comply with applicable law that governs their professional activities, Standard I(D) addresses *all* conduct that reflects poorly on the professional integrity, good reputation, or competence of members and candidates. Any act that involves lying, cheating, stealing, or other dishonest conduct is a violation of this standard if the offense reflects adversely on a member's or candidate's professional activities. Although CFA Institute discourages any sort of unethical behavior by members and candidates, the Code and Standards are primarily aimed at conduct and actions related to a member's or candidate's professional life.

Conduct that damages trustworthiness or competence may include behavior that, although not illegal, nevertheless negatively affects a member's or candidate's ability to perform his or her responsibilities. For example, abusing alcohol during business hours might constitute a violation of this standard because it could have a detrimental effect on the member's or candidate's ability to fulfill his or her professional responsibilities. Personal bankruptcy may not reflect on the integrity or trustworthiness of the person declaring bankruptcy, but if the circumstances of the bankruptcy involve fraudulent or deceitful business conduct, the bankruptcy may be a violation of this standard.

In some cases, the absence of appropriate conduct or the lack of sufficient effort may be a violation of Standard I(D). The integrity of the investment profession is built on trust. A member or candidate—whether an investment banker, rating or research analyst, or portfolio manager—is expected to conduct the necessary due diligence to properly understand the nature and risks of an investment before making an investment recommendation. By not taking these steps and, instead, relying on someone else in the process to perform them, members or candidates may violate the trust their clients have placed in them. This loss of trust may have a significant impact on the reputation of the member or candidate and the operations of the financial market as a whole.

Individuals may attempt to abuse the CFA Institute Professional Conduct Program by actively seeking CFA Institute enforcement of the Code and Standards, and Standard I(D) in particular, as a method of settling personal, political, or other disputes unrelated to professional ethics. CFA Institute is aware of this issue, and appropriate disciplinary policies, procedures, and enforcement mechanisms are in place to address misuse of the Code and Standards and the Professional Conduct Program in this way.

STANDARD I(D): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

In addition to ensuring that their own behavior is consistent with Standard I(D), to prevent general misconduct, members and candidates should encourage their firms to adopt the following policies and procedures to support the principles of Standard I(D):

- *Code of ethics*: Develop and/or adopt a code of ethics to which every employee must subscribe, and make clear that any personal behavior that reflects poorly on the individual involved, the institution as a whole, or the investment industry will not be tolerated.
- *List of violations*: Disseminate to all employees a list of potential violations and associated disciplinary sanctions, up to and including dismissal from the firm.
- *Employee references*: Check references of potential employees to ensure that they are of good character and not ineligible to work in the investment industry because of past infractions of the law.

STANDARD I(D): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Professionalism and Competence):

Simon Sasserman is a trust investment officer at a bank in a small affluent town. He enjoys lunching every day with friends at the country club, where his clients have observed him having numerous drinks. Back at work after lunch, he clearly is intoxicated while making investment decisions. His colleagues make a point of handling any business with Sasserman in the morning because they distrust his judgment after lunch.

> *Comment*: Sasserman's excessive drinking at lunch and subsequent intoxication at work constitute a violation of Standard I(D) because this conduct has raised questions about his professionalism and competence. His behavior reflects poorly on him, his employer, and the investment industry.

Example 2 (Fraud and Deceit):

Howard Hoffman, a security analyst at ATZ Brothers, Inc., a large brokerage house, submits reimbursement forms over a two-year period to ATZ's self-funded health insurance program for more than two dozen bills, most of which have been altered to increase the amount due. An investigation by the firm's director of employee benefits uncovers the inappropriate conduct. ATZ subsequently terminates Hoffman's employment and notifies CFA Institute.

> *Comment*: Hoffman violated Standard I(D) because he engaged in intentional conduct involving fraud and deceit in the workplace that adversely reflected on his integrity.

Example 3 (Fraud and Deceit):

Jody Brink, an analyst covering the automotive industry, volunteers much of her spare time to local charities. The board of one of the charitable institutions decides to buy five new vans to deliver hot lunches to low-income elderly people. Brink offers to donate her time to handle purchasing agreements. To pay a long-standing debt to a friend who operates an automobile dealership—and to compensate herself for her trouble—she agrees to a price 20% higher than normal and splits the surcharge with her friend. The director of the charity ultimately discovers the scheme and tells Brink that her services, donated or otherwise, are no longer required.

> *Comment*: Brink engaged in conduct involving dishonesty, fraud, and misrepresentation and has violated Standard I(D).

Example 4 (Personal Actions and Integrity):

Carmen Garcia manages a mutual fund dedicated to socially responsible investing. She is also an environmental activist. As the result of her participation in nonviolent protests, Garcia has been arrested on numerous occasions for trespassing on the property of a large petrochemical plant that is accused of damaging the environment.

> *Comment*: Generally, Standard I(D) is not meant to cover legal transgressions resulting from acts of civil disobedience in support of personal beliefs because such conduct does not reflect poorly on the member's or candidate's professional reputation, integrity, or competence.

Example 5 (Professional Misconduct):

Meredith Rasmussen works on a buy-side trading desk of an investment management firm and concentrates on in-house trades for a hedge fund subsidiary managed by a team at the investment management firm. The hedge fund has been very successful and is marketed globally by the firm. From her experience as the trader for much of the activity of the fund, Rasmussen has become quite knowledgeable about the hedge fund's strategy, tactics, and performance. When a distinct break in the market occurs and many of the securities involved in the hedge fund's strategy decline markedly in value, Rasmussen observes that the reported performance of the hedge fund does not reflect this decline. In her experience, the lack of effect is a very unlikely occurrence. She approaches the head of trading about her concern and is told that she should not ask any questions and that the fund is big and successful and is not her concern. She is fairly sure something is not right, so she contacts the compliance officer, who also tells her to stay away from the issue of the hedge fund's reporting.

> *Comment*: Rasmussen has clearly come across an error in policies, procedures, and compliance practices within the firm's operations. According to the firm's procedures for reporting potentially unethical activity, she should pursue the issue by gathering some proof of her reason for doubt. Should all internal communications within the firm not satisfy her concerns, Rasmussen should consider reporting the potential unethical activity to the appropriate regulator.

See also Standard IV(A) for guidance on whistleblowing and Standard IV(C) for the duties of a supervisor.

STANDARD II(A): INTEGRITY OF CAPITAL MARKETS - MATERIAL NONPUBLIC INFORMATION

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Standard II(A) Material Nonpublic Information

Members and Candidates who possess material nonpublic information that could affect the value of an investment must not act or cause others to act on the information.

Guidance

Highlights:

- *What Is "Material" Information?*
- *What Constitutes "Nonpublic" Information?*
- *Mosaic Theory*
- *Social Media*
- *Using Industry Experts*
- *Investment Research Reports*

Trading or inducing others to trade on material nonpublic information erodes confidence in capital markets, institutions, and investment professionals by supporting the idea that those with inside information and special access can take unfair advantage of the general investing public. Although trading on inside information may lead to short-term profits, in the long run, individuals and the profession as a whole suffer from such trading. These actions have caused and will continue to cause investors to avoid capital markets because the markets are perceived to be "rigged" in favor of the knowledgeable insider. When the investing public avoids capital markets, the markets and capital allocation become less efficient and less supportive of strong and vibrant economies. Standard II(A) promotes and maintains a high level of confidence in market integrity, which is one of the foundations of the investment profession.

The prohibition on using this information goes beyond the direct buying and selling of individual securities or bonds. Members and candidates must not use material nonpublic information to influence their investment actions related to derivatives (e.g., swaps or option contracts), mutual funds, or other alternative investments. *Any* trading based on material nonpublic information constitutes a violation of Standard II(A). The expansion of financial products and the increasing interconnectivity of financial markets globally have resulted in new potential opportunities for trading on material nonpublic information.

What Is "Material" Information?

Information is "material" if its disclosure would probably have an impact on the price of a security or if reasonable investors would want to know the information before making an investment decision. In other words, information is material if it would significantly alter the total mix of information currently available about a security in such a way that the price of the security would be affected.

The specificity of the information, the extent of its difference from public information, its nature, and its reliability are key factors in determining whether a particular piece of information fits the definition of material. For example, material information may include, but is not limited to, information on the following:

- earnings;
- mergers, acquisitions, tender offers, or joint ventures;
- changes in assets or asset quality;
- innovative products, processes, or discoveries (e.g., new product trials or research efforts);
- new licenses, patents, registered trademarks, or regulatory approval/rejection of a product;
- developments regarding customers or suppliers (e.g., the acquisition or loss of a contract);
- changes in management;
- change in auditor notification or the fact that the issuer may no longer rely on an auditor's report or qualified opinion;
- events regarding the issuer's securities (e.g., defaults on senior securities, calls of securities for redemption, repurchase plans, stock splits, changes in dividends, changes to the rights of security holders, and public or private sales of additional securities);
- bankruptcies;
- significant legal disputes;
- government reports of economic trends (employment, housing starts, currency information, etc.);
- orders for large trades before they are executed; and
- new or changing equity or debt ratings issued by a third party (e.g., sell-side recommendations and credit ratings).

In addition to the substance and specificity of the information, the source or relative reliability of the information also determines materiality. The less reliable a source, the less likely the information provided would be considered material. For example, factual information from a corporate insider regarding a significant new contract for a company is likely to be material, whereas an assumption based on speculation by a competitor about the same contract is likely to be less reliable and, therefore, not material. Additionally, information about trials of a new drug, product, or service under development from qualified personnel involved in the trials is likely to be material, whereas educated conjecture by subject experts not connected to the trials is unlikely to be material.

Also, the more ambiguous the effect of the information on price, the less material that information is considered. If it is unclear whether and to what extent the information will affect the price of a security, the information may not be considered material. The passage of time may also render information that was once important immaterial.

What Constitutes "Nonpublic" Information?

Information is "nonpublic" until it has been disseminated or is available to the marketplace in general (as opposed to a select group of investors). "Disseminated" can be defined as "made known." For example, a company report of profits that is posted on the internet and distributed widely through a press release or accompanied by a filing has been effectively disseminated to the marketplace. Members and candidates must have a reasonable expectation that people have received the information before it can be considered public. It is not necessary, however, to wait for the slowest method of delivery. Once the information is disseminated to the market, it is public information that is no longer covered by this standard.

Members and candidates must be particularly aware of information that is selectively disclosed by corporations to a small group of investors, analysts, or other market participants. Information that is made available to analysts remains nonpublic until it is made available to investors in general. Corporations that disclose information on a limited basis create the potential for insider-trading violations.

Issues of selective disclosure often arise when a corporate insider provides material information to analysts in a briefing or conference call before that information is released to the public. Analysts must be aware that a disclosure made to a room full of analysts does not necessarily make the disclosed information "public." Analysts should also be alert to the possibility that they are selectively receiving material nonpublic information when a company provides them with guidance or interpretation of such publicly available information as financial statements or regulatory filings.

A member or candidate may use insider information provided legitimately by the source company for the specific purpose of conducting due diligence according to the business agreement between the parties for such activities as mergers, loan underwriting, credit ratings, and offering engagements. In such instances, the investment professional would not be considered in violation of Standard II(A) by using the material information. However, the use of insider information provided by the source company for other purposes, especially to trade or entice others to trade the securities of the firm, conflicts with this standard.

Mosaic Theory

A financial analyst gathers and interprets large quantities of information from many sources. The analyst may use significant conclusions derived from the analysis of public and nonmaterial nonpublic information as the basis for investment recommendations and decisions even if those conclusions would have been material inside information had they been communicated directly to the analyst by a company. Under the "mosaic theory," financial analysts are free to act on this collection, or mosaic, of information without risking violation.

The practice of financial analysis depends on the free flow of information. For the fair and efficient operation of the capital markets, analysts and investors must have the greatest amount of information possible to facilitate making well-informed investment decisions about how and where to invest capital. Accurate, timely, and intelligible communication is essential if analysts and investors are to obtain the data needed to make informed decisions about how and where to invest capital. These disclosures must go beyond the information mandated by the reporting requirements of the securities laws and should include specific business information about items used to guide a company's future growth, such as new products, capital projects, and the competitive environment. Analysts seek and use such information to compare and contrast investment alternatives.

Much of the information used by analysts comes directly from companies. Analysts often receive such information through contacts with corporate insiders, especially investor-relations staff and financial officers. Information may be disseminated in the form of press releases, through oral presentations by company executives in analysts' meetings or conference calls, or during analysts' visits to company premises. In seeking to develop the most accurate and complete picture of a company, analysts should also reach beyond contacts with companies themselves and collect information from other sources, such as customers, contractors, suppliers, and the companies' competitors.

Analysts are in the business of formulating opinions and insights that are not obvious to the general investing public about the attractiveness of particular securities. In the course of their work, analysts actively seek out corporate information not generally known to the market for the express purpose of analyzing that information, forming an opinion on its significance, and informing their clients, who can be expected to trade on the basis of the recommendation. Analysts' initiatives to discover and analyze information and communicate their findings to their clients significantly enhance market efficiency, thus benefiting all investors (see *Dirks v. Securities and Exchange Commission*). Accordingly, violations of Standard II(A) will *not* result when a perceptive analyst reaches a conclusion about a corporate action or event through an analysis of public information and items of nonmaterial nonpublic information.

Investment professionals should note, however, that although analysts are free to use mosaic information in their research reports, they should save and document all their research [see Standard V(C)–Record Retention]. Evidence of the analyst's knowledge of public and nonmaterial nonpublic information about a corporation strengthens the assertion that the analyst reached his or her conclusions solely through appropriate methods rather than through the use of material nonpublic information.

Social Media

The continuing advancement in technology allows members, candidates, and the industry at large to exchange information at rates not previously available. It is important for investment professionals to understand the implications of using information from the internet and social media platforms because all such information may not actually be considered public.

Some social media platforms require membership in specific groups in order to access the published content. Members and candidates participating in groups with membership limitations should verify that material information obtained from these sources can also be accessed from a source that would be considered available to the public (e.g., company filings, webpages, and press releases).

Members and candidates may use social media platforms to communicate with clients or investors without conflicting with this standard. As long as the information reaches all clients or is open to the investing public, the use of these platforms would be comparable with other traditional forms of communications, such as e-mails and press releases. Members and candidates, as required by Standard I(A), should also complete all appropriate regulatory filings related to information distributed through social media platforms.

Using Industry Experts

The increased demand for insights for understanding the complexities of some industries has led to an expansion of engagement with outside experts. As the level of engagement increased, new businesses formed to connect analysts and investors with individuals who have specialized knowledge of their industry (e.g., technology or pharmaceuticals). These networks offer investors the opportunity to reach beyond their usual business circles to speak with experts regarding economic conditions, industry trends, and technical issues relating to specific products and services.

Members and candidates may provide compensation to individuals for their insights without violating this standard. However, members and candidates are ultimately responsible for ensuring that they are not requesting or acting on confidential information received from external experts, which is in violation of security regulations and laws or duties to others. As the recent string of insider-trading cases displayed, some experts are willing to provide confidential and protected information for the right incentive.

Firms connecting experts with members or candidates often require both parties to sign agreements concerning the disclosure of material nonpublic information. Even with the protections from such compliance practices, if an expert provides material nonpublic information, members and candidates would be prohibited from taking investment actions on the associated firm until the information became publicly known to the market.

Investment Research Reports

When a particularly well-known or respected analyst issues a report or makes changes to his or her recommendation, that information alone may have an effect on the market and thus may be considered material. Theoretically, under Standard II(A), such a report would have to be made public at the time it was distributed to clients. The analyst is not a company insider, however, and does not have access to inside information. Presumably, the analyst created the report from information available to the public (mosaic theory) and by using his or her expertise to interpret the information. The analyst's hard work, paid for by the client, generated the conclusions.

Simply because the public in general would find the conclusions material does not require that the analyst make his or her work public. Investors who are not clients of the analyst can either do the work themselves or become clients of the analyst to gain access to the analyst's expertise.

STANDARD II(A): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Achieve Public Dissemination

If a member or candidate determines that information is material, the member or candidate should make reasonable efforts to achieve public dissemination of the information. These efforts usually entail encouraging the issuing company to make the information public. If public dissemination is not possible, the member or candidate must communicate the information only to the designated supervisory and compliance personnel within the member's or candidate's firm and must not take investment action or alter current investment recommendations on the basis of the information. Moreover, members and candidates must not knowingly engage in any conduct that may induce company insiders to privately disclose material nonpublic information.

Adopt Compliance Procedures

Members and candidates should encourage their firms to adopt compliance procedures to prevent the misuse of material nonpublic information. Particularly important is improving compliance in such areas as the review of employee and proprietary trading, the review of investment recommendations, documentation of firm procedures, and the supervision of interdepartmental communications in multiservice firms. Compliance procedures should suit the particular characteristics of a firm, including its size and the nature of its business.

Members and candidates are encouraged to inform their supervisor and compliance personnel of suspected inappropriate use of material nonpublic information as the basis for security trading activities or recommendations being made within their firm.

Adopt Disclosure Procedures

Members and candidates should encourage their firms to develop and follow disclosure policies designed to ensure that information is disseminated to the marketplace in an equitable manner. For example, analysts from small firms should receive the same information and attention from a company as analysts from large firms receive. Similarly, companies should not provide certain information to buy-side analysts but not to sell-side analysts, or vice versa. Furthermore, a company should not discriminate among analysts in the provision of information or “blackball” particular analysts who have given negative reports on the company in the past.

Within investment and research firms, members and candidates should encourage the development of and compliance with procedures for distributing new and updated investment opinions to clients. Recommendations of this nature may represent material market-moving information that needs to be communicated to all clients fairly.

Issue Press Releases

Companies should consider issuing press releases prior to analyst meetings and conference calls and scripting those meetings and calls to decrease the chance that further information will be disclosed. If material nonpublic information is disclosed for the first time in an analyst meeting or call, the company should promptly issue a press release or otherwise make the information publicly available.

Firewall Elements

An information barrier commonly referred to as a “firewall” is the most widely used approach for preventing the communication of material nonpublic information within firms. It restricts the flow of confidential information to those who need to know the information to perform their jobs effectively. The minimum elements of such a system include, but are not limited to, the following:

- substantial control of relevant interdepartmental communications, preferably through a clearance area within the firm in either the compliance or legal department;
- review of employee trading through the maintenance of “watch,” “restricted,” and “rumor” lists;
- documentation of the procedures designed to limit the flow of information between departments and of the actions taken to enforce those procedures; and
- heightened review or restriction of proprietary trading while a firm is in possession of material nonpublic information.

Appropriate Interdepartmental Communications

Although documentation requirements must, for practical reasons, take into account the differences between the activities of small firms and those of large, multiservice firms, firms of all sizes and types benefit by improving the documentation of their internal enforcement of firewall procedures. Therefore, even at small firms, procedures concerning interdepartmental communication, the review of trading activity, and the investigation of possible violations should be compiled and formalized.

Physical Separation of Departments

As a practical matter, to the greatest extent possible, firms should consider the physical separation of departments and files to prevent the communication of sensitive information that should not be shared. For example, the investment banking and corporate finance areas of a brokerage firm should be separated from the sales and research departments, and a bank's commercial lending department should be segregated from its trust and research departments.

Prevention of Personnel Overlap

There should be no overlap of personnel between the investment banking and corporate finance areas of a brokerage firm and the sales and research departments or between a bank's commercial lending department and its trust and research departments. For a firewall to be effective in a multiservice firm, an employee should be on only one side of the firewall at any time. Inside knowledge may not be limited to information about a specific offering or the current financial condition of a company. Analysts may be exposed to much information about the company, including new product developments or future budget projections that clearly constitute inside knowledge and thus preclude the analyst from returning to his or her research function. For example, an analyst who follows a particular company may provide limited assistance to the investment bankers under carefully controlled circumstances when the firm's investment banking department is involved in a deal with the company. That analyst must then be treated as though he or she were an investment banker; the analyst must remain on the investment banking side of the wall until any information he or she learns is publicly disclosed. In short, the analyst cannot use any information learned in the course of the project for research purposes and cannot share that information with colleagues in the research department.

A Reporting System

A primary objective of an effective firewall procedure is to establish a reporting system in which authorized people review and approve communications between departments. If an employee behind a firewall believes that he or she needs to share confidential information with someone on the other side of the wall, the employee should consult a designated compliance officer to determine whether sharing the information is necessary and how much information should be shared. If the sharing is necessary, the compliance officer should coordinate the process of "looking over the wall" so that the necessary information will be shared and the integrity of the procedure will be maintained.

A single supervisor or compliance officer should have the specific authority and responsibility of deciding whether information is material and whether it is sufficiently public to be used as the basis for investment decisions. Ideally, the supervisor or compliance officer responsible for communicating information to a firm's research or brokerage area would not be a member of that area.

Personal Trading Limitations

Firms should consider restrictions or prohibitions on personal trading by employees and should carefully monitor both proprietary trading and personal trading by employees. Firms should require employees to make periodic reports (to the extent that such reporting is not already required by securities laws) of their own transactions and transactions made for the benefit of family members. Securities should be placed on a restricted list when a firm has or may have material nonpublic information. The broad distribution of a restricted list often triggers the sort of trading the list was developed to avoid. Therefore, a watch list shown to only the few people responsible for compliance should be used to monitor transactions in specified securities. The use of a watch list in combination with a restricted list is an increasingly common means of ensuring effective control of personal trading.

Record Maintenance

Multiservice firms should maintain written records of the communications between various departments. Firms should place a high priority on training and should consider instituting comprehensive training programs, particularly for employees in sensitive areas.

Proprietary Trading Procedures

Procedures concerning the restriction or review of a firm's proprietary trading while the firm possesses material nonpublic information will necessarily depend on the types of proprietary trading in which the firm may engage. A prohibition on all types of proprietary activity when a firm comes into possession of material nonpublic information is *not* appropriate. For example, when a firm acts as a market maker, a prohibition on proprietary trading may be counterproductive to the goals of maintaining the confidentiality of information and market liquidity. This concern is particularly important in the relationships between small, regional broker/dealers and small issuers. In many situations, a firm will take a small issuer public with the understanding that the firm will continue to be a market maker in the stock. In such instances, a withdrawal by the firm from market-making activities would be a clear tip to outsiders. Firms that continue market-making activity while in the possession of material nonpublic information should, however, instruct their market makers to remain passive with respect to the market—that is, to take only the contra side of unsolicited customer trades.

In risk-arbitrage trading, the case for a trading prohibition is more compelling than it is in the case of market making. The impetus for arbitrage trading is neither passive nor reactive, and the potential for illegal profits is greater than in market making. The most prudent course for firms is to suspend arbitrage activity when a security is placed on the watch list. Those firms that continue arbitrage activity face a high hurdle in proving the adequacy of their internal procedures for preventing trading on material nonpublic information and must demonstrate a stringent review and documentation of firm trades.

Communication to All Employees

Members and candidates should encourage their employers to circulate written compliance policies and guidelines to all employees. Policies and guidelines should be used in conjunction with training programs aimed at enabling employees to recognize material nonpublic information. Such information is not always clearly identifiable.

Employees must be given sufficient training to either make an informed decision or to realize they need to consult a supervisor or compliance officer before engaging in questionable transactions. Appropriate policies reinforce that using material nonpublic information is illegal in many countries. Such trading activities based on material nonpublic information undermine the integrity of the individual, the firm, and the capital markets.

STANDARD II(A): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Acting on Nonpublic Information):

Frank Barnes, the president and controlling shareholder of the SmartTown clothing chain, decides to accept a tender offer and sell the family business at a price almost double the market price of its shares. He describes this decision to his sister (SmartTown's treasurer), who conveys it to her daughter (who owns no stock in the family company at present), who tells her husband, Staple. Staple, however, tells his stockbroker, Alex Halsey, who immediately buys SmartTown stock for himself.

> *Comment*: The information regarding the pending sale is both material and nonpublic. Staple has violated Standard II(A) by communicating the inside information to his broker. Halsey also has violated the standard by buying the shares on the basis of material nonpublic information.

Example 2 (Controlling Nonpublic Information):

Samuel Peter, an analyst with Scotland and Pierce Incorporated, is assisting his firm with a secondary offering for Bright Ideas Lamp Company. Peter participates, via telephone conference call, in a meeting with Scotland and Pierce investment banking employees and Bright Ideas' CEO. Peter is advised that the company's earnings projections for the next year have significantly dropped. Throughout the telephone conference call, several Scotland and Pierce salespeople and portfolio managers walk in and out of Peter's office, where the telephone call is taking place. As a result, they are aware of the drop in projected earnings for Bright Ideas. Before the conference call is concluded, the salespeople trade the stock of the company on behalf of the firm's clients and other firm personnel trade the stock in a firm proprietary account and in employees' personal accounts.

> *Comment*: Peter has violated Standard II(A) because he failed to prevent the transfer and misuse of material nonpublic information to others in his firm. Peter's firm should have adopted information barriers to prevent the communication of nonpublic information between departments of the firm. The salespeople and portfolio managers who traded on the information have also violated Standard II(A) by trading on inside information.

Example 3 (Selective Disclosure of Material Information):

Elizabeth Levenson is based in Hanoi and covers the Vietnamese market for her firm, which is based in Singapore. She is invited, together with the other 10 largest shareholders of a manufacturing company, to meet the finance director of that company. During the meeting, the finance director states that the company expects its workforce to strike next Friday, which will cripple productivity and distribution. Can Levenson use this information as a basis to change her rating on the company from "buy" to "sell"?

> *Comment*: Levenson must first determine whether the material information is public. According to Standard II(A), if the company has not made this information public (a small group forum does not qualify as a method of public dissemination), she cannot use the information.

Example 4 (Determining Materiality):

Leah Fechtman is trying to decide whether to hold or sell shares of an oil-and-gas exploration company that she owns in several of the funds she manages. Although the company has underperformed the index for some time already, the trends in the industry sector signal that companies of this type might become takeover targets. While she is considering her decision, her doctor, who casually follows the markets, mentions that she thinks that the company in question will soon be bought out by a large multinational conglomerate and that it would be a good idea to buy the stock right now. After talking to various investment professionals and checking their opinions on the company as well as checking industry trends, Fechtman decides the next day to accumulate more stock in the oil-and-gas exploration company.

> *Comment*: Although information on an expected takeover bid may be of the type that is generally material and nonpublic, in this case, the source of information is unreliable, so the information cannot be considered material. Therefore, Fechtman is not prohibited from trading the stock on the basis of this information.

Example 5 (Applying the Mosaic Theory):

Jagdish Teja is a buy-side analyst covering the furniture industry. Looking for an attractive company to recommend as a buy, he analyzes several furniture makers by studying their financial reports and visiting their operations. He also talks to some designers and retailers to find out which furniture styles are trendy and popular. Although none of the companies that he analyzes are a clear buy, he discovers that one of them, Swan Furniture Company (SFC), may be in financial trouble. SFC's extravagant new designs have been introduced at substantial cost. Even though these designs initially attracted attention, the public is now buying more conservative furniture from other makers. Based on this information and on a profit-and-loss analysis, Teja believes that SFC's next quarter earnings will drop substantially. He issues a sell recommendation for SFC. Immediately after receiving that recommendation, investment managers start reducing the SFC stock in their portfolios.

> *Comment*: Information on quarterly earnings data is material and nonpublic. Teja arrived at his conclusion about the earnings drop on the basis of public information and on pieces of nonmaterial nonpublic information (such as opinions of designers and retailers). Therefore, trading based on Teja's correct conclusion is not prohibited by Standard II(A).

Example 6 (Applying the Mosaic Theory):

Roger Clement is a senior financial analyst who specializes in the European automobile sector at Rivoli Capital. Because he has been repeatedly nominated by many leading industry magazines and newsletters as a "best analyst" for the automobile industry, he is widely regarded as an authority on the sector. After speaking with representatives of Turgot Chariots—a European auto manufacturer with sales primarily in South Korea—and after conducting interviews with salespeople, labor leaders, his firm's Korean currency analysts, and banking officials, Clement analyzed Turgot Chariots and concluded that (1) its newly introduced model will probably not meet sales expectations, (2) its corporate restructuring strategy may well face serious opposition from unions, (3) the depreciation of the Korean won should lead to pressure on margins for the industry in general and Turgot's market segment in particular, and (4) banks could take a tougher-than-expected stance in the upcoming round of credit renegotiations with the company. For these reasons, he changes his conclusion about the company from "market outperform" to "market underperform." Clement retains the support material used to reach his conclusion in case questions later arise.

> *Comment*: To reach a conclusion about the value of the company, Clement has pieced together a number of nonmaterial or public bits of information that affect Turgot Chariots. Therefore, under the mosaic theory, Clement has not violated Standard II(A) in drafting the report.

Example 7 (Analyst Recommendations as Material Nonpublic Information):

The next day, Clement is preparing to be interviewed on a global financial news television program where he will discuss his changed recommendation on Turgot Chariots for the first time in public. While preparing for the program, he mentions to the show's producers and Mary Zito, the journalist who will be interviewing him, the information he will be discussing. Just prior to going on the air, Zito sells her holdings in Turgot Chariots. She also phones her father with the information because she knows that he and other family members have investments in Turgot Chariots.

> *Comment*: When Zito receives advance notice of Clement's change of opinion, she knows it will have a material impact on the stock price, even if she is not totally aware of Clement's underlying reasoning. She is not a client of Clement but obtains early access to the material nonpublic information prior to publication. Her trades are thus based on material nonpublic information and violate Standard II(A).
>
> Zito further violates the Standard by relaying the information to her father. It would not matter if he or any other family member traded; the act of providing the information violates Standard II(A). The fact that the information is provided to a family member does not absolve someone of the prohibition of using or communicating material nonpublic information.

Example 8 (Acting on Nonpublic Information):

Ashton Kellogg is a retired investment professional who manages his own portfolio. He owns shares in National Savings, a large local bank. A close friend and golfing buddy, John Mayfield, is a senior executive at National. National has seen its stock price drop considerably, and the news and outlook are not good. In a conversation about the economy and the banking industry on the golf course, Mayfield relays the information that National will surprise the investment community in a few days when it announces excellent earnings for the quarter. Kellogg is pleasantly surprised by this information, and thinking that Mayfield, as a senior executive, knows the law and would not disclose inside information, he doubles his position in the bank. Subsequently, National announces that it had good operating earnings but had to set aside reserves for anticipated significant losses on its loan portfolio. The combined news causes the stock to go down 60%.

> *Comment*: Even though Kellogg believes that Mayfield would not break the law by disclosing inside information and money was lost on the purchase, Kellogg should not have purchased additional shares of National. It is the member's or candidate's responsibility to make sure, before executing investment actions, that comments about earnings are not material non-public information. Kellogg has violated Standard II(A).

Example 9 (Mosaic Theory):

John Doll is a research analyst for a hedge fund that also sells its research to a select group of paying client investment firms. Doll's focus is medical technology companies and products, and he has been in the business long enough and has been successful enough to build up a very credible network of friends and experts in the business. Doll has been working on a major research report recommending Boyce Health, a medical device manufacturer. He recently ran into an old acquaintance at a wedding who is a senior executive at Boyce, and Doll asked about the business. Doll was drawn to a statement that the executive, who has responsibilities in the new products area, made about a product: "I would not get too excited about the medium-term prospects; we have a lot of work to do first." Doll incorporated this and other information about the new Boyce product in his long-term recommendation of Boyce.

> *Comment*: Doll's conversation with the senior executive is part of the mosaic of information used in recommending Boyce. When holding discussions with a firm executive, Doll would need to guard against soliciting or obtaining material nonpublic information. Before issuing the report, the executive's statement about the continuing development of the product would need to be weighed against the other known public facts to determine whether it would be considered material.

Example 10 (Materiality Determination):

Larry Nadler, a trader for a mutual fund, gets a text message from another firm's trader, whom he has known for years. The message indicates a software company is going to report strong earnings when the firm publicly announces in two days. Nadler has a buy order from a portfolio manager within his firm to purchase several hundred thousand shares of the stock. Nadler is aggressive in placing the portfolio manager's order and completes the purchases by the following morning, a day ahead of the firm's planned earnings announcement.

> *Comment*: There are often rumors and whisper numbers before a release of any kind. The text message from the other trader would most likely be considered market noise. Unless Nadler knew that the trader had an ongoing business relationship with the public firm, he had no reason to suspect he was receiving material nonpublic information that would prevent him from completing the trading request of the portfolio manager.

Example 11 (Using an Expert Network):

Mary McCoy is the senior drug analyst at a mutual fund. Her firm hires a service that connects her to experts in the treatment of cancer. Through various phone conversations, McCoy enhances her understanding of the latest therapies for successful treatment. This information is critical to Mary making informed recommendations of the companies producing these drugs.

> *Comment*: McCoy is appropriately using the expert networks to enhance her evaluation process. She has neither asked for nor received information that may be considered material and nonpublic, such as preliminary trial results. McCoy is allowed to seek advice from professionals within the industry that she follows.

Example 12 (Using an Expert Network):

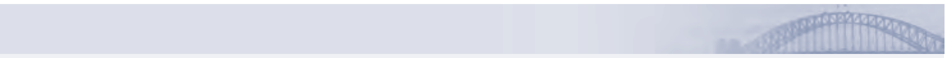
Tom Watson is a research analyst working for a hedge fund. To stay informed, Watson relies on outside experts for information on such industries as technology and pharmaceuticals, where new advancements occur frequently. The meetings with the industry experts often are arranged through networks or placement agents that have specific policies and procedures in place to deter the exchange of material nonpublic information.

Watson arranges a call to discuss future prospects for one of the fund's existing technology company holdings, a company that was testing a new semiconductor product. The scientist leading the tests indicates his disappointment with the performance of the new semiconductor. Following the call, Watson relays the insights he received to others at the fund. The fund sells its current position in the company and buys many put options because the market is anticipating the success of the new semiconductor and the share price reflects the market's optimism.

> *Comment*: Watson has violated Standard II(A) by passing along material nonpublic information concerning the ongoing product tests, which the fund used to trade in the securities and options of the related company. Watson cannot simply rely on the agreements signed by individuals who participate in expert networks that state that he has not received information that would prohibit his trading activity. He must make his own determination whether information he received through these arrangements reaches a materiality threshold that would affect his trading abilities.

STANDARD II(B): INTEGRITY OF CAPITAL MARKETS - MARKET MANIPULATION

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;



Members and Candidates must not engage in practices that distort prices or artificially inflate trading volume with the intent to mislead market participants.

Guidance

Highlights:

- *Information-Based Manipulation*
- *Transaction-Based Manipulation*

Standard II(B) requires that members and candidates uphold market integrity by prohibiting market manipulation. Market manipulation includes practices that distort security prices or trading volume with the intent to deceive people or entities that rely on information in the market. Market manipulation damages the interests of all investors by disrupting the smooth functioning of financial markets and lowering investor confidence.

Market manipulation may lead to a lack of trust in the fairness of the capital markets, resulting in higher risk premiums and reduced investor participation. A reduction in the efficiency of a local capital market may negatively affect the growth and economic health of the country and may also influence the operations of the globally interconnected capital markets. Although market manipulation may be less likely to occur in mature financial markets than in emerging markets, cross-border investing increasingly exposes all global investors to the potential for such practices.

Market manipulation includes (1) the dissemination of false or misleading information and (2) transactions that deceive or would be likely to mislead market participants by distorting the price-setting mechanism of financial instruments. The development of new products and technologies increases the incentives, means, and opportunities for market manipulation. Additionally, the increasing complexity and sophistication of the technologies used for communicating with market participants have created new avenues for manipulation.

Information-Based Manipulation

Information-based manipulation includes, but is not limited to, spreading false rumors to induce trading by others. For example, members and candidates must refrain from "pumping up" the price of an investment by issuing misleading positive information or overly optimistic projections of a security's worth only to later "dump" the investment (i.e., sell it) once the price, fueled by the misleading information's effect on other market participants, reaches an artificially high level.

Transaction-Based Manipulation

Transaction-based manipulation involves instances where a member or candidate knew or should have known that his or her actions could affect the pricing of a security. This type of manipulation includes, but is not limited to, the following:

- transactions that artificially affect prices or volume to give the impression of activity or price movement in a financial instrument, which represent a diversion from the expectations of a fair and efficient market, and
- securing a controlling, dominant position in a financial instrument to exploit and manipulate the price of a related derivative and/or the underlying asset.

Standard II(B) is not intended to preclude transactions undertaken on legitimate trading strategies based on perceived market inefficiencies. The intent of the action is critical to determining whether it is a violation of this standard.

STANDARD II(B): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Independent Analysis and Company Promotion):

The principal owner of Financial Information Services (FIS) entered into an agreement with two microcap companies to promote the companies' stock in exchange for stock and cash compensation. The principal owner caused FIS to disseminate e-mails, design and maintain several websites, and distribute an online investment newsletter—all of which recommended investment in the two companies. The systematic publication of purportedly independent analyses and recommendations containing inaccurate and highly promotional and speculative statements increased public investment in the companies and led to dramatically higher stock prices.

> *Comment*: The principal owner of FIS violated Standard II(B) by using inaccurate reporting and misleading information under the guise of independent analysis to artificially increase the stock price of the companies. Furthermore, the principal owner violated Standard V(A)–Diligence and Reasonable Basis by not having a reasonable and adequate basis for recommending the two companies and violated Standard VI(A)–Disclosure of Conflicts by not disclosing to investors the compensation agreements (which constituted a conflict of interest).

Example 2 (Personal Trading Practices and Price):

John Gray is a private investor in Belgium who bought a large position several years ago in Fame Pharmaceuticals, a German small-cap security with limited average trading volume. He has now decided to significantly reduce his holdings owing to the poor price performance. Gray is worried that the low trading volume for the stock may cause the price to decline further as he attempts to sell his large position.

Gray devises a plan to divide his holdings into multiple accounts in different brokerage firms and private banks in the names of family members, friends, and even a private religious institution. He then creates a rumor campaign on various blogs and social media outlets promoting the company.

Gray begins to buy and sell the stock using the accounts in hopes of raising the trading volume and the price. He conducts the trades through multiple brokers, selling slightly larger positions than he bought on a tactical schedule, and over time, he is able to reduce his holding as desired without negatively affecting the sale price.

> *Comment*: John violated Standard II(B) by fraudulently creating the appearance that there was a greater investor interest in the stock through the online rumors. Additionally, through his trading strategy, he created the appearance that there was greater liquidity in the stock than actually existed. He was able to manipulate the price through both misinformation and trading practices.

Example 3 (Creating Artificial Price Volatility):

Matthew Murphy is an analyst at Divisadero Securities & Co., which has a significant number of hedge funds among its most important brokerage clients. Some of the hedge funds hold short positions on Wirewolf Semiconductor. Two trading days before the publication of a quarter-end report, Murphy alerts his sales force that he is about to issue a research report on Wirewolf that will include the following opinions:

- quarterly revenues are likely to fall short of management's guidance,
- earnings will be as much as 5 cents per share (or more than 10%) below consensus, and
- Wirewolf's highly respected chief financial officer may be about to join another company.

Knowing that Wirewolf has already entered its declared quarter-end "quiet period" before reporting earnings (and thus would be reluctant to respond to rumors), Murphy times the release of his research report specifically to sensationalize the negative aspects of the message in order to create significant downward pressure on Wirewolf's stock—to the distinct advantage of Divisadero's hedge fund clients. The report's conclusions are based on speculation, not on fact. The next day, the research report is broadcast to all of Divisadero's clients and to the usual newswire services.

Before Wirewolf's investor-relations department can assess the damage on the final trading day of the quarter and refute Murphy's report, its stock opens trading sharply lower, allowing Divisadero's clients to cover their short positions at substantial gains.

> *Comment*: Murphy violated Standard II(B) by aiming to create artificial price volatility designed to have a material impact on the price of an issuer's stock. Moreover, by lacking an adequate basis for the recommendation, Murphy also violated Standard V(A)–Diligence and Reasonable Basis.

Example 4 (Personal Trading and Volume):

Rajesh Sekar manages two funds—an equity fund and a balanced fund—whose equity components are supposed to be managed in accordance with the same model. According to that model, the funds' holdings in stock of Digital Design Inc. (DD) are excessive. Reduction of the DD holdings would not be easy, however, because the stock has low liquidity in the stock market. Sekar decides to start trading larger portions of DD stock back and forth between his two funds to slowly increase the price; he believes market participants will see growing volume and increasing price and become interested in the stock. If other investors are willing to buy the DD stock because of such interest, then Sekar will be able to get rid of at least some of his overweight position without inducing price decreases. In this way, the whole transaction will be for the benefit of fund participants, even if additional brokers' commissions are incurred.

> *Comment*: Sekar's plan would be beneficial for his funds' participants but is based on artificial distortion of both trading volume and the price of the DD stock and thus constitutes a violation of Standard II(B).

Example 5 ("Pump-Priming" Strategy):

ACME Futures Exchange is launching a new bond futures contract. To convince investors, traders, arbitrageurs, hedgers, and so on, to use its contract, the exchange attempts to demonstrate that it has the best liquidity. To do so, it enters into agreements with members in which they commit to a substantial minimum trading volume on the new contract over a specific period in exchange for substantial reductions of their regular commissions.

> *Comment*: The formal liquidity of a market is determined by the obligations set on market makers, but the actual liquidity of a market is better estimated by the actual trading volume and bid–ask spreads. Attempts to mislead participants about the actual liquidity of the market constitute a violation of Standard II(B). In this example, investors have been intentionally misled to believe they chose the most liquid instrument for some specific purpose, but they could eventually see the actual liquidity of the contract significantly reduced after the term of the agreement expires. If the ACME Futures Exchange fully discloses its agreement with members to boost transactions over some initial launch period, it will not violate Standard II(B). ACME's intent is not to harm investors but, on the contrary, to give them a better service. For that purpose, it may engage in a liquidity-pumping strategy, but the strategy must be disclosed.

Example 6 (Creating Artificial Price Volatility):

Emily Gordon, an analyst of household products companies, is employed by a research boutique, Picador & Co. Based on information that she has gathered during a trip through Latin America, she believes that Hygene, Inc., a major marketer of personal care products, has generated better-than-expected sales from its new product initiatives in South America. After modestly boosting her projections for revenue and for gross profit margin in her worksheet models for Hygene, Gordon estimates that her earnings projection of US$2.00 per diluted share for the current year may be as much as 5% too low. She contacts the chief financial officer (CFO) of Hygene to try to gain confirmation of her findings from her trip and to get some feedback regarding her revised models. The CFO declines to comment and reiterates management's most recent guidance of US$1.95–US$2.05 for the year.

Gordon decides to try to force a comment from the company by telling Picador & Co. clients who follow a momentum investment style that consensus earnings projections for Hygene are much too low; she explains that she is considering raising her published estimate by an ambitious US$0.15 to US$2.15 per share. She believes that when word of an unrealistically high earnings projection filters back to Hygene's investor-relations department, the company will feel compelled to update its earnings guidance. Meanwhile, Gordon hopes that she is at least correct with respect to the earnings direction and that she will help clients who act on her insights to profit from a quick gain by trading on her advice.

> *Comment*: By exaggerating her earnings projections in order to try to fuel a quick gain in Hygene's stock price, Gordon is in violation of Standard II(B). Furthermore, by virtue of previewing her intentions of revising upward her earnings projections to only a select group of clients, she is in violation of Standard III(B)–Fair Dealing. However, it would have been acceptable for Gordon to write a report that
>
> - framed her earnings projection in a range of possible outcomes,
> - outlined clearly the assumptions used in her Hygene models that took into consideration the findings from her trip through Latin America, and
> - was distributed to all Picador & Co. clients in an equitable manner.

Example 7 (Pump and Dump Strategy):

In an effort to pump up the price of his holdings in Moosehead & Belfast Railroad Company, Steve Weinberg logs on to several investor chat rooms on the internet to start rumors that the company is about to expand its rail network in anticipation of receiving a large contract for shipping lumber.

> *Comment*: Weinberg has violated Standard II(B) by disseminating false information about Moosehead & Belfast with the intent to mislead market participants.

Example 8 (Manipulating Model Inputs):

Bill Mandeville supervises a structured financing team for Superior Investment Bank. His responsibilities include packaging new structured investment products and managing Superior's relationship with relevant rating agencies. To achieve the best rating possible, Mandeville uses mostly positive scenarios as model inputs—scenarios that reflect minimal downside risk in the assets underlying the structured products. The resulting output statistics in the rating request and underwriting prospectus support the idea that the new structured products have minimal potential downside risk. Additionally, Mandeville's compensation from Superior is partially based on both the level of the rating assigned and the successful sale of new structured investment products but does not have a link to the long-term performance of the instruments.

Mandeville is extremely successful and leads Superior as the top originator of structured investment products for the next two years. In the third year, the economy experiences difficulties and the values of the assets underlying structured products significantly decline. The subsequent defaults lead to major turmoil in the capital markets, the demise of Superior Investment Bank, and the loss of Mandeville's employment.

> *Comment*: Mandeville manipulates the inputs of a model to minimize associated risk to achieve higher ratings. His understanding of structured products allows him to skillfully decide which inputs to include in support of the desired rating and price. This information manipulation for short-term gain, which is in violation of Standard II(B), ultimately causes significant damage to many parties and the capital markets as a whole. Mandeville should have realized that promoting a rating and price with inaccurate information could cause not only a loss of price confidence in the particular structured product but also a loss of investor trust in the system. Such loss of confidence affects the ability of the capital markets to operate efficiently.

Example 9 (Information Manipulation):

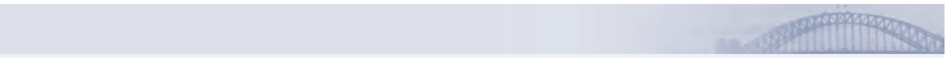
Allen King is a performance analyst for Torrey Investment Funds. King believes that the portfolio manager for the firm's small- and microcap equity fund dislikes him because the manager never offers him tickets to the local baseball team's games but does offer tickets to other employees. To incite a potential regulatory review of the manager, King creates user profiles on several online forums under the portfolio manager's name and starts rumors about potential mergers for several of the smaller companies in the portfolio. As the prices of these companies' stocks increase, the portfolio manager sells the position, which leads to an investigation by the regulator as King desired.

> *Comment*: King has violated Standard II(B) even though he did not personally profit from the market's reaction to the rumor. In posting the false information, King misleads others into believing the companies were likely to be acquired. Although his intent was to create trouble for the portfolio manager, his actions clearly manipulated the factual information that was available to the market.

STANDARD III(A): DUTIES TO CLIENTS - LOYALTY, PRUDENCE, AND CARE

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Standard III(A) Loyalty, Prudence, and Care



Members and Candidates have a duty of loyalty to their clients and must act with reasonable care and exercise prudent judgment. Members and Candidates must act for the benefit of their clients and place their clients' interests before their employer's or their own interests.

Guidance

Highlights:

- *Understanding the Application of Loyalty, Prudence, and Care*
- *Identifying the Actual Investment Client*
- *Developing the Client's Portfolio*
- *Soft Commission Policies*
- *Proxy Voting Policies*

Standard III(A) clarifies that client interests are paramount. A member's or candidate's responsibility to a client includes a duty of loyalty and a duty to exercise reasonable care. Investment actions must be carried out for the sole benefit of the client and in a manner the member or candidate believes, given the known facts and circumstances, to be in the best interest of the client. Members and candidates must exercise the same level of prudence, judgment, and care that they would apply in the management and disposition of their own interests in similar circumstances.

Prudence requires caution and discretion. The exercise of prudence by investment professionals requires that they act with the care, skill, and diligence that a reasonable person acting in a like capacity and familiar with such matters would use. In the context of managing a client's portfolio, prudence requires following the investment parameters set forth by the client and balancing risk and return. Acting with care requires members and candidates to act in a prudent and judicious manner in avoiding harm to clients.

Standard III(A) sets minimum expectations for members and candidates when fulfilling their responsibilities to their clients. Regulatory and legal requirements for such duties can vary across the investment industry depending on a variety of factors, including job function of the investment professional, the existence of an adviser/client relationship, and the nature of the recommendations being offered. From the perspective of the end user of financial services, these different standards can be arcane and confusing, leaving investors unsure of what level of service to expect from investment professionals they employ. The single standard of conduct described in Standard III(A) benefits investors by establishing a benchmark for the duties of loyalty, prudence, and care and clarifies that all CFA Institute members and candidates, regardless of job title, local laws, or cultural differences, are required to comply with these fundamental responsibilities. Investors hiring members or candidates who must adhere to the duty of loyalty, prudence, and care set forth in this standard can be confident that these responsibilities are a requirement regardless of any legally imposed fiduciary duties.

Standard III(A), however, is not a substitute for a member's or candidate's legal or regulatory obligations. As stated in Standard I(A), members and candidates must abide by the most strict requirements imposed on them by regulators or the Code and Standards, including any legally imposed fiduciary duty. Members and candidates must also be aware of whether they have "custody" or effective control of client assets. If so, a heightened level of responsibility arises. Members and candidates are considered to have custody if they have any direct or indirect access to client funds. Members and candidates must manage any pool of assets in their control in accordance with the terms of the governing documents (such as trust documents and investment management agreements), which are the primary determinant of the manager's powers and duties. Whenever their actions are contrary to provisions of those instruments or applicable law, members and candidates are at risk of violating Standard III(A).

Understanding the Application of Loyalty, Prudence, and Care

Standard III(A) establishes a minimum benchmark for the duties of loyalty, prudence, and care that are required of all members and candidates regardless of whether a legal fiduciary duty applies. Although fiduciary duty often encompasses the principles of loyalty, prudence, and care, Standard III(A) does not render all members and candidates fiduciaries. The responsibilities of members and candidates for fulfilling their obligations under this standard depend greatly on the nature of their professional responsibilities and the relationships they have with clients. The conduct of members and candidates may or may not rise to the level of being a fiduciary, depending on the type of client, whether the member or candidate is giving investment advice, and the many facts and circumstances surrounding a particular transaction or client relationship.

Fiduciary duties are often imposed by law or regulation when an individual or institution is charged with the duty of acting for the benefit of another party, such as managing investment assets. The duty required in fiduciary relationships exceeds what is acceptable in many other business relationships because a fiduciary is in an enhanced position of trust. Although members and candidates must comply with any legally imposed fiduciary duty, the Code and Standards neither impose such a legal responsibility nor require all members or candidates to act as fiduciaries. However, Standard III(A) requires members and candidates to work in the client's best interest no matter what the job function.

A member or candidate who does not provide advisory services to a client but who acts only as a trade execution professional must prudently work in the client's interest when completing requested trades. Acting in the client's best interest requires these professionals to use their skills and diligence to execute trades in the most favorable terms that can be achieved. Members and candidates operating in such positions must use care to operate within the parameters set by the client's trading instructions.

Members and candidates may also operate in a blended environment where they execute client trades and offer advice on a limited set of investment options. The extent of the advisory arrangement and limitations should be outlined in the agreement with the client at the outset of the relationship. For instance, members and candidates should inform clients that the advice provided will be limited to the propriety products of the firm and not include other products available on the market. Clients who want access to a wider range of investment products would have the information necessary to decide not to engage with members or candidates working under these restrictions.

Members and candidates operating in this blended context would comply with their obligations by recommending the allowable products that are consistent with the client's objectives and risk tolerance. They would exercise care through diligently aligning the client's needs with the attributes of the products being recommended. Members and candidates should place the client's interests first by disregarding any firm or personal interest in motivating a recommended transaction.

There is a large variety of professional relationships that members and candidates have with their clients. Standard III(A) requires them to fulfill the obligations outlined explicitly or implicitly in the client agreements to the best of their abilities and with loyalty, prudence, and care. Whether a member or candidate is structuring a new securitization transaction, completing a credit rating analysis, or leading a public company, he or she must work with prudence and care in delivering the agreed-on services.

Identifying the Actual Investment Client

The first step for members and candidates in fulfilling their duty of loyalty to clients is to determine the identity of the "client" to whom the duty of loyalty is owed. In the context of an investment manager managing the personal assets of an individual, the client is easily identified. When the manager is responsible for the portfolios of pension plans or trusts, however, the client is not the person or entity who hires the manager but, rather, the beneficiaries of the plan or trust. The duty of loyalty is owed to the ultimate beneficiaries.

In some situations, an actual client or group of beneficiaries may not exist. Members and candidates managing a fund to an index or an expected mandate owe the duty of loyalty, prudence, and care to invest in a manner consistent with the stated mandate. The decisions of a fund's manager, although benefiting all fund investors, do not have to be based on an individual investor's requirements and risk profile. Client loyalty and care for those investing in the fund are the responsibility of members and candidates who have an advisory relationship with those individuals.

Situations involving potential conflicts of interest with respect to responsibilities to clients may be extremely complex because they may involve a number of competing interests. The duty of loyalty, prudence, and care applies to a large number of persons in varying capacities, but the exact duties may differ in many respects in accord with the relationship with each client or each type of account in which the assets are managed. Members and candidates must not only put their obligations to clients first in all dealings but also endeavor to avoid all real or potential conflicts of interest.

Members and candidates with positions whose responsibilities do not include direct investment management also have "clients" that must be considered. Just as there are various types of advisory relationships, members and candidates must look at their roles and responsibilities when making a determination of who their clients are. Sometimes the client is easily identifiable; such is the case in the relationship between a company executive and the firm's public shareholders. At other times, the client may be the investing public as a whole, in which case the goals of independence and objectivity of research surpass the goal of loyalty to a single organization.

Developing the Client's Portfolio

The duty of loyalty, prudence, and care owed to the individual client is especially important because the professional investment manager typically possesses greater knowledge in the investment arena than the client does. This disparity places the individual client in a vulnerable position; the client must trust the manager. The manager in these situations should ensure that the client's objectives and expectations for the performance of the account are realistic and suitable to the client's circumstances and that the risks involved are appropriate. In most circumstances, recommended investment strategies should relate to the long-term objectives and circumstances of the client.

Particular care must be taken to detect whether the goals of the investment manager or the firm in conducting business, selling products, and executing security transactions potentially conflict with the best interests and objectives of the client. When members and candidates cannot avoid potential conflicts between their firm and clients' interests, they must provide clear and factual disclosures of the circumstances to the clients.

Members and candidates must follow any guidelines set by their clients for the management of their assets. Some clients, such as charitable organizations and pension plans, have strict investment policies that limit investment options to certain types or classes of investment or prohibit investment in certain securities. Other organizations have aggressive policies that do not prohibit investments by type but, instead, set criteria on the basis of the portfolio's total risk and return.

Investment decisions must be judged in the context of the total portfolio rather than by individual investment within the portfolio. The member's or candidate's duty is satisfied with respect to a particular investment if the individual has thoroughly considered the investment's place in the overall portfolio, the risk of loss and opportunity for gains, tax implications, and the diversification, liquidity, cash flow, and overall return requirements of the assets or the portion of the assets for which the manager is responsible.

Soft Commission Policies

An investment manager often has discretion over the selection of brokers executing transactions. Conflicts may arise when an investment manager uses client brokerage to purchase research services, a practice commonly called "soft dollars" or "soft commissions." A member or candidate who pays a higher brokerage commission than he or she would normally pay to allow for the purchase of goods or services, without corresponding benefit to the client, violates the duty of loyalty to the client.

From time to time, a client will direct a manager to use the client's brokerage to purchase goods or services for the client, a practice that is commonly called "directed brokerage." Because brokerage commission is an asset of the client and is used to benefit that client, not the manager, such a practice does not violate any duty of loyalty. However, a member or candidate is obligated to seek "best price" and "best execution" and be assured by the client that the goods or services purchased from the brokerage will benefit the account beneficiaries. "Best execution" refers to a trading process that seeks to maximize the value of the client's portfolio within the client's stated investment objectives and constraints. In addition, the member or candidate should disclose to the client that the client may not be getting best execution from the directed brokerage.

Proxy Voting Policies

The duty of loyalty, prudence, and care may apply in a number of situations facing the investment professional besides those related directly to investing assets.

Part of a member's or candidate's duty of loyalty includes voting proxies in an informed and responsible manner. Proxies have economic value to a client, and members and candidates must ensure that they properly safeguard and maximize this value. An investment manager who fails to vote, casts a vote without considering the impact of the question, or votes blindly with management on nonroutine governance issues (e.g., a change in company capitalization) may violate this standard. Voting of proxies is an integral part of the management of investments.

A cost–benefit analysis may show that voting all proxies may not benefit the client, so voting proxies may not be necessary in all instances. Members and candidates should disclose to clients their proxy voting policies.

STANDARD III(A): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Regular Account Information

Members and candidates with control of client assets (1) should submit to each client, at least quarterly, an itemized statement showing the funds and securities in the custody or possession of the member or candidate plus all debits, credits, and transactions that occurred during the period, (2) should disclose to the client where the assets are to be maintained, as well as where or when they are moved, and (3) should separate the client's assets from any other party's assets, including the member's or candidate's own assets.

Client Approval

If a member or candidate is uncertain about the appropriate course of action with respect to a client, the member or candidate should consider what he or she would expect or demand if the member or candidate were the client. If in doubt, a member or candidate should disclose the questionable matter in writing to the client and obtain client approval.

Firm Policies

Members and candidates should address and encourage their firms to address the following topics when drafting the statements or manuals containing their policies and procedures regarding responsibilities to clients:

- *Follow all applicable rules and laws*: Members and candidates must follow all legal requirements and applicable provisions of the Code and Standards.
- *Establish the investment objectives of the client*: Make a reasonable inquiry into a client's investment experience, risk and return objectives, and financial constraints prior to making investment recommendations or taking investment actions.

- *Consider all the information when taking actions*: When taking investment actions, members and candidates must consider the appropriateness and suitability of the investment relative to (1) the client's needs and circumstances, (2) the investment's basic characteristics, and (3) the basic characteristics of the total portfolio.
- *Diversify*: Members and candidates should diversify investments to reduce the risk of loss, unless diversification is not consistent with plan guidelines or is contrary to the account objectives.
- *Carry out regular reviews*: Members and candidates should establish regular review schedules to ensure that the investments held in the account adhere to the terms of the governing documents.
- *Deal fairly with all clients with respect to investment actions*: Members and candidates must not favor some clients over others and should establish policies for allocating trades and disseminating investment recommendations.
- *Disclose conflicts of interest*: Members and candidates must disclose all actual and potential conflicts of interest so that clients can evaluate those conflicts.
- *Disclose compensation arrangements*: Members and candidates should make their clients aware of all forms of manager compensation.
- *Vote proxies*: In most cases, members and candidates should determine who is authorized to vote shares and vote proxies in the best interests of the clients and ultimate beneficiaries.
- *Maintain confidentiality*: Members and candidates must preserve the confidentiality of client information.
- *Seek best execution*: Unless directed by the client as ultimate beneficiary, members and candidates must seek best execution for their clients. (Best execution is defined in the preceding text.)
- *Place client interests first*: Members and candidates must serve the best interests of clients.

STANDARD III(A): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Identifying the Client—Plan Participants):

First Country Bank serves as trustee for the Miller Company's pension plan. Miller is the target of a hostile takeover attempt by Newton, Inc. In attempting to ward off Newton, Miller's managers persuade Julian Wiley, an investment manager at First Country Bank, to purchase Miller common stock in the open market for the employee pension plan. Miller's officials indicate that such action would be favorably received and would probably result in other accounts being placed with the bank. Although Wiley believes the stock is overvalued and would not ordinarily buy it, he purchases the stock to support Miller's managers, to maintain Miller's good favor toward the bank, and to realize additional new business. The heavy stock purchases cause Miller's market price to rise to such a level that Newton retracts its takeover bid.

> *Comment*: Standard III(A) requires that a member or candidate, in evaluating a takeover bid, act prudently and solely in the interests of plan participants and beneficiaries. To meet this requirement, a member or candidate must carefully evaluate the long-term prospects of the company against the short-term prospects presented by the takeover offer and by the ability to invest elsewhere. In this instance, Wiley, acting on behalf of his employer, which was the trustee for a pension plan, clearly violated Standard III(A). He used the pension plan to perpetuate existing management, perhaps to the detriment of plan participants and the company's shareholders, and to benefit himself. Wiley's responsibilities to the plan participants and beneficiaries should have taken precedence over any ties of his bank to corporate managers and over his self-interest. Wiley had a duty to examine the takeover offer on its own merits and to make an independent decision. The guiding principle is the appropriateness of the investment decision to the pension plan, not whether the decision benefited Wiley or the company that hired him.

Example 2 (Client Commission Practices):

JNI, a successful investment counseling firm, serves as investment manager for the pension plans of several large regionally based companies. Its trading activities generate a significant amount of commission-related business. JNI uses the brokerage and research services of many firms, but most of its trading activity is handled through a large brokerage company, Thompson, Inc., because the executives of the two firms have a close friendship. Thompson's commission structure is high in comparison with charges for similar brokerage services from other firms. JNI considers Thompson's research services and execution capabilities average. In exchange for JNI directing its brokerage to Thompson, Thompson absorbs a number of JNI overhead expenses, including those for rent.

> *Comment*: JNI executives are breaching their responsibilities by using client brokerage for services that do not benefit JNI clients and by not obtaining best price and best execution for their clients. Because JNI executives are not upholding their duty of loyalty, they are violating Standard III(A).

Example 3 (Brokerage Arrangements):

Charlotte Everett, a struggling independent investment adviser, serves as investment manager for the pension plans of several companies. One of her brokers, Scott Company, is close to consummating management agreements with prospective new clients whereby Everett would manage the new client accounts and trade the accounts exclusively through Scott. One of Everett's existing clients, Crayton Corporation, has directed Everett to place securities transactions for Crayton's account exclusively through Scott. But to induce Scott to exert efforts to send more new accounts to her, Everett also directs transactions to Scott from other clients without their knowledge.

> *Comment*: Everett has an obligation at all times to seek best price and best execution on all trades. Everett may direct new client trades exclusively through Scott Company as long as Everett receives best price and execution on the trades or receives a written statement from new clients that she is *not* to seek best price and execution and that they are aware of the consequence for their accounts. Everett may trade other accounts through Scott as a

> reward for directing clients to Everett only if the accounts receive best price and execution and the practice is disclosed to the accounts. Because Everett does not disclose the directed trading, Everett has violated Standard III(A).

Example 4 (Brokerage Arrangements):

Emilie Rome is a trust officer for Paget Trust Company. Rome's supervisor is responsible for reviewing Rome's trust account transactions and her monthly reports of personal stock transactions. Rome has been using Nathan Gray, a broker, almost exclusively for trust account brokerage transactions. When Gray makes a market in stocks, he has been giving Rome a lower price for personal purchases and a higher price for sales than he gives to Rome's trust accounts and other investors.

> *Comment*: Rome is violating her duty of loyalty to the bank's trust accounts by using Gray for brokerage transactions simply because Gray trades Rome's personal account on favorable terms. Rome is placing her own interests before those of her clients.

Example 5 (Client Commission Practices):

Lauren Parker, an analyst with Provo Advisors, covers South American equities for her firm. She likes to travel to the markets for which she is responsible and decides to go on a trip to Chile, Argentina, and Brazil. The trip is sponsored by SouthAM, Inc., a research firm with a small broker/dealer affiliate that uses the clearing facilities of a larger New York brokerage house. SouthAM specializes in arranging South American trips for analysts during which they can meet with central bank officials, government ministers, local economists, and senior executives of corporations. SouthAM accepts commission dollars at a ratio of 2 to 1 against the hard-dollar costs of the research fee for the trip. Parker is not sure that SouthAM's execution is competitive, but without informing her supervisor, she directs the trading desk at Provo to start giving commission business to SouthAM so she can take the trip. SouthAM has conveniently timed the briefing trip to coincide with the beginning of Carnival season, so Parker also decides to spend five days of vacation in Rio de Janeiro at the end of the trip. Parker uses commission dollars to pay for the five days of hotel expenses.

> *Comment*: Parker is violating Standard III(A) by not exercising her duty of loyalty to her clients. She should have determined whether the commissions charged by SouthAM are reasonable in relation to the benefit of the research provided by the trip. She also should have determined whether best execution and prices could be received from SouthAM. In addition, the five extra days are not part of the research effort because they do not assist in the investment decision making. Thus, the hotel expenses for the five days should not be paid for with client assets.

Example 6 (Excessive Trading):

Vida Knauss manages the portfolios of a number of high-net-worth individuals. A major part of her investment management fee is based on trading commissions. Knauss engages in extensive trading for each of her clients to ensure that she attains the minimum commission level set by her firm. Although the securities purchased and sold for the clients are appropriate and fall within the acceptable asset classes for the clients, the amount of trading for each account exceeds what is necessary to accomplish the client's investment objectives.

> *Comment*: Knauss has violated Standard III(A) because she is using the assets of her clients to benefit her firm and herself.

Example 7 (Managing Family Accounts):

Adam Dill recently joined New Investments Asset Managers. To assist Dill in building a book of clients, both his father and brother opened new fee-paying accounts. Dill followed all the firm's procedures in noting his relationships with these clients and in developing their investment policy statements.

After several years, the number of Dill's clients has grown, but he still manages the original accounts of his family members. An IPO is coming to market that is a suitable investment for many of his clients, including his brother. Dill does not receive the amount of stock he requested, so to avoid any appearance of a conflict of interest, he does not allocate any shares to his brother's account.

> *Comment*: Dill has violated Standard III(A) because he is not acting for the benefit of his brother's account as well as his other accounts. The brother's account is a regular fee-paying account comparable to the accounts of his other clients. By not allocating the shares proportionately across *all* accounts for which he thought the IPO was suitable, Dill is disadvantaging specific clients.
>
> Dill would have been correct in not allocating shares to his brother's account if that account was being managed outside the normal fee structure of the firm.

Example 8 (Identifying the Client):

Donna Hensley has been hired by a law firm to testify as an expert witness. Although the testimony is intended to represent impartial advice, she is concerned that her work may have negative consequences for the law firm. If the law firm is Hensley's client, how does she ensure that her testimony will not violate the required duty of loyalty, prudence, and care to one's client?

> *Comment*: In this situation, the law firm represents Hensley's employer and the aspect of "who is the client" is not well defined. When acting as an expert witness, Hensley is bound by the standard of independence and objectivity in the same manner as an independent research analyst would be bound. Hensley must not let the law firm influence the testimony she provides in the legal proceedings.

Example 9 (Identifying the Client):

Jon Miller is a mutual fund portfolio manager. The fund is focused on the global financial services sector. Wanda Spears is a private wealth manager in the same city as Miller and is a friend of Miller. At a local CFA Institute society meeting, Spears mentions to Miller that her new client is an investor in Miller's fund. She states that the two of them now share a responsibility to this client.

> *Comment*: Spears' statement is not totally correct. Because she provides the advisory services to her new client, she alone is bound by the duty of loyalty to this client. Miller's responsibility is to manage the fund according to the investment policy statement of the fund. His actions should not be influenced by the needs of any particular fund investor.

Example 10 (Client Loyalty):

After providing client account investment performance to the external-facing departments but prior to it being finalized for release to clients, Teresa Nguyen, an investment performance analyst, notices the reporting system missed a trade. Correcting the omission resulted in a large loss for a client that had previously placed the firm on “watch” for potential termination owing to underperformance in prior periods. Nguyen knows this news is unpleasant but informs the appropriate individuals that the report needs to be updated before releasing it to the client.

> *Comment*: Nguyen’s actions align with the requirements of Standard III(A). Even though the correction may lead to the firm’s termination by the client, withholding information on errors would not be in the best interest of the client.

Example 11 (Execution-Only Responsibilities):

Baftija Sulejman recently became a candidate in the CFA Program. He is a broker who executes client-directed trades for several high-net-worth individuals. Sulejman does not provide any investment advice and only executes the trading decisions made by clients. He is concerned that the Code and Standards impose a fiduciary duty on him in his dealing with clients and sends an e-mail to the CFA Ethics Helpdesk (ethics@cfainstitute.org) to seek guidance on this issue.

> *Comment*: In this instance, Sulejman serves in an execution-only capacity and his duty of loyalty, prudence, and care is centered on the skill and diligence used when executing trades—namely, by seeking best execution and making trades within the parameters set by the clients (instructions on quantity, price, timing, etc.). Acting in the best interests of the client dictates that trades are executed on the most favorable terms that can be achieved for the client. Given this job function, the requirements of the Code and Standards for loyalty, prudence, and care clearly do not impose a fiduciary duty.

STANDARD III(B): DUTIES TO CLIENTS - FAIR DEALING

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Members and Candidates must deal fairly and objectively with all clients when providing investment analysis, making investment recommendations, taking investment action, or engaging in other professional activities.

Guidance

Highlights:

- *Investment Recommendations*
- *Investment Action*

Standard III(B) requires members and candidates to treat all clients fairly when disseminating investment recommendations or making material changes to prior investment recommendations or when taking investment action with regard to general purchases, new issues, or secondary offerings. Only through the fair treatment of all parties can the investment management profession maintain the confidence of the investing public.

When an investment adviser has multiple clients, the potential exists for the adviser to favor one client over another. This favoritism may take various forms—from the quality and timing of services provided to the allocation of investment opportunities.

The term "fairly" implies that the member or candidate must take care not to discriminate against any clients when disseminating investment recommendations or taking investment action. Standard III(B) does not state "equally" because members and candidates could not possibly reach all clients at exactly the same time—whether by printed mail, telephone (including text messaging), computer (including internet updates and e-mail distribution), facsimile (fax), or wire. Each client has unique needs, investment criteria, and investment objectives, so not all investment opportunities are suitable for all clients. In addition, members and candidates may provide more personal, specialized, or in-depth service to clients who are willing to pay for premium services through higher management fees or higher levels of brokerage. Members and candidates may differentiate their services to clients, but different levels of service must not disadvantage or negatively affect clients. In addition, the different service levels should be disclosed to clients and prospective clients and should be available to everyone (i.e., different service levels should not be offered selectively).

Standard III(B) covers conduct in two broadly defined categories—investment recommendations and investment action.

Investment Recommendations

The first category of conduct involves members and candidates whose primary function is the preparation of investment recommendations to be disseminated either to the public or within a firm for the use of others in making investment decisions. This group includes members and candidates employed by investment counseling, advisory, or consulting firms as well as banks, brokerage firms, and insurance companies. The criterion is that the member's or candidate's primary responsibility is the preparation of recommendations to be acted on by others, including those in the member's or candidate's organization.

An investment recommendation is any opinion expressed by a member or candidate in regard to purchasing, selling, or holding a given security or other investment. The opinion may be disseminated to customers or clients through an initial detailed research report, through a brief update report, by addition to or deletion from a list of recommended securities, or simply by oral communication. A recommendation that is distributed to anyone outside the organization is considered a communication for general distribution under Standard III(B).

Standard III(B) addresses the manner in which investment recommendations or changes in prior recommendations are disseminated to clients. Each member or candidate is obligated to ensure that information is disseminated in such a manner that all clients have a fair opportunity to act on every recommendation. Communicating with all clients on a uniform basis presents practical problems for members and candidates because of differences in timing and methods of communication with various types of customers and clients. Members and candidates should encourage their firms to design an equitable system to prevent selective or discriminatory disclosure and should inform clients about what kind of communications they will receive.

The duty to clients imposed by Standard III(B) may be more critical when members or candidates change their recommendations than when they make initial recommendations. Material changes in a member's or candidate's prior investment recommendations because of subsequent research should be communicated to all current clients; particular care should be taken that the information reaches those clients who the member or candidate knows have acted on or been affected by the earlier advice. Clients who do not know that the member or candidate has changed a recommendation and who, therefore, place orders contrary to a current recommendation should be advised of the changed recommendation before the order is accepted.

Investment Action

The second category of conduct includes those members and candidates whose primary function is taking investment action (portfolio management) on the basis of recommendations prepared internally or received from external sources. Investment action, like investment recommendations, can affect market value. Consequently, Standard III(B) requires that members or candidates treat all clients fairly in light of their investment objectives and circumstances. For example, when making investments in new offerings or in secondary financings, members and candidates should distribute the issues to all customers for whom the investments are appropriate in a manner consistent with the policies of the firm for allocating blocks of stock. If the issue is oversubscribed, then the issue should be prorated to all subscribers. This action should be taken on a round-lot basis to avoid odd-lot distributions. In addition, if the issue is oversubscribed, members and candidates should forgo any sales to themselves or their immediate families in order to free up additional shares for clients. If the investment professional's family-member accounts are managed similarly to the accounts of other clients of the firm, however, the family-member accounts should not be excluded from buying such shares.

Members and candidates must make every effort to treat all individual and institutional clients in a fair and impartial manner. A member or candidate may have multiple relationships with an institution; for example, the member or candidate may be a corporate trustee, pension fund manager, manager of funds for individuals employed by the customer, loan originator, or creditor. A member or candidate must exercise care to treat all clients fairly.

Members and candidates should disclose to clients and prospective clients the documented allocation procedures they or their firms have in place and how the procedures would affect the client or prospect. The disclosure should be clear and complete so that the client can make an informed investment decision. Even when complete disclosure is made, however, members and candidates must put client interests ahead of their own. A member's or candidate's duty of fairness and loyalty to clients can never be overridden by client consent to patently unfair allocation procedures.

Treating clients fairly also means that members and candidates should not take advantage of their position in the industry to the detriment of clients. For instance, in the context of IPOs, members and candidates must make bona fide public distributions of "hot issue" securities (defined as securities of a public offering that are trading at a premium in the secondary market whenever such trading commences because of the great demand for the securities). Members and candidates are prohibited from withholding such securities for their own benefit and must not use such securities as a reward or incentive to gain benefit.

STANDARD III(B): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Develop Firm Policies

Although Standard III(B) refers to a member's or candidate's responsibility to deal fairly and objectively with clients, members and candidates should also encourage their firms to establish compliance procedures requiring all employees who disseminate investment recommendations or take investment actions to treat customers and clients fairly. At the very least, a member or candidate should recommend appropriate procedures to management if none are in place. And the member or candidate should make management aware of possible violations of fair-dealing practices within the firm when they come to the attention of the member or candidate.

The extent of the formality and complexity of such compliance procedures depends on the nature and size of the organization and the type of securities involved. An investment adviser who is a sole proprietor and handles only discretionary accounts might not disseminate recommendations to the public, but that adviser should have formal written procedures to ensure that all clients receive fair investment action.

Good business practice dictates that initial recommendations be made available to all customers who indicate an interest. Although a member or candidate need not communicate a recommendation to all customers, the selection process by which customers receive information should be based on suitability and known interest, not on any preferred or favored status. A common practice to assure fair dealing is to communicate recommendations simultaneously within the firm and to customers.

Members and candidates should consider the following points when establishing fair-dealing compliance procedures:

- *Limit the number of people involved*: Members and candidates should make reasonable efforts to limit the number of people who are privy to the fact that a recommendation is going to be disseminated.
- *Shorten the time frame between decision and dissemination*: Members and candidates should make reasonable efforts to limit the amount of time that elapses between the decision to make an investment recommendation and the time the actual recommendation is disseminated. If a detailed institutional recommendation that might take two or three weeks to publish is in preparation, a short summary report including the conclusion might be published in advance. In an organization where both a research committee and an investment policy committee must approve a recommendation, the meetings should be held on the same day if possible. The process of reviewing reports and printing and mailing them, faxing them, or distributing them by e-mail necessarily involves the passage of time, sometimes long periods of time. In large firms with extensive review processes, the time factor is usually not within the control of the analyst who prepares the report. Thus, many firms and their analysts communicate to customers and firm personnel the new or changed recommendations by an update or "flash" report. The communication technique might be fax, e-mail, wire, or short written report.
- *Publish guidelines for pre-dissemination behavior*: Members and candidates should encourage firms to develop guidelines that prohibit personnel who have prior knowledge of an investment recommendation from discussing or taking any action on the pending recommendation.

- *Simultaneous dissemination*: Members and candidates should establish procedures for the timing of dissemination of investment recommendations so that all clients are treated fairly—that is, are informed at approximately the same time. For example, if a firm is going to announce a new recommendation, supervisory personnel should time the announcement to avoid placing any client or group of clients at an unfair advantage relative to other clients. A communication to all branch offices should be sent at the time of the general announcement. (When appropriate, the firm should accompany the announcement of a new recommendation with a statement that trading restrictions for the firm's employees are now in effect. The trading restrictions should stay in effect until the recommendation is widely distributed to all relevant clients.) Once this distribution has occurred, the member or candidate may follow up separately with individual clients, but members and candidates should not give favored clients advance information when such advance notification may disadvantage other clients.
- *Maintain a list of clients and their holdings*: Members and candidates should maintain a list of all clients and the securities or other investments each client holds in order to facilitate notification of customers or clients of a change in an investment recommendation. If a particular security or other investment is to be sold, such a list can be used to ensure that all holders are treated fairly in the liquidation of that particular investment.
- *Develop and document trade allocation procedures*: When formulating procedures for allocating trades, members and candidates should develop a set of guiding principles that ensure
    - fairness to advisory clients, both in priority of execution of orders and in the allocation of the price obtained in execution of block orders or trades,
    - timeliness and efficiency in the execution of orders, and
    - accuracy of the member's or candidate's records as to trade orders and client account positions.

With these principles in mind, members and candidates should develop or encourage their firm to develop written allocation procedures, with particular attention to procedures for block trades and new issues. Procedures to consider are as follows:

- requiring orders and modifications or cancellations of orders to be documented and time stamped;
- processing and executing orders on a first-in, first-out basis with consideration of bundling orders for efficiency as appropriate for the asset class or the security;
- developing a policy to address such issues as calculating execution prices and "partial fills" when trades are grouped, or in a block, for efficiency;
- giving all client accounts participating in a block trade the same execution price and charging the same commission;
- when the full amount of the block order is not executed, allocating partially executed orders among the participating client accounts pro rata on the basis of order size while not going below an established minimum lot size for some securities (e.g., bonds); and
- when allocating trades for new issues, obtaining advance indications of interest, allocating securities by client (rather than portfolio manager), and providing a method for calculating allocations.

Disclose Trade Allocation Procedures

Members and candidates should disclose to clients and prospective clients how they select accounts to participate in an order and how they determine the amount of securities each account will buy or sell. Trade allocation procedures must be fair and equitable, and disclosure of inequitable allocation methods does not relieve the member or candidate of this obligation.

Establish Systematic Account Review

Member and candidate supervisors should review each account on a regular basis to ensure that no client or customer is being given preferential treatment and that the investment actions taken for each account are suitable for each account's objectives. Because investments should be based on individual needs and circumstances, an investment manager may have good reasons for placing a given security or other investment in one account while selling it from another account and should fully document the reasons behind both sides of the transaction. Members and candidates should encourage firms to establish review procedures, however, to detect whether trading in one account is being used to benefit a favored client.

Disclose Levels of Service

Members and candidates should disclose to all clients whether the organization offers different levels of service to clients for the same fee or different fees. Different levels of service should not be offered to clients selectively.

STANDARD III(B): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Selective Disclosure):

Bradley Ames, a well-known and respected analyst, follows the computer industry. In the course of his research, he finds that a small, relatively unknown company whose shares are traded over the counter has just signed significant contracts with some of the companies he follows. After a considerable amount of investigation, Ames decides to write a research report on the small company and recommend purchase of its shares. While the report is being reviewed by the company for factual accuracy, Ames schedules a luncheon with several of his best clients to discuss the company. At the luncheon, he mentions the purchase recommendation scheduled to be sent early the following week to all the firm's clients.

> *Comment*: Ames has violated Standard III(B) by disseminating the purchase recommendation to the clients with whom he has lunch a week before the recommendation is sent to all clients.

Example 2 (Fair Dealing between Funds):

Spencer Rivers, president of XYZ Corporation, moves his company's growth-oriented pension fund to a particular bank primarily because of the excellent investment performance achieved by the bank's commingled fund for the prior five-year period. Later, Rivers compares the results of his pension fund with those of the bank's commingled fund. He is startled to learn that, even though the two accounts have the same investment objectives and similar portfolios, his company's pension fund has significantly underperformed the bank's commingled fund. Questioning this result at his next meeting with the pension fund's manager, Rivers is told that, as a matter of policy, when a new security is placed on the recommended list, Morgan Jackson, the pension fund manager, first purchases the security for the commingled account and then purchases it on a pro rata basis for all other pension fund accounts. Similarly, when a sale is recommended, the security is sold first from the commingled account and then sold on a pro rata basis from all other accounts. Rivers also learns that if the bank cannot get enough shares (especially of hot issues) to be meaningful to all the accounts, its policy is to place the new issues only in the commingled account.

Seeing that Rivers is neither satisfied nor pleased by the explanation, Jackson quickly adds that nondiscretionary pension accounts and personal trust accounts have a lower priority on purchase and sale recommendations than discretionary pension fund accounts. Furthermore, Jackson states, the company's pension fund had the opportunity to invest up to 5% in the commingled fund.

> *Comment*: The bank's policy does not treat all customers fairly, and Jackson has violated her duty to her clients by giving priority to the growth-oriented commingled fund over all other funds and to discretionary accounts over nondiscretionary accounts. Jackson must execute orders on a systematic basis that is fair to all clients. In addition, trade allocation procedures should be disclosed to all clients when they become clients. Of course, in this case, disclosure of the bank's policy would not change the fact that the policy is unfair.

Example 3 (Fair Dealing and IPO Distribution):

Dominic Morris works for a small regional securities firm. His work consists of corporate finance activities and investing for institutional clients. Arena, Ltd., is planning to go public. The partners have secured rights to buy an arena football league franchise and are planning to use the funds from the issue to complete the purchase. Because arena football is the current rage, Morris believes he has a hot issue on his hands. He has quietly negotiated some options for himself for helping convince Arena to do the financing through his securities firm. When he seeks expressions of interest, the institutional buyers oversubscribe the issue. Morris, assuming that the institutions have the financial clout to drive the stock up, then fills all orders (including his own) and decreases the institutional blocks.

> *Comment*: Morris has violated Standard III(B) by not treating all customers fairly. He should not have taken any shares himself and should have prorated the shares offered among all clients. In addition, he should have disclosed to his firm and to his clients that he received options as part of the deal [see Standard VI(A)–Disclosure of Conflicts].

Example 4 (Fair Dealing and Transaction Allocation):

Eleanor Preston, the chief investment officer of Porter Williams Investments (PWI), a medium-size money management firm, has been trying to retain a client, Colby Company. Management at Colby, which accounts for almost half of PWI's revenues, recently told Preston that if the performance of its account did not improve, it would find a new money manager. Shortly after this threat, Preston purchases mortgage-backed securities (MBSs) for several accounts, including Colby's. Preston is busy with a number of transactions that day, so she fails to allocate the trades immediately or write up the trade tickets. A few days later, when Preston is allocating trades, she notes that some of the MBSs have significantly increased in price and some have dropped. Preston decides to allocate the profitable trades to Colby and spread the losing trades among several other PWI accounts.

> *Comment*: Preston has violated Standard III(B) by failing to deal fairly with her clients in taking these investment actions. Preston should have allocated the trades prior to executing the orders, or she should have had a systematic approach to allocating the trades, such as pro rata, as soon as practical after they were executed. Among other things, Preston must disclose to the client that the adviser may act as broker for, receive commissions from, and have a potential conflict of interest regarding both parties in agency cross-transactions. After the disclosure, she should obtain from the client consent authorizing such transactions in advance.

Example 5 (Selective Disclosure):

Saunders Industrial Waste Management (SIWM) publicly indicates to analysts that it is comfortable with the somewhat disappointing earnings-per-share projection of US$1.16 for the quarter. Bernard Roberts, an analyst at Coffey Investments, is confident that SIWM management has understated the forecasted earnings so that the real announcement will cause an "upside surprise" and boost the price of SIWM stock. The "whisper number" (rumored) estimate based on extensive research and discussed among knowledgeable analysts is higher than US$1.16. Roberts repeats the US$1.16 figure in his research report to all Coffey clients but informally tells his large clients that he expects the earnings per share to be higher, making SIWM a good buy.

> *Comment*: By not sharing his opinion regarding the potential for a significant upside earnings surprise with all clients, Roberts is not treating all clients fairly and has violated Standard III(B).

Example 6 (Additional Services for Select Clients):

Jenpin Weng uses e-mail to issue a new recommendation to all his clients. He then calls his three largest institutional clients to discuss the recommendation in detail.

> *Comment*: Weng has not violated Standard III(B) because he widely disseminated the recommendation and provided the information to all his clients prior to discussing it with a select few. Weng's largest clients received additional personal service because they presumably pay higher fees or because they have a large amount of assets under Weng's management. If Weng had discussed the report with a select group of clients prior to distributing it to all his clients, he would have violated Standard III(B).

Example 7 (Minimum Lot Allocations):

Lynn Hampton is a well-respected private wealth manager in her community with a diversified client base. She determines that a new 10-year bond being offered by Healthy Pharmaceuticals is appropriate for five of her clients. Three clients request to purchase US$10,000 each, and the other two request US$50,000 each. The minimum lot size is established at US$5,000, and the issue is oversubscribed at the time of placement. Her firm's policy is that odd-lot allocations, especially those below the minimum, should be avoided because they may affect the liquidity of the security at the time of sale.

Hampton is informed she will receive only US$55,000 of the offering for all accounts. Hampton distributes the bond investments as follows: The three accounts that requested US$10,000 are allocated US$5,000 each, and the two accounts that requested US$50,000 are allocated US$20,000 each.

> *Comment*: Hampton has not violated Standard III(B), even though the distribution is not on a completely pro rata basis because of the required minimum lot size. With the total allocation being significantly below the amount requested, Hampton ensured that each client received at least the minimum lot size of the issue. This approach allowed the clients to efficiently sell the bond later if necessary.

Example 8 (Excessive Trading):

Ling Chan manages the accounts for many pension plans, including the plan of his father's employer. Chan developed similar but not identical investment policies for each client, so the investment portfolios are rarely the same. To minimize the cost to his father's pension plan, he intentionally trades more frequently in the accounts of other clients to ensure the required brokerage is incurred to continue receiving free research for use by all the pensions.

> *Comment*: Chan is violating Standard III(B) because his trading actions are disadvantaging his clients to enhance a relationship with a preferred client. All clients are benefiting from the research being provided and should incur their fair portion of the costs. This does not mean that additional trading should occur if a client has not paid an equal portion of the commission; trading should occur only as required by the strategy.

Example 9 (Limited Social Media Disclosures):

Mary Burdette was recently hired by Fundamental Investment Management (FIM) as a junior auto industry analyst. Burdette is expected to expand the social media presence of the firm because she is active with various networks, including Facebook, LinkedIn, and Twitter. Although Burdette's supervisor, Joe Graf, has never used social media, he encourages Burdette to explore opportunities to increase FIM's online presence and ability to share content, communicate, and broadcast information to clients. In response to Graf's encouragement, Burdette is working on a proposal detailing the advantages of getting FIM onto Twitter in addition to launching a company Facebook page.

As part of her auto industry research for FIM, Burdette is completing a report on the financial impact of Sun Drive Auto Ltd.'s new solar technology for compact automobiles. This research report will be her first for FIM, and she believes Sun Drive's technology could revolutionize the auto industry. In her excitement, Burdette sends a quick tweet to FIM Twitter followers summarizing her "buy" recommendation for Sun Drive Auto stock.

> *Comment*: Burdette has violated Standard III(B) by sending an investment recommendation to a select group of contacts prior to distributing it to all clients. Burdette must make sure she has received the appropriate training about FIM's policies and procedures, including the appropriate business use of personal social media networks before engaging in such activities.
>
> See Standard IV(C) for guidance related to the duties of the supervisor.

Example 10 (Fair Dealing between Clients):

Paul Rove, performance analyst for Alpha-Beta Investment Management, is describing to the firm's chief investment officer (CIO) two new reports he would like to develop to assist the firm in meeting its obligations to treat clients fairly. Because many of the firm's clients have similar investment objectives and portfolios, Rove suggests a report detailing securities owned across several clients and the percentage of the portfolio the security represents. The second report would compare the monthly performance of portfolios with similar strategies. The outliers within each report would be submitted to the CIO for review.

> *Comment*: As a performance analyst, Rove likely has little direct contact with clients and thus has limited opportunity to treat clients differently. The recommended reports comply with Standard III(B) while helping the firm conduct after-the-fact reviews of how effectively the firm's advisers are dealing with their clients' portfolios. Reports that monitor the fair treatment of clients are an important oversight tool to ensure that clients are treated fairly.

STANDARD III(C): DUTIES TO CLIENTS – SUITABILITY

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

1 When Members and Candidates are in an advisory relationship with a client, they must:

- **a** Make a reasonable inquiry into a client's or prospective client's investment experience, risk and return objectives, and financial constraints prior to making any investment recommendation or taking investment action and must reassess and update this information regularly.
- **b** Determine that an investment is suitable to the client's financial situation and consistent with the client's written objectives, mandates, and constraints before making an investment recommendation or taking investment action.
- **c** Judge the suitability of investments in the context of the client's total portfolio.

2 When Members and Candidates are responsible for managing a portfolio to a specific mandate, strategy, or style, they must make only investment recommendations or take only investment actions that are consistent with the stated objectives and constraints of the portfolio.

Guidance

Highlights:

- *Developing an Investment Policy*
- *Understanding the Client's Risk Profile*
- *Updating an Investment Policy*
- *The Need for Diversification*
- *Addressing Unsolicited Trading Requests*
- *Managing to an Index or Mandate*

Standard III(C) requires that members and candidates who are in an investment advisory relationship with clients consider carefully the needs, circumstances, and objectives of the clients when determining the appropriateness and suitability of a given investment or course of investment action. An appropriate suitability determination will not, however, prevent some investments or investment actions from losing value.

In judging the suitability of a potential investment, the member or candidate should review many aspects of the client's knowledge, experience related to investing, and financial situation. These aspects include, but are not limited to, the risk profile of the investment as compared with the constraints of the client, the impact of the investment on the diversity of the portfolio, and whether the client has the means or net worth to assume the associated risk. The investment professional's determination of suitability should reflect only the investment recommendations or actions that a prudent person would be willing to undertake. Not every investment opportunity will be suitable for every portfolio, regardless of the potential return being offered.

The responsibilities of members and candidates to gather information and make a suitability analysis prior to making a recommendation or taking investment action fall on those members and candidates who provide investment advice in the course of an advisory relationship with a client. Other members and candidates may be simply executing specific instructions for retail clients when buying or selling securities, such as shares in mutual funds. These members and candidates and some others, such as sell-side analysts, may not have the opportunity to judge the suitability of a particular investment for the ultimate client.

Developing an Investment Policy

When an advisory relationship exists, members and candidates must gather client information at the inception of the relationship. Such information includes the client's financial circumstances, personal data (such as age and occupation) that are relevant to investment decisions, attitudes toward risk, and objectives in investing. This information should be incorporated into a written investment policy statement (IPS) that addresses the client's risk tolerance, return requirements, and all investment constraints (including time horizon, liquidity needs, tax concerns, legal and regulatory factors, and unique circumstances). Without identifying such client factors, members and candidates cannot judge whether a particular investment or strategy is suitable for a particular client. The IPS also should identify and describe the roles and responsibilities of the parties to the advisory relationship and investment process, as well as schedules for review and evaluation of the IPS. After formulating long-term capital market expectations, members and candidates can assist in developing an appropriate strategic asset allocation and investment program for the client, whether these are presented in separate documents or incorporated in the IPS or in appendices to the IPS.

Understanding the Client's Risk Profile

One of the most important factors to be considered in matching appropriateness and suitability of an investment with a client's needs and circumstances is measuring that client's tolerance for risk. The investment professional must consider the possibilities of rapidly changing investment environments and their likely impact on a client's holdings, both individual securities and the collective portfolio. The risk of many investment strategies can and should be analyzed and quantified in advance.

The use of synthetic investment vehicles and derivative investment products has introduced particular issues of risk. Members and candidates should pay careful attention to the leverage inherent in many of these vehicles or products when considering them for use in a client's investment program. Such leverage and limited liquidity, depending on the degree to which they are hedged, bear directly on the issue of suitability for the client.

Updating an Investment Policy

Updating the IPS should be repeated at least annually and also prior to material changes to any specific investment recommendations or decisions on behalf of the client. The effort to determine the needs and circumstances of each client is not a one-time occurrence. Investment recommendations or decisions are usually part of an ongoing process that takes into account the diversity and changing nature of portfolio and client characteristics. The passage of time is bound to produce changes that are important with respect to investment objectives.

For an individual client, important changes might include the number of dependents, personal tax status, health, liquidity needs, risk tolerance, amount of wealth beyond that represented in the portfolio, and extent to which compensation and other income provide for current income needs. With respect to an institutional client, such changes might relate to the magnitude of unfunded liabilities in a pension fund, the withdrawal privileges in an employee savings plan, or the distribution requirements of a charitable foundation. Without efforts to update information concerning client factors, one or more factors could change without the investment manager's knowledge.

Suitability review can be done most effectively when the client fully discloses his or her complete financial portfolio, including those portions not managed by the member or candidate. If clients withhold information about their financial portfolios, the suitability analysis conducted by members and candidates cannot be expected to be complete; it must be based on the information provided.

The Need for Diversification

The investment profession has long recognized that combining several different investments is likely to provide a more acceptable level of risk exposure than having all assets in a single investment. The unique characteristics (or risks) of an individual investment may become partially or entirely neutralized when it is combined with other individual investments within a portfolio. Some reasonable amount of diversification is thus the norm for many portfolios, especially those managed by individuals or institutions that have some degree of legal fiduciary responsibility.

An investment with high relative risk on its own may be a suitable investment in the context of the entire portfolio or when the client's stated objectives contemplate speculative or risky investments. The manager may be responsible for only a portion of the client's total portfolio, or the client may not have provided a full financial picture. Members and candidates can be responsible for assessing the suitability of an investment only on the basis of the information and criteria actually provided by the client.

Addressing Unsolicited Trading Requests

Members and candidates may receive requests from a client for trades that do not properly align with the risk and return objectives outlined in the client's investment policy statement. These transaction requests may be based on the client's individual biases or professional experience. Members and candidates will need to make reasonable efforts to balance their clients' trading requests with their responsibilities to follow the agreed-on investment policy statement.

In cases of unsolicited trade requests that a member or candidate knows are unsuitable for a client, the member or candidate should refrain from making the trade until he or she discusses the concerns with the client. The discussions and resulting actions may encompass a variety of scenarios depending on how the requested unsuitable investment relates to the client's full portfolio.

Many times, an unsolicited request may be expected to have only a minimum impact on the entire portfolio because the size of the requested trade is small or the trade would result in a limited change to the portfolio's risk profile. In discussing the trade, the member or candidate should focus on educating the investor on how the request deviates from the current policy statement. Following the discussion, the member or candidate may follow his or her firm's policies regarding the necessary client approval for executing unsuitable trades. At a minimum, the client should acknowledge the discussion and accept the conditions that make the recommendation unsuitable.

Should the unsolicited request be expected to have a material impact on the portfolio, the member or candidate should use this opportunity to update the investment policy statement. Doing so would allow the client to fully understand the potential effect of the requested trade on his or her current goals or risk levels.

Members and candidates may have some clients who decline to modify their policy statements while insisting an unsolicited trade be made. In such instances, members or candidates will need to evaluate the effectiveness of their services to the client. The options available to the members or candidates will depend on the services provided by their employer. Some firms may allow for the trade to be executed in a new unmanaged account. If alternative options are not available, members and candidates ultimately will need to determine whether they should continue the advisory arrangement with the client.

Managing to an Index or Mandate

Some members and candidates do not manage money for individuals but are responsible for managing a fund to an index or an expected mandate. The responsibility of these members and candidates is to invest in a manner consistent with the stated mandate. For example, a member or candidate who serves as the fund manager for a large-cap income fund would not be following the fund mandate by investing heavily in small-cap or start-up companies whose stock is speculative in nature. Members and candidates who manage pooled assets to a specific mandate are not responsible for determining the suitability of the *fund* as an investment for investors who may be purchasing shares in the fund. The responsibility for determining the suitability of an investment for clients can be conferred only on members and candidates who have an advisory relationship with clients.

STANDARD III(C): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Investment Policy Statement

To fulfill the basic provisions of Standard III(C), a member or candidate should put the needs and circumstances of each client and the client's investment objectives into a written investment policy statement. In formulating an investment policy for the client, the member or candidate should take the following into consideration:

- client identification—(1) type and nature of client, (2) the existence of separate beneficiaries, and (3) approximate portion of total client assets that the member or candidate is managing;
- investor objectives—(1) return objectives (income, growth in principal, maintenance of purchasing power) and (2) risk tolerance (suitability, stability of values);
- investor constraints—(1) liquidity needs, (2) expected cash flows (patterns of additions and/or withdrawals), (3) investable funds (assets and liabilities or other commitments), (4) time horizon, (5) tax considerations, (6) regulatory and legal circumstances, (7) investor preferences, prohibitions, circumstances, and unique needs, and (8) proxy voting responsibilities and guidance; and
- performance measurement benchmarks.

Regular Updates

The investor's objectives and constraints should be maintained and reviewed periodically to reflect any changes in the client's circumstances. Members and candidates should regularly compare client constraints with capital market expectations to arrive at an appropriate asset allocation. Changes in either factor may result in a fundamental change in asset allocation. Annual review is reasonable unless business or other reasons, such as a major change in market conditions, dictate more frequent review. Members and candidates should document attempts to carry out such a review if circumstances prevent it.

Suitability Test Policies

With the increase in regulatory required suitability tests, members and candidates should encourage their firms to develop related policies and procedures. The procedures will differ according to the size of the firm and the scope of the services offered to its clients.

The test procedures should require the investment professional to look beyond the potential return of the investment and include the following:

- an analysis of the impact on the portfolio's diversification,
- a comparison of the investment risks with the client's assessed risk tolerance, and
- the fit of the investment with the required investment strategy.

STANDARD III(C): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Investment Suitability—Risk Profile):

Caleb Smith, an investment adviser, has two clients: Larry Robertson, 60 years old, and Gabriel Lanai, 40 years old. Both clients earn roughly the same salary, but Robertson has a much higher risk tolerance because he has a large asset base. Robertson is willing to invest part of his assets very aggressively; Lanai wants only to achieve a steady rate of return with low volatility to pay for his children's education. Smith recommends investing 20% of both portfolios in zero-yield, small-cap, high-technology equity issues.

> *Comment*: In Robertson's case, the investment may be appropriate because of his financial circumstances and aggressive investment position, but this investment is not suitable for Lanai. Smith is violating Standard III(C) by applying Robertson's investment strategy to Lanai because the two clients' financial circumstances and objectives differ.

Example 2 (Investment Suitability—Entire Portfolio):

Jessica McDowell, an investment adviser, suggests to Brian Crosby, a risk-averse client, that covered call options be used in his equity portfolio. The purpose would be to enhance Crosby's income and partially offset any untimely depreciation in the portfolio's value should the stock market or other circumstances affect his holdings unfavorably. McDowell educates Crosby about all possible outcomes, including the risk of incurring an added tax liability if a stock rises in price and is called away and, conversely, the risk of his holdings losing protection on the downside if prices drop sharply.

> *Comment*: When determining suitability of an investment, the primary focus should be the characteristics of the client's entire portfolio, not the characteristics of single securities on an issue-by-issue basis. The basic characteristics of the entire portfolio will largely determine whether investment recommendations are taking client factors into account. Therefore, the most important aspects of a particular investment are those that will affect the characteristics of the total portfolio. In this case, McDowell properly considers the investment in the context of the entire portfolio and thoroughly explains the investment to the client.

Example 3 (IPS Updating):

In a regular meeting with client Seth Jones, the portfolio managers at Blue Chip Investment Advisors are careful to allow some time to review his current needs and circumstances. In doing so, they learn that some significant changes have recently taken place in his life. A wealthy uncle left Jones an inheritance that increased his net worth fourfold, to US$1 million.

> *Comment*: The inheritance has significantly increased Jones's ability (and possibly his willingness) to assume risk and has diminished the average yield required to meet his current income needs. Jones's financial circumstances have definitely changed, so Blue Chip managers must update Jones's investment policy statement to reflect how his investment objectives have changed. Accordingly, the Blue Chip portfolio managers should consider a somewhat higher equity ratio for his portfolio than was called for by the previous circumstances, and the managers' specific common stock recommendations might be heavily tilted toward low-yield, growth-oriented issues.

Example 4 (Following an Investment Mandate):

Louis Perkowski manages a high-income mutual fund. He purchases zero-dividend stock in a financial services company because he believes the stock is undervalued and is in a potential growth industry, which makes it an attractive investment.

> *Comment*: A zero-dividend stock does not seem to fit the mandate of the fund that Perkowski is managing. Unless Perkowski's investment fits within the mandate or is within the realm of allowable investments the fund has made clear in its disclosures, Perkowski has violated Standard III(C).

Example 5 (IPS Requirements and Limitations):

Max Gubler, chief investment officer of a property/casualty insurance subsidiary of a large financial conglomerate, wants to improve the diversification of the subsidiary's investment portfolio and increase its returns. The subsidiary's investment policy statement provides for highly liquid investments, such as large-cap equities and government, supranational, and corporate bonds with a minimum credit rating of AA and maturity of no more than five years. In a recent presentation, a venture capital group offered very attractive prospective returns on some of its private equity funds that provide seed capital to ventures. An exit strategy was already contemplated, but investors would have to observe a minimum three-year lockup period and a subsequent laddered exit option for a maximum of one-third of their shares per year. Gubler does not want to miss this opportunity. After extensive analysis, with the intent to optimize the return on the equity assets within the subsidiary's current portfolio, he invests 4% in this seed fund, leaving the portfolio's total equity exposure still well below its upper limit.

> *Comment*: Gubler is violating Standard III(A)–Loyalty, Prudence, and Care as well as Standard III(C). His new investment locks up part of the subsidiary's assets for at least three years and up to as many as five years and possibly beyond. The IPS requires investments in highly liquid investments and describes accepted asset classes; private equity investments with a lockup period certainly do not qualify. Even without a lockup period, an asset class with only an occasional, and thus implicitly illiquid, market may not be suitable for the portfolio. Although an IPS typically describes objectives and constraints in great detail, the manager must also make every effort to understand the client's business and circumstances. Doing so should enable the manager to recognize, understand, and discuss with the client other factors that may be or may become material in the investment management process.

Example 6 (Submanager and IPS Reviews):

Paul Ostrowski's investment management business has grown significantly over the past couple of years, and some clients want to diversify internationally. Ostrowski decides to find a submanager to handle the expected international investments. Because this will be his first subadviser, Ostrowski uses the CFA Institute model "request for proposal" to design a questionnaire for his search. By his deadline, he receives seven completed questionnaires from a variety of domestic and international firms trying to gain his business. Ostrowski reviews all the applications in detail and decides to select the firm that charges the lowest fees because doing so will have the least impact on his firm's bottom line.

> *Comment*: When selecting an external manager or subadviser, Ostrowski needs to ensure that the new manager's services are appropriate for his clients. This due diligence includes comparing the risk profile of the clients with the investment strategy of the manager. In basing the decision on the fee structure alone, Ostrowski may be violating Standard III(C).
>
> When clients ask to diversify into international products, it is an appropriate time to review and update the clients' IPSs. Ostrowski's review may determine that the risk of international investments modifies the risk profiles of the clients or does not represent an appropriate investment.
>
> See also Standard V(A)–Diligence and Reasonable Basis for further discussion of the review process needed in selecting appropriate submanagers.

Example 7 (Investment Suitability—Risk Profile):

Samantha Snead, a portfolio manager for Thomas Investment Counsel, Inc., specializes in managing public retirement funds and defined benefit pension plan accounts, all of which have long-term investment objectives. A year ago, Snead's employer, in an attempt to motivate and retain key investment professionals, introduced a bonus compensation system that rewards portfolio managers on the basis of quarterly performance relative to their peers and to certain benchmark indexes. In an attempt to improve the short-term performance of her accounts, Snead changes her investment strategy and purchases several high-beta stocks for client portfolios. These purchases are seemingly contrary to the clients' investment policy statements. Following their purchase, an officer of Griffin Corporation, one of Snead's pension fund clients, asks why Griffin Corporation's portfolio seems to be dominated by high-beta stocks of companies that often appear among the most actively traded issues. No change in objective or strategy has been recommended by Snead during the year.

> *Comment*: Snead violated Standard III(C) by investing the clients' assets in high-beta stocks. These high-risk investments are contrary to the long-term risk profile established in the clients' IPSs. Snead has changed the investment strategy of the clients in an attempt to reap short-term rewards offered by her firm's new compensation arrangement, not in response to changes in clients' investment policy statements.
>
> See also Standard VI(A)–Disclosure of Conflicts.

Example 8 (Investment Suitability):

Andre Shrub owns and operates Conduit, an investment advisory firm. Prior to opening Conduit, Shrub was an account manager with Elite Investment, a hedge fund managed by his good friend Adam Reed. To attract clients to a new Conduit fund, Shrub offers lower-than-normal management fees. He can do so because the fund consists of two top-performing funds managed by Reed. Given his personal friendship with Reed and the prior performance record of these two funds, Shrub believes this new fund is a winning combination for all parties. Clients quickly invest with Conduit to gain access to the Elite funds. No one is turned away because Conduit is seeking to expand its assets under management.

> *Comment*: Shrub has violated Standard III(C) because the risk profile of the new fund may not be suitable for every client. As an investment adviser, Shrub needs to establish an investment policy statement for each client and recommend only investments that match each client's risk and return profile in the IPS. Shrub is required to act as more than a simple sales agent for Elite.
>
> Although Shrub cannot disobey the direct request of a client to purchase a specific security, he should fully discuss the risks of a planned purchase and provide reasons why it might not be suitable for a client. This requirement may lead members and candidates to decline new customers if those customers' requested investment decisions are significantly out of line with their stated requirements.
>
> See also Standard V(A)–Diligence and Reasonable Basis.

STANDARD III(D): DUTIES TO CLIENTS - PERFORMANCE PRESENTATION

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

When communicating investment performance information, Members and Candidates must make reasonable efforts to ensure that it is fair, accurate, and complete.

Guidance

Standard III(D) requires members and candidates to provide credible performance information to clients and prospective clients and to avoid misstating performance or misleading clients and prospective clients about the investment performance of members or candidates or their firms. This standard encourages full disclosure of investment performance data to clients and prospective clients.

Standard III(D) covers any practice that would lead to misrepresentation of a member's or candidate's performance record, whether the practice involves performance presentation or performance measurement. This standard prohibits misrepresentations of past performance or reasonably expected performance. A member or candidate must give a fair and complete presentation of performance information whenever communicating data with respect to the performance history of individual accounts, composites or groups of accounts, or composites of an analyst's or firm's performance results. Furthermore, members and candidates should not state or imply that clients will obtain or benefit from a rate of return that was generated in the past.

The requirements of this standard are not limited to members and candidates managing separate accounts. Whenever a member or candidate provides performance information for which the manager is claiming responsibility, such as for pooled funds, the history must be accurate. Research analysts promoting the success or accuracy of their recommendations must ensure that their claims are fair, accurate, and complete.

If the presentation is brief, the member or candidate must make available to clients and prospects, on request, the detailed information supporting that communication. Best practice dictates that brief presentations include a reference to the limited nature of the information provided.

STANDARD III(D): RECOMMENDED PROCEDURES

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Apply the GIPS Standards

For members and candidates who are showing the performance history of the assets they manage, compliance with the GIPS standards is the best method to meet their obligations under Standard III(D). Members and candidates should encourage their firms to comply with the GIPS standards.

Compliance without Applying GIPS Standards

Members and candidates can also meet their obligations under Standard III(D) by

- considering the knowledge and sophistication of the audience to whom a performance presentation is addressed,
- presenting the performance of the weighted composite of similar portfolios rather than using a single representative account,
- including terminated accounts as part of performance history with a clear indication of when the accounts were terminated,
- including disclosures that fully explain the performance results being reported (for example, stating, when appropriate, that results are simulated when model results are used, clearly indicating when the performance record is that of a prior entity, or disclosing whether the performance is gross of fees, net of fees, or after tax), and
- maintaining the data and records used to calculate the performance being presented.

STANDARD III(D): APPLICATION OF THE STANDARD

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Example 1 (Performance Calculation and Length of Time):

Kyle Taylor of Taylor Trust Company, noting the performance of Taylor's common trust fund for the past two years, states in a brochure sent to his potential clients, "You can expect steady 25% annual compound growth of the value of your investments over the year." Taylor Trust's common trust fund did increase at the rate of 25% per year for the past year, which mirrored the increase of the entire market. The fund has never averaged that growth for more than one year, however, and the average rate of growth of all of its trust accounts for five years is 5% per year.

> *Comment*: Taylor's brochure is in violation of Standard III(D). Taylor should have disclosed that the 25% growth occurred only in one year. Additionally, Taylor did not include client accounts other than those in the firm's common trust fund. A general claim of firm performance should take into account the performance of all categories of accounts. Finally, by stating that clients can expect a steady 25% annual compound growth rate, Taylor is also violating Standard I(C)–Misrepresentation, which prohibits assurances or guarantees regarding an investment.

Example 2 (Performance Calculation and Asset Weighting):

Anna Judd, a senior partner of Alexander Capital Management, circulates a performance report for the capital appreciation accounts for the years 1988 through 2004. The firm claims compliance with the GIPS standards. Returns are not calculated in accordance with the requirements of the GIPS standards, however, because the composites are not asset weighted.

> *Comment*: Judd is in violation of Standard III(D). When claiming compliance with the GIPS standards, firms must meet *all* of the requirements, make mandatory disclosures, and meet any other requirements that apply to that firm's specific situation. Judd's violation is not from any misuse of the data but from a false claim of GIPS compliance.

Example 3 (Performance Presentation and Prior Fund/Employer):

Aaron McCoy is vice president and managing partner of the equity investment group of Mastermind Financial Advisors, a new business. Mastermind recruited McCoy because he had a proven six-year track record with G&P Financial. In developing Mastermind's advertising and marketing campaign, McCoy prepares an advertisement that includes the equity investment performance he achieved at G&P Financial. The advertisement for Mastermind does not identify the equity performance as being earned while at G&P. The advertisement is distributed to existing clients and prospective clients of Mastermind.

> *Comment*: McCoy has violated Standard III(D) by distributing an advertisement that contains material misrepresentations about the historical performance of Mastermind. Standard III(D) requires that members and candidates make every reasonable effort to ensure that performance information is a fair, accurate, and complete representation of an individual's or firm's performance. As a general matter, this standard does not prohibit showing past performance of funds managed at a prior firm as part of a performance track record as long as showing that record is accompanied by appropriate disclosures about where the performance took place and

> the person's specific role in achieving that performance. If McCoy chooses to use his past performance from G&P in Mastermind's advertising, he should make full disclosure of the source of the historical performance.

Example 4 (Performance Presentation and Simulated Results):

Jed Davis has developed a mutual fund selection product based on historical information from the 1990–95 period. Davis tested his methodology by applying it retroactively to data from the 1996–2003 period, thus producing simulated performance results for those years. In January 2004, Davis's employer decided to offer the product and Davis began promoting it through trade journal advertisements and direct dissemination to clients. The advertisements included the performance results for the 1996–2003 period but did not indicate that the results were simulated.

> *Comment*: Davis violated Standard III(D) by failing to clearly identify simulated performance results. Standard III(D) prohibits members and candidates from making any statements that misrepresent the performance achieved by them or their firms and requires members and candidates to make every reasonable effort to ensure that performance information presented to clients is fair, accurate, and complete. Use of simulated results should be accompanied by full disclosure as to the source of the performance data, including the fact that the results from 1995 through 2003 were the result of applying the model retroactively to that time period.

Example 5 (Performance Calculation and Selected Accounts Only):

In a presentation prepared for prospective clients, William Kilmer shows the rates of return realized over a five-year period by a "composite" of his firm's discretionary accounts that have a "balanced" objective. This composite, however, consisted of only a few of the accounts that met the balanced criterion set by the firm, excluded accounts under a certain asset level without disclosing the fact of their exclusion, and included accounts that did not have the balanced mandate because those accounts would boost the investment results. In addition, to achieve better results, Kilmer manipulated the narrow range of accounts included in the composite by changing the accounts that made up the composite over time.

> *Comment*: Kilmer violated Standard III(D) by misrepresenting the facts in the promotional material sent to prospective clients, distorting his firm's performance record, and failing to include disclosures that would have clarified the presentation.

Example 6 (Performance Attribution Changes):

Art Purell is reviewing the quarterly performance attribution reports for distribution to clients. Purell works for an investment management firm with a bottom-up, fundamentals-driven investment process that seeks to add value through stock selection. The attribution methodology currently compares each stock with its sector. The attribution report indicates that the value added this quarter came from asset allocation and that stock selection contributed negatively to the calculated return.

Through running several different scenarios, Purell discovers that calculating attribution by comparing each stock with its industry and then rolling the effect to the sector level improves the appearance of the manager's stock selection activities. Because the firm defines the attribution terms and the results better reflect the stated strategy, Purell recommends that the client reports should use the revised methodology.

> *Comment*: Modifying the attribution methodology without proper notifications to clients would fail to meet the requirements of Standard III(D). Purrell's recommendation is being done solely for the interest of the firm to improve its perceived ability to meet the stated investment strategy. Such changes are unfair to clients and obscure the facts regarding the firm's abilities.
>
> Had Purell believed the new methodology offered improvements to the original model, then he would have needed to report the results of both calculations to the client. The report should also include the reasons why the new methodology is preferred, which would allow the client to make a meaningful comparison to prior results and provide a basis for comparing future attributions.

Example 7 (Performance Calculation Methodology Disclosure):

While developing a new reporting package for existing clients, Alisha Singh, a performance analyst, discovers that her company's new system automatically calculates both time-weighted and money-weighted returns. She asks the head of client services and retention which value would be preferred given that the firm has various investment strategies that include bonds, equities, securities without leverage, and alternatives. Singh is told not to label the return value so that the firm may show whichever value is greatest for the period.

> *Comment*: Following these instructions would lead to Singh violating Standard III(D). In reporting inconsistent return values, Singh would not be providing complete information to the firm's clients. Full information is provided when clients have sufficient information to judge the performance generated by the firm.

Example 8 (Performance Calculation Methodology Disclosure):

Richmond Equity Investors manages a long–short equity fund in which clients can trade once a week (on Fridays). For transparency reasons, a daily net asset value of the fund is calculated by Richmond. The monthly fact sheets of the fund report month-to-date and year-to-date performance. Richmond publishes the performance based on the higher of the last trading day of the month (typically, not the last business day) or the last business day of the month as determined by Richmond. The fact sheet mentions only that the data are as of the end of the month, without giving the exact date. Maggie Clark, the investment performance analyst in charge of the calculations, is concerned about the frequent changes and asks her supervisor whether they are appropriate.

> *Comment*: Clark's actions in questioning the changing performance metric comply with Standard III(D). She has shown concern that these changes are not presenting an accurate and complete picture of the performance generated.

STANDARD III(E): DUTIES TO CLIENTS - PRESERVATION OF CONFIDENTIALITY

Members and Candidates must keep information about current, former, and prospective clients confidential unless:

1. The information concerns illegal activities on the part of the client;
2. Disclosure is required by law; or
3. The client or prospective client permits disclosure of the information.

Guidance

Highlights:

- *Status of Client*
- *Compliance with Laws*
- *Electronic Information and Security*
- *Professional Conduct Investigations by CFA Institute*

Standard III(E) requires that members and candidates preserve the confidentiality of information communicated to them by their clients, prospective clients, and former clients. This standard is applicable when (1) the member or candidate receives information because of his or her special ability to conduct a portion of the client's business or personal affairs and (2) the member or candidate receives information that arises from or is relevant to that portion of the client's business that is the subject of the special or confidential relationship. If disclosure of the information is required by law or the information concerns illegal activities by the client, however, the member or candidate may have an obligation to report the activities to the appropriate authorities.

Status of Client

This standard protects the confidentiality of client information even if the person or entity is no longer a client of the member or candidate. Therefore, members and candidates must continue to maintain the confidentiality of client records even after the client relationship has ended. If a client or former client expressly authorizes the member or candidate to disclose information, however, the member or candidate may follow the terms of the authorization and provide the information.

Compliance with Laws

As a general matter, members and candidates must comply with applicable law. If applicable law requires disclosure of client information in certain circumstances, members and candidates must comply with the law. Similarly, if applicable law requires members and candidates to maintain confidentiality, even if the information concerns illegal activities on the part of the client, members and candidates should not disclose such information. Additionally, applicable laws, such as inter-departmental communication restrictions within financial institutions, can impose limitations on information flow about a client within an entity that may lead to a violation of confidentiality. When in doubt, members and candidates should consult with their employer's compliance personnel or legal counsel before disclosing confidential information about clients.

Electronic Information and Security

Because of the ever-increasing volume of electronically stored information, members and candidates need to be particularly aware of possible accidental disclosures. Many employers have strict policies about how to electronically communicate sensitive client information and store client information on personal laptops, mobile devices, or portable disk/flash drives. In recent years, regulatory authorities have imposed stricter data security laws applying to the use of mobile remote digital communication, including the use of social media, that must be considered. Standard III(E) does not require members or candidates to become experts in information security technology, but they should have a thorough understanding of the policies of their employer. The size and operations of the firm will lead to differing policies for ensuring the security of confidential information maintained within the firm. Members and candidates should encourage their firm to conduct regular periodic training on confidentiality procedures for all firm personnel, including portfolio associates, receptionists, and other non-investment staff who have routine direct contact with clients and their records.

Professional Conduct Investigations by CFA Institute

The requirements of Standard III(E) are not intended to prevent members and candidates from cooperating with an investigation by the CFA Institute Professional Conduct Program (PCP). When permissible under applicable law, members and candidates shall consider the PCP an extension of themselves when requested to provide information about a client in support of a PCP investigation into their own conduct. Members and candidates are encouraged to cooperate with investigations into the conduct of others. Any information turned over to the PCP is kept in the strictest confidence. Members and candidates will not be considered in violation of this standard by forwarding confidential information to the PCP.

STANDARD III(E): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

The simplest, most conservative, and most effective way to comply with Standard III(E) is to avoid disclosing any information received from a client except to authorized fellow employees who are also working for the client. In some instances, however, a member or candidate may want to disclose information received from clients that is outside the scope of the confidential relationship and does not involve illegal activities. Before making such a disclosure, a member or candidate should ask the following:

- In what context was the information disclosed? If disclosed in a discussion of work being performed for the client, is the information relevant to the work?
- Is the information background material that, if disclosed, will enable the member or candidate to improve service to the client?

Members and candidates need to understand and follow their firm's electronic information communication and storage procedures. If the firm does not have procedures in place, members and candidates should encourage the development of procedures that appropriately reflect the firm's size and business operations.

Communicating with Clients

Technological changes are constantly enhancing the methods that are used to communicate with clients and prospective clients. Members and candidates should make reasonable efforts to ensure that firm-supported communication methods and compliance procedures follow practices designed for preventing accidental distribution of confidential information. Given the rate at which technology changes, a regular review of privacy protection measures is encouraged.

Members and candidates should be diligent in discussing with clients the appropriate methods for providing confidential information. It is important to convey to clients that not all firm-sponsored resources may be appropriate for such communications.

STANDARD III(E): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Possessing Confidential Information):

Sarah Connor, a financial analyst employed by Johnson Investment Counselors, Inc., provides investment advice to the trustees of City Medical Center. The trustees have given her a number of internal reports concerning City Medical's needs for physical plant renovation and expansion. They have asked Connor to recommend investments that would generate capital appreciation in endowment funds to meet projected capital expenditures. Connor is approached by a local businessman, Thomas Kasey, who is considering a substantial contribution either to City Medical Center or to another local hospital. Kasey wants to find out the building plans of both institutions before making a decision, but he does not want to speak to the trustees.

> *Comment*: The trustees gave Connor the internal reports so she could advise them on how to manage their endowment funds. Because the information in the reports is clearly both confidential and within the scope of the confidential relationship, Standard III(E) requires that Connor refuse to divulge information to Kasey.

Example 2 (Disclosing Confidential Information):

Lynn Moody is an investment officer at the Lester Trust Company. She has an advisory customer who has talked to her about giving approximately US$50,000 to charity to reduce her income taxes. Moody is also treasurer of the Home for Indigent Widows (HIW), which is planning its annual giving campaign. HIW hopes to expand its list of prospects, particularly those capable of substantial gifts. Moody recommends that HIW's vice president for corporate gifts call on her customer and ask for a donation in the US$50,000 range.

> *Comment*: Even though the attempt to help the Home for Indigent Widows was well intended, Moody violated Standard III(E) by revealing confidential information about her client.

Example 3 (Disclosing Possible Illegal Activity):

Government officials approach Casey Samuel, the portfolio manager for Garcia Company's pension plan, to examine pension fund records. They tell her that Garcia's corporate tax returns are being audited and the pension fund is being reviewed. Two days earlier, Samuel had learned in a regular investment review with Garcia officers that potentially excessive and improper charges were being made to the pension plan by Garcia. Samuel consults her employer's general counsel and is advised that Garcia has probably violated tax and fiduciary regulations and laws.

> *Comment*: Samuel should inform her supervisor of these activities, and her employer should take steps, with Garcia, to remedy the violations. If that approach is not successful, Samuel and her employer should seek advice of legal counsel to determine the appropriate steps to be taken. Samuel may well have a duty to disclose the evidence she has of the continuing legal violations and to resign as asset manager for Garcia.

Example 4 (Disclosing Possible Illegal Activity):

David Bradford manages money for a family-owned real estate development corporation. He also manages the individual portfolios of several of the family members and officers of the corporation, including the chief financial officer (CFO). Based on the financial records of the corporation and some questionable practices of the CFO that Bradford has observed, Bradford believes that the CFO is embezzling money from the corporation and putting it into his personal investment account.

> *Comment*: Bradford should check with his firm's compliance department or appropriate legal counsel to determine whether applicable securities regulations require reporting the CFO's financial records.

Example 5 (Accidental Disclosure of Confidential Information):

Lynn Moody is an investment officer at the Lester Trust Company (LTC). She has stewardship of a significant number of individually managed taxable accounts. In addition to receiving quarterly written reports, about a dozen high-net-worth individuals have indicated to Moody a willingness to receive communications about overall economic and financial market outlooks directly from her by way of a social media platform. Under the direction of her firm's technology and compliance departments, she established a new group page on an existing social media platform specifically for her clients. In the instructions provided to clients, Moody asked them to "join" the group so they may be granted access to the posted content. The instructions also advised clients that all comments posted would be available to the public and thus the platform was not an appropriate method for communicating personal or confidential information.

Six months later, in early January, Moody posted LTC's year-end "Market Outlook." The report outlined a new asset allocation strategy that the firm is adding to its recommendations in the new year. Moody introduced the publication with a note informing her clients that she would be discussing the changes with them individually in their upcoming meetings.

One of Moody's clients responded directly on the group page that his family recently experienced a major change in their financial profile. The client described highly personal and confidential details of the event. Unfortunately, all clients that were part of the group were also able to read the detailed posting until Moody was able to have the comment removed.

Comment: Moody has taken reasonable steps for protecting the confidentiality of client information while using the social media platform. She provided instructions clarifying that all information posted to the site would be publically viewable to all group members and warned against using this method for communicating confidential information. The accidental disclosure of confidential information by a client is not under Moody's control. Her actions to remove the information promptly once she became aware further align with Standard III(E).

In understanding the potential sensitivity clients express surrounding the confidentiality of personal information, this event highlights a need for further training. Moody might advocate for additional warnings or controls for clients when they consider using social media platforms for two-way communications.

STANDARD IV(A): DUTIES TO EMPLOYERS – LOYALTY

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Standard IV(A) Loyalty

In matters related to their employment, Members and Candidates must act for the benefit of their employer and not deprive their employer of the advantage of their skills and abilities, divulge confidential information, or otherwise cause harm to their employer.

Guidance

Highlights:

- *Employer Responsibilities*
- *Independent Practice*
- *Leaving an Employer*
- *Use of Social Media*
- *Whistleblowing*
- *Nature of Employment*

Standard IV(A) requires members and candidates to protect the interests of their firm by refraining from any conduct that would injure the firm, deprive it of profit, or deprive it of the member's or candidate's skills and ability. Members and candidates must always place the interests of clients above the interests of their employer but should also consider the effects of their conduct on the sustainability and integrity of the employer firm. In matters related to their employment, members and candidates must not engage in conduct that harms the interests of their employer. Implicit in this standard is the obligation of members and candidates to comply with the policies and procedures established by their employers that govern the employer–employee relationship—to the extent that such policies and procedures do not conflict with applicable laws, rules, or regulations or the Code and Standards.

This standard is not meant to be a blanket requirement to place employer interests ahead of personal interests in all matters. The standard does not require members and candidates to subordinate important personal and family obligations to their work. Members and candidates should enter into a dialogue with their employer about balancing personal and employment obligations when personal matters may interfere with their work on a regular or significant basis.

Employer Responsibilities

The employer–employee relationship imposes duties and responsibilities on both parties. Employers must recognize the duties and responsibilities that they owe to their employees if they expect to have content and productive employees.

Members and candidates are encouraged to provide their employer with a copy of the Code and Standards. These materials will inform the employer of the responsibilities of a CFA Institute member or a candidate in the CFA Program. The Code and Standards also serve as a basis for questioning employer policies and practices that conflict with these responsibilities.

Employers are not obligated to adhere to the Code and Standards. In expecting to retain competent employees who are members and candidates, however, they should not develop conflicting policies and procedures. The employer is responsible for a positive working environment, which includes an ethical workplace. Senior management has the additional responsibility to devise compensation structures and incentive arrangements that do not encourage unethical behavior.

Independent Practice

Included in Standard IV(A) is the requirement that members and candidates abstain from independent competitive activity that could conflict with the interests of their employer. Although Standard IV(A) does not preclude members or candidates from entering into an independent business while still employed, members and candidates who plan to engage in independent practice for compensation must notify their employer and describe the types of services they will render to prospective independent clients, the expected duration of the services, and the compensation for the services. Members and candidates should not render services until they receive consent from their employer to all of the terms of the arrangement. “Practice” means any service that the employer currently makes available for remuneration. “Undertaking independent practice” means engaging in competitive business, as opposed to making preparations to begin such practice.

Leaving an Employer

When members and candidates are planning to leave their current employer, they must continue to act in the employer’s best interest. They must not engage in any activities that would conflict with this duty until their resignation becomes effective. It is difficult to define specific guidelines for those members and candidates who are planning to compete with their employer as part of a new venture. The circumstances of each situation must be reviewed to distinguish permissible preparations from violations of duty. Activities that might constitute a violation, especially in combination, include the following:

- misappropriation of trade secrets,
- misuse of confidential information,
- solicitation of the employer’s clients prior to cessation of employment,

- self-dealing (appropriating for one's own property a business opportunity or information belonging to one's employer), and
- misappropriation of clients or client lists.

A departing employee is generally free to make arrangements or preparations to go into a competitive business before terminating the relationship with his or her employer as long as such preparations do not breach the employee's duty of loyalty. A member or candidate who is contemplating seeking other employment must not contact existing clients or potential clients prior to leaving his or her employer for purposes of soliciting their business for the new employer. Once notice is provided to the employer of the intent to resign, the member or candidate must follow the employer's policies and procedures related to notifying clients of his or her planned departure. In addition, the member or candidate must not take records or files to a new employer without the written permission of the previous employer.

Once an employee has left the firm, the skills and experience that an employee obtained while employed are not "confidential" or "privileged" information. Similarly, simple knowledge of the names and existence of former clients is generally not confidential information unless deemed such by an agreement or by law. Standard IV(A) does not prohibit experience or knowledge gained at one employer from being used at another employer. Firm records or work performed on behalf of the firm that is stored in paper copy or electronically for the member's or candidate's convenience while employed, however, should be erased or returned to the employer unless the firm gives permission to keep those records after employment ends.

The standard does not prohibit former employees from contacting clients of their previous firm as long as the contact information does not come from the records of the former employer or violate an applicable "noncompete agreement." Members and candidates are free to use public information after departing to contact former clients without violating Standard IV(A) as long as there is no specific agreement not to do so.

Employers often require employees to sign noncompete agreements that preclude a departing employee from engaging in certain conduct. Members and candidates should take care to review the terms of any such agreement when leaving their employer to determine what, if any, conduct those agreements may prohibit.

In some markets, there are agreements between employers within an industry that outline information that departing employees are permitted to take upon resignation, such as the "Protocol for Broker Recruiting" in the United States. These agreements ease individuals' transition between firms that have agreed to follow the outlined procedures. Members and candidates who move between firms that sign such agreements may rely on the protections provided as long as they faithfully adhere to all the procedures outlined.

For example, under the agreement between many US brokers, individuals are allowed to take some general client contact information when departing. To be protected, a copy of the information the individual is taking must be provided to the local management team for review. Additionally, the specific client information may only be used by the departing employee and not others employed by the new firm.

Use of Social Media

The growth in various online networking platforms, such as LinkedIn, Twitter, and Facebook (commonly referred to as social media platforms), is providing new opportunities and challenges for businesses. Members and candidates should understand and abide by all applicable firm policies and regulations as to the acceptable use of social media platforms to interact with clients and prospective clients. This is especially important when a member or candidate is planning to leave an employer.

Social media use makes determining how and when departure notification is delivered to clients more complex. Members and candidates may have developed profiles on these platforms that include connections with individuals who are clients of the firm, as well as individuals unrelated to their employer. Communications through social media platforms that potentially reach current clients should adhere to the employer's policies and procedures regarding notification of departing employees.

Social media connections with clients are also raising questions concerning the differences between public information and firm property. Specific accounts and user profiles of members and candidates may be created for solely professional reasons, including firm-approved accounts for client engagements. Such firm-approved business-related accounts would be considered part of the firm's assets, thus requiring members and candidates to transfer or delete the accounts as directed by their firm's policies and procedures. Best practice for members and candidates is to maintain separate accounts for their personal and professional social media activities. Members and candidates should discuss with their employers how profiles should be treated when a single account includes personal connections and also is used to conduct aspects of their professional activities.

Whistleblowing

A member's or candidate's personal interests, as well as the interests of his or her employer, are secondary to protecting the integrity of capital markets and the interests of clients. Therefore, circumstances may arise (e.g., when an employer is engaged in illegal or unethical activity) in which members and candidates must act contrary to their employer's interests in order to comply with their duties to the market and clients. In such instances, activities that would normally violate a member's or candidate's duty to his or her employer (such as contradicting employer instructions, violating certain policies and procedures, or preserving a record by copying employer records) may be justified. Such action would be permitted only if the intent is clearly aimed at protecting clients or the integrity of the market, not for personal gain.

Nature of Employment

A wide variety of business relationships exists within the investment industry. For instance, a member or candidate may be an employee or an independent contractor. Members and candidates must determine whether they are employees or independent contractors in order to determine the applicability of Standard IV(A). This issue will be decided largely by the degree of control exercised by the employing entity over the member or candidate. Factors determining control include whether the member's or candidate's hours, work location, and other parameters of the job are set; whether facilities are provided to the member or candidate; whether the member's or candidate's expenses are reimbursed; whether the member or candidate seeks work from other employers; and the number of clients or employers the member or candidate works for.

A member's or candidate's duties within an independent contractor relationship are governed by the oral or written agreement between the member and the client. Members and candidates should take care to define clearly the scope of their responsibilities and the expectations of each client within the context of each relationship. Once a member or candidate establishes a relationship with a client, the member or candidate has a duty to abide by the terms of the agreement.

STANDARD IV(A): RECOMMENDED PROCEDURES

- recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Employers may establish codes of conduct and operating procedures for their employees to follow. Members and candidates should fully understand the policies to ensure that they are not in conflict with the Code and Standards. The following topics identify policies that members and candidates should encourage their firms to adopt if the policies are not currently in place.

Competition Policy

A member or candidate must understand any restrictions placed by the employer on offering similar services outside the firm while employed by the firm. The policy may outline the procedures for requesting approval to undertake the outside service or may be a strict prohibition of such service. If a member's or candidate's employer elects to have its employees sign a noncompete agreement as part of the employment agreement, the member or candidate should ensure that the details are clear and fully explained prior to signing the agreement.

Termination Policy

Members and candidates should clearly understand the termination policies of their employer. Termination policies should establish clear procedures regarding the resignation process, including addressing how the termination will be disclosed to clients and staff and whether updates posted through social media platforms will be allowed. The firm's policy may also outline the procedures for transferring ongoing research and account management responsibilities. Finally, the procedures should address agreements that allow departing employees to remove specific client-related information upon resignation.

Incident-Reporting Procedures

Members and candidates should be aware of their firm's policies related to whistleblowing and encourage their firm to adopt industry best practices in this area. Many firms are required by regulatory mandates to establish confidential and anonymous reporting procedures that allow employees to report potentially unethical and illegal activities in the firm.

Employee Classification

Members and candidates should understand their status within their employer firm. Firms are encouraged to adopt a standardized classification structure (e.g., part time, full time, outside contractor) for their employees and indicate how each of the firm's policies applies to each employee class.

STANDARD IV(A): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Soliciting Former Clients):

Samuel Magee manages pension accounts for Trust Assets, Inc., but has become frustrated with the working environment and has been offered a position with Fiduciary Management. Before resigning from Trust Assets, Magee asks four big accounts to leave that firm and open accounts with Fiduciary. Magee also persuades several prospective clients to sign agreements with Fiduciary Management. Magee had previously made presentations to these prospects on behalf of Trust Assets.

> *Comment*: Magee violated the employee–employer principle requiring him to act solely for his employer's benefit. Magee's duty is to Trust Assets as long as he is employed there. The solicitation of Trust Assets' current clients and prospective clients is unethical and violates Standard IV(A).

Example 2 (Former Employer's Documents and Files):

James Hightower has been employed by Jason Investment Management Corporation for 15 years. He began as an analyst but assumed increasing responsibilities and is now a senior portfolio manager and a member of the firm's investment policy committee. Hightower has decided to leave Jason Investment and start his own investment management business. He has been careful not to tell any of Jason's clients that he is leaving; he does not want to be accused of breaching his duty to Jason by soliciting Jason's clients before his departure. Hightower is planning to copy and take with him the following documents and information he developed or worked on while at Jason: (1) the client list, with addresses, telephone numbers, and other pertinent client information; (2) client account statements; (3) sample marketing presentations to prospective clients containing Jason's performance record; (4) Jason's recommended list of securities; (5) computer models to determine asset allocations for accounts with various objectives; (6) computer models for stock selection; and (7) personal computer spreadsheets for Hightower's major corporate recommendations, which he developed when he was an analyst.

> *Comment*: Except with the consent of their employer, departing members and candidates may not take employer property, which includes books, records, reports, and other materials, because taking such materials may interfere with their employer's business opportunities. Taking any employer records, even those the member or candidate prepared, violates Standard IV(A). Employer records include items stored in hard copy or any other medium (e.g., home computers, portable storage devices, cell phones).

Example 3 (Addressing Rumors):

Reuben Winston manages all-equity portfolios at Target Asset Management (TAM), a large, established investment counselor. Ten years previously, Philpott & Company, which manages a family of global bond mutual funds, acquired TAM in a diversification move. After the merger, the combined operations prospered in the fixed-income business but the equity management business at TAM languished. Lately, a few of the equity pension accounts that had been with TAM before the merger have terminated their relationships with TAM. One day, Winston finds on his voice mail the following message from a concerned client: "Hey! I just heard that Philpott is close to announcing the sale of your firm's equity management business to Rugged Life. What is going on?" Not being aware of any such deal, Winston and his associates are stunned. Their internal inquiries are met with denials from Philpott management, but the rumors persist. Feeling left in the dark, Winston contemplates leading an employee buyout of TAM's equity management business.

> *Comment*: An employee-led buyout of TAM's equity asset management business would be consistent with Standard IV(A) because it would rest on the permission of the employer and, ultimately, the clients. In this case, however, in which employees suspect the senior managers or principals are not truthful or forthcoming, Winston should consult legal counsel to determine appropriate action.

Example 4 (Ownership of Completed Prior Work):

Laura Clay, who is unemployed, wants part-time consulting work while seeking a full-time analyst position. During an interview at Bradley Associates, a large institutional asset manager, Clay is told that the firm has no immediate research openings but would be willing to pay her a flat fee to complete a study of the wireless communications industry within a given period of time. Clay would be allowed unlimited access to Bradley's research files and would be welcome to come to the offices and use whatever support facilities are available during normal working hours. Bradley's research director does not seek any exclusivity for Clay's output, and the two agree to the arrangement on a handshake. As Clay nears completion of the study, she is offered an analyst job in the research department of Winston & Company, a brokerage firm, and she is pondering submitting the draft of her wireless study for publication by Winston.

> *Comment*: Although she is under no written contractual obligation to Bradley, Clay has an obligation to let Bradley act on the output of her study before Winston & Company or Clay uses the information to their advantage. That is, unless Bradley gives permission to Clay and waives its rights to her wireless report, Clay would be in violation of Standard IV(A) if she were to immediately recommend to Winston the same transactions recommended in the report to Bradley. Furthermore, Clay must not take from Bradley any research file material or other property that she may have used.

Example 5 (Ownership of Completed Prior Work):

Emma Madeline, a recent college graduate and a candidate in the CFA Program, spends her summer as an unpaid intern at Murdoch and Lowell. The senior managers at Murdoch are attempting to bring the firm into compliance with the GIPS standards, and Madeline is assigned to assist in its efforts. Two months into her internship, Madeline applies for a job at McMillan & Company, which has plans to become GIPS compliant. Madeline accepts the job with McMillan. Before leaving Murdoch, she copies the firm's software that she helped develop because she believes this software will assist her in her new position.

> *Comment*: Even though Madeline does not receive monetary compensation for her services at Murdoch, she has used firm resources in creating the software and is considered an employee because she receives compensation and benefits in the form of work experience and knowledge. By copying the software, Madeline violated Standard IV(A) because she misappropriated Murdoch's property without permission.

Example 6 (Soliciting Former Clients):

Dennis Elliot has hired Sam Chisolm, who previously worked for a competing firm. Chisolm left his former firm after 18 years of employment. When Chisolm begins working for Elliot, he wants to contact his former clients because he knows them well and is certain that many will follow him to his new employer. Is Chisolm in violation of Standard IV(A) if he contacts his former clients?

> *Comment*: Because client records are the property of the firm, contacting former clients for any reason through the use of client lists or other information taken from a former employer without permission would be a violation of Standard IV(A). In addition, the nature and extent of the contact with former clients may be governed by the terms of any noncompete agreement signed by the employee and the former employer that covers contact with former clients after employment.
>
> Simple knowledge of the names and existence of former clients is not confidential information, just as skills or experience that an employee obtains while employed are not "confidential" or "privileged" information. The Code and Standards do not impose a prohibition on the use of experience or knowledge gained at one employer from being used at another employer. The Code and Standards also do not prohibit former employees from contacting clients of their previous firm, in the absence of a noncompete agreement. Members and candidates are free to use public information about their former firm after departing to contact former clients without violating Standard IV(A).
>
> In the absence of a noncompete agreement, as long as Chisolm maintains his duty of loyalty to his employer before joining Elliot's firm, does not take steps to solicit clients until he has left his former firm, and does not use material from his former employer without its permission after he has left, he is not in violation of the Code and Standards.

Example 7 (Starting a New Firm):

Geraldine Allen currently works at a registered investment company as an equity analyst. Without notice to her employer, she registers with government authorities to start an investment company that will compete with her employer, but she does not actively seek clients. Does registration of this competing company with the appropriate regulatory authorities constitute a violation of Standard IV(A)?

> *Comment*: Allen's preparation for the new business by registering with the regulatory authorities does not conflict with the work for her employer if the preparations have been done on Allen's own time outside the office and if Allen will not be soliciting clients for the business or otherwise operating the new company until she has left her current employer.

Example 8 (Competing with Current Employer):

Several employees are planning to depart their current employer within a few weeks and have been careful to not engage in any activities that would conflict with their duty to their current employer. They have just learned that one of their employer's clients has undertaken a request for proposal (RFP) to review and possibly hire a new investment consultant. The RFP has been sent to the employer and all of its competitors. The group believes that the new entity to be formed would be qualified to respond to the RFP and be eligible for the business. The RFP submission period is likely to conclude before the employees' resignations are effective. Is it permissible for the group of departing employees to respond to the RFP for their anticipated new firm?

> *Comment*: A group of employees responding to an RFP that their employer is also responding to would lead to direct competition between the employees and the employer. Such conduct violates Standard IV(A) unless the group of employees receives permission from their employer as well as the entity sending out the RFP.

Example 9 (Externally Compensated Assignments):

Alfonso Mota is a research analyst with Tyson Investments. He works part time as a mayor for his hometown, a position for which he receives compensation. Must Mota seek permission from Tyson to serve as mayor?

> *Comment*: If Mota's mayoral duties are so extensive and time-consuming that they might detract from his ability to fulfill his responsibilities at Tyson, he should discuss his outside activities with his employer and come to a mutual agreement regarding how to manage his personal commitments with his responsibilities to his employer.

Example 10 (Soliciting Former Clients):

After leaving her employer, Shawna McQuillen establishes her own money management business. While with her former employer, she did not sign a noncompete agreement that would have prevented her from soliciting former clients. Upon her departure, she does not take any of her client lists or contact information and she clears her personal computer of any employer records, including client contact information. She obtains the phone numbers of her former clients through public records and contacts them to solicit their business.

> *Comment*: McQuillen is not in violation of Standard IV(A) because she has not used information or records from her former employer and is not prevented by an agreement with her former employer from soliciting her former clients.

Example 11 (Whistleblowing Actions):

Meredith Rasmussen works on a buy-side trading desk and concentrates on in-house trades for a hedge fund subsidiary managed by a team at the investment management firm. The hedge fund has been very successful and is marketed globally by the firm. From her experience as the trader for much of the activity of the fund, Rasmussen has become quite knowledgeable about the hedge fund's strategy, tactics, and performance. When a distinct break in the market occurs, however, and many of the securities involved in the hedge fund's strategy decline markedly in value, Rasmussen observes that the reported performance of the hedge fund does not reflect this decline. In her experience, the lack of any effect is a very unlikely occurrence. She approaches the head of trading about her concern and is told that she should not ask any questions and that the fund is big and successful and is not her concern. She is fairly sure something is not right, so she contacts the compliance officer, who also tells her to stay away from the issue of this hedge fund's reporting.

> *Comment*: Rasmussen has clearly come upon an error in policies, procedures, and compliance practices in the firm's operations. Having been unsuccessful in finding a resolution with her supervisor and the compliance officer, Rasmussen should consult the firm's whistleblowing policy to determine the appropriate next step toward informing management of her concerns. The potentially unethical actions of the investment management division are appropriate grounds for further disclosure, so Rasmussen's whistleblowing would not represent a violation of Standard IV(A).
>
> See also Standard I(D)–Misconduct and Standard IV(C)–Responsibilities of Supervisors.

Example 12 (Soliciting Former Clients):

Angel Crome has been a private banker for YBSafe Bank for the past eight years. She has been very successful and built a considerable client portfolio during that time but is extremely frustrated by the recent loss of reputation by her current employer and subsequent client insecurity. A locally renowned headhunter contacted Crome a few days ago and offered her an interesting job with a competing private bank. This bank offers a substantial signing bonus for advisers with their own client portfolios. Crome figures that she can solicit at least 70% of her clients to follow her and gladly enters into the new employment contract.

> *Comment*: Crome may contact former clients upon termination of her employment with YBSafe Bank, but she is prohibited from using client records built by and kept with her in her capacity as an employee of YBSafe Bank. Client lists are proprietary information of her former employer and must not be used for her or her new employer's benefit. The use of written, electronic, or any other form of records other than publicly available information to contact her former clients at YBSafe Bank will be a violation of Standard IV(A).

Example 13 (Notification of Code and Standards):

Krista Smith is a relatively new assistant trader for the fixed-income desk of a major investment bank. She is on a team responsible for structuring collateralized debt obligations (CDOs) made up of securities in the inventory of the trading desk. At a meeting of the team, senior executives explain the opportunity to eventually separate the CDO into various risk-rated tranches to be sold to the clients of the firm. After the senior executives leave the meeting, the head trader announces various responsibilities of each member of the team and then says, “This is a good time to unload some of the junk we have been stuck with for a while and disguise it with ratings and a thick, unreadable prospectus, so don’t be shy in putting this CDO together. Just kidding.” Smith is worried by this remark and asks some of her colleagues what the head trader meant. They all respond that he was just kidding but that there is some truth in the remark because the CDO is seen by management as an opportunity to improve the quality of the securities in the firm’s inventory.

Concerned about the ethical environment of the workplace, Smith decides to talk to her supervisor about her concerns and provides the head trader with a copy of the Code and Standards. Smith discusses the principle of placing the client above the interest of the firm and the possibility that the development of the new CDO will not adhere to this responsibility. The head trader assures Smith that the appropriate analysis will be conducted when determining the appropriate securities for collateral. Furthermore, the ratings are assigned by an independent firm and the prospectus will include full and factual disclosures. Smith is reassured by the meeting, but she also reviews the company’s procedures and requirements for reporting potential violations of company policy and securities laws.

> *Comment*: Smith’s review of the company policies and procedures for reporting violations allows her to be prepared to report through the appropriate whistleblower process if she decides that the CDO development process involves unethical actions by others. Smith’s actions comply with the Code and Standards principles of placing the client’s interests first and being loyal to her employer. In providing her supervisor with a copy of the Code and Standards, Smith is highlighting the high level of ethical conduct she is required to adhere to in her professional activities.

Example 14 (Leaving an Employer):

Laura Webb just left her position as portfolio analyst at Research Systems, Inc. (RSI). Her employment contract included a non-solicitation agreement that requires her to wait two years before soliciting RSI clients for any investment-related services. Upon leaving, Webb was informed that RSI would contact clients immediately about her departure and introduce her replacement.

While working at RSI, Webb connected with clients, other industry associates, and friends through her LinkedIn network. Her business and personal relationships were intermingled because she considered many of her clients to be personal friends. Realizing that her LinkedIn network would be a valuable resource for new employment opportunities, she updated her profile several days following her departure from RSI. LinkedIn automatically sent a notification to Webb’s entire network that her employment status had been changed in her profile.

> *Comment*: Prior to her departure, Webb should have discussed any client information contained in her social media networks. By updating her LinkedIn profile after RSI notified clients and after her employment ended, she has appropriately placed her employer’s interests ahead of her own personal interests. In addition, she has not violated the non-solicitation agreement with RSI, unless it prohibited any contact with clients during the two-year period.

Example 15 (Confidential Firm Information):

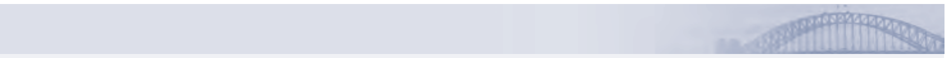
Sanjay Gupta is a research analyst at Naram Investment Management (NIM). NIM uses a team-based research process to develop recommendations on investment opportunities covered by the team members. Gupta, like others, provides commentary for NIM's clients through the company blog, which is posted weekly on the NIM password-protected website. According to NIM's policy, every contribution to the website must be approved by the company's compliance department before posting. Any opinions expressed on the website are disclosed as representing the perspective of NIM.

Gupta also writes a personal blog to share his experiences with friends and family. As with most blogs, Gupta's personal blog is widely available to interested readers through various internet search engines. Occasionally, when he disagrees with the team-based research opinions of NIM, Gupta uses his personal blog to express his own opinions as a counterpoint to the commentary posted on the NIM website. Gupta believes this provides his readers with a more complete perspective on these investment opportunities.

> *Comment*: Gupta is in violation of Standard IV(A) for disclosing confidential firm information through his personal blog. The recommendations on the firm's blog to clients are not freely available across the internet, but his personal blog post indirectly provides the firm's recommendations.
>
> Additionally, by posting research commentary on his personal blog, Gupta is using firm resources for his personal advantage. To comply with Standard IV(A), members and candidates must receive consent from their employer prior to using company resources.

STANDARD IV(B): DUTIES TO EMPLOYERS - ADDITIONAL COMPENSATION ARRANGEMENTS

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;



Members and Candidates must not accept gifts, benefits, compensation, or consideration that competes with or might reasonably be expected to create a conflict of interest with their employer's interest unless they obtain written consent from all parties involved.

Guidance

Standard IV(B) requires members and candidates to obtain permission from their employer before accepting compensation or other benefits from third parties for the services rendered to the employer or for any services that might create a conflict with their employer's interest. Compensation and benefits include direct compensation by the client and any indirect compensation or other benefits received from third parties. "Written consent" includes any form of communication that can be documented (for example, communication via e-mail that can be retrieved and documented).

Members and candidates must obtain permission for additional compensation/benefits because such arrangements may affect loyalties and objectivity and create potential conflicts of interest. Disclosure allows an employer to consider the outside arrangements when evaluating the actions and motivations of members and candidates. Moreover, the employer is entitled to have full knowledge of all compensation/benefit arrangements so as to be able to assess the true cost of the services members or candidates are providing.

There may be instances in which a member or candidate is hired by an employer on a “part-time” basis. “Part-time” status applies to employees who do not commit the full number of hours required for a normal work week. Members and candidates should discuss possible limitations to their abilities to provide services that may be competitive with their employer during the negotiation and hiring process. The requirements of Standard IV(B) would be applicable to limitations identified at that time.

STANDARD IV(B): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Members and candidates should make an immediate written report to their supervisor and compliance officer specifying any compensation they propose to receive for services in addition to the compensation or benefits received from their employer. The details of the report should be confirmed by the party offering the additional compensation, including performance incentives offered by clients. This written report should state the terms of any agreement under which a member or candidate will receive additional compensation; “terms” include the nature of the compensation, the approximate amount of compensation, and the duration of the agreement.

STANDARD IV(B): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Notification of Client Bonus Compensation):

Geoff Whitman, a portfolio analyst for Adams Trust Company, manages the account of Carol Cochran, a client. Whitman is paid a salary by his employer, and Cochran pays the trust company a standard fee based on the market value of assets in her portfolio. Cochran proposes to Whitman that “any year that my portfolio achieves at least a 15% return before taxes, you and your wife can fly to Monaco at my expense and use my condominium during the third week of January.” Whitman does not inform his employer of the arrangement and vacations in Monaco the following January as Cochran’s guest.

> *Comment*: Whitman violated Standard IV(B) by failing to inform his employer in writing of this supplemental, contingent compensation arrangement. The nature of the arrangement could have resulted in partiality to

> Cochran's account, which could have detracted from Whitman's performance with respect to other accounts he handles for Adams Trust. Whitman must obtain the consent of his employer to accept such a supplemental benefit.

Example 2 (Notification of Outside Compensation):

Terry Jones sits on the board of directors of Exercise Unlimited, Inc. In return for his services on the board, Jones receives unlimited membership privileges for his family at all Exercise Unlimited facilities. Jones purchases Exercise Unlimited stock for the client accounts for which it is appropriate. Jones does not disclose this arrangement to his employer because he does not receive monetary compensation for his services to the board.

> *Comment*: Jones has violated Standard IV(B) by failing to disclose to his employer benefits received in exchange for his services on the board of directors. The nonmonetary compensation may create a conflict of interest in the same manner as being paid to serve as a director.

Example 3 (Prior Approval for Outside Compensation):

Jonathan Hollis is an analyst of oil-and-gas companies for Specialty Investment Management. He is currently recommending the purchase of ABC Oil Company shares and has published a long, well-thought-out research report to substantiate his recommendation. Several weeks after publishing the report, Hollis receives a call from the investor-relations office of ABC Oil saying that Thomas Andrews, CEO of the company, saw the report and really liked the analyst's grasp of the business and his company. The investor-relations officer invites Hollis to visit ABC Oil to discuss the industry further. ABC Oil offers to send a company plane to pick Hollis up and arrange for his accommodations while visiting. Hollis, after gaining the appropriate approvals, accepts the meeting with the CEO but declines the offered travel arrangements.

Several weeks later, Andrews and Hollis meet to discuss the oil business and Hollis's report. Following the meeting, Hollis joins Andrews and the investment relations officer for dinner at an upscale restaurant near ABC Oil's headquarters.

Upon returning to Specialty Investment Management, Hollis provides a full review of the meeting to the director of research, including a disclosure of the dinner attended.

> *Comment*: Hollis's actions did not violate Standard IV(B). Through gaining approval before accepting the meeting and declining the offered travel arrangements, Hollis sought to avoid any potential conflicts of interest between his company and ABC Oil. Because the location of the dinner was not available prior to arrival and Hollis notified his company of the dinner upon his return, accepting the dinner should not impair his objectivity. By disclosing the dinner, Hollis has enabled Specialty Investment Management to assess whether it has any impact on future reports and recommendations by Hollis related to ABC Oil.

STANDARD IV(C): DUTIES TO EMPLOYERS - RESPONSIBILITIES OF SUPERVISORS

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Members and Candidates must make reasonable efforts to ensure that anyone subject to their supervision or authority complies with applicable laws, rules, regulations, and the Code and Standards.

Guidance

Highlights:

- *System for Supervision*
- *Supervision Includes Detection*

Standard IV(C) states that members and candidates must promote actions by all employees under their supervision and authority to comply with applicable laws, rules, regulations, and firm policies and the Code and Standards.

Any investment professional who has employees subject to her or his control or influence—whether or not the employees are CFA Institute members, CFA charterholders, or candidates in the CFA Program—exercises supervisory responsibility. Members and candidates acting as supervisors must also have in-depth knowledge of the Code and Standards so that they can apply this knowledge in discharging their supervisory responsibilities.

The conduct that constitutes reasonable supervision in a particular case depends on the number of employees supervised and the work performed by those employees. Members and candidates with oversight responsibilities for large numbers of employees may not be able to personally evaluate the conduct of these employees on a continuing basis. These members and candidates may delegate supervisory duties to subordinates who directly oversee the other employees. A member's or candidate's responsibilities under Standard IV(C) include instructing those subordinates to whom supervision is delegated about methods to promote compliance, including preventing and detecting violations of laws, rules, regulations, firm policies, and the Code and Standards.

At a minimum, Standard IV(C) requires that members and candidates with supervisory responsibility make reasonable efforts to prevent and detect violations by ensuring the establishment of effective compliance systems. However, an effective compliance system goes beyond enacting a code of ethics, establishing policies and procedures to achieve compliance with the code and applicable law, and reviewing employee actions to determine whether they are following the rules.

To be effective supervisors, members and candidates should implement education and training programs on a recurring or regular basis for employees under their supervision. Such programs will assist the employees with meeting their professional obligations to practice in an ethical manner within the applicable legal system. Further, establishing incentives—monetary or otherwise—for employees not only to meet business goals but also to reward ethical behavior offers supervisors another way to assist employees in complying with their legal and ethical obligations.

Often, especially in large organizations, members and candidates may have supervisory responsibility but not the authority to establish or modify firm-wide compliance policies and procedures or incentive structures. Such limitations should not prevent a member or candidate from working with his or her own superiors and within the firm structure to develop and implement effective compliance tools, including but not limited to:

- a code of ethics,
- compliance policies and procedures,
- education and training programs,
- an incentive structure that rewards ethical conduct, and
- adoption of firm-wide best practice standards (e.g., the GIPS standards, the CFA Institute Asset Manager Code of Professional Conduct).

A member or candidate with supervisory responsibility should bring an inadequate compliance system to the attention of the firm's senior managers and recommend corrective action. If the member or candidate clearly cannot discharge supervisory responsibilities because of the absence of a compliance system or because of an inadequate compliance system, the member or candidate should decline in writing to accept supervisory responsibility until the firm adopts reasonable procedures to allow adequate exercise of supervisory responsibility.

System for Supervision

Members and candidates with supervisory responsibility must understand what constitutes an adequate compliance system for their firms and make reasonable efforts to see that appropriate compliance procedures are established, documented, communicated to covered personnel, and followed. "Adequate" procedures are those designed to meet industry standards, regulatory requirements, the requirements of the Code and Standards, and the circumstances of the firm. Once compliance procedures are established, the supervisor must also make reasonable efforts to ensure that the procedures are monitored and enforced.

To be effective, compliance procedures must be in place prior to the occurrence of a violation of the law or the Code and Standards. Although compliance procedures cannot be designed to anticipate every potential violation, they should be designed to anticipate the activities most likely to result in misconduct. Compliance programs must be appropriate for the size and nature of the organization. The member or candidate should review model compliance procedures or other industry programs to ensure that the firm's procedures meet the minimum industry standards.

Once a supervisor learns that an employee has violated or may have violated the law or the Code and Standards, the supervisor must promptly initiate an assessment to determine the extent of the wrongdoing. Relying on an employee's statements about the extent of the violation or assurances that the wrongdoing will not reoccur is not enough. Reporting the misconduct up the chain of command and warning the employee to cease the activity are also not enough. Pending the outcome of the investigation, a supervisor should take steps to ensure that the violation will not be repeated, such as placing limits on the employee's activities or increasing the monitoring of the employee's activities.

Supervision Includes Detection

Members and candidates with supervisory responsibility must also make reasonable efforts to detect violations of laws, rules, regulations, firm policies, and the Code and Standards. The supervisors exercise reasonable supervision by establishing and implementing written compliance procedures and ensuring that those procedures are followed through periodic review. If a member or candidate has adopted reasonable procedures and taken steps to institute an effective compliance program, then the member or candidate may not be in violation of Standard IV(C) if he or she does not detect violations that occur despite these efforts. The fact that violations do occur may indicate, however, that the compliance procedures are inadequate. In addition, in some cases, merely enacting such procedures may not be sufficient to fulfill the duty required by Standard IV(C). A member or candidate may be in violation of Standard IV(C) if he or she knows or should know that the procedures designed to promote compliance, including detecting and preventing violations, are not being followed.

STANDARD IV(C): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Codes of Ethics or Compliance Procedures

Members and candidates are encouraged to recommend that their employers adopt a code of ethics. Adoption of a code of ethics is critical to establishing a strong ethical foundation for investment advisory firms and their employees. Codes of ethics formally emphasize and reinforce the client loyalty responsibilities of investment firm personnel, protect investing clients by deterring misconduct, and protect the firm's reputation for integrity.

There is a distinction, however, between codes of ethics and the specific policies and procedures needed to ensure compliance with the codes and with securities laws and regulations. Although both are important, codes of ethics should consist of fundamental, principle-based ethical and fiduciary concepts that are applicable to all of the firm's employees. In this way, firms can best convey to employees and clients the ethical ideals that investment advisers strive to achieve. These concepts need to be implemented, however, by detailed, firm-wide compliance policies and procedures. Compliance procedures assist the firm's personnel in fulfilling the responsibilities enumerated in the code of ethics and make probable that the ideals expressed in the code of ethics will be adhered to in the day-to-day operation of the firm.

Stand-alone codes of ethics should be written in plain language and should address general fiduciary concepts. They should be unencumbered by numerous detailed procedures. Codes presented in this way are the most effective in stressing to employees that they are in positions of trust and must act with integrity at all times. Mingling compliance procedures in the firm's code of ethics goes against the goal of reinforcing the ethical obligations of employees.

Separating the code of ethics from compliance procedures will also reduce, if not eliminate, the legal terminology and "boilerplate" language that can make the underlying ethical principles incomprehensible to the average person. Above all, to ensure the creation of a culture of ethics and integrity rather than one that merely focuses on following the rules, the principles in the code of ethics must be stated in a way that is accessible and understandable to everyone in the firm.

Members and candidates should encourage their employers to provide their codes of ethics to clients. In this case also, a simple, straightforward code of ethics will be best understood by clients. Unencumbered by the compliance procedures, the code of ethics will be effective in conveying that the firm is committed to conducting business in an ethical manner and in the best interests of the clients.

Adequate Compliance Procedures

A supervisor complies with Standard IV(C) by identifying situations in which legal violations or violations of the Code and Standards are likely to occur and by establishing and enforcing compliance procedures to prevent such violations. Adequate compliance procedures should

- be contained in a clearly written and accessible manual that is tailored to the firm's operations,
- be drafted so that the procedures are easy to understand,
- designate a compliance officer whose authority and responsibility are clearly defined and who has the necessary resources and authority to implement the firm's compliance procedures,
- describe the hierarchy of supervision and assign duties among supervisors,
- implement a system of checks and balances,
- outline the scope of the procedures,
- outline procedures to document the monitoring and testing of compliance procedures,
- outline permissible conduct, and
- delineate procedures for reporting violations and sanctions.

Once a compliance program is in place, a supervisor should

- disseminate the contents of the program to appropriate personnel,
- periodically update procedures to ensure that the measures are adequate under the law,
- continually educate personnel regarding the compliance procedures,
- issue periodic reminders of the procedures to appropriate personnel,
- incorporate a professional conduct evaluation as part of an employee's performance review,
- review the actions of employees to ensure compliance and identify violators, and
- take the necessary steps to enforce the procedures once a violation has occurred.

Once a violation is discovered, a supervisor should

- respond promptly,
- conduct a thorough investigation of the activities to determine the scope of the wrongdoing,
- increase supervision or place appropriate limitations on the wrongdoer pending the outcome of the investigation, and
- review procedures for potential changes necessary to prevent future violations from occurring.

Implementation of Compliance Education and Training

No amount of ethics education and awareness will deter someone determined to commit fraud for personal enrichment. But the vast majority of investment professionals strive to achieve personal success with dedicated service to their clients and employers.

Regular ethics and compliance training, in conjunction with adoption of a code of ethics, is critical to investment firms seeking to establish a strong culture of integrity and to provide an environment in which employees routinely engage in ethical conduct in compliance with the law. Training and education assist individuals in both recognizing areas that are prone to ethical and legal pitfalls and identifying those circumstances and influences that can impair ethical judgment.

By implementing educational programs, supervisors can train their subordinates to put into practice what the firm's code of ethics requires. Education helps employees make the link between legal and ethical conduct and the long-term success of the business; a strong culture of compliance signals to clients and potential clients that the firm has truly embraced ethical conduct as fundamental to the firm's mission to serve its clients.

Establish an Appropriate Incentive Structure

Even if individuals want to make the right choices and follow an ethical course of conduct and are aware of the obstacles that may trip them up, they can still be influenced to act improperly by a corporate culture that embraces a "succeed at all costs" mentality, stresses results regardless of the methods used to achieve those results, and does not reward ethical behavior. Supervisors can reinforce an individual's natural desire to "do the right thing" by building a culture of integrity in the workplace.

Supervisors and firms must look closely at their incentive structure to determine whether the structure encourages profits and returns at the expense of ethically appropriate conduct. Reward structures may turn a blind eye to how desired outcomes are achieved and encourage dysfunctional or counterproductive behavior. Only when compensation and incentives are firmly tied to client interests and *how* outcomes are achieved, rather than *how much* is generated for the firm, will employees work to achieve a culture of integrity.

STANDARD IV(C): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Supervising Research Activities):

Jane Mattock, senior vice president and head of the research department of H&V, Inc., a regional brokerage firm, has decided to change her recommendation for Timber Products from buy to sell. In line with H&V's procedures, she orally advises certain other H&V executives of her proposed actions before the report is prepared for publication. As a result of Mattock's conversation with Dieter Frampton, one of the H&V executives accountable to Mattock, Frampton immediately sells Timber's stock from his own account and from certain discretionary client accounts. In addition, other personnel inform certain institutional customers of the changed recommendation before it is printed and disseminated to all H&V customers who have received previous Timber reports.

> *Comment*: Mattock has violated Standard IV(C) by failing to reasonably and adequately supervise the actions of those accountable to her. She did not prevent or establish reasonable procedures designed to prevent dissemination of or trading on the information by those who knew of her changed recommendation. She must ensure that her firm has procedures for reviewing or recording any trading in the stock of a corporation that has been the subject of an unpublished change in recommendation. Adequate procedures would have informed the subordinates of their duties and detected sales by Frampton and selected customers.

Example 2 (Supervising Research Activities):

Deion Miller is the research director for Jamestown Investment Programs. The portfolio managers have become critical of Miller and his staff because the Jamestown portfolios do not include any stock that has been the subject of a merger or tender offer. Georgia Ginn, a member of Miller's staff, tells Miller that she has been studying a local company, Excelsior, Inc., and recommends its purchase. Ginn adds that the company has been widely rumored to be the subject of a merger study by a well-known conglomerate and discussions between them are under way. At Miller's request, Ginn prepares a memo recommending the stock. Miller passes along Ginn's memo to the portfolio managers prior to leaving for vacation, and he notes that he has not reviewed the memo. As a result of the memo, the portfolio managers buy Excelsior stock immediately. The day Miller returns to the office, he learns that Ginn's only sources for the report were her brother, who is an acquisitions analyst with Acme Industries, the "well-known conglomerate," and that the merger discussions were planned but not held.

> *Comment*: Miller violated Standard IV(C) by not exercising reasonable supervision when he disseminated the memo without checking to ensure that Ginn had a reasonable and adequate basis for her recommendations and that Ginn was not relying on material nonpublic information.

Example 3 (Supervising Trading Activities):

David Edwards, a trainee trader at Wheeler & Company, a major national brokerage firm, assists a customer in paying for the securities of Highland, Inc., by using anticipated profits from the immediate sale of the same securities. Despite the fact that Highland is not on Wheeler's recommended list, a large volume of its stock is traded through Wheeler in this manner. Roberta Ann Mason is a Wheeler vice president responsible for supervising compliance with the securities laws in the trading department. Part of her compensation from Wheeler is based on commission revenues from the trading department. Although she notices the increased trading activity, she does nothing to investigate or halt it.

> *Comment*: Mason's failure to adequately review and investigate purchase orders in Highland stock executed by Edwards and her failure to supervise the trainee's activities violate Standard IV(C). Supervisors should be especially sensitive to actual or potential conflicts between their own self-interests and their supervisory responsibilities.

Example 4 (Supervising Trading Activities and Record Keeping):

Samantha Tabbing is senior vice president and portfolio manager for Crozet, Inc., a registered investment advisory and registered broker/dealer firm. She reports to Charles Henry, the president of Crozet. Crozet serves as the investment adviser and principal underwriter for ABC and XYZ public mutual funds. The two funds' prospectuses allow Crozet to trade financial futures for the funds for the limited purpose of hedging against market risks. Henry, extremely impressed by Tabbing's performance in the past two years, directs Tabbing to act as portfolio manager for the funds. For the benefit of its employees, Crozet has also organized the Crozet Employee Profit-Sharing Plan (CEPSP), a defined contribution retirement plan. Henry assigns Tabbing to manage 20% of the assets of CEPSP. Tabbing's investment objective for her portion of CEPSP's assets is aggressive growth. Unbeknownst to Henry, Tabbing frequently places S&P 500 Index purchase and sale orders for the funds and the CEPSP without providing the futures commission merchants (FCMs) who take the orders with any prior or simultaneous designation of the account for which the trade has been placed. Frequently, neither Tabbing nor anyone else at Crozet completes an internal trade ticket to record the time an order was placed or the specific account for which the order was intended. FCMs often designate a specific account only after the trade, when Tabbing provides such designation. Crozet has no written operating procedures or compliance manual concerning its futures trading, and its compliance department does not review such trading. After observing the market's movement, Tabbing assigns to CEPSP the S&P 500 positions with more favorable execution prices and assigns positions with less favorable execution prices to the funds.

> *Comment*: Henry violated Standard IV(C) by failing to adequately supervise Tabbing with respect to her S&P 500 trading. Henry further violated Standard IV(C) by failing to establish record-keeping and reporting procedures to prevent or detect Tabbing's violations. Henry must make a reasonable effort to determine that adequate compliance procedures covering all employee trading activity are established, documented, communicated, and followed.

Example 5 (Accepting Responsibility):

Meredith Rasmussen works on a buy-side trading desk and concentrates on in-house trades for a hedge fund subsidiary managed by a team at the investment management firm. The hedge fund has been very successful and is marketed globally by the firm. From her experience as the trader for much of the activity of the fund, Rasmussen has become quite knowledgeable about the hedge fund's strategy, tactics, and performance. When a distinct break in the market occurs and many of the securities involved in the hedge fund's strategy decline markedly in value, however, Rasmussen observes that the reported performance of the hedge fund does not at all reflect this decline. From her experience, this lack of an effect is a very unlikely occurrence. She approaches the head of trading about her concern and is told that she should not ask any questions and that the fund is too big and successful and is not her concern. She is fairly sure something is not right, so she contacts the compliance officer and is again told to stay away from the hedge fund reporting issue.

> *Comment*: Rasmussen has clearly come upon an error in policies, procedures, and compliance practices within the firm's operations. According to Standard IV(C), the supervisor and the compliance officer have the responsibility to review the concerns brought forth by Rasmussen. Supervisors

> have the responsibility of establishing and encouraging an ethical culture in the firm. The dismissal of Rasmussen's question violates Standard IV(C) and undermines the firm's ethical operations.
>
> See also Standard I(D)–Misconduct and, for guidance on whistleblowing, Standard IV(A)–Loyalty.

Example 6 (Inadequate Procedures):

Brendan Witt, a former junior sell-side technology analyst, decided to return to school to earn an MBA. To keep his research skills and industry knowledge sharp, Witt accepted a position with On-line and Informed, an independent internet-based research company. The position requires the publication of a recommendation and report on a different company every month. Initially, Witt is a regular contributor of new research and a participant in the associated discussion boards that generally have positive comments on the technology sector. Over time, his ability to manage his educational requirements and his work requirements begin to conflict with one another. Knowing a recommendation is due the next day for On-line, Witt creates a report based on a few news articles and what the conventional wisdom of the markets has deemed the "hot" security of the day.

> *Comment*: Allowing the report submitted by Witt to be posted highlights a lack of compliance procedures by the research firm. Witt's supervisor needs to work with the management of On-line to develop an appropriate review process to ensure that all contracted analysts comply with the requirements.
>
> See also Standard V(A)–Diligence and Reasonable Basis because it relates to Witt's responsibility for substantiating a recommendation.

Example 7 (Inadequate Supervision):

Michael Papis is the chief investment officer of his state's retirement fund. The fund has always used outside advisers for the real estate allocation, and this information is clearly presented in all fund communications. Thomas Nagle, a recognized sell-side research analyst and Papis's business school classmate, recently left the investment bank he worked for to start his own asset management firm, Accessible Real Estate. Nagle is trying to build his assets under management and contacts Papis about gaining some of the retirement fund's allocation. In the previous few years, the performance of the retirement fund's real estate investments was in line with the fund's benchmark but was not extraordinary. Papis decides to help out his old friend and also to seek better returns by moving the real estate allocation to Accessible. The only notice of the change in adviser appears in the next annual report in the listing of associated advisers.

> *Comment*: Papis's actions highlight the need for supervision and review at all levels in an organization. His responsibilities may include the selection of external advisers, but the decision to change advisers appears arbitrary. Members and candidates should ensure that their firm has appropriate policies and procedures in place to detect inappropriate actions, such as the action taken by Papis.
>
> See also Standard V(A)–Diligence and Reasonable Basis, Standard V(B)–Communication with Clients and Prospective Clients, and Standard VI(A)–Disclosure of Conflicts.

Example 8 (Supervising Research Activities):

Mary Burdette was recently hired by Fundamental Investment Management (FIM) as a junior auto industry analyst. Burdette is expected to expand the social media presence of the firm because she is active with various networks, including Facebook, LinkedIn, and Twitter. Although Burdette's supervisor, Joe Graf, has never used social media, he encourages Burdette to explore opportunities to increase FIM's online presence and ability to share content, communicate, and broadcast information to clients. In response to Graf's encouragement, Burdette is working on a proposal detailing the advantages of getting FIM onto Twitter in addition to launching a company Facebook page.

As part of her auto industry research for FIM, Burdette is completing a report on the financial impact of Sun Drive Auto Ltd.'s new solar technology for compact automobiles. This research report will be her first for FIM, and she believes Sun Drive's technology could revolutionize the auto industry. In her excitement, Burdette sends a quick tweet to FIM Twitter followers summarizing her "buy" recommendation for Sun Drive Auto stock.

> *Comment*: Graf has violated Standard IV(C) by failing to reasonably supervise Burdette with respect to the contents of her tweet. He did not establish reasonable procedures to prevent the unauthorized dissemination of company research through social media networks. Graf must make sure all employees receive regular training about FIM's policies and procedures, including the appropriate business use of personal social media networks.
>
> See Standard III(B) for additional guidance.

Example 9 (Supervising Research Activities):

Chen Wang leads the research department at YYRA Retirement Planning Specialists. Chen supervises a team of 10 analysts in a fast-paced and understaffed organization. He is responsible for coordinating the firm's approved process to review all reports before they are provided to the portfolio management team for use in rebalancing client portfolios.

One of Chen's direct reports, Huang Mei, covers the banking industry. Chen must submit the latest updates to the portfolio management team tomorrow morning. Huang has yet to submit her research report on ZYX Bank because she is uncomfortable providing a "buy" or "sell" opinion of ZYX on the basis of the completed analysis. Pressed for time and concerned that Chen will reject a "hold" recommendation, she researches various websites and blogs on the banking sector for whatever she can find on ZYX. One independent blogger provides a new interpretation of the recently reported data Huang has analyzed and concludes with a strong "sell" recommendation for ZYX. She is impressed by the originality and resourcefulness of this blogger's report.

Very late in the evening, Huang submits her report and "sell" recommendation to Chen without any reference to the independent blogger's report. Given the late time of the submission and the competence of Huang's prior work, Chen compiles this report with the recommendations from each of the other analysts and meets with the portfolio managers to discuss implementation.

> *Comment*: Chen has violated Standard IV(C) by neglecting to reasonably and adequately follow the firm's approved review process for Huang's research report. The delayed submission and the quality of prior work do not remove Chen's requirement to uphold the designated review process. A member or candidate with supervisory responsibility must make reasonable efforts to see that appropriate procedures are established, documented, communicated to covered personnel, and followed.

STANDARD V(A): INVESTMENT ANALYSIS, RECOMMENDATIONS, AND ACTIONS - DILIGENCE AND REASONABLE BASIS

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Standard V(A) Diligence and Reasonable Basis

Members and Candidates must:

1 Exercise diligence, independence, and thoroughness in analyzing investments, making investment recommendations, and taking investment actions.

2 Have a reasonable and adequate basis, supported by appropriate research and investigation, for any investment analysis, recommendation, or action.

Guidance

Highlights:

- *Defining Diligence and Reasonable Basis*
- *Using Secondary or Third-Party Research*
- *Using Quantitatively Oriented Research*
- *Developing Quantitatively Oriented Techniques*
- *Selecting External Advisers and Subadvisers*
- *Group Research and Decision Making*

The application of Standard V(A) depends on the investment philosophy the member, candidate, or firm is following, the role of the member or candidate in the investment decision-making process, and the support and resources provided by the member's or candidate's employer. These factors will dictate the nature of the diligence and thoroughness of the research and the level of investigation required by Standard V(A).

The requirements for issuing conclusions based on research will vary in relation to the member's or candidate's role in the investment decision-making process, but the member or candidate must make reasonable efforts to cover all pertinent issues when arriving at a recommendation. Members and candidates enhance transparency by providing or offering to provide supporting information to clients when recommending a purchase or sale or when changing a recommendation.

Defining Diligence and Reasonable Basis

Every investment decision is based on a set of facts known and understood at the time. Clients turn to members and candidates for advice and expect these advisers to have more information and knowledge than they do. This information and knowledge is the basis from which members and candidates apply their professional judgment in taking investment actions and making recommendations.

At a basic level, clients want assurance that members and candidates are putting forth the necessary effort to support the recommendations they are making. Communicating the level and thoroughness of the information reviewed before the member or candidate makes a judgment allows clients to understand the reasonableness of the recommended investment actions.

As with determining the suitability of an investment for the client, the necessary level of research and analysis will differ with the product, security, or service being offered. In providing an investment service, members and candidates typically use a variety of resources, including company reports, third-party research, and results from quantitative models. A reasonable basis is formed through a balance of these resources appropriate for the security or decision being analyzed.

The following list provides some, but definitely not all, examples of attributes to consider while forming the basis for a recommendation:

- global, regional, and country macroeconomic conditions,
- a company's operating and financial history,
- the industry's and sector's current conditions and the stage of the business cycle,
- a mutual fund's fee structure and management history,
- the output and potential limitations of quantitative models,
- the quality of the assets included in a securitization, and
- the appropriateness of selected peer-group comparisons.

Even though an investment recommendation may be well informed, downside risk remains for any investment. Members and candidates can base their decisions only on the information available at the time decisions are made. The steps taken in developing a diligent and reasonable recommendation should minimize unexpected downside events.

Using Secondary or Third-Party Research

If members and candidates rely on secondary or third-party research, they must make reasonable and diligent efforts to determine whether such research is sound. Secondary research is defined as research conducted by someone else in the member's or candidate's firm. Third-party research is research conducted by entities outside the member's or candidate's firm, such as a brokerage firm, bank, or research firm. If a member or candidate has reason to suspect that either secondary or third-party research or information comes from a source that lacks a sound basis, the member or candidate must not rely on that information.

Members and candidates should make reasonable enquiries into the source and accuracy of all data used in completing their investment analysis and recommendations. The sources of the information and data will influence the level of the review a member or candidate must undertake. Information and data taken from internet sources, such as personal blogs, independent research aggregation websites, or social media websites, likely require a greater level of review than information from more established research organizations.

Criteria that a member or candidate can use in forming an opinion on whether research is sound include the following:

- assumptions used,
- rigor of the analysis performed,
- date/timeliness of the research, and
- evaluation of the objectivity and independence of the recommendations.

A member or candidate may rely on others in his or her firm to determine whether secondary or third-party research is sound and use the information in good faith unless the member or candidate has reason to question its validity or the processes and procedures used by those responsible for the research. For example, a portfolio manager may not have a choice of a data source because the firm's senior managers conducted due diligence to determine which vendor would provide services; the member or candidate can use the information in good faith assuming the due diligence process was deemed adequate.

A member or candidate should verify that the firm has a policy about the timely and consistent review of approved research providers to ensure that the quality of the research continues to meet the necessary standards. If such a policy is not in place at the firm, the member or candidate should encourage the development and adoption of a formal review practice.

Using Quantitatively Oriented Research

Standard V(A) applies to the rapidly expanding use of quantitatively oriented research models and processes, such as computer-generated modeling, screening, and ranking of investment securities; the creation or valuation of derivative instruments; and quantitative portfolio construction techniques. These models and processes are being used for much more than the back testing of investment strategies, especially with continually advancing technology and techniques. The continued broad development of quantitative methods and models is an important part of capital market developments.

Members and candidates need to have an understanding of the parameters used in models and quantitative research that are incorporated into their investment recommendations. Although they are not required to become experts in every technical aspect of the models, they must understand the assumptions and limitations inherent in any model and how the results were used in the decision-making process.

The reliance on and potential limitations of financial models became clear through the investment crisis that unfolded in 2007 and 2008. In some cases, the financial models used to value specific securities and related derivative products did not adequately demonstrate the level of associated risks. Members and candidates should make reasonable efforts to test the output of investment models and other pre-programed analytical tools they use. Such validation should occur before incorporating the process into their methods, models, or analyses.

Although not every model can test for every factor or outcome, members and candidates should ensure that their analyses incorporate a broad range of assumptions sufficient to capture the underlying characteristics of investments. The omission from the analysis of potentially negative outcomes or of levels of risk outside the norm may misrepresent the true economic value of an investment. The possible scenarios for analysis should include factors that are likely to have a substantial influence on the investment value and may include extremely positive and negative scenarios.

Developing Quantitatively Oriented Techniques

Individuals who create new quantitative models and services must exhibit a higher level of diligence in reviewing new products than the individuals who ultimately use the analytical output. Members and candidates involved in the development and oversight of quantitatively oriented models, methods, and algorithms must understand the technical aspects of the products they provide to clients. A thorough testing of the model and resulting analysis should be completed prior to product distribution.

Members and candidates need to consider the source and time horizon of the data used as inputs in financial models. The information from many commercially available databases may not effectively incorporate both positive and negative market cycles. In the development of a recommendation, the member or candidate may need to test the models by using volatility and performance expectations that represent scenarios outside the observable databases. In reviewing the computer models or the resulting output, members and candidates need to pay particular attention to the assumptions used in the analysis and the rigor of the analysis to ensure that the model incorporates a wide range of possible input expectations, including negative market events.

Selecting External Advisers and Subadvisers

Financial instruments and asset allocation techniques continue to develop and evolve. This progression has led to the use of specialized managers to invest in specific asset classes or diversification strategies that complement a firm's in-house expertise. Standard V(A) applies to the level of review necessary in selecting an external adviser or subadviser to manage a specifically mandated allocation. Members and candidates must review managers as diligently as they review individual funds and securities.

Members and candidates who are directly involved with the use of external advisers need to ensure that their firms have standardized criteria for reviewing these selected external advisers and managers. Such criteria would include, but would not be limited to, the following:

- reviewing the adviser's established code of ethics,
- understanding the adviser's compliance and internal control procedures,
- assessing the quality of the published return information, and
- reviewing the adviser's investment process and adherence to its stated strategy.

Codes, standards, and guides to best practice published by CFA Institute provide members and candidates with examples of acceptable practices for external advisers and advice in selecting a new adviser. The following guides are available at the CFA Institute website (www.cfainstitute.org): Asset Manager Code of Professional Conduct, Global Investment Performance Standards, and Model Request for Proposal (for equity, credit, or real estate managers).

Group Research and Decision Making

Commonly, members and candidates are part of a group or team that is collectively responsible for producing investment analysis or research. The conclusions or recommendations of the group report represent the consensus of the group and are not necessarily the views of the member or candidate, even though the name of the member or candidate is included on the report. In some instances, a member or candidate will not agree with the view of the group. If, however, the member or candidate believes that the consensus opinion has a reasonable and adequate basis and is independent and objective, the member or candidate need not decline to be identified with the report. If the member or candidate is confident in the process, the member or candidate does not need to dissociate from the report even if it does not reflect his or her opinion.

STANDARD V(A): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Members and candidates should encourage their firms to consider the following policies and procedures to support the principles of Standard V(A):

- Establish a policy requiring that research reports, credit ratings, and investment recommendations have a basis that can be substantiated as reasonable and adequate. An individual employee (a supervisory analyst) or a group of employees (a review committee) should be appointed to review and approve such items prior to external circulation to determine whether the criteria established in the policy have been met.

- Develop detailed, written guidance for analysts (research, investment, or credit), supervisory analysts, and review committees that establishes the due diligence procedures for judging whether a particular recommendation has a reasonable and adequate basis.
- Develop measurable criteria for assessing the quality of research, the reasonableness and adequacy of the basis for any recommendation or rating, and the accuracy of recommendations over time. In some cases, firms may consider implementing compensation arrangements that depend on these measurable criteria and that are applied consistently to all related analysts.
- Develop detailed, written guidance that establishes minimum levels of scenario testing of all computer-based models used in developing, rating, and evaluating financial instruments. The policy should contain criteria related to the breadth of the scenarios tested, the accuracy of the output over time, and the analysis of cash flow sensitivity to inputs.
- Develop measurable criteria for assessing outside providers, including the quality of information being provided, the reasonableness and adequacy of the provider's collection practices, and the accuracy of the information over time. The established policy should outline how often the provider's products are reviewed.
- Adopt a standardized set of criteria for evaluating the adequacy of external advisers. The policy should include how often and on what basis the allocation of funds to the adviser will be reviewed.

STANDARD V(A): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Sufficient Due Diligence):

Helen Hawke manages the corporate finance department of Sarkozi Securities, Ltd. The firm is anticipating that the government will soon close a tax loophole that currently allows oil-and-gas exploration companies to pass on drilling expenses to holders of a certain class of shares. Because market demand for this tax-advantaged class of stock is currently high, Sarkozi convinces several companies to undertake new equity financings at once, before the loophole closes. Time is of the essence, but Sarkozi lacks sufficient resources to conduct adequate research on all the prospective issuing companies. Hawke decides to estimate the IPO prices on the basis of the relative size of each company and to justify the pricing later when her staff has time.

> *Comment*: Sarkozi should have taken on only the work that it could adequately handle. By categorizing the issuers by general size, Hawke has bypassed researching all the other relevant aspects that should be considered

> when pricing new issues and thus has not performed sufficient due diligence. Such an omission can result in investors purchasing shares at prices that have no actual basis. Hawke has violated Standard V(A).

Example 2 (Sufficient Scenario Testing):

Babu Dhaliwal works for Heinrich Brokerage in the corporate finance group. He has just persuaded Feggans Resources, Ltd., to allow his firm to do a secondary equity financing at Feggans Resources' current stock price. Because the stock has been trading at higher multiples than similar companies with equivalent production, Dhaliwal presses the Feggans Resources managers to project what would be the maximum production they could achieve in an optimal scenario. Based on these numbers, he is able to justify the price his firm will be asking for the secondary issue. During a sales pitch to the brokers, Dhaliwal then uses these numbers as the base-case production levels that Feggans Resources will achieve.

> *Comment*: When presenting information to the brokers, Dhaliwal should have given a range of production scenarios and the probability of Feggans Resources achieving each level. By giving the maximum production level as the likely level of production, he has misrepresented the chances of achieving that production level and seriously misled the brokers. Dhaliwal has violated Standard V(A).

Example 3 (Developing a Reasonable Basis):

Brendan Witt, a former junior sell-side technology analyst, decided to return to school to earn an MBA. To keep his research skills and industry knowledge sharp, Witt accepted a position with On-line and Informed, an independent internet-based research company. The position requires the publication of a recommendation and report on a different company every month. Initially, Witt is a regular contributor of new research and a participant in the associated discussion boards that generally have positive comments on the technology sector. Over time, his ability to manage his educational requirements and his work requirements begin to conflict with one another. Knowing a recommendation is due the next day for On-line, Witt creates a report based on a few news articles and what the conventional wisdom of the markets has deemed the "hot" security of the day.

> *Comment*: Witt's knowledge of and exuberance for technology stocks, a few news articles, and the conventional wisdom of the markets do not constitute, without more information, a reasonable and adequate basis for a stock recommendation that is supported by appropriate research and investigation. Therefore, Witt has violated Standard V(A).
>
> See also Standard IV(C)–Responsibilities of Supervisors because it relates to the firm's inadequate procedures.

Example 4 (Timely Client Updates):

Kristen Chandler is an investment consultant in the London office of Dalton Securities, a major global investment consultant firm. One of her UK pension funds has decided to appoint a specialist US equity manager. Dalton's global manager of research relies on local consultants to cover managers within their regions and, after conducting thorough due diligence, puts their views and ratings in Dalton's manager database. Chandler accesses Dalton's global manager research database and conducts a screen of all US equity managers on the basis of a match with the client's desired philosophy/style, performance, and tracking-error targets. She selects the five managers that meet these criteria and puts them in a briefing report that is delivered to the client 10 days later. Between the time of Chandler's database search and the delivery of the report to the client, Chandler is told that Dalton has updated the database with the information that one of the firms that Chandler has recommended for consideration lost its chief investment officer, the head of its US equity research, and the majority of its portfolio managers on the US equity product—all of whom have left to establish their own firm. Chandler does not revise her report with this updated information.

> *Comment*: Chandler has failed to satisfy the requirement of Standard V(A). Although Dalton updated the manager ratings to reflect the personnel turnover at one of the firms, Chandler did not update her report to reflect the new information.

Example 5 (Group Research Opinions):

Evelyn Mastakis is a junior analyst who has been asked by her firm to write a research report predicting the expected interest rate for residential mortgages over the next six months. Mastakis submits her report to the fixed-income investment committee of her firm for review, as required by firm procedures. Although some committee members support Mastakis's conclusion, the majority of the committee disagrees with her conclusion, and the report is significantly changed to indicate that interest rates are likely to increase more than originally predicted by Mastakis. Should Mastakis ask that her name be taken off the report when it is disseminated?

> *Comment*: The results of research are not always clear, and different people may have different opinions based on the same factual evidence. In this case, the committee may have valid reasons for issuing a report that differs from the analyst's original research. The firm can issue a report that is different from the original report of an analyst as long as there is a reasonable and adequate basis for its conclusions.
>
> Generally, analysts must write research reports that reflect their own opinion and can ask the firm not to put their name on reports that ultimately differ from that opinion. When the work is a group effort, however, not all members of the team may agree with all aspects of the report. Ultimately, members and candidates can ask to have their names removed from the report, but if they are satisfied that the process has produced results or conclusions that have a reasonable and adequate basis, members and candidates do not have to dissociate from the report even when they do not agree with its contents. If Mastakis is confident in the process, she does not need to dissociate from the report even if it does not reflect her opinion.

Example 6 (Reliance on Third-Party Research):

Gary McDermott runs a two-person investment management firm. McDermott's firm subscribes to a service from a large investment research firm that provides research reports. McDermott's firm makes investment recommendations on the basis of these reports.

> *Comment*: Members and candidates can rely on third-party research but must make reasonable and diligent efforts to determine that such research is sound. If McDermott undertakes due diligence efforts on a regular basis

> to ensure that the research produced by the large firm is objective and reasonably based, McDermott can rely on that research when making investment recommendations to clients.

Example 7 (Due Diligence in Submanager Selection):

Paul Ostrowski's business has grown significantly over the past couple of years, and some clients want to diversify internationally. Ostrowski decides to find a submanager to handle the expected international investments. Because this will be his first subadviser, Ostrowski uses the CFA Institute model "request for proposal" to design a questionnaire for his search. By his deadline, he receives seven completed questionnaires from a variety of domestic and international firms trying to gain his business. Ostrowski reviews all the applications in detail and decides to select the firm that charges the lowest fees because doing so will have the least impact on his firm's bottom line.

> *Comment*: The selection of an external adviser or subadviser should be based on a full and complete review of the adviser's services, performance history, and cost structure. In basing the decision on the fee structure alone, Ostrowski may be violating Standard V(A).
>
> See also Standard III(C)–Suitability because it relates to the ability of the selected adviser to meet the needs of the clients.

Example 8 (Sufficient Due Diligence):

Michael Papis is the chief investment officer of his state's retirement fund. The fund has always used outside advisers for the real estate allocation, and this information is clearly presented in all fund communications. Thomas Nagle, a recognized sell-side research analyst and Papis's business school classmate, recently left the investment bank he worked for to start his own asset management firm, Accessible Real Estate. Nagle is trying to build his assets under management and contacts Papis about gaining some of the retirement fund's allocation. In the previous few years, the performance of the retirement fund's real estate investments was in line with the fund's benchmark but was not extraordinary. Papis decides to help out his old friend and also to seek better returns by moving the real estate allocation to Accessible. The only notice of the change in adviser appears in the next annual report in the listing of associated advisers.

> *Comment*: Papis violated Standard V(A). His responsibilities may include the selection of the external advisers, but the decision to change advisers appears to have been arbitrary. If Papis was dissatisfied with the current real estate adviser, he should have conducted a proper solicitation to select the most appropriate adviser.
>
> See also Standard IV(C)–Responsibilities of Supervisors, Standard V(B)–Communication with Clients and Prospective Clients, and Standard VI(A)–Disclosure of Conflicts.

Example 9 (Sufficient Due Diligence):

Andre Shrub owns and operates Conduit, an investment advisory firm. Prior to opening Conduit, Shrub was an account manager with Elite Investment, a hedge fund managed by his good friend Adam Reed. To attract clients to a new Conduit fund, Shrub offers lower-than-normal management fees. He can do so because the fund consists of two top-performing funds managed by Reed. Given his personal friendship with

Reed and the prior performance record of these two funds, Shrub believes this new fund is a winning combination for all parties. Clients quickly invest with Conduit to gain access to the Elite funds. No one is turned away because Conduit is seeking to expand its assets under management.

> *Comment*: Shrub violated Standard V(A) by not conducting a thorough analysis of the funds managed by Reed before developing the new Conduit fund. Shrub's reliance on his personal relationship with Reed and his prior knowledge of Elite are insufficient justification for the investments. The funds may be appropriately considered, but a full review of their operating procedures, reporting practices, and transparency are some elements of the necessary due diligence.
>
> See also Standard III(C)–Suitability.

Example 10 (Sufficient Due Diligence):

Bob Thompson has been doing research for the portfolio manager of the fixed-income department. His assignment is to do sensitivity analysis on securitized subprime mortgages. He has discussed with the manager possible scenarios to use to calculate expected returns. A key assumption in such calculations is housing price appreciation (HPA) because it drives "prepays" (prepayments of mortgages) and losses. Thompson is concerned with the significant appreciation experienced over the previous five years as a result of the increased availability of funds from subprime mortgages. Thompson insists that the analysis should include a scenario run with −10% for Year 1, −5% for Year 2, and then (to project a worst-case scenario) 0% for Years 3 through 5. The manager replies that these assumptions are too dire because there has never been a time in their available database when HPA was negative.

Thompson conducts his research to better understand the risks inherent in these securities and evaluates these securities in the worst-case scenario, a less likely but possible environment. Based on the results of the enhanced scenarios, Thompson does not recommend the purchase of the securitization. Against the general market trends, the manager follows Thompson's recommendation and does not invest. The following year, the housing market collapses. In avoiding the subprime investments, the manager's portfolio outperforms its peer group that year.

> *Comment*: Thompson's actions in running the scenario test with inputs beyond the historical trends available in the firm's databases adhere to the principles of Standard V(A). His concerns over recent trends provide a sound basis for further analysis. Thompson understands the limitations of his model, when combined with the limited available historical information, to accurately predict the performance of the funds if market conditions change negatively.
>
> See also Standard I(B)–Independence and Objectivity.

Example 11 (Use of Quantitatively Oriented Models):

Espacia Liakos works in sales for Hellenica Securities, a firm specializing in developing intricate derivative strategies to profit from particular views on market expectations. One of her clients is Eugenie Carapalis, who has become convinced that commodity prices will become more volatile over the coming months. Carapalis asks Liakos to quickly engineer a strategy that will benefit from this expectation. Liakos turns to Hellenica's modeling group to fulfill this request. Because of the tight deadline, the modeling group outsources parts of the work to several trusted third parties. Liakos implements the disparate components of the strategy as the firms complete them.

Within a month, Carapalis is proven correct: Volatility across a range of commodities increases sharply. But her derivatives position with Hellenica returns huge losses, and the losses increase daily. Liakos investigates and realizes that although each of the various components of the strategy had been validated, they had never been evaluated as an integrated whole. In extreme conditions, portions of the model worked at cross-purposes with other portions, causing the overall strategy to fail dramatically.

> *Comment*: Liakos violated Standard V(A). Members and candidates must understand the statistical significance of the results of the models they recommend and must be able to explain them to clients. Liakos did not take adequate care to ensure a thorough review of the whole model; its components were evaluated only individually. Because Carapalis clearly intended to implement the strategy as a whole rather than as separate parts, Liakos should have tested how the components of the strategy interacted as well as how they performed individually.

Example 12 (Successful Due Diligence/Failed Investment):

Alton Newbury is an investment adviser to high-net-worth clients. A client with an aggressive risk profile in his investment policy statement asks about investing in the Top Shelf hedge fund. This fund, based in Calgary, Alberta, Canada, has reported 20% returns for the first three years. The fund prospectus states that its strategy involves long and short positions in the energy sector and extensive leverage. Based on his analysis of the fund's track record, the principals involved in managing the fund, the fees charged, and the fund's risk profile, Newbury recommends the fund to the client and secures a position in it. The next week, the fund announces that it has suffered a loss of 60% of its value and is suspending operations and redemptions until after a regulatory review. Newbury's client calls him in a panic and asks for an explanation.

> *Comment*: Newbury's actions were consistent with Standard V(A). Analysis of an investment that results in a reasonable basis for recommendation does not guarantee that the investment has no downside risk. Newbury should discuss the analysis process with the client while reminding him or her that past performance does not lead to guaranteed future gains and that losses in an aggressive investment portfolio should be expected.

Example 13 (Quantitative Model Diligence):

Barry Cannon is the lead quantitative analyst at CityCenter Hedge Fund. He is responsible for the development, maintenance, and enhancement of the proprietary models the fund uses to manage its investors' assets. Cannon reads several high-level mathematical publications and blogs to stay informed of current developments. One blog, run by Expert CFA, presents some intriguing research that may benefit one of CityCenter's current models. Cannon is under pressure from firm executives to improve the model's predictive abilities, and he incorporates the factors discussed in the online research. The updated output recommends several new investments to the fund's portfolio managers.

> *Comment*: Cannon has violated Standard V(A) by failing to have a reasonable basis for the new recommendations made to the portfolio managers. He needed to diligently research the effect of incorporating the new factors before offering the output recommendations. Cannon may use the blog for ideas, but it is his responsibility to determine the effect on the firm's proprietary models.

> See Standard VII(B) regarding the violation by "Expert CFA" in the use of the CFA designation.

Example 14 (Selecting a Service Provider):

Ellen Smith is a performance analyst at Artic Global Advisors, a firm that manages global equity mandates for institutional clients. She was asked by her supervisor to review five new performance attribution systems and recommend one that would more appropriately explain the firm's investment strategy to clients. On the list was a system she recalled learning about when visiting an exhibitor booth at a recent conference. The system is highly quantitative and something of a "black box" in how it calculates the attribution values. Smith recommended this option without researching the others because the sheer complexity of the process was sure to impress the clients.

> *Comment*: Smith's actions do not demonstrate a sufficient level of diligence in reviewing this product to make a recommendation for selecting the service. Besides not reviewing or considering the other four potential systems, she did not determine whether the "black box" attribution process aligns with the investment practices of the firm, including its investments in different countries and currencies. Smith must review and understand the process of any software or system before recommending its use as the firm's attribution system.

Example 15 (Subadviser Selection):

Craig Jackson is working for Adams Partners, Inc., and has been assigned to select a hedge fund subadviser to improve the diversification of the firm's large fund-of-funds product. The allocation must be in place before the start of the next quarter. Jackson uses a consultant database to find a list of suitable firms that claim compliance with the GIPS standards. He calls more than 20 firms on the list to confirm their potential interest and to determine their most recent quarterly and annual total return values. Because of the short turnaround, Jackson recommends the firm with the greatest total return values for selection.

> *Comment*: By considering only performance and GIPS compliance, Jackson has not conducted sufficient review of potential firms to satisfy the requirements of Standard V(A). A thorough investigation of the firms and their operations should be conducted to ensure that their addition would increase the diversity of clients' portfolios and that they are suitable for the fund-of-funds product.

Example 16 (Manager Selection):

Timothy Green works for Peach Asset Management, where he creates proprietary models that analyze data from the firm request for proposal questionnaires to identify managers for possible inclusion in the firm's fund-of-funds investment platform. Various criteria must be met to be accepted to the platform. Because of the number of respondents to the questionnaires, Green uses only the data submitted to make a recommendation for adding a new manager.

> *Comment*: By failing to conduct any additional outside review of the information to verify what was submitted through the request for proposal, Green has likely not satisfied the requirements of Standard V(A). The amount of

> information requested from outside managers varies among firms. Although the requested information may be comprehensive, Green should ensure sufficient effort is undertaken to verify the submitted information before recommending a firm for inclusion. This requires that he goes beyond the information provided by the manager on the request for proposal questionnaire and may include interviews with interested managers, reviews of regulatory filings, and discussions with the managers' custodian or auditor.

Example 17 (Technical Model Requirements):

Jérôme Dupont works for the credit research group of XYZ Asset Management, where he is in charge of developing and updating credit risk models. In order to perform accurately, his models need to be regularly updated with the latest market data.

Dupont does not interact with or manage money for any of the firm's clients. He is in contact with the firm's US corporate bond fund manager, John Smith, who has only very superficial knowledge of the model and who from time to time asks very basic questions regarding the output recommendations. Smith does not consult Dupont with respect to finalizing his clients' investment strategies.

Dupont's recently assigned objective is to develop a new emerging market corporate credit risk model. The firm is planning to expand into emerging credit, and the development of such a model is a critical step in this process. Because Smith seems to follow the model's recommendations without much concern for its quality as he develops his clients' investment strategies, Dupont decides to focus his time on the development of the new emerging market model and neglects to update the US model.

After several months without regular updates, Dupont's diagnostic statistics start to show alarming signs with respect to the quality of the US credit model. Instead of conducting the long and complicated data update, Dupont introduces new codes into his model with some limited new data as a quick "fix." He thinks this change will address the issue without needing to complete the full data update, so he continues working on the new emerging market model.

Several months following the quick "fix," another set of diagnostic statistics reveals nonsensical results and Dupont realizes that his earlier change contained an error. He quickly corrects the error and alerts Smith. Smith realizes that some of the prior trades he performed were due to erroneous model results. Smith rebalances the portfolio to remove the securities purchased on the basis of the questionable results without reporting the issue to anyone else.

> *Comment*: Smith violated standard V(A) because exercising "diligence, independence, and thoroughness in analyzing investments, making investment recommendations, and taking investment actions" means that members and candidates must understand the technical aspects of the products they provide to clients. Smith does not understand the model he is relying on to manage money. Members and candidates should also make reasonable enquiries into the source and accuracy of all data used in completing their investment analysis and recommendations.
>
> Dupont violated V(A) even if he does not trade securities or make investment decisions. Dupont's models give investment recommendations, and Dupont is accountable for the quality of those recommendations. Members and candidates should make reasonable efforts to test the output of preprogramed analytical tools they use. Such validation should occur before incorporating the tools into their decision-making process.
>
> See also Standard V(B)–Communication with Clients and Prospective Clients.

STANDARD V(B): INVESTMENT ANALYSIS, RECOMMENDATIONS, AND ACTIONS - COMMUNICATION WITH CLIENTS AND PROSPECTIVE CLIENTS

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Members and Candidates must:

1. Disclose to clients and prospective clients the basic format and general principles of the investment processes they use to analyze investments, select securities, and construct portfolios and must promptly disclose any changes that might materially affect those processes.
2. Disclose to clients and prospective clients significant limitations and risks associated with the investment process.
3. Use reasonable judgment in identifying which factors are important to their investment analyses, recommendations, or actions and include those factors in communications with clients and prospective clients.
4. Distinguish between fact and opinion in the presentation of investment analyses and recommendations.

Guidance

Highlights:

- *Informing Clients of the Investment Process*
- *Different Forms of Communication*
- *Identifying Risk and Limitations*
- *Report Presentation*
- *Distinction between Facts and Opinions in Reports*

Standard V(B) addresses member and candidate conduct with respect to communicating with clients. Developing and maintaining clear, frequent, and thorough communication practices is critical to providing high-quality financial services to clients. When clients understand the information communicated to them, they also can understand exactly how members and candidates are acting on their behalf, which gives clients the opportunity to make well-informed decisions about their investments. Such understanding can be accomplished only through clear communication.

Standard V(B) states that members and candidates should communicate in a recommendation the factors that were instrumental in making the investment recommendation. A critical part of this requirement is to distinguish clearly between opinions and facts. In preparing a research report, the member or candidate must present the basic characteristics of the security(ies) being analyzed, which will allow the reader to evaluate the report and incorporate information the reader deems relevant to his or her investment decision-making process.

Similarly, in preparing a recommendation about, for example, an asset allocation strategy, alternative investment vehicle, or structured investment product, the member or candidate should include factors that are relevant to the asset classes that are being discussed. Follow-up communication of significant changes in the risk characteristics of a security or asset strategy is required. Providing regular updates to any changes in the risk characteristics is recommended.

Informing Clients of the Investment Process

Members and candidates must adequately describe to clients and prospective clients the manner in which they conduct the investment decision-making process. Such disclosure should address factors that have positive and negative influences on the recommendations, including significant risks and limitations of the investment process used. The member or candidate must keep clients and other interested parties informed on an ongoing basis about changes to the investment process, especially newly identified significant risks and limitations. Only by thoroughly understanding the nature of the investment product or service can a client determine whether changes to that product or service could materially affect his or her investment objectives.

Understanding the basic characteristics of an investment is of great importance in judging the suitability of that investment on a standalone basis, but it is especially important in determining the impact each investment will have on the characteristics of a portfolio. Although the risk and return characteristics of a common stock might seem to be essentially the same for any investor when the stock is viewed in isolation, the effects of those characteristics greatly depend on the other investments held. For instance, if the particular stock will represent 90% of an individual's investments, the stock's importance in the portfolio is vastly different from what it would be to an investor with a highly diversified portfolio for whom the stock will represent only 2% of the holdings.

A firm's investment policy may include the use of outside advisers to manage various portions of clients' assets under management. Members and candidates should inform the clients about the specialization or diversification expertise provided by the external adviser(s). This information allows clients to understand the full mix of products and strategies being applied that may affect their investment objectives.

Different Forms of Communication

For purposes of Standard V(B), communication is not confined to a written report of the type traditionally generated by an analyst researching a security, company, or industry. A presentation of information can be made via any means of communication, including in-person recommendation or description, telephone conversation, media broadcast, or transmission by computer (e.g., on the internet).

Computer and mobile device communications have rapidly evolved over the past few years. Members and candidates using any social media service to communicate business information must be diligent in their efforts to avoid unintended problems because these services may not be available to all clients. When providing information to clients through new technologies, members and candidates should take reasonable steps to ensure that such delivery would treat all clients fairly and, if necessary, be considered publicly disseminated.

The nature of client communications is highly diverse—from one word ("buy" or "sell") to in-depth reports of more than 100 pages. A communication may contain a general recommendation about the market, asset allocations, or classes of investments (e.g., stocks, bonds, real estate) or may relate to a specific security. If recommendations are contained in capsule form (such as a recommended stock list), members and candidates should notify clients that additional information and analyses are available from the producer of the report.

Identifying Risks and Limitations

Members and candidates must outline to clients and prospective clients significant risks and limitations of the analysis contained in their investment products or recommendations. The type and nature of significant risks will depend on the investment process that members and candidates are following and on the personal circumstances of the client. In general, the use of leverage constitutes a significant risk and should be disclosed.

Members and candidates must adequately disclose the general market-related risks and the risks associated with the use of complex financial instruments that are deemed significant. Other types of risks that members and candidates may consider disclosing include, but are not limited to, counterparty risk, country risk, sector or industry risk, security-specific risk, and credit risk.

Investment securities and vehicles may have limiting factors that influence a client's or potential client's investment decision. Members and candidates must report to clients and prospective clients the existence of limitations significant to the decision-making process. Examples of such factors and attributes include, but are not limited to, investment liquidity and capacity. Liquidity is the ability to liquidate an investment on a timely basis at a reasonable cost. Capacity is the investment amount beyond which returns will be negatively affected by new investments.

The appropriateness of risk disclosure should be assessed on the basis of what was known at the time the investment action was taken (often called an *ex ante* basis). Members and candidates must disclose significant risks known to them at the time of the disclosure. Members and candidates cannot be expected to disclose risks they are unaware of at the time recommendations or investment actions are made. In assessing compliance with Standard V(B), it is important to establish knowledge of a purported significant risk or limitation. A one-time investment loss that occurs after the disclosure does not constitute a pertinent factor in assessing whether significant risks and limitations were properly disclosed. Having no knowledge of a risk or limitation that subsequently triggers a loss may reveal a deficiency in the diligence and reasonable basis of the research of the member or candidate but may not reveal a breach of Standard V(B).

Report Presentation

Once the analytical process has been completed, the member or candidate who prepares the report must include those elements that are important to the analysis and conclusions of the report so that the reader can follow and challenge the report's reasoning. A report writer who has done adequate investigation may emphasize certain areas, touch briefly on others, and omit certain aspects deemed unimportant. For instance, a report may dwell on a quarterly earnings release or new-product introduction and omit other matters as long as the analyst clearly stipulates the limits to the scope of the report.

Investment advice based on quantitative research and analysis must be supported by readily available reference material and should be applied in a manner consistent with previously applied methodology. If changes in methodology are made, they should be highlighted.

Distinction between Facts and Opinions in Reports

Standard V(B) requires that opinion be separated from fact. Violations often occur when reports fail to separate the past from the future by not indicating that earnings estimates, changes in the outlook for dividends, or future market price information are *opinions* subject to future circumstances.

In the case of complex quantitative analyses, members and candidates must clearly separate fact from statistical conjecture and should identify the known limitations of an analysis. Members and candidates may violate Standard V(B) by failing to identify the limits of statistically developed projections because such omission leaves readers unaware of the limits of the published projections.

Members and candidates should explicitly discuss with clients and prospective clients the assumptions used in the investment models and processes to generate the analysis. Caution should be used in promoting the perceived accuracy of any model or process to clients because the ultimate output is merely an estimate of future results and not a certainty.

STANDARD V(B): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Because the selection of relevant factors is an analytical skill, determination of whether a member or candidate has used reasonable judgment in excluding and including information in research reports depends heavily on case-by-case review rather than a specific checklist.

Members and candidates should encourage their firms to have a rigorous methodology for reviewing research that is created for publication and dissemination to clients.

To assist in the after-the-fact review of a report, the member or candidate must maintain records indicating the nature of the research and should, if asked, be able to supply additional information to the client (or any user of the report) covering factors not included in the report.

STANDARD V(B): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Sufficient Disclosure of Investment System):

Sarah Williamson, director of marketing for Country Technicians, Inc., is convinced that she has found the perfect formula for increasing Country Technicians' income and diversifying its product base. Williamson plans to build on Country Technicians' reputation as a leading money manager by marketing an exclusive and expensive investment advice letter to high-net-worth individuals. One hitch in the plan is the complexity of Country Technicians' investment system—a combination of technical trading rules (based on historical price and volume fluctuations) and portfolio construction rules designed to minimize risk. To simplify the newsletter, she decides to include only each week's top five "buy" and "sell" recommendations and to leave out details of the valuation models and the portfolio structuring scheme.

> *Comment*: Williamson's plans for the newsletter violate Standard V(B). Williamson need not describe the investment system in detail in order to implement the advice effectively, but she must inform clients of Country Technicians' basic process and logic. Without understanding the basis for a recommendation, clients cannot possibly understand its limitations or its inherent risks.

Example 2 (Providing Opinions as Facts):

Richard Dox is a mining analyst for East Bank Securities. He has just finished his report on Boisy Bay Minerals. Included in his report is his own assessment of the geological extent of mineral reserves likely to be found on the company's land. Dox completed this calculation on the basis of the core samples from the company's latest drilling. According to Dox's calculations, the company has more than 500,000 ounces of gold on the property. Dox concludes his research report as follows: "Based on the fact that the company has 500,000 ounces of gold to be mined, I recommend a strong BUY."

> *Comment*: If Dox issues the report as written, he will violate Standard V(B). His calculation of the total gold reserves for the property based on the company's recent sample drilling is a quantitative opinion, not a fact. Opinion must be distinguished from fact in research reports.

Example 3 (Proper Description of a Security):

Olivia Thomas, an analyst at Government Brokers, Inc., which is a brokerage firm specializing in government bond trading, has produced a report that describes an investment strategy designed to benefit from an expected decline in US interest rates. The firm's derivative products group has designed a structured product that will allow the firm's clients to benefit from this strategy. Thomas's report describing the strategy indicates that high returns are possible if various scenarios for declining interest rates are assumed. Citing the proprietary nature of the structured product underlying the strategy, the report does not describe in detail how the firm is able to offer such returns or the related risks in the scenarios, nor does the report address the likely returns of the strategy if, contrary to expectations, interest rates rise.

> *Comment*: Thomas has violated Standard V(B) because her report fails to describe properly the basic characteristics of the actual and implied risks of the investment strategy, including how the structure was created and the degree to which leverage was embedded in the structure. The report should include a balanced discussion of how the strategy would perform in the case of rising as well as falling interest rates, preferably illustrating how the strategies might be expected to perform in the event of a reasonable variety of interest rate and credit risk–spread scenarios. If liquidity issues are relevant with regard to the valuation of either the derivatives or the underlying securities, provisions the firm has made to address those risks should also be disclosed.

Example 4 (Notification of Fund Mandate Change):

May & Associates is an aggressive growth manager that has represented itself since its inception as a specialist at investing in small-cap US stocks. One of May's selection criteria is a maximum capitalization of US$250 million for any given company. After a string of successful years of superior performance relative to its peers, May has expanded its client base significantly, to the point at which assets under management now exceed US$3 billion. For liquidity purposes, May's chief investment officer (CIO) decides to lift the maximum permissible market-cap ceiling to US$500 million and change the firm's sales and marketing literature accordingly to inform prospective clients and third-party consultants.

> *Comment*: Although May's CIO is correct about informing potentially interested parties as to the change in investment process, he must also notify May's existing clients. Among the latter group might be a number of clients who not only retained May as a small-cap manager but also retained mid-cap and large-cap specialists in a multiple-manager approach. Such clients could regard May's change of criteria as a style change that distorts their overall asset allocations.

Example 5 (Notification of Fund Mandate Change):

Rather than lifting the ceiling for its universe from US$250 million to US$500 million, May & Associates extends its small-cap universe to include a number of non-US companies.

> *Comment*: Standard V(B) requires that May's CIO advise May's clients of this change because the firm may have been retained by some clients specifically for its prowess at investing in US small-cap stocks. Other changes that require client notification are introducing derivatives to emulate a certain market sector or relaxing various other constraints, such as portfolio beta. In all such cases, members and candidates must disclose changes to all interested parties.

Example 6 (Notification of Changes to the Investment Process):

RJZ Capital Management is an active value-style equity manager that selects stocks by using a combination of four multifactor models. The firm has found favorable results when back testing the most recent 10 years of available market data in a new dividend discount model (DDM) designed by the firm. This model is based on projected inflation rates, earnings growth rates, and interest rates. The president of RJZ decides to replace its simple model that uses price to trailing 12-month earnings with the new DDM.

> *Comment*: Because the introduction of a new and different valuation model represents a material change in the investment process, RJZ's president must communicate the change to the firm's clients. RJZ is moving away from a model based on hard data toward a new model that is at least partly dependent on the firm's forecasting skills. Clients would likely view such a model as a significant change rather than a mere refinement of RJZ's process.

Example 7 (Notification of Changes to the Investment Process):

RJZ Capital Management loses the chief architect of its multifactor valuation system. Without informing its clients, the president of RJZ decides to redirect the firm's talents and resources toward developing a product for passive equity management—a product that will emulate the performance of a major market index.

> *Comment*: By failing to disclose to clients a substantial change to its investment process, the president of RJZ has violated Standard V(B).

Example 8 (Notification of Changes to the Investment Process):

At Fundamental Asset Management, Inc., the responsibility for selecting stocks for addition to the firm's "approved" list has just shifted from individual security analysts to a committee consisting of the research director and three senior portfolio managers. Eleanor Morales, a portfolio manager with Fundamental Asset Management, thinks this change is not important enough to communicate to her clients.

> *Comment*: Morales must disclose the process change to all her clients. Some of Fundamental's clients might be concerned about the morale and motivation among the firm's best research analysts after such a change. Moreover, clients might challenge the stock-picking track record of the portfolio managers and might even want to monitor the situation closely.

Example 9 (Sufficient Disclosure of Investment System):

Amanda Chinn is the investment director for Diversified Asset Management, which manages the endowment of a charitable organization. Because of recent staff departures, Diversified has decided to limit its direct investment focus to large-cap securities and supplement the needs for small-cap and mid-cap management by hiring outside fund managers. In describing the planned strategy change to the charity, Chinn's update letter states, "As investment director, I will directly oversee the investment team managing the endowment's large-capitalization allocation. I will coordinate the selection and ongoing review of external managers responsible for allocations to other classes." The letter also describes the reasons for the change and the characteristics external managers must have to be considered.

> *Comment*: Standard V(B) requires the disclosure of the investment process used to construct the portfolio of the fund. Changing the investment process from managing all classes of investments within the firm to the use of external managers is one example of information that needs to be communicated to clients. Chinn and her firm have embraced the principles of Standard V(B) by providing their client with relevant information. The charity can now make a reasonable decision about whether Diversified Asset Management remains the appropriate manager for its fund.

Example 10 (Notification of Changes to the Investment Process):

Michael Papis is the chief investment officer of his state's retirement fund. The fund has always used outside advisers for the real estate allocation, and this information is clearly presented in all fund communications. Thomas Nagle, a recognized sell-side research analyst and Papis's business school classmate, recently left the investment bank he worked for to start his own asset management firm, Accessible Real Estate. Nagle is trying to build his assets under management and contacts Papis about gaining some of the retirement fund's allocation. In the previous few years, the performance of the retirement fund's real estate investments was in line with the fund's benchmark but was not extraordinary. Papis decides to help out his old friend and also to seek better returns by moving the real estate allocation to Accessible. The only notice of the change in adviser appears in the next annual report in the listing of associated advisers.

> *Comment*: Papis has violated Standard V(B). He attempted to hide the nature of his decision to change external managers by making only a limited disclosure. The plan recipients and the fund's trustees need to be aware when changes are made to ensure that operational procedures are being followed.
>
> See also Standard IV(C)–Responsibilities of Supervisors, Standard V(A)–Diligence and Reasonable Basis, and Standard VI(A)–Disclosure of Conflicts.

Example 11 (Notification of Errors):

Jérôme Dupont works for the credit research group of XYZ Asset Management, where he is in charge of developing and updating credit risk models. In order to perform accurately, his models need to be regularly updated with the latest market data.

Dupont does not interact with or manage money for any of the firm's clients. He is in contact with the firm's US corporate bond fund manager, John Smith, who has only very superficial knowledge of the model and who from time to time asks very basic questions regarding the output recommendations. Smith does not consult Dupont with respect to finalizing his clients' investment strategies.

Dupont's recently assigned objective is to develop a new emerging market corporate credit risk model. The firm is planning to expand into emerging credit, and the development of such a model is a critical step in this process. Because Smith seems to follow the model's recommendations without much concern for its quality as he develops his clients' investment strategies, Dupont decides to focus his time on the development of the new emerging market model and neglects to update the US model.

After several months without regular updates, Dupont's diagnostic statistics start to show alarming signs with respect to the quality of the US credit model. Instead of conducting the long and complicated data update, Dupont introduces new codes into his model with some limited new data as a quick "fix." He thinks this change will address the issue without needing to complete the full data update, so he continues working on the new emerging market model.

Several months following the quick "fix," another set of diagnostic statistics reveals nonsensical results and Dupont realizes that his earlier change contained an error. He quickly corrects the error and alerts Smith. Smith realizes that some of the prior trades he performed were due to erroneous model results. Smith rebalances the portfolio to remove the securities purchased on the basis of the questionable results without reporting the issue to anyone else.

> *Comment*: Smith violated V(B) by not disclosing a material error in the investment process. Clients should have been informed about the error and the corrective actions the firm was undertaking on their behalf.
>
> See also Standard V(A)–Diligence and Reasonable Basis.

Example 12 (Notification of Risks and Limitations):

Quantitative analyst Yuri Yakovlev has developed an investment strategy that selects small-cap stocks on the basis of quantitative signals. Yakovlev's strategy typically identifies only a small number of stocks (10–20) that tend to be illiquid, but according to his backtests, the strategy generates significant risk-adjusted returns. The partners at Yakovlev's firm, QSC Capital, are impressed by these results. After a thorough examination of the strategy's risks, stress testing, historical back testing, and scenario analysis, QSC decides to seed the strategy with US$10 million of internal capital in order for Yakovlev to create a track record for the strategy.

After two years, the strategy has generated performance returns greater than the appropriate benchmark and the Sharpe ratio of the fund is close to 1.0. On the basis of these results, QSC decides to actively market the fund to large institutional investors. While creating the offering materials, Yakovlev informs the marketing team that the capacity of the strategy is limited. The extent of the limitation is difficult to ascertain with precision; it depends on market liquidity and other factors in his model that can evolve over time. Yakovlev indicates that given the current market conditions, investments in the fund beyond US$100 million of capital could become more difficult and negatively affect expected fund returns.

Alan Wellard, the manager of the marketing team, is a partner with 30 years of marketing experience and explains to Yakovlev that these are complex technical issues that will muddy the marketing message. According to Wellard, the offering material should focus solely on the great track record of the fund. Yakovlev does not object because the fund has only US$12 million of capital, very far from the US$100 million threshold.

> *Comment*: Yakovlev and Wellard have not appropriately disclosed a significant limitation associated with the investment product. Yakovlev believes this limitation, once reached, will materially affect the returns of the fund. Although the fund is currently far from the US$100 million mark, current and prospective investors must be made aware of this capacity issue. If significant limitations are complicated to grasp and clients do not have the technical background required to understand them, Yakovlev and Wellard should either educate the clients or ascertain whether the fund is suitable for each client.

Example 13 (Notification of Risks and Limitations):

Brickell Advisers offers investment advisory services mainly to South American clients. Julietta Ramon, a risk analyst at Brickell, describes to clients how the firm uses value at risk (VaR) analysis to track the risk of its strategies. Ramon assures clients that calculating a VaR at a 99% confidence level, using a 20-day holding period, and applying a methodology based on an *ex ante* Monte Carlo simulation is extremely effective. The firm has never had losses greater than those predicted by this VaR analysis.

> *Comment*: Ramon has not sufficiently communicated the risks associated with the investment process to satisfy the requirements of Standard V(B). The losses predicted by a VaR analysis depend greatly on the inputs used in the model. The size and probability of losses can differ significantly from what an individual model predicts. Ramon must disclose how the inputs were selected and the potential limitations and risks associated with the investment strategy.

Example 14 (Notification of Risks and Limitations):

Lily Smith attended an industry conference and noticed that John Baker, an investment manager with Baker Associates, attracted a great deal of attention from the conference participants. On the basis of her knowledge of Baker's reputation and the interest he received at the conference, Smith recommends adding Baker Associates to the approved manager platform. Her recommendation to the approval committee included the statement "John Baker is well respected in the industry, and his insights are consistently sought after by investors. Our clients are sure to benefit from investing with Baker Associates."

Comment: Smith is not appropriately separating facts from opinions in her recommendation to include the manager within the platform. Her actions conflict with the requirements of Standard V(B). Smith is relying on her opinions about Baker's reputation and the fact that many attendees were talking with him at the conference. Smith should also review the requirements of Standard V(A) regarding reasonable basis to determine the level of review necessary to recommend Baker Associates.

STANDARD V(C): INVESTMENT ANALYSIS, RECOMMENDATIONS, AND ACTIONS - RECORD RETENTION

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Members and Candidates must develop and maintain appropriate records to support their investment analyses, recommendations, actions, and other investment-related communications with clients and prospective clients.

Guidance

Highlights:

- *New Media Records*
- *Records Are Property of the Firm*
- *Local Requirements*

Members and candidates must retain records that substantiate the scope of their research and reasons for their actions or conclusions. The retention requirement applies to decisions to buy or sell a security as well as reviews undertaken that do not lead to a change in position. Which records are required to support recommendations or investment actions depends on the role of the member or candidate in the investment decision-making process. Records may be maintained either in hard copy or electronic form.

Some examples of supporting documentation that assists the member or candidate in meeting the requirements for retention are as follows:

- personal notes from meetings with the covered company,
- press releases or presentations issued by the covered company,
- computer-based model outputs and analyses,
- computer-based model input parameters,
- risk analyses of securities' impacts on a portfolio,
- selection criteria for external advisers,
- notes from clients from meetings to review investment policy statements, and
- outside research reports.

New Media Records

The increased use of new and evolving technological formats (e.g., social media) for gathering and sharing information creates new challenges in maintaining the appropriate records and files. The nature or format of the information does not remove a member's or candidate's responsibility to maintain a record of information used in his or her analysis or communicated to clients.

Members and candidates should understand that although employers and local regulators are developing digital media retention policies, these policies may lag behind the advent of new communication channels. Such lag places greater responsibility on the individual for ensuring that all relevant information is retained. Examples of non-print media formats that should be retained include, but are not limited to,

- e-mails,
- text messages,
- blog posts, and
- Twitter posts.

Records Are Property of the Firm

As a general matter, records created as part of a member's or candidate's professional activity on behalf of his or her employer are the property of the firm. When a member or candidate leaves a firm to seek other employment, the member or candidate cannot take the property of the firm, including original forms or copies of supporting records of the member's or candidate's work, to the new employer without the express consent of the previous employer. The member or candidate cannot use historical recommendations or research reports created at the previous firm because the supporting documentation is unavailable. For future use, the member or candidate must re-create the supporting records at the new firm with information gathered through public sources or directly from the covered company and not from memory or sources obtained at the previous employer.

Local Requirements

Local regulators often impose requirements on members, candidates, and their firms related to record retention that must be followed. Firms may also implement policies detailing the applicable time frame for retaining research and client communication records. Fulfilling such regulatory and firm requirements satisfies the requirements of Standard V(C). In the absence of regulatory guidance or firm policies, CFA Institute recommends maintaining records for at least seven years.

STANDARD V(C): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

The responsibility to maintain records that support investment action generally falls with the firm rather than individuals. Members and candidates must, however, archive research notes and other documents, either electronically or in hard copy, that support their current investment-related communications. Doing so will assist their firms in complying with requirements for preservation of internal or external records.

STANDARD V(C): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Record Retention and IPS Objectives and Recommendations):

One of Nikolas Lindstrom's clients is upset by the negative investment returns of his equity portfolio. The investment policy statement for the client requires that the portfolio manager follow a benchmark-oriented approach. The benchmark for the client includes a 35% investment allocation in the technology sector. The client acknowledges that this allocation was appropriate, but over the past three years, technology stocks have suffered severe losses. The client complains to the investment manager for allocating so much money to this sector.

> *Comment*: For Lindstrom, having appropriate records is important to show that over the past three years, the portion of technology stocks in the benchmark index was 35%, as called for in the IPS. Lindstrom should also have the client's IPS stating that the benchmark was appropriate for the client's investment objectives. He should also have records indicating that the investment has been explained appropriately to the client and that the IPS was updated on a regular basis. Taking these actions, Lindstrom would be in compliance with Standard V(C).

Example 2 (Record Retention and Research Process):

Malcolm Young is a research analyst who writes numerous reports rating companies in the luxury retail industry. His reports are based on a variety of sources, including interviews with company managers, manufacturers, and economists; on-site company visits; customer surveys; and secondary research from analysts covering related industries.

> *Comment*: Young must carefully document and keep copies of all the information that goes into his reports, including the secondary or third-party research of other analysts. Failure to maintain such files would violate Standard V(C).

Example 3 (Records as Firm, Not Employee, Property):

Martin Blank develops an analytical model while he is employed by Green Partners Investment Management, LLP (GPIM). While at the firm, he systematically documents the assumptions that make up the model as well as his reasoning behind the assumptions. As a result of the success of his model, Blank is hired to be the head of the research department of one of GPIM's competitors. Blank takes copies of the records supporting his model to his new firm.

> *Comment*: The records created by Blank supporting the research model he developed at GPIM are the records of GPIM. Taking the documents with him to his new employer without GPIM's permission violates Standard V(C). To use the model in the future, Blank must re-create the records supporting his model at the new firm.

STANDARD VI(A): CONFLICTS OF INTEREST - DISCLOSURE OF CONFLICTS

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Standard VI(A) Disclosure of Conflicts

Members and Candidates must make full and fair disclosure of all matters that could reasonably be expected to impair their independence and objectivity or interfere with respective duties to their clients, prospective clients, and employer. Members and Candidates must ensure that such disclosures are prominent, are delivered in plain language, and communicate the relevant information effectively.

Guidance

Highlights:

- *Disclosure of Conflicts to Employers*
- *Disclosure to Clients*
- *Cross-Departmental Conflicts*
- *Conflicts with Stock Ownership*
- *Conflicts as a Director*

Best practice is to avoid actual conflicts or the appearance of conflicts of interest when possible. Conflicts of interest often arise in the investment profession. Conflicts can occur between the interests of clients, the interests of employers, and the member's or candidate's own personal interests. Common sources for conflict are compensation structures, especially incentive and bonus structures that provide immediate returns for members and candidates with little or no consideration of long-term value creation.

Identifying and managing these conflicts is a critical part of working in the investment industry and can take many forms. When conflicts cannot be reasonably avoided, clear and complete disclosure of their existence is necessary.

Standard VI(A) protects investors and employers by requiring members and candidates to fully disclose to clients, potential clients, and employers all actual and potential conflicts of interest. Once a member or candidate has made full disclosure, the member's or candidate's employer, clients, and prospective clients will have the information needed to evaluate the objectivity of the investment advice or action taken on their behalf.

To be effective, disclosures must be prominent and must be made in plain language and in a manner designed to effectively communicate the information. Members and candidates have the responsibility of determining how often, in what manner, and in what particular circumstances the disclosure of conflicts must be made. Best practices dictate updating disclosures when the nature of a conflict of interest changes materially—for example, if the nature of a conflict of interest worsens through the introduction of bonuses based on each quarter's profits as to opposed annual profits. In making and updating disclosures of conflicts of interest, members and candidates should err on the side of caution to ensure that conflicts are effectively communicated.

Disclosure of Conflicts to Employers

Disclosure of conflicts to employers may be appropriate in many instances. When reporting conflicts of interest to employers, members and candidates must give their employers enough information to assess the impact of the conflict. By complying with employer guidelines, members and candidates allow their employers to avoid potentially embarrassing and costly ethical or regulatory violations.

Reportable situations include conflicts that would interfere with rendering unbiased investment advice and conflicts that would cause a member or candidate to act not in the employer's best interest. The same circumstances that generate conflicts to be reported to clients and prospective clients also would dictate reporting to employers. Ownership of stocks analyzed or recommended, participation on outside boards, and financial or other pressures that could influence a decision are to be promptly reported to the employer so that their impact can be assessed and a decision on how to resolve the conflict can be made.

The mere appearance of a conflict of interest may create problems for members, candidates, and their employers. Therefore, many of the conflicts previously mentioned could be explicitly prohibited by an employer. For example, many employers restrict personal trading, outside board membership, and related activities to prevent situations that might not normally be considered problematic from a conflict-of-interest point of view but that could give the appearance of a conflict of interest. Members and candidates must comply with these restrictions. Members and candidates must take reasonable steps to avoid conflicts and, if they occur inadvertently, must report them promptly so that the employer and the member or candidate can resolve them as quickly and effectively as possible.

Standard VI(A) also deals with a member's or candidate's conflicts of interest that might be detrimental to the employer's business. Any potential conflict situation that could prevent clear judgment about or full commitment to the execution of a member's or candidate's duties to the employer should be reported to the member's or candidate's employer and promptly resolved.

Disclosure to Clients

Members and candidates must maintain their objectivity when rendering investment advice or taking investment action. Investment advice or actions may be perceived to be tainted in numerous situations. Can a member or candidate remain objective if, on behalf of the firm, the member or candidate obtains or assists in obtaining fees for services? Can a member or candidate give objective advice if he or she owns stock in the company that is the subject of an investment recommendation or if the member or candidate has a close personal relationship with the company managers? Requiring members and candidates to disclose all matters that reasonably could be expected to impair the member's or candidate's objectivity allows clients and prospective clients to judge motives and possible biases for themselves.

Often in the investment industry, a conflict, or the perception of a conflict, cannot be avoided. The most obvious conflicts of interest, which should always be disclosed, are relationships between an issuer and the member, the candidate, or his or her firm (such as a directorship or consultancy by a member; investment banking, underwriting, and financial relationships; broker/dealer market-making activities; and material beneficial ownership of stock). For the purposes of Standard VI(A), members and candidates beneficially own securities or other investments if they have a direct or indirect pecuniary interest in the securities, have the power to vote or direct the voting of the shares of the securities or investments, or have the power to dispose or direct the disposition of the security or investment.

A member or candidate must take reasonable steps to determine whether a conflict of interest exists and disclose to clients any known conflicts of the member's or candidate's firm. Disclosure of broker/dealer market-making activities alerts clients that a purchase or sale might be made from or to the firm's principal account and that the firm has a special interest in the price of the stock.

Additionally, disclosures should be made to clients regarding fee arrangements, subadvisory agreements, or other situations involving nonstandard fee structures. Equally important is the disclosure of arrangements in which the firm benefits directly from investment recommendations. An obvious conflict of interest is the rebate of a portion of the service fee some classes of mutual funds charge to investors. Members and candidates should ensure that their firms disclose such relationships so clients can fully understand the costs of their investments and the benefits received by their investment manager's employer.

Cross-Departmental Conflicts

Other circumstances can give rise to actual or potential conflicts of interest. For instance, a sell-side analyst working for a broker/dealer may be encouraged, not only by members of her or his own firm but by corporate issuers themselves, to write research reports about particular companies. The buy-side analyst is likely to be faced with similar conflicts as banks exercise their underwriting and security-dealing powers. The marketing division may ask an analyst to recommend the stock of a certain company in order to obtain business from that company.

The potential for conflicts of interest also exists with broker-sponsored limited partnerships formed to invest venture capital. Increasingly, members and candidates are expected not only to follow issues from these partnerships once they are offered to the public but also to promote the issues in the secondary market after public offerings. Members, candidates, and their firms should attempt to resolve situations presenting potential conflicts of interest or disclose them in accordance with the principles set forth in Standard VI(A).

Conflicts with Stock Ownership

The most prevalent conflict requiring disclosure under Standard VI(A) is a member's or candidate's ownership of stock in companies that he or she recommends to clients or that clients hold. Clearly, the easiest method for preventing a conflict is to prohibit members and candidates from owning any such securities, but this approach is overly burdensome and discriminates against members and candidates.

Therefore, sell-side members and candidates should disclose any materially beneficial ownership interest in a security or other investment that the member or candidate is recommending. Buy-side members and candidates should disclose their procedures for reporting requirements for personal transactions. Conflicts arising from personal investing are discussed more fully in the guidance for Standard VI(B).

Conflicts as a Director

Service as a director poses three basic conflicts of interest. First, a conflict may exist between the duties owed to clients and the duties owed to shareholders of the company. Second, investment personnel who serve as directors may receive the securities or options to purchase securities of the company as compensation for serving on the board, which could raise questions about trading actions that might increase the value of those securities. Third, board service creates the opportunity to receive material nonpublic information involving the company. Even though the information is confidential, the perception could be that information not available to the public is being communicated to a director's firm—whether a broker, investment adviser, or other type of organization. When members or candidates providing investment services also serve as directors, they should be isolated from those making investment decisions by the use of firewalls or similar restrictions.

STANDARD VI(A): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Members or candidates should disclose special compensation arrangements with the employer that might conflict with client interests, such as bonuses based on short-term performance criteria, commissions, incentive fees, performance fees, and referral fees. If the member's or candidate's firm does not permit such disclosure, the member or candidate should document the request and may consider dissociating from the activity.

Members' and candidates' firms are encouraged to include information on compensation packages in firms' promotional literature. If a member or candidate manages a portfolio for which the fee is based on capital gains or capital appreciation (a performance fee), this information should be disclosed to clients. If a member, a candidate, or a member's or candidate's firm has outstanding agent options to buy stock as part of the compensation package for corporate financing activities, the amount and expiration date of these options should be disclosed as a footnote to any research report published by the member's or candidate's firm.

STANDARD VI(A): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Conflict of Interest and Business Relationships):

Hunter Weiss is a research analyst with Farmington Company, a broker and investment banking firm. Farmington's merger and acquisition department has represented Vimco, a conglomerate, in all of Vimco's acquisitions for 20 years. From time to time, Farmington officers sit on the boards of directors of various Vimco subsidiaries. Weiss is writing a research report on Vimco.

> *Comment*: Weiss must disclose in his research report Farmington's special relationship with Vimco. Broker/dealer management of and participation in public offerings must be disclosed in research reports. Because the position of underwriter to a company entails a special past and potential

> future relationship with a company that is the subject of investment advice, it threatens the independence and objectivity of the report writer and must be disclosed.

Example 2 (Conflict of Interest and Business Stock Ownership):

The investment management firm of Dover & Roe sells a 25% interest in its partnership to a multinational bank holding company, First of New York. Immediately after the sale, Margaret Hobbs, president of Dover & Roe, changes her recommendation for First of New York's common stock from "sell" to "buy" and adds First of New York's commercial paper to Dover & Roe's approved list for purchase.

> *Comment*: Hobbs must disclose the new relationship with First of New York to all Dover & Roe clients. This relationship must also be disclosed to clients by the firm's portfolio managers when they make specific investment recommendations or take investment actions with respect to First of New York's securities.

Example 3 (Conflict of Interest and Personal Stock Ownership):

Carl Fargmon, a research analyst who follows firms producing office equipment, has been recommending purchase of Kincaid Printing because of its innovative new line of copiers. After his initial report on the company, Fargmon's wife inherits from a distant relative US$3 million of Kincaid stock. He has been asked to write a follow-up report on Kincaid.

> *Comment*: Fargmon must disclose his wife's ownership of the Kincaid stock to his employer and in his follow-up report. Best practice would be to avoid the conflict by asking his employer to assign another analyst to draft the follow-up report.

Example 4 (Conflict of Interest and Personal Stock Ownership):

Betty Roberts is speculating in penny stocks for her own account and purchases 100,000 shares of Drew Mining, Inc., for US$0.30 a share. She intends to sell these shares at the sign of any substantial upward price movement of the stock. A week later, her employer asks her to write a report on penny stocks in the mining industry to be published in two weeks. Even without owning the Drew stock, Roberts would recommend it in her report as a "buy." A surge in the price of the stock to the US$2 range is likely to result once the report is issued.

> *Comment*: Although this holding may not be material, Roberts must disclose it in the report and to her employer before writing the report because the gain for her will be substantial if the market responds strongly to her recommendation. The fact that she has only recently purchased the stock adds to the appearance that she is not entirely objective.

Example 5 (Conflict of Interest and Compensation Arrangements):

Samantha Snead, a portfolio manager for Thomas Investment Counsel, Inc., specializes in managing public retirement funds and defined benefit pension plan accounts, all of which have long-term investment objectives. A year ago, Snead's employer, in an attempt to motivate and retain key investment professionals, introduced a bonus compensation system that rewards portfolio managers on the basis of quarterly performance relative to their peers and to certain benchmark indexes. In an attempt to improve the short-term performance of her accounts, Snead changes her investment strategy and purchases several high-beta stocks for client portfolios. These purchases are seemingly contrary to the clients' investment policy statements. Following their purchase, an officer of Griffin Corporation, one of Snead's pension fund clients, asks why Griffin Corporation's portfolio seems to be dominated by high-beta stocks of companies that often appear among the most actively traded issues. No change in objective or strategy has been recommended by Snead during the year.

> *Comment*: Snead has violated Standard VI(A) by failing to inform her clients of the changes in her compensation arrangement with her employer, which created a conflict of interest between her compensation and her clients' IPSs. Firms may pay employees on the basis of performance, but pressure by Thomas Investment Counsel to achieve short-term performance goals is in basic conflict with the objectives of Snead's accounts.
>
> See also Standard III(C)–Suitability.

Example 6 (Conflict of Interest, Options, and Compensation Arrangements):

Wayland Securities works with small companies doing IPOs or secondary offerings. Typically, these deals are in the US$10 million to US$50 million range, and as a result, the corporate finance fees are quite small. To compensate for the small fees, Wayland Securities usually takes "agent options"—that is, rights (exercisable within a two-year time frame) to acquire up to an additional 10% of the current offering. Following an IPO performed by Wayland for Falk Resources, Ltd., Darcy Hunter, the head of corporate finance at Wayland, is concerned about receiving value for her Falk Resources options. The options are due to expire in one month, and the stock is not doing well. She contacts John Fitzpatrick in the research department of Wayland Securities, reminds him that he is eligible for 30% of these options, and indicates that now would be a good time to give some additional coverage to Falk Resources. Fitzpatrick agrees and immediately issues a favorable report.

> *Comment*: For Fitzpatrick to avoid being in violation of Standard VI(A), he must indicate in the report the volume and expiration date of agent options outstanding. Furthermore, because he is personally eligible for some of the options, Fitzpatrick must disclose the extent of this compensation. He also must be careful to not violate his duty of independence and objectivity under Standard I(B).

Example 7 (Conflict of Interest and Compensation Arrangements):

Gary Carter is a representative with Bengal International, a registered broker/dealer. Carter is approached by a stock promoter for Badger Company, who offers to pay Carter additional compensation for sales of Badger Company's stock to Carter's clients. Carter accepts the stock promoter's offer but does not disclose the arrangements to his clients or to his employer. Carter sells shares of the stock to his clients.

> *Comment*: Carter has violated Standard VI(A) by failing to disclose to clients that he is receiving additional compensation for recommending and selling Badger stock. Because he did not disclose the arrangement with Badger to his clients, the clients were unable to evaluate whether Carter's recommendations to buy Badger were affected by this arrangement. Carter's conduct also violated Standard VI(A) by failing to disclose to his employer monetary compensation received in addition to the compensation and benefits conferred by his employer. Carter was required by Standard VI(A) to disclose the arrangement with Badger to his employer so that his employer could evaluate whether the arrangement affected Carter's objectivity and loyalty.

Example 8 (Conflict of Interest and Directorship):

Carol Corky, a senior portfolio manager for Universal Management, recently became involved as a trustee with the Chelsea Foundation, a large not-for-profit foundation in her hometown. Universal is a small money manager (with assets under management of approximately US$100 million) that caters to individual investors. Chelsea has assets in excess of US$2 billion. Corky does not believe informing Universal of her involvement with Chelsea is necessary.

> *Comment*: By failing to inform Universal of her involvement with Chelsea, Corky violated Standard VI(A). Given the large size of the endowment at Chelsea, Corky's new role as a trustee can reasonably be expected to be time consuming, to the possible detriment of Corky's portfolio responsibilities with Universal. Also, as a trustee, Corky may become involved in the investment decisions at Chelsea. Therefore, Standard VI(A) obligates Corky to discuss becoming a trustee at Chelsea with her compliance officer or supervisor at Universal before accepting the position, and she should have disclosed the degree to which she would be involved in investment decisions at Chelsea.

Example 9 (Conflict of Interest and Personal Trading):

Bruce Smith covers eastern European equities for Marlborough Investments, an investment management firm with a strong presence in emerging markets. While on a business trip to Russia, Smith learns that investing in Russian equities directly is difficult but that equity-linked notes that replicate the performance of underlying Russian equities can be purchased from a New York–based investment bank. Believing that his firm would not be interested in such a security, Smith purchases a note linked to a Russian telecommunications company for his own account without informing Marlborough. A month later, Smith decides that the firm should consider investing in Russian equities by way of the equity-linked notes. He prepares a write-up on the market that concludes with a recommendation to purchase several of the notes. One note he recommends is linked to the same Russian telecom company that Smith holds in his personal account.

> *Comment*: Smith has violated Standard VI(A) by failing to disclose his purchase and ownership of the note linked to the Russian telecom company. Smith is required by the standard to disclose the investment opportunity to his employer and look to his company's policies on personal trading to determine whether it was proper for him to purchase the note for his own account. By purchasing the note, Smith may or may not have impaired his ability to make an unbiased and objective assessment of the appropriateness of the derivative instrument for his firm, but Smith's failure to disclose the purchase to his employer impaired his employer's ability to decide whether his ownership of the security is a conflict of interest that might affect Smith's future recommendations. Then, when he recommended the particular telecom notes to his firm, Smith compounded his problems by not disclosing that he owned the notes in his personal account—a clear conflict of interest.

Example 10 (Conflict of Interest and Requested Favors):

Michael Papis is the chief investment officer of his state's retirement fund. The fund has always used outside advisers for the real estate allocation, and this information is clearly presented in all fund communications. Thomas Nagle, a recognized sell-side research analyst and Papis's business school classmate, recently left the investment bank he worked for to start his own asset management firm, Accessible Real Estate. Nagle is trying to build his assets under management and contacts Papis about gaining some of the retirement fund's allocation. In the previous few years, the performance of the retirement fund's real estate investments was in line with the fund's benchmark but was not extraordinary. Papis decides to help out his old friend and also to seek better returns by moving the real estate allocation to Accessible. The only notice of the change in adviser appears in the next annual report in the listing of associated advisers.

> *Comment*: Papis has violated Standard VI(A) by not disclosing to his employer his personal relationship with Nagle. Disclosure of his past history with Nagle would allow his firm to determine whether the conflict may have impaired Papis's independence in deciding to change managers.
>
> See also Standard IV(C)–Responsibilities of Supervisors, Standard V(A)–Diligence and Reasonable Basis, and Standard V(B)–Communication with Clients and Prospective Clients.

Example 11 (Conflict of Interest and Business Relationships):

Bob Wade, trust manager for Central Midas Bank, was approached by Western Funds about promoting its family of funds, with special interest in the service-fee class. To entice Central to promote this class, Western Funds offered to pay the bank a service fee of 0.25%. Without disclosing the fee being offered to the bank, Wade asked one of the investment managers to review the Western Funds family of funds to determine whether they were suitable for clients of Central. The manager completed the normal due diligence review and determined that the funds were fairly valued in the market with fee structures on a par with their competitors. Wade decided to accept Western's offer and instructed the team of portfolio managers to exclusively promote these funds and the service-fee class to clients seeking to invest new funds or transfer from their current investments. So as to not influence the investment managers, Wade did not disclose the fee offer and allowed that income to flow directly to the bank.

> *Comment*: Wade is violating Standard VI(A) by not disclosing the portion of the service fee being paid to Central. Although the investment managers may not be influenced by the fee, neither they nor the client have the proper information about Wade's decision to exclusively market this fund family and class of investments. Central may come to rely on the new fee as a component of the firm's profitability and may be unwilling to offer other products in the future that could affect the fees received.
>
> See also Standard I(B)–Independence and Objectivity.

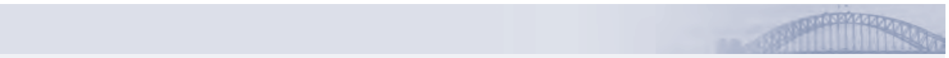
Example 12 (Disclosure of Conflicts to Employers):

Yehudit Dagan is a portfolio manager for Risk Management Bank (RMB), whose clients include retirement plans and corporations. RMB provides a defined contribution retirement plan for its employees that offers 20 large diversified mutual fund investment options, including a mutual fund managed by Dagan's RMB colleagues. After being employed for six months, Dagan became eligible to participate in the retirement plan, and she intends to allocate her retirement plan assets in six of the investment options, including the fund managed by her RMB colleagues. Dagan is concerned that joining the plan will lead to a potentially significant amount of paperwork for her (e.g., disclosure of her retirement account holdings and needing preclearance for her transactions), especially with her investing in the in-house fund.

> *Comment*: Standard VI(A) would not require Dagan to disclose her personal or retirement investments in large diversified mutual funds, unless specifically required by her employer. For practical reasons, the standard does not require Dagan to gain preclearance for ongoing payroll deduction contributions to retirement plan account investment options.
>
> Dagan should ensure that her firm does not have a specific policy regarding investment—whether personal or in the retirement account—for funds managed by the company's employees. These mutual funds may be subject to the company's disclosure, preclearance, and trading restriction procedures to identify possible conflicts prior to the execution of trades.

STANDARD VI(B): CONFLICTS OF INTEREST - PRIORITY OF TRANSACTIONS

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;



Investment transactions for clients and employers must have priority over investment transactions in which a Member or Candidate is the beneficial owner.

Guidance

Highlights:

- *Avoiding Potential Conflicts*
- *Personal Trading Secondary to Trading for Clients*
- *Standards for Nonpublic Information*
- *Impact on All Accounts with Beneficial Ownership*

Standard VI(B) reinforces the responsibility of members and candidates to give the interests of their clients and employers priority over their personal financial interests. This standard is designed to prevent any potential conflict of interest or the appearance of a conflict of interest with respect to personal transactions. Client interests have priority. Client transactions must take precedence over transactions made on behalf of the member's or candidate's firm or personal transactions.

Avoiding Potential Conflicts

Conflicts between the client's interest and an investment professional's personal interest may occur. Although conflicts of interest exist, nothing is inherently unethical about individual managers, advisers, or mutual fund employees making money from personal investments as long as (1) the client is not disadvantaged by the trade, (2) the investment professional does not benefit personally from trades undertaken for clients, and (3) the investment professional complies with applicable regulatory requirements.

Some situations occur where a member or candidate may need to enter a personal transaction that runs counter to current recommendations or what the portfolio manager is doing for client portfolios. For example, a member or candidate may be required at some point to sell an asset to make a college tuition payment or a down payment on a home, to meet a margin call, or so on. The sale may be contrary to the long-term advice the member or candidate is currently providing to clients. In these situations, the same three criteria given in the preceding paragraph should be applied in the transaction so as to not violate Standard VI(B).

Personal Trading Secondary to Trading for Clients

Standard VI(B) states that transactions for clients and employers must have priority over transactions in securities or other investments for which a member or candidate is the beneficial owner. The objective of the standard is to prevent personal transactions from adversely affecting the interests of clients or employers. A member or candidate having the same investment positions or being co-invested with clients does not always create a conflict. Some clients in certain investment situations require members or candidates to have aligned interests. Personal investment positions or transactions of members or candidates or their firm should never, however, adversely affect client investments.

Standards for Nonpublic Information

Standard VI(B) covers the activities of members and candidates who have knowledge of pending transactions that may be made on behalf of their clients or employers, who have access to nonpublic information during the normal preparation of research recommendations, or who take investment actions. Members and candidates are prohibited from conveying nonpublic information to any person whose relationship to the member or candidate makes the member or candidate a beneficial owner of the person's securities. Members and candidates must not convey this information to any other person if the nonpublic information can be deemed material.

Impact on All Accounts with Beneficial Ownership

Members or candidates may undertake transactions in accounts for which they are a beneficial owner only after their clients and employers have had adequate opportunity to act on a recommendation. Personal transactions include those made for the member's or candidate's own account, for family (including spouse, children, and other immediate family members) accounts, and for accounts in which the member or candidate has a direct or indirect pecuniary interest, such as a trust or retirement account. Family accounts that are client accounts should be treated like any other firm account and should neither be given special treatment nor be disadvantaged because of the family relationship. If a member or candidate has a beneficial ownership in the account, however, the member or candidate may be subject to preclearance or reporting requirements of the employer or applicable law.

STANDARD VI(B): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Policies and procedures designed to prevent potential conflicts of interest, and even the appearance of a conflict of interest, with respect to personal transactions are critical to establishing investor confidence in the securities industry. Therefore, members and candidates should urge their firms to establish such policies and procedures. Because investment firms vary greatly in assets under management, types of clients, number of employees, and so on, each firm should have policies regarding personal investing that are best suited to the firm. Members and candidates should then prominently disclose these policies to clients and prospective clients.

The specific provisions of each firm's standards will vary, but all firms should adopt certain basic procedures to address the conflict areas created by personal investing. These procedures include the following:

- *Limited participation in equity IPOs*: Some eagerly awaited IPOs rise significantly in value shortly after the issue is brought to market. Because the new issue may be highly attractive and sought after, the opportunity to participate in the IPO may be limited. Therefore, purchases of IPOs by investment personnel create conflicts of interest in two principal ways. First, participation in an IPO may have the appearance of taking away an attractive investment opportunity from clients for personal gain—a clear breach of the duty of loyalty to clients. Second, personal purchases in IPOs may have the appearance that the investment opportunity is being bestowed as an incentive to make future investment decisions for the benefit of the party providing the opportunity. Members and candidates can avoid these conflicts or appearances of conflicts of interest by not participating in IPOs.

  Reliable and systematic review procedures should be established to ensure that conflicts relating to IPOs are identified and appropriately dealt with by supervisors. Members and candidates should preclear their participation in IPOs, even in situations without any conflict of interest between a member's or candidate's participation in an IPO and the client's interests. Members and candidates should not benefit from the position that their clients occupy in the marketplace—through preferred trading, the allocation of limited offerings, or oversubscription.

- *Restrictions on private placements*: Strict limits should be placed on investment personnel acquiring securities in private placements, and appropriate supervisory and review procedures should be established to prevent noncompliance.

  Firms do not routinely use private placements for clients (e.g., venture capital deals) because of the high risk associated with them. Conflicts related to private placements are more significant to members and candidates who manage large pools of assets or act as plan sponsors because these managers may be offered special opportunities, such as private placements, as a reward or an enticement for continuing to do business with a particular broker.

  Participation in private placements raises conflict-of-interest issues that are similar to issues surrounding IPOs. Investment personnel should not be involved in transactions, including (but not limited to) private placements, that could be perceived as favors or gifts that seem designed to influence future judgment or to reward past business deals.

  Whether the venture eventually proves to be good or bad, managers have an immediate conflict concerning private placement opportunities. If and when the investments go public, participants in private placements have an incentive to recommend the investments to clients regardless of the suitability of the investments for their clients. Doing so increases the value of the participants' personal portfolios.

- *Establish blackout/restricted periods*: Investment personnel involved in the investment decision-making process should establish blackout periods prior to trades for clients so that managers cannot take advantage of their knowledge of client activity by "front-running" client trades (trading for one's personal account before trading for client accounts).

  Individual firms must decide who within the firm should be required to comply with the trading restrictions. At a minimum, all individuals who are involved in the investment decision-making process should be subject to the same restricted period. Each firm must determine specific requirements related to blackout and restricted periods that are most relevant to the firm while ensuring that the procedures are governed by the guiding principles set forth in the Code and Standards. Size of firm and type of securities purchased are relevant factors. For example, in a large firm, a blackout requirement is, in effect, a total trading ban because the firm is continually trading in most securities. In a small firm, the blackout period is more likely to prevent the investment manager from front-running.

- *Reporting requirements*: Supervisors should establish reporting procedures for investment personnel, including disclosure of personal holdings/beneficial ownerships, confirmations of trades to the firm and the employee, and preclearance procedures. Once trading restrictions are in place, they must be enforced. The best method for monitoring and enforcing procedures to eliminate conflicts of interest in personal trading is through reporting requirements, including the following:

  - **Disclosure of holdings in which the employee has a beneficial interest**. Disclosure by investment personnel to the firm should be made upon commencement of the employment relationship and at least annually thereafter. To address privacy considerations, disclosure of personal holdings should be handled in a confidential manner by the firm.

  - **Providing duplicate confirmations of transactions**. Investment personnel should be required to direct their brokers to supply to firms duplicate copies or confirmations of all their personal securities transactions and copies of periodic statements for all securities accounts. The duplicate confirmation requirement has two purposes: (1) The requirement sends a message that there is independent verification, which reduces the likelihood of unethical behavior, and (2) it enables verification of the accounting of the flow of personal investments that cannot be determined from merely looking at holdings.

- **Preclearance procedures**. Investment personnel should examine all planned personal trades to identify possible conflicts prior to the execution of the trades. Preclearance procedures are designed to identify possible conflicts before a problem arises.

- *Disclosure of policies*: Members and candidates should fully disclose to investors their firm's policies regarding personal investing. The information about employees' personal investment activities and policies will foster an atmosphere of full and complete disclosure and calm the public's legitimate concerns about the conflicts of interest posed by investment personnel's personal trading. The disclosure must provide helpful information to investors; it should not be simply boilerplate language, such as "investment personnel are subject to policies and procedures regarding their personal trading."

STANDARD VI(B): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Personal Trading):

Research analyst Marlon Long does not recommend purchase of a common stock for his employer's account because he wants to purchase the stock personally and does not want to wait until the recommendation is approved and the stock is purchased by his employer.

> *Comment*: Long has violated Standard VI(B) by taking advantage of his knowledge of the stock's value before allowing his employer to benefit from that information.

Example 2 (Trading for Family Member Account):

Carol Baker, the portfolio manager of an aggressive growth mutual fund, maintains an account in her husband's name at several brokerage firms with which the fund and a number of Baker's other individual clients do a substantial amount of business. Whenever a hot issue becomes available, she instructs the brokers to buy it for her husband's account. Because such issues normally are scarce, Baker often acquires shares in hot issues but her clients are not able to participate in them.

> *Comment*: To avoid violating Standard VI(B), Baker must acquire shares for her mutual fund first and acquire them for her husband's account only after doing so, even though she might miss out on participating in new

> issues via her husband's account. She also must disclose the trading for her husband's account to her employer because this activity creates a conflict between her personal interests and her employer's interests.

Example 3 (Family Accounts as Equals):

Erin Toffler, a portfolio manager at Esposito Investments, manages the retirement account established with the firm by her parents. Whenever IPOs become available, she first allocates shares to all her other clients for whom the investment is appropriate; only then does she place any remaining portion in her parents' account, if the issue is appropriate for them. She has adopted this procedure so that no one can accuse her of favoring her parents.

> *Comment*: Toffler has violated Standard VI(B) by breaching her duty to her parents by treating them differently from her other accounts simply because of the family relationship. As fee-paying clients of Esposito Investments, Toffler's parents are entitled to the same treatment as any other client of the firm. If Toffler has beneficial ownership in the account, however, and Esposito Investments has preclearance and reporting requirements for personal transactions, she may have to preclear the trades and report the transactions to Esposito.

Example 4 (Personal Trading and Disclosure):

Gary Michaels is an entry-level employee who holds a low-paying job serving both the research department and the investment management department of an active investment management firm. He purchases a sports car and begins to wear expensive clothes after only a year of employment with the firm. The director of the investment management department, who has responsibility for monitoring the personal stock transactions of all employees, investigates and discovers that Michaels has made substantial investment gains by purchasing stocks just before they were put on the firm's recommended "buy" list. Michaels was regularly given the firm's quarterly personal transaction form but declined to complete it.

> *Comment*: Michaels violated Standard VI(B) by placing personal transactions ahead of client transactions. In addition, his supervisor violated Standard IV(C)–Responsibilities of Supervisors by permitting Michaels to continue to perform his assigned tasks without having signed the quarterly personal transaction form. Note also that if Michaels had communicated information about the firm's recommendations to a person who traded the security, that action would be a misappropriation of the information and a violation of Standard II(A)–Material Nonpublic Information.

Example 5 (Trading Prior to Report Dissemination):

A brokerage's insurance analyst, Denise Wilson, makes a closed-circuit TV report to her firm's branches around the country. During the broadcast, she includes negative comments about a major company in the insurance industry. The following day, Wilson's report is printed and distributed to the sales force and public customers. The report recommends that both short-term traders and intermediate investors take profits by selling that insurance company's stock. Seven minutes after the broadcast, however, Ellen Riley, head of the firm's trading department, had closed out a long "call" position in the stock. Shortly thereafter, Riley established a sizable "put" position in the stock. When asked about her activities, Riley claimed she took the actions to facilitate anticipated sales by institutional clients.

> *Comment*: Riley did not give customers an opportunity to buy or sell in the options market before the firm itself did. By taking action before the report was disseminated, Riley's firm may have depressed the price of the calls and increased the price of the puts. The firm could have avoided a conflict of interest if it had waited to trade for its own account until its clients had an opportunity to receive and assimilate Wilson's recommendations. As it is, Riley's actions violated Standard VI(B).

STANDARD VI(C): CONFLICTS OF INTEREST - REFERRAL FEES

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Members and Candidates must disclose to their employer, clients, and prospective clients, as appropriate, any compensation, consideration, or benefit received from or paid to others for the recommendation of products or services.

Guidance

Standard VI(C) states the responsibility of members and candidates to inform their employer, clients, and prospective clients of any benefit received for referrals of customers and clients. Such disclosures allow clients or employers to evaluate (1) any partiality shown in any recommendation of services and (2) the full cost of the services. Members and candidates must disclose when they pay a fee or provide compensation to others who have referred prospective clients to the member or candidate.

Appropriate disclosure means that members and candidates must advise the client or prospective client, before entry into any formal agreement for services, of any benefit given or received for the recommendation of any services provided by the member or candidate. In addition, the member or candidate must disclose the nature of the consideration or benefit—for example, flat fee or percentage basis, one-time or continuing benefit, based on performance, benefit in the form of provision of research or other noncash benefit—together with the estimated dollar value. Consideration includes all fees, whether paid in cash, in soft dollars, or in kind.

STANDARD VI(C): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Members and candidates should encourage their employers to develop procedures related to referral fees. The firm may completely restrict such fees. If the firm does not adopt a strict prohibition of such fees, the procedures should indicate the appropriate steps for requesting approval.

Employers should have investment professionals provide to the clients notification of approved referral fee programs and provide the employer regular (at least quarterly) updates on the amount and nature of compensation received.

STANDARD VI(C): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Disclosure of Referral Arrangements and Outside Parties):

Brady Securities, Inc., a broker/dealer, has established a referral arrangement with Lewis Brothers, Ltd., an investment counseling firm. In this arrangement, Brady Securities refers all prospective tax-exempt accounts, including pension, profit-sharing, and endowment accounts, to Lewis Brothers. In return, Lewis Brothers makes available to Brady Securities on a regular basis the security recommendations and reports of its research staff, which registered representatives of Brady Securities use in serving customers. In addition, Lewis Brothers conducts monthly economic and market reviews for Brady Securities personnel and directs all stock commission business generated by referral accounts to Brady Securities.

Willard White, a partner in Lewis Brothers, calculates that the incremental costs involved in functioning as the research department of Brady Securities are US$20,000 annually.

Referrals from Brady Securities last year resulted in fee income of US$200,000 for Lewis Brothers, and directing all stock trades through Brady Securities resulted in additional costs to Lewis Brothers' clients of US$10,000.

Diane Branch, the chief financial officer of Maxwell Inc., contacts White and says that she is seeking an investment manager for Maxwell's profit-sharing plan. She adds, "My friend Harold Hill at Brady Securities recommended your firm without qualification, and that's good enough for me. Do we have a deal?" White accepts the new account but does not disclose his firm's referral arrangement with Brady Securities.

> *Comment*: White has violated Standard VI(C) by failing to inform the prospective customer of the referral fee payable in services and commissions for an indefinite period to Brady Securities. Such disclosure could have caused Branch to reassess Hill's recommendation and make a more critical evaluation of Lewis Brothers' services.

Example 2 (Disclosure of Interdepartmental Referral Arrangements):

James Handley works for the trust department of Central Trust Bank. He receives compensation for each referral he makes to Central Trust's brokerage department and personal financial management department that results in a sale. He refers several of his clients to the personal financial management department but does not disclose the arrangement within Central Trust to his clients.

> *Comment*: Handley has violated Standard VI(C) by not disclosing the referral arrangement at Central Trust Bank to his clients. Standard VI(C) does not distinguish between referral payments paid by a third party for referring clients to the third party and internal payments paid within the firm to attract new business to a subsidiary. Members and candidates must disclose all such referral fees. Therefore, Handley is required to disclose, at the time of referral, any referral fee agreement in place among Central Trust Bank's departments. The disclosure should include the nature and the value of the benefit and should be made in writing.

Example 3 (Disclosure of Referral Arrangements and Informing Firm):

Katherine Roberts is a portfolio manager at Katama Investments, an advisory firm specializing in managing assets for high-net-worth individuals. Katama's trading desk uses a variety of brokerage houses to execute trades on behalf of its clients. Roberts asks the trading desk to direct a large portion of its commissions to Naushon, Inc., a small broker/dealer run by one of Roberts' business school classmates. Katama's traders have found that Naushon is not very competitive on pricing, and although Naushon generates some research for its trading clients, Katama's other analysts have found most of Naushon's research to be not especially useful. Nevertheless, the traders do as Roberts asks, and in return for receiving a large portion of Katama's business, Naushon recommends the investment services of Roberts and Katama to its wealthiest clients. This arrangement is not disclosed to either Katama or the clients referred by Naushon.

> *Comment*: Roberts is violating Standard VI(C) by failing to inform her employer of the referral arrangement.

Example 4 (Disclosure of Referral Arrangements and Outside Organizations):

Alex Burl is a portfolio manager at Helpful Investments, a local investment advisory firm. Burl is on the advisory board of his child's school, which is looking for ways to raise money to purchase new playground equipment for the school. Burl discusses a plan with his supervisor in which he will donate to the school a portion of his service fee from new clients referred by the parents of students at the school. Upon getting the approval from Helpful, Burl presents the idea to the school's advisory board and directors. The school agrees to announce the program at the next parent event and asks Burl to provide the appropriate written materials to be distributed. A week following the distribution of the flyers, Burl receives the first school-related referral. In establishing the client's investment policy statement, Burl clearly discusses the school's referral and outlines the plans for distributing the donation back to the school.

> *Comment*: Burl has not violated Standard VI(C) because he secured the permission of his employer, Helpful Investments, and the school prior to beginning the program and because he discussed the arrangement with the client at the time the investment policy statement was designed.

Example 5 (Disclosure of Referral Arrangements and Outside Parties):

The sponsor of a state employee pension is seeking to hire a firm to manage the pension plan's emerging market allocation. To assist in the review process, the sponsor has hired Thomas Arrow as a consultant to solicit proposals from various advisers. Arrow is contracted by the sponsor to represent its best interest in selecting the most appropriate new manager. The process runs smoothly, and Overseas Investments is selected as the new manager.

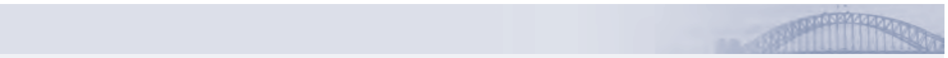
The following year, news breaks that Arrow is under investigation by the local regulator for accepting kickbacks from investment managers after they are awarded new pension allocations. Overseas Investments is included in the list of firms allegedly making these payments. Although the sponsor is happy with the performance of Overseas since it has been managing the pension plan's emerging market funds, the sponsor still decides to have an independent review of the proposals and the selection process to ensure that Overseas was the appropriate firm for its needs. This review confirms that, even though Arrow was being paid by both parties, the recommendation of Overseas appeared to be objective and appropriate.

> *Comment*: Arrow has violated Standard VI(C) because he did not disclose the fee being paid by Overseas. Withholding this information raises the question of a potential lack of objectivity in the recommendation of Overseas by Arrow; this aspect is in addition to questions about the legality of having firms pay to be considered for an allocation.
>
> Regulators and governmental agencies may adopt requirements concerning allowable consultant activities. Local regulations sometimes include having a consultant register with the regulatory agency's ethics board. Regulator policies may include a prohibition on acceptance of payments from investment managers receiving allocations and require regular reporting of contributions made to political organizations and candidates. Arrow would have to adhere to these requirements as well as the Code and Standards.

STANDARD VII(A): RESPONSIBILITIES AS A CFA INSTITUTE MEMBER OR CFA CANDIDATE - CONDUCT AS PARTICIPANTS IN CFA INSTITUTE PROGRAMS

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

Standard VII(A) Conduct as Participants in CFA Institute Programs



Members and Candidates must not engage in any conduct that compromises the reputation or integrity of CFA Institute or the CFA designation or the integrity, validity, or security of CFA Institute programs.

Guidance

Highlights:

- *Confidential Program Information*
- *Additional CFA Program Restrictions*
- *Expressing an Opinion*

Standard VII(A) covers the conduct of CFA Institute members and candidates involved with the CFA Program and prohibits any conduct that undermines the public's confidence that the CFA charter represents a level of achievement based on merit and ethical conduct. There is an array of CFA Institute programs beyond the CFA Program that provide additional educational and credentialing opportunities, including the Certificate in Investment Performance Measurement (CIPM) Program and the CFA Institute Investment Foundations™ Program. The standard's function is to hold members and candidates to a high ethical criterion while they are participating in or involved with any CFA Institute program. Conduct covered includes but is not limited to

- giving or receiving assistance (cheating) on any CFA Institute examinations;
- violating the rules, regulations, and testing policies of CFA Institute programs;
- providing confidential program or exam information to candidates or the public;
- disregarding or attempting to circumvent security measures established for any CFA Institute examinations;
- improperly using an association with CFA Institute to further personal or professional goals; and
- misrepresenting information on the Professional Conduct Statement or in the CFA Institute Continuing Education Program.

Confidential Program Information

CFA Institute is vigilant about protecting the integrity of CFA Institute programs' content and examination processes. CFA Institute program rules, regulations, and policies prohibit candidates from disclosing confidential material gained during the exam process.

Examples of information that cannot be disclosed by candidates sitting for an exam include but are not limited to

- specific details of questions appearing on the exam and
- broad topical areas and formulas tested or not tested on the exam.

All aspects of the exam, including questions, broad topical areas, and formulas, tested or not tested, are considered confidential until such time as CFA Institute elects to release them publicly. This confidentiality requirement allows CFA Institute to maintain the integrity and rigor of exams for future candidates. Standard VII(A) does not prohibit candidates from discussing nonconfidential information or curriculum material with others or in study groups in preparation for the exam.

Candidates increasingly use online forums and new technology as part of their exam preparations. CFA Institute actively polices blogs, forums, and related social networking groups for information considered confidential. The organization works with both individual candidates and the sponsors of online or offline services to promptly remove any and all violations. As noted in the discussion of Standard I(A)–Knowledge of the Law, candidates, members, and the public are encouraged to report suspected violations to CFA Institute.

Additional CFA Program Restrictions

The CFA Program rules, regulations, and policies define additional allowed and disallowed actions concerning the exams. Violating any of the testing policies, such as the calculator policy, personal belongings policy, or the Candidate Pledge, constitutes a violation of Standard VII(A). Candidates will find all of these policies on the CFA Program portion of the CFA Institute website (www.cfainstitute.org). Exhibit 2 provides the Candidate Pledge, which highlights the respect candidates must have for the integrity, validity, and security of the CFA exam.

Members may participate as volunteers in various aspects of the CFA Program. Standard VII(A) prohibits members from disclosing and/or soliciting confidential material gained prior to or during the exam and grading processes with those outside the CFA exam development process.

Examples of information that cannot be shared by members involved in developing, administering, or grading the exams include but are not limited to

- questions appearing on the exam or under consideration,
- deliberation related to the exam process, and
- information related to the scoring of questions.

Members may also be asked to offer assistance with other CFA Institute programs, including but not limited to the CIPM and Investment Foundations programs. Members participating in any CFA Institute program should do so with the same level of integrity and confidentiality as is required of participation in the CFA Program.

Expressing an Opinion

Standard VII(A) does *not* cover expressing opinions regarding CFA Institute, the CFA Program, or other CFA Institute programs. Members and candidates are free to disagree and express their disagreement with CFA Institute on its policies, its procedures, or any advocacy positions taken by the organization. When expressing a personal opinion, a candidate is prohibited from disclosing content-specific information, including any actual exam question and the information as to subject matter covered or not covered in the exam.

Exhibit 2 Sample of CFA Program Testing Policies

Candidate Pledge	As a candidate in the CFA Program, I am obligated to follow Standard VII(A) of the CFA Institute Standards of Professional Conduct, which states that members and candidates must not engage in any conduct that compromises the reputation or integrity of CFA Institute or the CFA designation or the integrity, validity, or security of the CFA exam. ■ Prior to this exam, I have not given or received information regarding the content of this exam. During this exam, I will not give or receive any information regarding the content of this exam. ■ After this exam, I will not disclose **ANY** portion of this exam and I will not remove **ANY** exam materials from the testing room in original or copied form. I understand that all exam materials, including my answers, are the property of CFA Institute and will not be returned to me in any form. ■ I will follow **ALL** rules of the CFA Program as stated on the CFA Institute website and the back cover of the exam book. My violation of any rules of the CFA Program will result in CFA Institute voiding my exam results and may lead to suspension or termination of my candidacy in the CFA Program.

STANDARD VII(A): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Sharing Exam Questions):

Travis Nero serves as a proctor for the administration of the CFA examination in his city. In the course of his service, he reviews a copy of the Level II exam on the evening prior to the exam's administration and provides information concerning the exam questions to two candidates who use it to prepare for the exam.

> *Comment*: Nero and the two candidates have violated Standard VII(A). By giving information about the exam questions to two candidates, Nero provided an unfair advantage to the two candidates and undermined the integrity and validity of the Level II exam as an accurate measure of the knowledge, skills, and abilities necessary to earn the right to use the CFA designation. By accepting the information, the candidates also compromised the integrity and validity of the Level II exam and undermined the ethical framework that is a key part of the designation.

Example 2 (Bringing Written Material into Exam Room):

Loren Sullivan is enrolled to take the Level II CFA examination. He has been having difficulty remembering a particular formula, so prior to entering the exam room, he writes the formula on the palm of his hand. During the afternoon section of the exam, a proctor notices Sullivan looking at the palm of his hand. She asks to see his hand and finds the formula.

> *Comment*: Because Sullivan wrote down information from the Candidate Body of Knowledge (CBOK) and took that written information into the exam room, his conduct compromised the validity of his exam performance and violated Standard VII(A). Sullivan's conduct was also in direct contradiction with the rules and regulations of the CFA Program, the Candidate Pledge, and the CFA Institute Code and Standards.

Example 3 (Writing after Exam Period End):

At the conclusion of the morning section of the Level I CFA examination, the proctors announce, "Stop writing now." John Davis has not completed the exam, so he continues to randomly fill in ovals on his answer sheet. A proctor approaches Davis's desk and reminds him that he should stop writing immediately. Davis, however, continues to complete the answer sheet. After the proctor asks him to stop writing two additional times, Davis finally puts down his pencil.

> *Comment*: By continuing to complete his exam after time was called, Davis has violated Standard VII(A). By continuing to write, Davis took an unfair advantage over other candidates, and his conduct compromised the validity of his exam performance. Additionally, by not heeding the proctor's repeated instructions, Davis violated the rules and regulations of the CFA Program.

Example 4 (Sharing Exam Content):

After completing Level II of the CFA exam, Annabelle Rossi posts on her blog about her experience. She posts the following: "Level II is complete! I think I did fairly well on the exam. It was really difficult, but fair. I think I did especially well on the derivatives questions. And there were tons of them! I think I counted 18! The ethics questions were really hard. I'm glad I spent so much time on the Code and Standards. I was surprised to see there were no questions at all about IPO allocations. I expected there to be a couple. Well, off to celebrate getting through it. See you tonight?"

> *Comment*: Rossi did not violate Standard VII(A) when she wrote about how difficult she found the exam or how well she thinks she may have done. By revealing portions of the CBOK covered on the exam and areas not covered, however, she did violate Standard VII(A) and the Candidate Pledge. Depending on the time frame in which the comments were posted, Rossi not only may have assisted future candidates but also may have provided an unfair advantage to candidates yet to sit for the same exam, thereby undermining the integrity and validity of the Level II exam.

Example 5 (Sharing Exam Content):

Level I candidate Etienne Gagne has been a frequent visitor to an internet forum designed specifically for CFA Program candidates. The week after completing the Level I examination, Gagne and several others begin a discussion thread on the forum about the most challenging questions and attempt to determine the correct answers.

> *Comment*: Gagne has violated Standard VII(A) by providing and soliciting confidential exam information, which compromises the integrity of the exam process and violates the Candidate Pledge. In trying to determine correct answers to specific questions, the group's discussion included question-specific details considered to be confidential to the CFA Program.

Example 6 (Sharing Exam Content):

CFA4Sure is a company that produces test-preparation materials for CFA Program candidates. Many candidates register for and use the company's products. The day after the CFA examination, CFA4Sure sends an e-mail to all its customers asking them to share with the company the hardest questions from the exam so that CFA4Sure can better prepare its customers for the next exam administration. Marisol Pena e-mails a summary of the questions she found most difficult on the exam.

> *Comment*: Pena has violated Standard VII(A) by disclosing a portion of the exam questions. The information provided is considered confidential until publicly released by CFA Institute. CFA4Sure is likely to use such feedback to refine its review materials for future candidates. Pena's sharing of the specific questions undermines the integrity of the exam while potentially making the exam easier for future candidates.
>
> If the CFA4Sure employees who participated in the solicitation of confidential CFA Program information are CFA Institute members or candidates, they also have violated Standard VII(A).

Example 7 (Discussion of Exam Grading Guidelines and Results):

Prior to participating in grading CFA examinations, Wesley Whitcomb is required to sign a CFA Institute Grader Agreement. As part of the Grader Agreement, Whitcomb agrees not to reveal or discuss the exam materials with anyone except CFA Institute staff or other graders. Several weeks after the conclusion of the CFA exam grading, Whitcomb tells several colleagues who are candidates in the CFA Program which question he graded. He also discusses the guideline answer and adds that few candidates scored well on the question.

> *Comment*: Whitcomb violated Standard VII(A) by breaking the Grader Agreement and disclosing information related to a specific question on the exam, which compromised the integrity of the exam process.

Example 8 (Compromising CFA Institute Integrity as a Volunteer):

Jose Ramirez is an investor-relations consultant for several small companies that are seeking greater exposure to investors. He is also the program chair for the CFA Institute society in the city where he works. Ramirez schedules only companies that are his clients to make presentations to the society and excludes other companies.

> *Comment*: Ramirez, by using his volunteer position at CFA Institute to benefit himself and his clients, compromises the reputation and integrity of CFA Institute and thus violates Standard VII(A).

Example 9 (Compromising CFA Institute Integrity as a Volunteer):

Marguerite Warrenski is a member of the CFA Institute GIPS Executive Committee, which oversees the creation, implementation, and revision of the GIPS standards. As a member of the Executive Committee, she has advance knowledge of confidential information regarding the GIPS standards, including any new or revised standards the committee is considering. She tells her clients that her Executive Committee membership will allow her to better assist her clients in keeping up with changes to the Standards and facilitating their compliance with the changes.

> *Comment*: Warrenski is using her association with the GIPS Executive Committee to promote her firm's services to clients and potential clients. In defining her volunteer position at CFA Institute as a strategic business advantage over competing firms and implying to clients that she would use confidential information to further their interests, Warrenski is compromising the reputation and integrity of CFA Institute and thus violating Standard VII(A). She may factually state her involvement with the Executive Committee but cannot infer any special advantage to her clients from such participation.

STANDARD VII(B): RESPONSIBILITIES AS A CFA INSTITUTE MEMBER OR CFA CANDIDATE - REFERENCE TO CFA INSTITUTE, THE CFA DESIGNATION, AND THE CFA PROGRAM

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

When referring to CFA Institute, CFA Institute membership, the CFA designation, or candidacy in the CFA Program, Members and Candidates must not misrepresent or exaggerate the meaning or implications of membership in CFA Institute, holding the CFA designation, or candidacy in the CFA Program.

Guidance

Highlights:

- *CFA Institute Membership*
- *Using the CFA Designation*
- *Referring to Candidacy in the CFA Program*

Standard VII(B) is intended to prevent promotional efforts that make promises or guarantees that are tied to the CFA designation. Individuals must not exaggerate the meaning or implications of membership in CFA Institute, holding the CFA designation, or candidacy in the CFA Program.

Standard VII(B) is not intended to prohibit factual statements related to the positive benefit of earning the CFA designation. However, statements referring to CFA Institute, the CFA designation, or the CFA Program that overstate the competency of an individual or imply, either directly or indirectly, that superior performance can be expected from someone with the CFA designation are not allowed under the standard.

Statements that highlight or emphasize the commitment of CFA Institute members, CFA charterholders, and CFA candidates to ethical and professional conduct or mention the thoroughness and rigor of the CFA Program are appropriate. Members and candidates may make claims about the relative merits of CFA Institute, the CFA Program, or the Code and Standards as long as those statements are implicitly or explicitly stated as the opinion of the speaker. Statements that do not express opinions have to be supported by facts.

Standard VII(B) applies to any form of communication, including but not limited to communications made in electronic or written form (such as on firm letterhead, business cards, professional biographies, directory listings, printed advertising, firm brochures, or personal resumes) and oral statements made to the public, clients, or prospects.

CFA Institute Membership

The term “CFA Institute member” refers to “regular” and “affiliate” members of CFA Institute who have met the membership requirements as defined in the CFA Institute Bylaws. Once accepted as a CFA Institute member, the member must satisfy the following requirements to maintain his or her status:

- remit annually to CFA Institute a completed Professional Conduct Statement, which renews the commitment to abide by the requirements of the Code and Standards and the CFA Institute Professional Conduct Program, and
- pay applicable CFA Institute membership dues on an annual basis.

If a CFA Institute member fails to meet any of these requirements, the individual is no longer considered an active member. Until membership is reactivated, individuals must not present themselves to others as active members. They may state, however, that they were CFA Institute members in the past or refer to the years when their membership was active.

Using the CFA Designation

Those who have earned the right to use the Chartered Financial Analyst designation are encouraged to do so but only in a manner that does not misrepresent or exaggerate the meaning or implications of the designation. The use of the designation may be accompanied by an accurate explanation of the requirements that have been met to earn the right to use the designation.

"CFA charterholders" are those individuals who have earned the right to use the CFA designation granted by CFA Institute. These people have satisfied certain requirements, including completion of the CFA Program and required years of acceptable work experience. Once granted the right to use the designation, individuals must also satisfy the CFA Institute membership requirements (see above) to maintain their right to use the designation.

If a CFA charterholder fails to meet any of the membership requirements, he or she forfeits the right to use the CFA designation. Until membership is reactivated, individuals must not present themselves to others as CFA charterholders. They may state, however, that they were charterholders in the past.

Given the growing popularity of social media, where individuals may anonymously express their opinions, pseudonyms or online profile names created to hide a member's identity should not be tagged with the CFA designation.

Use of the CFA designation by a CFA charterholder is governed by the terms and conditions of the annual Professional Conduct Statement Agreement, entered into between CFA Institute and its membership prior to commencement of use of the CFA designation and reaffirmed annually.

Referring to Candidacy in the CFA Program

Candidates in the CFA Program may refer to their participation in the CFA Program, but such references must clearly state that an individual is a *candidate* in the CFA Program and must not imply that the candidate has achieved any type of partial designation. A person is a candidate in the CFA Program if

- the person's application for registration in the CFA Program has been accepted by CFA Institute, as evidenced by issuance of a notice of acceptance, and the person is enrolled to sit for a specified examination or
- the registered person has sat for a specified examination but exam results have not yet been received.

If an individual is registered for the CFA Program but declines to sit for an exam or otherwise does not meet the definition of a candidate as described in the CFA Institute Bylaws, then that individual is no longer considered an active candidate. Once the person is enrolled to sit for a future examination, his or her CFA Program candidacy resumes.

CFA Program candidates must never state or imply that they have a partial designation as a result of passing one or more levels or cite an expected completion date of any level of the CFA Program. Final award of the charter is subject to meeting the CFA Program requirements and approval by the CFA Institute Board of Governors.

If a candidate passes each level of the exam in consecutive years and wants to state that he or she did so, that is not a violation of Standard VII(B) because it is a statement of fact. If the candidate then goes on to claim or imply superior ability by obtaining the designation in only three years, however, he or she is in violation of Standard VII(B).

Exhibit 3 provides examples of proper and improper references to the CFA designation.

Exhibit 3 Proper and Improper References to the CFA Designation

Proper References	Improper References
“Completion of the CFA Program has enhanced my portfolio management skills.”	“CFA charterholders achieve better performance results.”
“John Smith passed all three CFA Program examinations in three consecutive years.”	“John Smith is among the elite, having passed all three CFA examinations in three consecutive attempts.”
“The CFA designation is globally recognized and attests to a charterholder’s success in a rigorous and comprehensive study program in the field of investment management and research analysis.”	“As a CFA charterholder, I am the most qualified to manage client investments.”
“The credibility that the CFA designation affords and the skills the CFA Program cultivates are key assets for my future career development.”	“As a CFA charterholder, Jane White provides the best value in trade execution.”
“I enrolled in the CFA Program to obtain the highest set of credentials in the global investment management industry.”	“Enrolling as a candidate in the CFA Program ensures one of becoming better at valuing debt securities.”
“I passed Level I of the CFA Program.”	“CFA, Level II”
“I am a 2010 Level III candidate in the CFA Program.”	“CFA, Expected 2011”
“I passed all three levels of the CFA Program and may be eligible for the CFA charter upon completion of the required work experience.”	“CFA, Expected 2011” “John Smith, Charter Pending”

STANDARD VII(B): RECOMMENDED PROCEDURES

c recommend practices and procedures designed to prevent violations of the Code of Ethics and Standards of Professional Conduct.

Misuse of a member’s CFA designation or CFA candidacy or improper reference to it is common by those in a member’s or candidate’s firm who do not possess knowledge of the requirements of Standard VII(B). As an appropriate step to reduce this risk, members and candidates should disseminate written information about Standard VII(B) and the accompanying guidance to their firm’s legal, compliance, public relations, and marketing departments (see www.cfainstitute.org).

For materials that refer to employees’ affiliation with CFA Institute, members and candidates should encourage their firms to create templates that are approved by a central authority (such as the compliance department) as being consistent with Standard VII(B). This practice promotes consistency and accuracy in the firm of references to CFA Institute membership, the CFA designation, and CFA candidacy.

STANDARD VII(B): APPLICATION OF THE STANDARD

a demonstrate the application of the Code of Ethics and Standards of Professional Conduct to situations involving issues of professional integrity;

b identify conduct that conforms to the Code and Standards and conduct that violates the Code and Standards;

Example 1 (Passing Exams in Consecutive Years):

An advertisement for AZ Investment Advisors states that all the firm's principals are CFA charterholders and all passed the three examinations on their first attempt. The advertisement prominently links this fact to the notion that AZ's mutual funds have achieved superior performance.

> *Comment*: AZ may state that all principals passed the three examinations on the first try as long as this statement is true, but it must not be linked to performance or imply superior ability. Implying that (1) CFA charterholders achieve better investment results and (2) those who pass the exams on the first try may be more successful than those who do not violates Standard VII(B).

Example 2 (Right to Use CFA Designation):

Five years after receiving his CFA charter, Louis Vasseur resigns his position as an investment analyst and spends the next two years traveling abroad. Because he is not actively engaged in the investment profession, he does not file a completed Professional Conduct Statement with CFA Institute and does not pay his CFA Institute membership dues. At the conclusion of his travels, Vasseur becomes a self-employed analyst accepting assignments as an independent contractor. Without reinstating his CFA Institute membership by filing his Professional Conduct Statement and paying his dues, he prints business cards that display "CFA" after his name.

> *Comment*: Vasseur has violated Standard VII(B) because his right to use the CFA designation was suspended when he failed to file his Professional Conduct Statement and stopped paying dues. Therefore, he no longer is able to state or imply that he is an active CFA charterholder. When Vasseur files his Professional Conduct Statement, resumes paying CFA Institute dues to activate his membership, and completes the CFA Institute reinstatement procedures, he will be eligible to use the CFA designation.

Example 3 ("Retired" CFA Institute Membership Status):

After a 25-year career, James Simpson retires from his firm. Because he is not actively engaged in the investment profession, he does not file a completed Professional Conduct Statement with CFA Institute and does not pay his CFA Institute membership dues. Simpson designs a plain business card (without a corporate logo) to hand out to friends with his new contact details, and he continues to put "CFA" after his name.

> *Comment*: Simpson has violated Standard VII(B). Because he failed to file his Professional Conduct Statement and ceased paying dues, his membership has been suspended and he has given up the right to use the CFA designation. CFA Institute has procedures, however, for reclassifying a member

> and charterholder as “retired” and reducing the annual dues. If he wants to obtain retired status, he needs to file the appropriate paperwork with CFA Institute. When Simpson receives his notification from CFA Institute that his membership has been reclassified as retired and he resumes paying reduced dues, his membership will be reactivated and his right to use the CFA designation will be reinstated.

Example 4 (Stating Facts about CFA Designation and Program):

Rhonda Reese has been a CFA charterholder since 2000. In a conversation with a friend who is considering enrolling in the CFA Program, she states that she has learned a great deal from the CFA Program and that many firms require their employees to be CFA charterholders. She would recommend the CFA Program to anyone pursuing a career in investment management.

> *Comment*: Reese’s comments comply with Standard VII(B). Her statements refer to facts: The CFA Program enhanced her knowledge, and many firms require the CFA designation for their investment professionals.

Example 5 (Order of Professional and Academic Designations):

Tatiana Prittima has earned both her CFA designation and a PhD in finance. She would like to cite both her accomplishments on her business card but is unsure of the proper method for doing so.

> *Comment*: The order of designations cited on such items as resumes and business cards is a matter of personal preference. Prittima is free to cite the CFA designation either before or after citing her PhD. Multiple designations must be separated by a comma.

Example 6 (Use of Fictitious Name):

Barry Glass is the lead quantitative analyst at CityCenter Hedge Fund. Glass is responsible for the development, maintenance, and enhancement of the proprietary models the fund uses to manage its investors’ assets. Glass reads several high-level mathematical publications and blogs to stay informed on current developments. One blog, run by Expert CFA, presents some intriguing research that may benefit one of CityCenter’s current models. Glass is under pressure from firm executives to improve the model’s predictive abilities, and he incorporates the factors discussed in the online research. The updated output recommends several new investments to the fund’s portfolio managers.

> *Comment*: “Expert CFA” has violated Standard VII(B) by using the CFA designation inappropriately. As with any research report, authorship of online comments must include the charterholder’s full name along with any reference to the CFA designation.
>
> See also Standard V(A), which Glass has violated for guidance on diligence and reasonable basis.

PRACTICE PROBLEMS

Unless otherwise stated in the question, all individuals in the following questions are CFA Institute members or candidates in the CFA Program and, therefore, are subject to the CFA Institute Code of Ethics and Standards of Professional Conduct.

1 Smith, a research analyst with a brokerage firm, decides to change his recommendation for the common stock of Green Company, Inc., from a “buy” to a “sell.” He mails this change in investment advice to all the firm’s clients on Wednesday. The day after the mailing, a client calls with a buy order for 500 shares of Green Company. In this circumstance, Smith should:

A Accept the order.

B Advise the customer of the change in recommendation before accepting the order.

C Not accept the order because it is contrary to the firm’s recommendation.

2 Which statement about a manager’s use of client brokerage commissions violates the Code and Standards?

A A client may direct a manager to use that client’s brokerage commissions to purchase goods and services for that client.

B Client brokerage commissions should be used to benefit the client and should be commensurate with the value of the brokerage and research services received.

C Client brokerage commissions may be directed to pay for the investment manager’s operating expenses.

3 Jamison is a junior research analyst with Howard & Howard, a brokerage and investment banking firm. Howard & Howard’s mergers and acquisitions department has represented the Britland Company in all of its acquisitions for the past 20 years. Two of Howard & Howard’s senior officers are directors of various Britland subsidiaries. Jamison has been asked to write a research report on Britland. What is the best course of action for her to follow?

A Jamison may write the report but must refrain from expressing any opinions because of the special relationships between the two companies.

B Jamison should not write the report because the two Howard & Howard officers serve as directors for subsidiaries of Britland.

C Jamison may write the report if she discloses the special relationships with the company in the report.

4 Which of the following statements clearly *conflicts* with the recommended procedures for compliance presented in the CFA Institute *Standards of Practice Handbook*?

A Firms should disclose to clients the personal investing policies and procedures established for their employees.

B Prior approval must be obtained for the personal investment transactions of all employees.

C For confidentiality reasons, personal transactions and holdings should not be reported to employers unless mandated by regulatory organizations.

5 Bronson provides investment advice to the board of trustees of a private university endowment fund. The trustees have provided Bronson with the fund’s financial information, including planned expenditures. Bronson receives a

© 2014 CFA Institute. All rights reserved.

phone call on Friday afternoon from Murdock, a prominent alumnus, requesting that Bronson fax him comprehensive financial information about the fund. According to Murdock, he has a potential contributor but needs the information that day to close the deal and cannot contact any of the trustees. Based on the CFA Institute Standards, Bronson should:

A Send Murdock the information because disclosure would benefit the client.

B Not send Murdock the information to preserve confidentiality.

C Send Murdock the information, provided Bronson promptly notifies the trustees.

6 Willier is the research analyst responsible for following Company X. All the information he has accumulated and documented suggests that the outlook for the company's new products is poor, so the stock should be rated a weak "hold." During lunch, however, Willier overhears a financial analyst from another firm whom he respects offer opinions that conflict with Willier's forecasts and expectations. Upon returning to his office, Willier releases a strong "buy" recommendation to the public. Willier:

A Violated the Standards by failing to distinguish between facts and opinions in his recommendation.

B Violated the Standards because he did not have a reasonable and adequate basis for his recommendation.

C Was in full compliance with the Standards.

7 An investment management firm has been hired by ETV Corporation to work on an additional public offering for the company. The firm's brokerage unit now has a "sell" recommendation on ETV, but the head of the investment banking department has asked the head of the brokerage unit to change the recommendation from "sell" to "buy." According to the Standards, the head of the brokerage unit would be permitted to:

A Increase the recommendation by no more than one increment (in this case, to a "hold" recommendation).

B Place the company on a restricted list and give only factual information about the company.

C Assign a new analyst to decide if the stock deserves a higher rating.

8 Albert and Tye, who recently started their own investment advisory business, have registered to take the Level III CFA examination. Albert's business card reads, "Judy Albert, CFA Level II." Tye has not put anything about the CFA designation on his business card, but promotional material that he designed for the business describes the CFA requirements and indicates that Tye participates in the CFA Program and has completed Levels I and II. According to the Standards:

A Albert has violated the Standards, but Tye has not.

B Tye has violated the Standards, but Albert has not.

C Both Albert and Tye have violated the Standards.

9 Scott works for a regional brokerage firm. He estimates that Walkton Industries will increase its dividend by US$1.50 a share during the next year. He realizes that this increase is contingent on pending legislation that would, if enacted, give Walkton a substantial tax break. The US representative for Walkton's home district has told Scott that, although she is lobbying hard for the bill and prospects for its passage are favorable, concern of the US Congress over the federal deficit could cause the tax bill to be voted down. Walkton Industries has not made any statements about a change in dividend policy. Scott writes in his research report, "We expect Walkton's stock price to rise by at least US$8.00 a share by the end of the year because the dividend will increase by US$1.50 a share. Investors buying the stock at the current time should expect to realize a total return of at least 15% on the stock." According to the Standards:

A Scott violated the Standards because he used material inside information.

B Scott violated the Standards because he failed to separate opinion from fact.

C Scott violated the Standards by basing his research on uncertain predictions of future government action.

10 Which one of the following actions will help to ensure the fair treatment of brokerage firm clients when a new investment recommendation is made?

A Informing all people in the firm in advance that a recommendation is to be disseminated.

B Distributing recommendations to institutional clients prior to individual accounts.

C Minimizing the time between the decision and the dissemination of a recommendation.

11 The mosaic theory holds that an analyst:

A Violates the Code and Standards if the analyst fails to have knowledge of and comply with applicable laws.

B Can use material public information and nonmaterial nonpublic information in the analyst's analysis.

C Should use all available and relevant information in support of an investment recommendation.

12 Jurgen is a portfolio manager. One of her firm's clients has told Jurgen that he will compensate her beyond the compensation provided by her firm on the basis of the capital appreciation of his portfolio each year. Jurgen should:

A Turn down the additional compensation because it will result in conflicts with the interests of other clients' accounts.

B Turn down the additional compensation because it will create undue pressure on her to achieve strong short-term performance.

C Obtain permission from her employer prior to accepting the compensation arrangement.

13 One of the discretionary accounts managed by Farnsworth is the Jones Corporation employee profit-sharing plan. Jones, the company president, recently asked Farnsworth to vote the shares in the profit-sharing plan in favor of the slate of directors nominated by Jones Corporation and against the directors sponsored by a dissident stockholder group. Farnsworth does not want to lose this account because he directs all the account's trades to a brokerage firm that provides Farnsworth with useful information about tax-free investments. Although this information is not of value in managing the Jones Corporation account, it does help in managing several other accounts. The brokerage firm providing this information also offers the lowest commissions for trades and provides best execution. Farnsworth investigates the director issue, concludes that the management-nominated slate is better for the long-run performance of the company than the dissident group's slate, and votes accordingly. Farnsworth:

A Violated the Standards in voting the shares in the manner requested by Jones but not in directing trades to the brokerage firm.

B Did not violate the Standards in voting the shares in the manner requested by Jones or in directing trades to the brokerage firm.

C Violated the Standards in directing trades to the brokerage firm but not in voting the shares as requested by Jones.

14 Brown works for an investment counseling firm. Green, a new client of the firm, is meeting with Brown for the first time. Green used another counseling firm for financial advice for years, but she has switched her account to Brown's firm. After spending a few minutes getting acquainted, Brown explains to Green that she has discovered a highly undervalued stock that offers large potential gains. She recommends that Green purchase the stock. Brown has committed a violation of the Standards. What should she have done differently?

A Brown should have determined Green's needs, objectives, and tolerance for risk before making a recommendation of any type of security.

B Brown should have thoroughly explained the characteristics of the company to Green, including the characteristics of the industry in which the company operates.

C Brown should have explained her qualifications, including her education, training, and experience and the meaning of the CFA designation.

15 Grey recommends the purchase of a mutual fund that invests solely in long-term US Treasury bonds. He makes the following statements to his clients:

I. "The payment of the bonds is guaranteed by the US government; therefore, the default risk of the bonds is virtually zero."

II. "If you invest in the mutual fund, you will earn a 10% rate of return each year for the next several years based on historical performance of the market."

Did Grey's statements violate the CFA Institute Code and Standards?

A Neither statement violated the Code and Standards.

B Only statement I violated the Code and Standards.

C Only statement II violated the Code and Standards.

16 Anderb, a portfolio manager for XYZ Investment Management Company—a registered investment organization that advises investment firms and private accounts—was promoted to that position three years ago. Bates, her supervisor, is responsible for reviewing Anderb's portfolio account transactions and her required monthly reports of personal stock transactions. Anderb has been using Jonelli, a broker, almost exclusively for brokerage transactions for the portfolio account. For securities in which Jonelli's firm makes a market, Jonelli has been giving Anderb lower prices for personal purchases and higher prices for personal sales than Jonelli gives to Anderb's portfolio accounts and other investors. Anderb has been filing monthly reports with Bates only for those months in which she has no personal transactions, which is about every fourth month. Which of the following is *most likely* to be a violation of the Code and Standards?

A Anderb failed to disclose to her employer her personal transactions.

B Anderb owned the same securities as those of her clients.

C Bates allowed Anderb to use Jonelli as her broker for personal trades.

17 Which of the following is a correct statement of a member's or candidate's duty under the Code and Standards?

A In the absence of specific applicable law or other regulatory requirements, the Code and Standards govern the member's or candidate's actions.

- **B** A member or candidate is required to comply only with applicable local laws, rules, regulations, or customs, even though the Code and Standards may impose a higher degree of responsibility or a higher duty on the member or candidate.
- **C** A member or candidate who trades securities in a securities market where no applicable local laws or stock exchange rules regulate the use of material nonpublic information may take investment action based on material nonpublic information.

18 Ward is scheduled to visit the corporate headquarters of Evans Industries. Ward expects to use the information he obtains there to complete his research report on Evans stock. Ward learns that Evans plans to pay all of Ward's expenses for the trip, including costs of meals, hotel room, and air transportation. Which of the following actions would be the *best* course for Ward to take under the Code and Standards?

- **A** Accept the expense-paid trip and write an objective report.
- **B** Pay for all travel expenses, including costs of meals and incidental items.
- **C** Accept the expense-paid trip but disclose the value of the services accepted in the report.

19 Which of the following statements is *correct* under the Code and Standards?

- **A** CFA Institute members and candidates are prohibited from undertaking independent practice in competition with their employer.
- **B** Written consent from the employer is necessary to permit independent practice that could result in compensation or other benefits in competition with a member's or candidate's employer.
- **C** Members and candidates are prohibited from making arrangements or preparations to go into a competitive business before terminating their relationship with their employer.

20 Smith is a financial analyst with XYZ Brokerage Firm. She is preparing a purchase recommendation on JNI Corporation. Which of the following situations is *most likely* to represent a conflict of interest for Smith that would have to be disclosed?

- **A** Smith frequently purchases items produced by JNI.
- **B** XYZ holds for its own account a substantial common stock position in JNI.
- **C** Smith's brother-in-law is a supplier to JNI.

21 Michelieu tells a prospective client, "I may not have a long-term track record yet, but I'm sure that you'll be very pleased with my recommendations and service. In the three years that I've been in the business, my equity-oriented clients have averaged a total return of more than 26% a year." The statement is true, but Michelieu only has a few clients, and one of his clients took a large position in a penny stock (against Michelieu's advice) and realized a huge gain. This large return caused the average of all of Michelieu's clients to exceed 26% a year. Without this one investment, the average gain would have been 8% a year. Has Michelieu violated the Standards?

- **A** No, because Michelieu is not promising that he can earn a 26% return in the future.
- **B** No, because the statement is a true and accurate description of Michelieu's track record.
- **C** Yes, because the statement misrepresents Michelieu's track record.

22 An investment banking department of a brokerage firm often receives material nonpublic information that could have considerable value if used in advising the firm's brokerage clients. In order to conform to the Code and Standards, which one of the following is the best policy for the brokerage firm?

A Permanently prohibit both "buy" and "sell" recommendations of the stocks of clients of the investment banking department.

B Establish physical and informational barriers within the firm to prevent the exchange of information between the investment banking and brokerage operations.

C Monitor the exchange of information between the investment banking department and the brokerage operation.

23 Stewart has been hired by Goodner Industries, Inc., to manage its pension fund. Stewart's duty of loyalty, prudence, and care is owed to:

A The management of Goodner.

B The participants and beneficiaries of Goodner's pension plan.

C The shareholders of Goodner.

24 Which of the following statements is a stated purpose of disclosure in Standard VI(C)–Referral Fees?

A Disclosure will allow the client to request discounted service fees.

B Disclosure will help the client evaluate any possible partiality shown in the recommendation of services.

C Disclosure means advising a prospective client about the referral arrangement once a formal client relationship has been established.

25 Rose, a portfolio manager for a local investment advisory firm, is planning to sell a portion of his personal investment portfolio to cover the costs of his child's academic tuition. Rose wants to sell a portion of his holdings in Household Products, but his firm recently upgraded the stock to "strong buy." Which of the following describes Rose's options under the Code and Standards?

A Based on his firm's "buy" recommendation, Rose cannot sell the shares because he would be improperly prospering from the inflated recommendation.

B Rose is free to sell his personal holdings once his firm is properly informed of his intentions.

C Rose can sell his personal holdings but only when a client of the firm places an order to buy shares of Household.

26 A former hedge fund manager, Jackman, has decided to launch a new private wealth management firm. From his prior experiences, he believes the new firm needs to achieve US$1 million in assets under management in the first year. Jackman offers a $10,000 incentive to any adviser who joins his firm with the minimum of $200,000 in committed investments. Jackman places notice of the opening on several industry web portals and career search sites. Which of the following is *correct* according to the Code and Standards?

A A member or candidate is eligible for the new position and incentive if he or she can arrange for enough current clients to switch to the new firm and if the member or candidate discloses the incentive fee.

B A member or candidate may not accept employment with the new firm because Jackman's incentive offer violates the Code and Standards.

- **C** A member or candidate is not eligible for the new position unless he or she is currently unemployed because soliciting the clients of the member's or candidate's current employer is prohibited.

27 Carter works for Invest Today, a local asset management firm. A broker that provides Carter with proprietary research through client brokerage arrangements is offering a new trading service. The broker is offering low-fee, execution-only trades to complement its traditional full-service, execution-and-research trades. To entice Carter and other asset managers to send additional business its way, the broker will apply the commissions paid on the new service toward satisfying the brokerage commitment of the prior full-service arrangements. Carter has always been satisfied with the execution provided on the full-service trades, and the new low-fee trades are comparable to the fees of other brokers currently used for the accounts that prohibit soft dollar arrangements.

- **A** Carter can trade for his accounts that prohibit soft dollar arrangements under the new low-fee trading scheme.
- **B** Carter cannot use the new trading scheme because the commissions are prohibited by the soft dollar restrictions of the accounts.
- **C** Carter should trade only through the new low-fee scheme and should increase his trading volume to meet his required commission commitment.

28 Rule has worked as a portfolio manager for a large investment management firm for the past 10 years. Rule earned his CFA charter last year and has decided to open his own investment management firm. After leaving his current employer, Rule creates some marketing material for his new firm. He states in the material, "In earning the CFA charter, a highly regarded credential in the investment management industry, I further enhanced the portfolio management skills learned during my professional career. While completing the examination process in three consecutive years, I consistently received the highest possible scores on the topics of Ethics, Alternative Investments, and Portfolio Management." Has Rule violated Standard VII(B)–Reference to CFA Institute, the CFA Designation, and the CFA Program in his marketing material?

- **A** Rule violated Standard VII(B) in stating that he completed the exams in three consecutive years.
- **B** Rule violated Standard VII(B) in stating that he received the highest scores in the topics of Ethics, Alternative Investments, and Portfolio Management.
- **C** Rule did not violate Standard VII(B).

29 Stafford is a portfolio manager for a specialized real estate mutual fund. Her firm clearly describes in the fund's prospectus its soft dollar policies. Stafford decides that entering the CFA Program will enhance her investment decision-making skill and decides to use the fund's soft dollar account to pay the registration and exam fees for the CFA Program. Which of the following statements is *most likely* correct?

- **A** Stafford did not violate the Code and Standards because the prospectus informed investors of the fund's soft dollar policies.
- **B** Stafford violated the Code and Standards because improving her investment skills is not a reasonable use of the soft dollar account.
- **C** Stafford violated the Code and Standards because the CFA Program does not meet the definition of research allowed to be purchased with brokerage commissions.

30 Long has been asked to be the keynote speaker at an upcoming investment conference. The event is being hosted by one of the third-party investment managers currently used by his pension fund. The manager offers to cover all conference and travel costs for Long and make the conference registrations free for three additional members of his investment management team. To ensure that the conference obtains the best speakers, the host firm has arranged for an exclusive golf outing for the day following the conference on a local championship-caliber course. Which of the following is *least likely* to violate Standard I(B)?

A Long may accept only the offer to have his conference-related expenses paid by the host firm.

B Long may accept the offer to have his conference-related expenses paid and may attend the exclusive golf outing at the expense of the hosting firm.

C Long may accept the entire package of incentives offered to speak at this conference.

31 Andrews, a private wealth manager, is conducting interviews for a new research analyst for his firm. One of the candidates is Wright, an analyst with a local investment bank. During the interview, while Wright is describing his analytical skills, he mentions a current merger in which his firm is acting as the adviser. Andrews has heard rumors of a possible merger between the two companies, but no releases have been made by the companies concerned. Which of the following actions by Andrews is *least likely* a violation of the Code and Standards?

A Waiting until the next day before trading on the information to allow time for it to become public.

B Notifying all investment managers in his firm of the new information so none of their clients are disadvantaged.

C Placing the securities mentioned as part of the merger on the firm's restricted trading list.

32 Pietro, president of Local Bank, has hired the bank's market maker, Vogt, to seek a merger partner. Local is currently listed on a stock exchange and has not reported that it is seeking strategic alternatives. Vogt has discussed the possibility of a merger with several firms, but they have all decided to wait until after the next period's financial data are available. The potential buyers believe the results will be worse than the results of prior periods and will allow them to pay less for Local Bank.

Pietro wants to increase the likelihood of structuring a merger deal quickly. Which of the following actions would *most likely* be a violation of the Code and Standards?

A Pietro could instruct Local Bank to issue a press release announcing that it has retained Vogt to find a merger partner.

B Pietro could place a buy order for 2,000 shares (or four times the average weekly volume) through Vogt for his personal account.

C After confirming with Local's chief financial officer, Pietro could instruct Local to issue a press release reaffirming the firm's prior announced earnings guidance for the full fiscal year.

33 ABC Investment Management acquires a new, very large account with two concentrated positions. The firm's current policy is to add new accounts for the purpose of performance calculation after the first full month of management. Cupp is responsible for calculating the firm's performance returns. Before the end of the initial month, Cupp notices that one of the significant holdings of the new accounts is acquired by another company, causing the value of the investment to double. Because of this holding, Cupp decides to account for the new portfolio as of the date of transfer, thereby allowing ABC Investment to reap the positive impact of that month's portfolio return.

A Cupp did not violate the Code and Standards because the GIPS standards allow composites to be updated on the date of large external cash flows.

B Cupp did not violate the Code and Standards because companies are allowed to determine when to incorporate new accounts into their composite calculation.

C Cupp violated the Code and Standards because the inclusion of the new account produces an inaccurate calculation of the monthly results according to the firm's stated policies.

34 Cannan has been working from home on weekends and occasionally saves correspondence with clients and completed work on her home computer. Because of worsening market conditions, Cannan is one of several employees released by her firm. While Cannan is looking for a new job, she uses the files she saved at home to request letters of recommendation from former clients. She also provides to prospective clients some of the reports as examples of her abilities.

A Cannan violated the Code and Standards because she did not receive permission from her former employer to keep or use the files after her employment ended.

B Cannan did not violate the Code and Standards because the files were created and saved on her own time and computer.

C Cannan violated the Code and Standards because she is prohibited from saving files on her home computer.

35 Quinn sat for the Level III CFA exam this past weekend. He updates his resume with the following statement: "In finishing the CFA Program, I improved my skills related to researching investments and managing portfolios. I will be eligible for the CFA charter upon completion of the required work experience."

A Quinn violated the Code and Standards by claiming he improved his skills through the CFA Program.

B Quinn violated the Code and Standards by incorrectly stating that he is eligible for the CFA charter.

C Quinn did not violate the Code and Standards with his resume update.

36 During a round of golf, Rodriguez, chief financial officer of Mega Retail, mentions to Hart, a local investment adviser and long-time personal friend, that Mega is having an exceptional sales quarter. Rodriguez expects the results to be almost 10% above the current estimates. The next day, Hart initiates the purchase of a large stake in the local exchange-traded retail fund for her personal account.

A Hart violated the Code and Standards by investing in the exchange-traded fund that included Mega Retail.

B Hart did not violate the Code and Standards because she did not invest directly in securities of Mega Retail.

C Rodriguez did not violate the Code and Standards because the comments made to Hart were not intended to solicit an investment in Mega Retail.

37 Park is very frustrated after taking her Level II exam. While she was studying for the exam, to supplement the curriculum provided, she ordered and used study material from a third-party provider. Park believes the additional material focused her attention on specific topic areas that were not tested while ignoring other areas. She posts the following statement on the provider's discussion board: "I am very dissatisfied with your firm's CFA Program Level II material. I found the exam extremely difficult and myself unprepared for specific questions after using your product. How could your service provide such limited instructional resources on the analysis of inventories and taxes when the exam had multiple questions about them? I will not recommend your products to other candidates."

A Park violated the Code and Standards by purchasing third-party review material.

B Park violated the Code and Standards by providing her opinion on the difficulty of the exam.

C Park violated the Code and Standards by providing specific information on topics tested on the exam.

38 Paper was recently terminated as one of a team of five managers of an equity fund. The fund had two value-focused managers and terminated one of them to reduce costs. In a letter sent to prospective employers, Paper presents, with written permission of the firm, the performance history of the fund to demonstrate his past success.

A Paper did not violate the Code and Standards.

B Paper violated the Code and Standards by claiming the performance of the entire fund as his own.

C Paper violated the Code and Standards by including the historical results of his prior employer.

39 Townsend was recently appointed to the board of directors of a youth golf program that is the local chapter of a national not-for-profit organization. The program is beginning a new fund-raising campaign to expand the number of annual scholarships it provides. Townsend believes many of her clients make annual donations to charity. The next week in her regular newsletter to all clients, she includes a small section discussing the fund-raising campaign and her position on the organization's board.

A Townsend did not violate the Code and Standards.

B Townsend violated the Code and Standards by soliciting donations from her clients through the newsletter.

C Townsend violated the Code and Standards by not getting approval of the organization before soliciting her clients.

SOLUTIONS

1 The correct answer is B. This question involves Standard III(B)–Fair Dealing. Smith disseminated a change in the stock recommendation to his clients but then received a request contrary to that recommendation from a client who probably had not yet received the recommendation. Prior to executing the order, Smith should take additional steps to ensure that the customer has received the change of recommendation. Answer A is incorrect because the client placed the order prior to receiving the recommendation and, therefore, does not have the benefit of Smith's most recent recommendation. Answer C is also incorrect; simply because the client request is contrary to the firm's recommendation does not mean a member can override a direct request by a client. After Smith contacts the client to ensure that the client has received the changed recommendation, if the client still wants to place a buy order for the shares, Smith is obligated to comply with the client's directive.

2 The correct answer is C. This question involves Standard III(A)–Loyalty, Prudence, and Care and the specific topic of soft dollars or soft commissions. Answer C is the correct choice because client brokerage commissions may not be directed to pay for the investment manager's operating expenses. Answer B describes how members and candidates should determine how to use brokerage commissions—that is, if the use is in the best interests of clients and is commensurate with the value of the services provided. Answer A describes a practice that is commonly referred to as "directed brokerage." Because brokerage is an asset of the client and is used to benefit the client, not the manager, such practice does not violate a duty of loyalty to the client. Members and candidates are obligated in all situations to disclose to clients their practices in the use of client brokerage commissions.

3 The correct answer is C. This question involves Standard VI(A)–Disclosure of Conflicts. The question establishes a conflict of interest in which an analyst, Jamison, is asked to write a research report on a company that is a client of the analyst's employer. In addition, two directors of the company are senior officers of Jamison's employer. Both facts establish that there are conflicts of interest that must be disclosed by Jamison in her research report. Answer B is incorrect because an analyst is not prevented from writing a report simply because of the special relationship the analyst's employer has with the company as long as that relationship is disclosed. Answer A is incorrect because whether or not Jamison expresses any opinions in the report is irrelevant to her duty to disclose a conflict of interest. Not expressing opinions does not relieve the analyst of the responsibility to disclose the special relationships between the two companies.

4 The correct answer is C. This question asks about compliance procedures relating to personal investments of members and candidates. The statement in answer C clearly conflicts with the recommended procedures in the *Standards of Practice Handbook*. Employers should compare personal transactions of employees with those of clients on a regular basis regardless of the existence of a requirement by any regulatory organization. Such comparisons ensure that employees' personal trades do not conflict with their duty to their clients, and the comparisons can be conducted in a confidential manner. The statement in answer A does not conflict with the procedures in the *Handbook*. Disclosure of such policies will give full information to clients regarding potential conflicts of interest on the part of those entrusted to manage their money. Answer B is incorrect because firms are encouraged to establish policies whereby employees clear their personal holdings and transactions with their employers.

5 The correct answer is B. This question relates to Standard III(A)–Loyalty, Prudence, and Care and Standard III(E)–Preservation of Confidentiality. In this case, the member manages funds of a private endowment. Clients, who are, in this case, the trustees of the fund, must place some trust in members and candidates. Bronson cannot disclose confidential financial information to anyone without the permission of the fund, regardless of whether the disclosure may benefit the fund. Therefore, answer A is incorrect. Answer C is incorrect because Bronson must notify the fund and obtain the fund's permission before publicizing the information.

6 The correct answer is B. This question relates to Standard V(A)–Diligence and Reasonable Basis. The opinion of another financial analyst is not an adequate basis for Willier's action in changing the recommendation. Answer C is thus incorrect. So is answer A because, although it is true that members and candidates must distinguish between facts and opinions in recommendations, the question does not illustrate a violation of that nature. If the opinion overheard by Willier had sparked him to conduct additional research and investigation that justified a change of opinion, then a changed recommendation would be appropriate.

7 The correct answer is B. This question relates to Standard I(B)–Independence and Objectivity. When asked to change a recommendation on a company stock to gain business for the firm, the head of the brokerage unit must refuse in order to maintain his independence and objectivity in making recommendations. He must not yield to pressure by the firm's investment banking department. To avoid the appearance of a conflict of interest, the firm should discontinue issuing recommendations about the company. Answer A is incorrect; changing the recommendation in any manner that is contrary to the analyst's opinion violates the duty to maintain independence and objectivity. Answer C is incorrect because merely assigning a new analyst to decide whether the stock deserves a higher rating will not address the conflict of interest.

8 The correct answer is A. Standard VII(B)–Reference to CFA Institute, the CFA Designation, and the CFA Program is the subject of this question. The reference on Albert's business card implies that there is a "CFA Level II" designation; Tye merely indicates in promotional material that he is participating in the CFA Program and has completed Levels I and II. Candidates may not imply that there is some sort of partial designation earned after passing a level of the CFA exam. Therefore, Albert has violated Standard VII(B). Candidates may communicate that they are participating in the CFA Program, however, and may state the levels that they have completed. Therefore, Tye has not violated Standard VII(B).

9 The correct answer is B. This question relates to Standard V(B)–Communication with Clients and Prospective Clients. Scott has issued a research report stating that he expects the price of Walkton Industries stock to rise by US$8 a share "because the dividend will increase" by US$1.50 per share. He has made this statement knowing that the dividend will increase only if Congress enacts certain legislation, an uncertain prospect. By stating that the dividend will increase, Scott failed to separate fact from opinion.

The information regarding passage of legislation is not material nonpublic information because it is conjecture, and the question does not state whether the US representative gave Scott her opinion on the passage of the legislation in confidence. She could have been offering this opinion to anyone who asked. Therefore, statement A is incorrect. It may be acceptable to base a recommendation, in part, on an expectation of future events, even though they may be uncertain. Therefore, answer C is incorrect.

10 The correct answer is C. This question, which relates to Standard III(B)–Fair Dealing, tests the knowledge of the procedures that will assist members and candidates in treating clients fairly when making investment recommendations. The step listed in C will help ensure the fair treatment of clients. Answer A may have negative effects on the fair treatment of clients. The more people who know about a pending change, the greater the chance that someone will inform some clients before the information's release. The firm should establish policies that limit the number of people who are aware in advance that a recommendation is to be disseminated. Answer B, distributing recommendations to institutional clients before distributing them to individual accounts, discriminates among clients on the basis of size and class of assets and is a violation of Standard III(B).

11 The correct answer is B. This question deals with Standard II(A)–Material Nonpublic Information. The mosaic theory states that an analyst may use material public information and nonmaterial nonpublic information in creating a larger picture than shown by any individual piece of information and the conclusions the analyst reaches become material only after the pieces are assembled. Answers A and C are accurate statements relating to the Code and Standards but do not describe the mosaic theory.

12 The correct answer is C. This question involves Standard IV(B)–Additional Compensation Arrangements. The arrangement described in the question—whereby Jurgen would be compensated beyond the compensation provided by her firm, on the basis of an account's performance—is not a violation of the Standards as long as Jurgen discloses the arrangement in writing to her employer and obtains permission from her employer prior to entering into the arrangement. Answers A and B are incorrect; although the private compensation arrangement could conflict with the interests of other clients and lead to short-term performance pressures, members and candidates may enter into such agreements as long as they have disclosed the arrangements to their employer and obtained permission for the arrangement from their employer.

13 The correct answer is B. This question relates to Standard III(A)–Loyalty, Prudence, and Care—specifically, a member's or candidate's responsibility for voting proxies and the use of client brokerage. According to the facts stated in the question, Farnsworth did not violate Standard III(A). Although the company president asked Farnsworth to vote the shares of the Jones Corporation profit-sharing plan a certain way, Farnsworth investigated the issue and concluded, independently, the best way to vote. Therefore, even though his decision coincided with the wishes of the company president, Farnsworth is not in violation of his responsibility to be loyal and to provide care to his clients. In this case, the participants and the beneficiaries of the profit-sharing plan are the clients, not the company's management. Had Farnsworth not investigated the issue or had he yielded to the president's wishes and voted for a slate of directors that he had determined was not in the best interest of the company, Farnsworth would have violated his responsibilities to the beneficiaries of the plan. In addition, because the brokerage firm provides the lowest commissions and best execution for securities transactions, Farnsworth has met his obligations to the client in using this brokerage firm. It does not matter that the brokerage firm also provides research information that is not useful for the account generating the commission because Farnsworth is not paying extra money of the client's for that information.

14 The correct answer is A. In this question, Brown is providing investment recommendations before making inquiries about the client's financial situation, investment experience, or investment objectives. Brown is thus violating

Standard III(C)–Suitability. Answers B and C provide examples of information members and candidates should discuss with their clients at the outset of the relationship, but these answers do not constitute a complete list of those factors. Answer A is the best answer.

15 The correct answer i*s* C. This question involves Standard I(C)–Misrepresentation. Statement I is a factual statement that discloses to clients and prospects accurate information about the terms of the investment instrument. Statement II, which guarantees a specific rate of return for a mutual fund, is an opinion stated as a fact and, therefore, violates Standard I(C). If statement II were rephrased to include a qualifying statement, such as "in my opinion, investors may earn . . ., " it would not be in violation of the Standards.

16 The correct answer is A. This question involves three of the Standards. Anderb, the portfolio manager, has been obtaining more favorable prices for her personal securities transactions than she gets for her clients, which is a breach of Standard III(A)–Loyalty, Prudence, and Care. In addition, she violated Standard I(D)–Misconduct by failing to adhere to company policy and by hiding her personal transactions from her firm. Anderb's supervisor, Bates, violated Standard IV(C)–Responsibilities of Supervisors; although the company had requirements for reporting personal trading, Bates failed to adequately enforce those requirements. Answer B does not represent a violation because Standard VI(B)–Priority of Transactions requires that personal trading in a security be conducted after the trading in that security of clients and the employer. The Code and Standards do not prohibit owning such investments, although firms may establish policies that limit the investment opportunities of members and candidates. Answer C does not represent a violation because the Code and Standards do not contain a prohibition against employees using the same broker for their personal accounts that they use for their client accounts. This arrangement should be disclosed to the employer so that the employer may determine whether a conflict of interest exists.

17 The correct answer is A because this question relates to Standard I(A)–Knowledge of the Law—specifically, global application of the Code and Standards. Members and candidates who practice in multiple jurisdictions may be subject to various securities laws and regulations. If applicable law is more strict than the requirements of the Code and Standards, members and candidates must adhere to applicable law; otherwise, members and candidates must adhere to the Code and Standards. Therefore, answer A is correct. Answer B is incorrect because members and candidates must adhere to the higher standard set by the Code and Standards if local applicable law is less strict. Answer C is incorrect because when no applicable law exists, members and candidates are required to adhere to the Code and Standards, and the Code and Standards prohibit the use of material nonpublic information.

18 The correct answer is B. The best course of action under Standard I(B)–Independence and Objectivity is to avoid a conflict of interest whenever possible. Therefore, for Ward to pay for all his expenses is the correct answer. Answer C details a course of action in which the conflict would be disclosed, but the solution is not as appropriate as avoiding the conflict of interest. Answer A would not be the best course because it would not remove the appearance of a conflict of interest; even though the report would not be affected by the reimbursement of expenses, it could appear to be.

19 The correct answer is B. Under Standard IV(A)–Loyalty, members and candidates may undertake independent practice that may result in compensation or other benefit in competition with their employer as long as they obtain consent from their employer. Answer C is not consistent with the Standards because the Standards allow members and candidates to make arrangements or preparations to go into competitive business as long as those arrangements do not interfere with their duty to their current employer. Answer A is not consistent with the Standards because the Standards do not include a complete prohibition against undertaking independent practice.

20 The correct answer is B. This question involves Standard VI(A)–Disclosure of Conflicts—specifically, the holdings of an analyst's employer in company stock. Answers A and C do not describe conflicts of interest that Smith would have to disclose. Answer A describes the use of a firm's products, which would not be a required disclosure. In answer C, the relationship between the analyst and the company through a relative is so tangential that it does not create a conflict of interest necessitating disclosure.

21 The correct answer is C. This question relates to Standard I(C)–Misrepresentation. Although Michelieu's statement about the total return of his clients' accounts on average may be technically true, it is misleading because the majority of the gain resulted from one client's large position taken against Michelieu's advice. Therefore, this statement misrepresents the investment performance the member is responsible for. He has not taken steps to present a fair, accurate, and complete presentation of performance. Answer B is thus incorrect. Answer A is incorrect because although Michelieu is not guaranteeing future results, his words are still a misrepresentation of his performance history.

22 The correct answer is B. The best policy to prevent violation of Standard II(A)–Material Nonpublic Information is the establishment of firewalls in a firm to prevent exchange of insider information. The physical and informational barrier of a firewall between the investment banking department and the brokerage operation prevents the investment banking department from providing information to analysts on the brokerage side who may be writing recommendations on a company stock. Prohibiting recommendations of the stock of companies that are clients of the investment banking department is an alternative, but answer A states that this prohibition would be permanent, which is not the best answer. Once an offering is complete and the material nonpublic information obtained by the investment banking department becomes public, resuming publishing recommendations on the stock is not a violation of the Code and Standards because the information of the investment banking department no longer gives the brokerage operation an advantage in writing the report. Answer C is incorrect because no exchange of information should be occurring between the investment banking department and the brokerage operation, so monitoring of such exchanges is not an effective compliance procedure for preventing the use of material nonpublic information.

23 The correct answer is B. Under Standard III(A)–Loyalty, Prudence, and Care, members and candidates who manage a company's pension fund owe these duties to the participants and beneficiaries of the pension plan, not the management of the company or the company's shareholders.

24 The correct answer is B. Answer B gives one of the two primary reasons listed in the *Handbook* for disclosing referral fees to clients under Standard VI(C)–Referral Fees. (The other is to allow clients and employers to evaluate the full cost of the services.) Answer A is incorrect because Standard VI(C) does not require members or candidates to discount their fees when they receive referral fees. Answer C is inconsistent with Standard VI(C) because disclosure of referral fees, to be effective, should be made to prospective clients before entering into a formal client relationship with them.

25 The correct answer is B. Standard VI(B)–Priority of Transactions does not limit transactions of company employees that differ from current recommendations as long as the sale does not disadvantage current clients. Thus, answer A is incorrect. Answer C is incorrect because the Standard does not require the matching of personal and client trades.

26 Answer C is correct. Standard IV(A)–Loyalty discusses activities permissible to members and candidates when they are leaving their current employer; soliciting clients is strictly prohibited. Thus, answer A is inconsistent with the Code and Standards even with the required disclosure. Answer B is incorrect because the offer does not directly violate the Code and Standards. There may be out-of-work members and candidates who can arrange the necessary commitments without violating the Code and Standards.

27 Answer A is correct. The question relates to Standard III(A)–Loyalty, Prudence, and Care. Carter believes the broker offers effective execution at a fee that is comparable with those of other brokers, so he is free to use the broker for all accounts. Answer B is incorrect because the accounts that prohibit soft dollar arrangements do not want to fund the purchase of research by Carter. The new trading scheme does not incur additional commissions from clients, so it would not go against the prohibitions. Answer C is incorrect because Carter should not incur unnecessary or excessive "churning" of the portfolios (excessive trading) for the purpose of meeting the brokerage commitments of soft dollar arrangements.

28 Answer B is correct according to Standard VII(B)–Reference to CFA Institute, the CFA Designation, and the CFA Program. CFA Program candidates do not receive their actual scores on the exam. Topic and subtopic results are grouped into three broad categories, and the exam is graded only as "pass" or "fail." Although a candidate may have achieved a topical score of "above 70%," she or he cannot factually state that she or he received the highest possible score because that information is not reported. Thus, answer C is incorrect. Answer A is incorrect as long as the member or candidate actually completed the exams consecutively. Standard VII(B) does not prohibit the communication of factual information about completing the CFA Program in three consecutive years.

29 Answer C is correct. According to Standard III(A)–Loyalty, Prudence, and Care, the CFA Program would be considered a personal or firm expense and should not be paid for with the fund's brokerage commissions. Soft dollar accounts should be used only to purchase research services that directly assist the investment manager in the investment decision-making process, not to assist the management of the firm or to further education. Thus, answer A is incorrect. Answer B is incorrect because the reasonableness of how the money is used is not an issue; the issue is that educational expense is not research.

30 Answer A is correct. Standard I(B)–Independence and Objectivity emphasizes the need for members and candidates to maintain their independence and objectivity. Best practices dictate that firms adopt a strict policy not to accept compensation for travel arrangements. At times, however, accepting paid travel would not compromise one's independence and objectivity. Answers B and C are incorrect because the added benefits—free conference admission for additional staff members and an exclusive golf retreat for the speaker—could be viewed as inducements related to the firm's working arrangements and not solely related to the speaking engagement. Should Long wish to bring other team members or participate in the golf outing, he or his firm should be responsible for the associated fees.

31 Answer C is correct. The guidance to Standard II(A)–Material Nonpublic Information recommends adding securities to the firm's restricted list when the firm has or may have material nonpublic information. By adding these securities to this list, Andrews would uphold this standard. Because waiting until the next day will not ensure that news of the merger is made public, answer A is incorrect. Negotiations may take much longer between the two companies, and the merger may never happen. Andrews must wait until the information is disseminated to the market before he trades on that information. Answer B is incorrect because Andrews should not disclose the information to other managers; no trading is allowed on material nonpublic information.

32 Answer B is correct. Through placing a personal purchase order that is significantly greater than the average volume, Pietro is violating Standard IIB–Market Manipulation. He is attempting to manipulate an increase in the share price and thus bring a buyer to the negotiating table. The news of a possible merger and confirmation of the firm's earnings guidance may also have positive effects on the price of Local Bank, but Pietro's actions in instructing the release of the information does not represent a violation through market manipulation. Announcements of this nature are common and practical to keep investors informed. Thus, answers A and C are incorrect.

33 Answer C is correct. Cupp violated Standard III(D)–Performance Presentations when he deviated from the firm's stated policies solely to capture the gain from the holding being acquired. Answer A is incorrect because the firm does not claim GIPS compliance and the GIPS standards require external cash flows to be treated in a consistent manner with the firm's documented policies. Answer B is incorrect because the firm does not state that it is updating its composite policies. If such a change were to occur, all cash flows for the month would have to be reviewed to ensure their consistent treatment under the new policy.

34 Answer A is correct. According to Standard V(C)–Record Retention, Cannan needed the permission of her employer to maintain the files at home after her employment ended. Without that permission, she should have deleted the files. All files created as part of a member's or candidate's professional activity are the property of the firm, even those created outside normal work hours. Thus, answer B is incorrect. Answer C is incorrect because the Code and Standards do not prohibit using one's personal computer to complete work for one's employer.

35 Answer B is correct. According to Standard VII(B)–Reference to CFA Institute, the CFA Designation, and the CFA Program, Quinn cannot claim to have finished the CFA Program or be eligible for the CFA charter until he officially learns that he has passed the Level III exam. Until the results for the most recent exam are released, those who sat for the exam should continue to refer to themselves as "candidates." Thus, answer C is incorrect. Answer A is incorrect because members and candidates may discuss areas of practice in which they believe the CFA Program improved their personal skills.

36 Answer A is correct. Hart's decision to invest in the retail fund appears directly correlated with Rodriguez's statement about the successful quarter of Mega Retail and thus violates Standard II(A)–Material Nonpublic Information. Rodriguez's information would be considered material because it would influence the share price of Mega Retail and probably influence the price of the entire exchange-traded retail fund. Thus, answer B is incorrect. Answer C is also incorrect because Rodriguez shared information that was both material and nonpublic. Company officers regularly have such knowledge about their firms, which is not a violation. The sharing of such information, however, even in a conversation between friends, does violate Standard II(A).

37 Answer C is correct. Standard VII(A)–Conduct as Members and Candidates in the CFA Program prohibits providing information to candidates or the public that is considered confidential to the CFA Program. In revealing that questions related to the analysis of inventories and analysis of taxes were on the exam, Park has violated this standard. Answer B is incorrect because the guidance for the standard explicitly acknowledges that members and candidates are allowed to offer their opinions about the CFA Program. Answer A is incorrect because candidates are not prohibited from using outside resources.

38 Answer B is correct. Paper has violated Standard III(D)–Performance Presentation by not disclosing that he was part of a team of managers that achieved the results shown. If he had also included the return of the portion he directly managed, he would not have violated the standard. Thus, answer A is incorrect. Answer C is incorrect because Paper received written permission from his prior employer to include the results.

39 Answer A is correct. Townsend has not provided any information about her clients to the leaders or managers of the golf program; thus, she has not violated Standard III(E)–Preservation of Confidentiality. Providing contact information about her clients for a direct-mail solicitation would have been a violation. Answer B is incorrect because the notice in the newsletter does not violate Standard III(E). Answer C is incorrect because the golf program's fund-raising campaign had already begun, so discussing the opportunity to donate was appropriate.

READING

59

Introduction to the Global Investment Performance Standards (GIPS)

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** explain why the GIPS standards were created, what parties the GIPS standards apply to, and who is benefitted by the standards;
☐	**b.** describe the key concepts of the GIPS standards for firms;
☐	**c.** explain the purpose of composites in performance reporting;
☐	**d.** describe the fundamentals of compliance, including the recommendations of the GIPS Standards with respect to the definition of the firm and the firm's definition of discretion;
☐	**e.** describe the concept of independent verification.

The objective of this reading is to provide candidates with an orientation to the GIPS standards. It explains why the GIPS standards were created, who can claim compliance, and who benefits from compliance. It also covers key concepts of the GIPS standards—composites, the definition of the firm, and the definition of investment discretion. Finally, the reading briefly discusses the purpose and benefits of verification. Upon completion of this reading, candidates should appreciate the benefits of an industry-wide set of standards for calculating and presenting investment performance based on the principles of fair representation and full disclosure.

The 2020 Edition of the GIPS Standards has three chapters:

1. GIPS Standards for Firms
2. GIPS Standards for Asset Owners
3. GIPS Standards for Verifiers

Organizations that compete for business must comply with the GIPS Standards for Firms.

© 2020 CFA Institute. All rights reserved.

Candidates are also responsible for reading the sections of the GIPS standards for Firms specifically referenced in this reading. A complete copy of the 2020 GIPS standards for firms can be found here: https://www.cfainstitute.org/en/ethics/codes/gips-standards/firms.

I. WHY WERE THE GIPS CREATED, WHO CAN CLAIM COMPLIANCE & WHO BENEFITS FROM COMPLIANCE

a explain why the GIPS standards were created, what parties the GIPS standards apply to, and who is benefitted by the standards;

b describe the key concepts of the GIPS standards for firms;

The mission of the GIPS standards is to promote ethics and integrity and instill trust through the use of the GIPS standards by achieving universal demand for compliance by asset owners, adoption by asset managers, and support from regulators for the ultimate benefit of the global investment community.

Institutions and individuals are constantly scrutinizing past investment performance returns in search of the best manager to achieve their investment objectives.

In the past, the investment community had great difficulty making meaningful comparisons on the basis of accurate investment performance data. Several performance measurement practices hindered the comparability of performance returns from one firm to another, while others called into question the accuracy and credibility of performance reporting overall. Misleading practices included:

- *Representative Accounts*: Selecting a top-performing portfolio to represent the firm's overall investment results for a specific mandate.
- *Survivorship Bias*: Presenting an "average" performance history that *excludes* portfolios whose poor performance was weak enough to result in termination of the firm.
- *Varying Time Periods*: Presenting performance for a selected time period during which the mandate produced excellent returns or out-performed its benchmark—making comparison with other firms' results difficult or impossible.

Making a valid comparison of investment performance among even the most ethical investment management firms was problematic. For example, a pension fund seeking to hire an investment management firm might receive proposals from several firms, possibly from different countries, all using different methodologies for calculating their results.

The GIPS standards are a practitioner-driven set of ethical principles that establish a standardized, industry-wide approach for investment firms to follow in calculating and presenting their historical investment results to prospective clients. The GIPS standards ensure fair representation and full disclosure of investment performance. In other words, the GIPS standards lead investment management firms to avoid misrepresentations of performance and to communicate all relevant information that prospective clients should know in order to evaluate past results.

The objectives of the GIPS standards are as follows:

- Promote investor interests and instill investor confidence.
- Ensure accurate and consistent data.
- Obtain worldwide acceptance of a single standard for calculating and presenting performance.

- Promote fair, global competition among investment firms.
- Promote industry self-regulation on a global basis.

I.I Who Can Claim Compliance?

Any firm that *actually manages assets* may choose to comply with the GIPS standards. Consultants cannot make a claim of compliance unless they actually manage assets for which they are making a claim of compliance. They can claim to endorse the standards and/or require that their investment managers comply with the standards. Similarly, software (and the vendors who supply software) cannot be "compliant." Software can assist firms in achieving compliance with the GIPS standards (e.g., by calculating performance in a manner consistent with the calculation requirements of the Standards), but only a firm managing assets can claim compliance once the firm has satisfied all requirements of the standards.

Asset owners may comply with the GIPS standards in the same way as firms if they compete for business. If they don't compete for business but report their performance to an oversight body, asset owners may choose to comply with the GIPS standards for asset owners.

Compliance is a firm-wide process that cannot be achieved on a single product or composite. A firm has only two options with regard to compliance with the GIPS standards: fully comply with *all* requirements of the GIPS standards and claim compliance through the use of the GIPS Compliance Statement; or not comply with all requirements of the GIPS standards and not claim compliance with, or make any reference to, the GIPS standards.

Complying with the GIPS standards is voluntary. Compliance with the GIPS standards is not typically required by legal or regulatory authorities.

I.II Who Benefits from Compliance?

The GIPS standards benefit asset managers and their prospective clients and asset owners and their oversight bodies.

- By choosing to comply with the GIPS standards, firms assure prospective clients that the historical "track record" they report is both complete and fairly presented. Compliance enables the GIPS-compliant firm to participate in competitive bids against other compliant firms throughout the world. Achieving and maintaining compliance may also strengthen the firm's internal controls over performance-related processes and procedures.
- Investors have a greater level of confidence in the integrity of performance presentations of a GIPS-compliant firm and can more easily compare performance presentations from different investment management firms. The GIPS standards certainly do not eliminate the need for in-depth due diligence on the part of the investor, but compliance with the Standards enhances the credibility of investment management firms that have chosen to undertake this responsibility.
- Asset owners provide performance information to their oversight bodies that allows them to make investment decisions and evaluate the performance of the funds under their supervision. Particularly where asset owners require their external managers to comply with the GIPS standards, reporting to the oversight body using the same principles facilitates the understanding of the sources of risk and excess return in the funds under supervision.

Please read the Preface and the Introduction to the Global Investment Performance Standards for Firms for additional insight into the history, purpose, and key concepts of the GIPS standards.

II. COMPOSITES

c explain the purpose of composites in performance reporting;

One of the key concepts of the standards is the required use of composites. A composite is an aggregation of one or more portfolios managed according to a similar investment mandate, objective, or strategy. The requirement to create, use and maintain composites is designed to prevent firms from cherry-picking—using the best-performing accounts to represent the performance of an investment strategy. A composite must include all actual, fee-paying, discretionary portfolios managed in accordance with the same investment mandate, objective, or strategy. For example, if a GIPS-compliant firm presents its track record for a Global Equity Composite (the Composite), the Composite must include all fee-paying portfolios that are managed, or have historically been managed, in the firm's Global Equity strategy. The firm may not subjectively select which Global Equity portfolios will be included in or excluded from the calculation and presentation of the Global Equity Composite. The determination of which portfolios to include in the Composite should be done according to pre-established criteria (i.e., on an ex-ante basis), not after the fact. This prevents a firm from including only their best-performing portfolios in the Composite.

A claim of compliance requires that all fee-paying discretionary accounts managed by the firm be included in at least one composite.

Please read Section 3.A. of the Global Investment Performance Standards for Firms on composites.

III. FUNDAMENTALS OF COMPLIANCE

d describe the fundamentals of compliance, including the recommendations of the GIPS Standards with respect to the definition of the firm and the firm's definition of discretion;

Several core principles create the foundation for the GIPS standards, including properly defining the firm, providing GIPS Reports to all prospective clients and prospective pooled fund investors, adhering to applicable laws and regulations, and ensuring that information presented is not false or misleading. Two important issues that a firm must consider when becoming compliant with the GIPS standards are the definition of the firm and the firm's definition of discretion. The definition of the firm is the foundation for firm-wide compliance and creates defined boundaries whereby total firm assets can be determined.

The GIPS standards state "The firm should adopt the broadest, most meaningful definition of the firm. The scope of this definition should include all geographical (country, regional, etc.) offices operating under the same brand name, regardless of the actual name of the individual investment management company."

The firm's definition of discretion establishes criteria to judge which portfolios must be included in a composite and is based on the firm's ability to implement its investment strategies. If documented client-imposed restrictions interfere with the implementation of the intended strategy to the extent that the portfolio is no longer representative of the strategy, the firm may determine that the portfolio is non-discretionary. Non-discretionary portfolios must not be included in a firm's composites.

Section 1 of the 2020 GIPS Standards for Firms addresses the fundamentals of compliance in more detail.

VERIFICATION

IV.

e describe the concept of independent verification.

Firms that claim compliance with the GIPS standards are responsible for their claim of compliance and for maintaining that compliance. That is, firms self-regulate their claim of compliance. Once a firm claims compliance with the standards, they may voluntarily hire an independent third party to perform a verification in order to increase confidence in the firm's claim of compliance. Verification may also increase the knowledge of the firm's performance measurement team and improve the consistency and quality of the firm's compliant presentations.

Verification is a process by which an independent verification firm (verifier) conducts testing of a firm on a firm-wide basis in accordance with the required verification procedures of the GIPS standards. Verification provides assurance on whether the firm's policies and procedures related to composite and pooled fund maintenance, as well as the calculation, presentation, and distribution of performance, have been designed in compliance with the GIPS standards and have been implemented on a firm-wide basis.

Verification is performed with respect to an entire firm, not on specific composites. Verification does not ensure the accuracy of any specific composite presentation.

Verification must be performed by an independent third party. A firm cannot perform its own verification.

Third-party verification brings additional credibility to a firm's claim of compliance. A verified firm may provide existing and prospective clients with greater assurance about its claim of compliance with the GIPS standards. Verification may also provide improved internal processes and procedures as well as marketing advantages to the firm.

PRACTICE PROBLEMS

1 A firm that does not adopt the GIPS standards could mischaracterize its overall performance by presenting a performance history:

- **A** that includes terminated portfolios.
- **B** composed of a single top-performing portfolio.
- **C** for an investment mandate over all periods since the firm's inception.

2 Which of the following statements regarding GIPS compliance is correct?

- **A** Asset owners that manage assets can claim compliance with the GIPS Standards.
- **B** Software that calculates performance in a manner consistent with the GIPS standards can claim compliance with GIPS standards.
- **C** Firms can comply with the GIPS standards by limiting their compliance claims to the provisions they have chosen to follow.

3 Each composite of a GIPS-compliant firm must consist of:

- **A** multiple portfolios.
- **B** portfolios selected on an *ex post* basis.
- **C** portfolios managed according to a similar investment strategy.

4 Verification:

- **A** must be performed on a firm-wide basis.
- **B** may be provided by the firm's compliance department.
- **C** ensures the accuracy of a specific composite presentation.

© 2020 CFA Institute. All rights reserved.

SOLUTIONS

1 B is correct. Selecting a top-performing portfolio to represent a firm's overall investment results for a specific mandate, also known as using representative accounts, is a misleading practice that is not allowed under the GIPS standards. A is incorrect because including terminated portfolios is consistent with the GIPS standards. If the firm instead presented a performance history that excludes terminated portfolios, however, such a practice would be misleading and not allowed under the GIPS standards. C is incorrect because presenting performance for its mandate covering all periods since the firm's inception is consistent with the GIPS standards. If the firm instead presented performance for a selected period during which it produced excellent returns or outperformed its benchmark, however, such a practice would be misleading and not allowed under the GIPS standards.

2 A is correct. Asset owners can make a claim of compliance if they actually manage assets for which they are making a claim of compliance. B is incorrect because software (and the vendors that supply software) cannot be GIPS compliant. Software can assist firms in achieving compliance with the GIPS standards, but only an asset manager can claim compliance. C is incorrect because a firm has only two options regarding compliance with the GIPS standards: fully comply with all requirements of the GIPS standards and claim compliance through the use of the GIPS Compliance Statement; or not comply with all requirements of the GIPS standards and not claim compliance with, or make any reference to, the GIPS standards.

3 C is correct. A composite is an aggregation of one or more portfolios managed according to a similar investment mandate, objective, or strategy. For example, if a GIPS-compliant firm presents performance for a global equity composite (the composite), the composite must include portfolios that are managed, or have historically been managed, according to the firm's global equity strategy. A is incorrect because a composite is an aggregation of one or more portfolios managed according to a similar investment mandate, objective, or strategy. A composite may consist of a single portfolio when it is the only portfolio managed according to a particular mandate. B is incorrect because the determination of which portfolio(s) to include in a composite should be done according to pre-established criteria (*ex ante* basis), not after the fact (*ex post* basis).

4 A is correct. Verification is performed with respect to an entire firm, not on specific composites. B is incorrect because verification must be performed by an independent third party. C is incorrect because verification provides assurance on whether the firm's policies and procedures related to composite and pooled fund maintenance, as well as the calculation, presentation, and distribution of performance, have been designed in compliance with the GIPS standards and have been implemented on a firm-wide basis; it does not ensure the accuracy of a specific composite presentation.

READING

60

Ethics Application

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** evaluate practices, policies, and conduct relative to the CFA Institute Code of Ethics and Standards of Professional Conduct;
☐	**b.** explain how the practices, policies, and conduct do or do not violate the CFA Institute Code of Ethics and Standards of Professional Conduct.

INTRODUCTION

This reading presents a number of short vignettes, or scenarios, inspired by real-world situations and events.[1] After reading the facts of each scenario, use your knowledge of the CFA Institute Code of Ethics and Standards of Professional Conduct to choose the best response to the multiple-choice question. After making your choice, be sure to review the correct response and case analysis, which discusses the rationale for why or why not a violation of the Code and Standards might have taken place and conduct that would comply with the Code and Standards.

PROFESSIONALISM

I.

A. Knowledge of the Law

> Members and Candidates must understand and comply with all applicable laws, rules, and regulations (including the CFA Institute Code of Ethics and Standards of Professional Conduct) of any government, regulatory organization, licensing agency, or professional association governing their professional activities. In the event of conflict, Members and Candidates must comply with the more strict law, rule, or regulation. Members and Candidates must not knowingly participate or assist in and must dissociate from any violation of such laws, rules, or regulations.

---

[1] Based on cases from the *Ethics in Investment Management Casebook* (CFA Institute 2019).

© 2021 CFA Institute. All rights reserved.

Mandracken

SBS Bank (SBS) serves as a custody bank for a wide range of clients. SBS offers a variety of services to its clients, including custody, clearing, payment, settlement, and record keeping. SBS charges its clients an asset-based fee for these services. Pursuant to the bank's client agreement, custody clients agree to reimburse the bank for out-of-pocket expenses for items paid by the custodian on their behalf. The majority of these expenses are for messages sent via the Society for Worldwide Interbank Financial Telecommunication (SWIFT), a secure messaging network used by banks and other financial institutions. Although SBS charges custody clients an established rate for SWIFT messages, the rate is greater than the actual cost of providing this service.

Mandracken, CFA, a vice president at SBS who oversees client service responsibilities, recognizes this discrepancy and brings it to the attention of his supervisor. In an email, Mandracken states that "although disclosure of charging for SWIFT fees is noted in the clients' fee schedules, the fees have always included an increase over actual cost, so the charge to clients is not a true pass-through because we add a margin." Mandracken's supervisor instructs him to reduce the SWIFT fee rate for new clients and to revisit the rate for existing clients when their contracts are renewed.

To meet his obligations under the CFA Institute Code and Standards, Mandracken should

A implement the corrective procedures as directed by his supervisor.

B implement the corrective procedures as directed by his supervisor but report his objections to the bank's board of directors.

C refuse to participate in any client interactions using the fee schedule until the bank revises the SWIFT rate to reflect the actual cost of the service.

Analysis C is correct. This case involves how to appropriately address the misconduct of others in carrying out your professional responsibilities. Standard I(A): Professionalism, Knowledge of the Law prohibits CFA Institute members and candidates from knowingly participating or assisting in legal or ethical violations and requires them to dissociate from any such activity. SBS is misrepresenting its reimbursable expenses to its custody clients and overcharging them. Although Mandracken brings the issue to the attention of his supervisor, his supervisor's corrective measures are inadequate because they (1) address the issue only for new clients, (2) do not immediately address the issue for existing clients, and (3) do not address the misrepresentation and overcharges to past clients. Mandracken must thus refuse to participate in any client interactions that use the fee schedule until the bank revises the SWIFT fees charged to custody clients to reflect actual out-of-pocket costs.

A is incorrect. Under the Code and Standards, Mandracken cannot continue to participate in or be associated with this misconduct. Inaction by a member and continuing association with those involved in illegal or unethical conduct might be viewed as participation or assistance in the illegal or unethical conduct.

B is incorrect. Mandracken might need to take drastic measures to dissociate from the activity and to protect client interests, such as leaving the bank or reporting the misconduct to the bank's board of directors or regulators. Several interim steps should be considered before this action, such as talking to his supervisor or the bank's compliance department about the inadequacy of the proposed corrective measures, but Mandracken cannot continue to interact with clients using a fraudulent fee schedule.

Pellie

Pellie, CFA, is CEO of Kwaume Investment Group (KIG), an investment adviser that is a wholly owned subsidiary of Kwaume Bank. A longtime bank customer that Pellie and some of the bank's board members know personally recently opened an investment account at KIG with a stated investment objective of earning income. The client made a few investments over the course of the next year, but most of the activity in the account involved several hundred bank transactions that totaled $90 million in deposits and $84 million in withdrawals.

The transactions included electronic transfers to and from individuals and entities located in bank secrecy havens and countries identified by the government as at risk for money-laundering activity. Pellie knew that the client was engaged in international business pursuits involving transactions with a higher risk potential for corruption and bribery. Given the client's longstanding relationship with the bank, Pellie assumes the transactions have a legitimate business purpose and accepts vague descriptions, such as "for services provided," "consulting fees," and "commissions." When he receives the daily anti-money-laundering (AML) reports, which are required by law when transactions trigger red flags of potentially suspicious activity, he approves them without further inquiry.

Pellie's actions are

A a violation of the CFA Institute Code and Standards.

B appropriate because Pellie is protecting the confidentiality of client information.

C appropriate because Pellie can rely on the account's clearing firm to report suspicious activity for the account.

Analysis A is correct. Pellie's actions violate Standard I(A): Professionalism, Knowledge of the Law, which states that CFA Institute members and candidates "must understand and comply with all applicable laws, rules, and regulations . . . governing their professional activities." The facts presented should have raised questions regarding the legitimacy of the client's account. The high volume of deposits and withdrawals combined with the low number of investment transactions is inconsistent with an investment account having a stated objective of earning income. The transactions in the account appeared to be high-risk transactions for money-laundering activity and should have received a greater level of scrutiny by Pellie. But rather than investigate as required by law, Pellie did not inquire further because of the client's longstanding relationship with the bank and its board members. Pellie's failure to comply with AML requirements imposed by law violates Standard I(A).

B is incorrect. The duty of loyalty to clients and the preservation of confidentiality of client information cannot be used as justification to allow clients to violate the law or otherwise damage the integrity or viability of global capital markets.

C is incorrect. Pellie cannot rely on the clearing firm to meet KIG's obligation to review the transactions for suspicious activity because that is his responsibility.

Mwangi

Mwangi, CFA, works for a firm that sells insurance products. Three of Mwangi's clients purchase one type of product (Class A) but later change their minds and ask to switch to another, lower priced product (Class B). For Mwangi to complete the transaction, the law requires that she have her clients sign new sale and purchase documents for the Class B product. Given that not all the documents are ready for signing at the time of the clients' request, Mwangi advises her clients to wait to sign until the complete set of documents has been prepared. When the full set of documents is ready, Mwangi tries unsuccessfully to reach her clients for their signatures. Because of the missing signatures, Mwangi's manager threatens to cancel the exchange, which, because of other investment purchases, would place the clients' accounts into an overdraft position. Under the firm's policies, account shortfalls are covered by selling account assets once the shortfall has been outstanding for two weeks. To prevent this from happening, Mwangi signs the necessary documents on behalf of her clients.

Mwangi's actions are

A a violation of the CFA Institute Code and Standards.

B acceptable because her clients had already given their permission for the exchange to be made.

C acceptable if the clients gave Mwangi explicit permission to sign the documents on their behalf.

Analysis A is correct. This case involves Standard I(A): Professionalism, Knowledge of the Law, which requires CFA Institute members and candidates to "comply with all applicable laws, rules, and regulations . . . governing their professional activities." To complete the exchange from the Class A to the Class B product, Mwangi's clients were legally required to sign the new sale and purchase documents, which did not happen. By signing the documents on behalf of the clients, Mwangi forged their signatures. This violation of the law constitutes a violation of Standard I(A).

B is incorrect. General approval, written or verbal, of the transaction by the clients is insufficient to meet the legal requirement for client signatures.

C is incorrect. Even if the clients fully understand what is required to make the switch and explicitly give Mwangi permission to sign the forms on their behalf, the law requires that the genuine signatures of the clients be on the documents. Although the intent of the law is to protect the clients, and the clients would be waiving their rights, Mwangi is still not permitted to circumvent the legal requirements.

B. Independence and Objectivity

> Members and Candidates must use reasonable care and judgment to achieve and maintain independence and objectivity in their professional activities. Members and Candidates must not offer, solicit, or accept any gift, benefit, compensation, or consideration that reasonably could be expected to compromise their own or another's independence and objectivity.

Myers

Myers, CFA, is a partner at Corboba, a $3 billion hedge fund that focuses on environmental, social, and governance investments. To support upcoming public office elections, Myers wants to personally donate $10,000 to one of the candidates, DeFrietas. As a passionate climate advocate and an avid proponent of responsible investment, Myers supports DeFrietas's backing of environmental policies to reduce air pollution and mitigate the effects of climate change. Myers believes his political contribution might also be beneficial for Corboba because DeFrietas is running for a position that can influence which hedge funds receive investments from the state's pension plans.

Myers's best course of action is to

A refrain from donating to DeFrietas.

B donate to DeFrietas because he is using personal funds in an amount that is insignificant relative to the size of the hedge fund.

C donate to DeFrietas because he will be supporting a candidate whose environmental policies align with his beliefs.

Analysis A is correct. If Myers makes the donation to DeFrietas's campaign, his actions would potentially violate Standard I(B): Professionalism, Independence and Objectivity, which states that CFA Institute members and candidates "must not offer . . . any gift, benefit, compensation, or consideration that reasonably could be expected to compromise . . . another's independence and objectivity." Although Myers has a legitimate reason for donating to DeFrietas, the facts clearly indicate that he believes the donation might also influence DeFrietas to choose Corboba to manage money for the state's pension plans. Standard I(B) prohibits CFA Institute members and candidates from attempting to gain lucrative allocations from government-sponsored pension funds by making donations to the political campaigns of individuals directly responsible for the manager hiring decisions.

B is incorrect. The source and the amount of the funds are irrelevant if the donation is meant to influence DeFrietas. In addition, the sum of $10,000 might be significant in the context of the campaign and thus important to DeFrietas. The critical issue is whether the donation is reasonably designed to improperly affect DeFrietas's independence and objectivity and benefit Myers or Corboba directly. Many factors would go into this determination, including the size of the donation, Myers's intent in making the donation, and how influential DeFrietas's position would be in making the hiring decision. The facts state that one of the reasons for Myers's donation is to have influence over receiving investment management business from the state.

C is incorrect. Myers seems to have both proper and improper motivations for making a political contribution to DeFrietas. Myers supports DeFrietas's positions on protecting the environment and wants to further those goals. Standard I(B) is not meant to prevent investment professionals from participating in the political process through financial or other support for candidates. But Myers recognizes that financial support of DeFrietas could benefit Corboba by gaining favor with someone who might be in a position to determine whether to invest in Corboba's fund. Therefore, Myers' actions could be perceived as inappropriate, so the best course of action would be to avoid any potential conflict by not donating to DeFrietas's campaign and finding other ways to support environmental protection policies.

C. Misrepresentation

> Members and Candidates must not knowingly make any misrepresentations relating to investment analysis, recommendations, actions, or other professional activities.

Lee

Lee, CFA, is a financial planner for AKC. AKC compensates its planners based on the number of AKC products they sell. Lee advises a married couple to transfer their retirement funds totaling $125,000 into a single AKC investment fund that follows a large-cap equity strategy. Lee discloses to the couple that they will have to pay a penalty totaling $30,000 for closing their retirement accounts but claims they will make up the loss with better investment returns from the AKC product.

Lee's actions are

A acceptable if the AKC product is suitable for the couple.

B a violation of the CFA Institute Code and Standards because she is promising a specific rate of return.

C acceptable because she fully discloses the negative consequences of closing their retirement accounts.

Analysis B is correct. This case involves Standard I(C): Professionalism, Misrepresentation, which states that CFA Institute members and candidates must not knowingly make any misrepresentation related to investment analysis, recommendations, or actions. This standard prohibits members and candidates from making any statements that promise or guarantee a specific rate of return on volatile investments. Because the equity-based investment is inherently volatile, and the future return is unpredictable, Lee's promises about future returns making up for the penalty of withdrawing the funds violates Standard I(C). Trust is the foundation of the investment profession. Investment professionals who make false or misleading statements not only harm investors but also reduce the level of investor confidence in the investment profession and threaten the integrity of the capital markets.

A is incorrect. Even if the AKC product is suitable for the couple, it is an equity-based investment that is inherently volatile. Any statements that the investment will provide a certain rate of return or be worth a certain amount in the future is a misrepresentation. This case also raises questions about whether advising the couple to take such a significant loss in their retirement savings would be in their best interest and whether Lee's independence and objectivity are compromised because her recommendation is influenced by the compensation scheme of her employer.

C is incorrect. Although Lee fully discloses the negative consequences of transferring the couple's assets to the AKC product, that disclosure does not mitigate her inappropriate and misleading statement about future expected returns.

Andersen

Andersen, CFA, is the CEO of an asset management firm. Andersen and other senior investment managers at his firm make an in-person proposal to manage the investments of a large pension plan. In response to a request from the pension plan, Andersen provides a list of the key personnel who would manage the account. While waiting for the outcome of the evaluation, one of the key personnel that Andersen identified and who was part of the team that made the in-person presentation leaves the firm.

Andersen should

A do nothing because the pension plan is hiring the firm, not an individual.

B immediately inform the pension plan that one of the key personnel has left the firm.

C hire a competent replacement for the person who left and then inform the pension plan of the change.

Analysis B is correct. This case relates to Standard I(C): Professionalism, Misrepresentation, which prohibits members and candidates from knowingly misrepresenting anything relating to their professional activities. After the proposal is submitted to the pension plan for its investment management business, a staff member leaves Andersen's team. Andersen has identified the person leaving as a key member of the team that made the initial presentation. Given the pension plan asked for a list of individuals who would be handling the account, this information is clearly important in their decision to hire Andersen's firm. Withholding the information about the person leaving would be a misrepresentation by omission.

A is incorrect. Andersen obviously has confidence in the firm as a whole and will likely replace the person leaving with a competent professional with similar experience and talent so that the transition in service is seamless for clients. The pension plan is clearly concerned about the specific personnel involved in managing its assets because it asked for that information. If this were a junior employee, a staff member who had limited effect on the investment decision-making process, or someone who was not listed as a key employee or who had not been part of the team making the presentation, then Andersen might not need to provide an update to the pension plan. However, given the information in the case, to avoid a misrepresentation Andersen must tell the pension plan about the departure of the key staff member.

C is incorrect. Waiting until a replacement is found is too late because knowing who the managers are on the team is clearly important to the pension plan's hiring decision.

Brodeur

Brodeur, CFA, is CEO of LeTour, a global company that makes electric cars. Brodeur and LeTour disclose to the public and regulators that the company will use Brodeur's personal social media account to disseminate information to LeTour investors and the investing public. After the media reports that the company is having difficulty producing and delivering its cars to buyers, Brodeur posts on his social media that the company is "considering taking LeTour private at $420 a share. Funding is secure."

Previously, Brodeur had met with a large sovereign wealth fund (SWF) that expressed general interest in investing in the company and taking the company private. But Brodeur and the SWF had not reached any agreement or determined a share purchase price. Brodeur was also in discussions with investment banks but had not yet retained any advisers to assist with taking the company private. After the post, LeTour's stock price increased more than 6% on significantly increased volume and closed at $380 per share, 10% higher than on the previous day. When asked about the specific stock price in the post, Brodeur admitted that he had not discussed pricing with any potential investor but chose the price as a joke. Brodeur's actions are

A inappropriate because the post was a misrepresentation of the facts.

B inappropriate because not all investors use social media, so Brodeur is selectively disclosing information and putting some investors at a disadvantage.

C appropriate because his post said only that he was "considering" taking the company private and thus contained only speculative, nonmaterial information.

Analysis A is correct. This case relates to Standard I(C): Professionalism, Misrepresentation, which states that CFA Institute members and candidates must not knowingly make any misrepresentation relating to professional or investment activities. Brodeur's post was based on assumptions and contrary to the facts. Among other things, he (1) had not reached any agreement with the SWF, (2) had never discussed a share price of $420, (3) had set the price as a joke, and (4) had not formally retained any legal or financial advisers to assist with taking the company private. Unlike market participants reading his post, Brodeur knew that his supposed "secure" funding was based on a general conversation regarding a potential investment of an unspecified amount in the context of an undefined transaction structure. Because of the uncertainties that would need to be resolved before going private could be possible, Brodeur knew or should have known that his statements were false and misleading. The online or interactive aspects of social media do not remove the need to be open and honest about the information being distributed.

B is incorrect. Disseminating information to investors using social media might be appropriate and ethical under certain conditions. Distribution channels to make information public do not need to be guaranteed to reach all investors, but they must be designed to effectively make the information public. As long as the information shared reaches all clients or is available to the investing public, social media platforms would be comparable with other traditional forms of communication, such as press releases or email communication. But because the information was misleading, Brodeur's use of social media was not appropriate.

C is incorrect. The post can be considered "material" information and not speculative because (1) the source of the post was the company's CEO, (2) the subject matter was dramatic and elemental, and (3) investors would want to know the information before making an investment decision.

D. Misconduct

> Members and Candidates must not engage in any professional conduct involving dishonesty, fraud, or deceit or commit any act that reflects adversely on their professional reputation, integrity, or competence.

Hanse

Hanse, CFA, is a portfolio manager employed by a global investment bank. She manages an environmental, social, and governance (ESG) investment fund. In her free time, Hanse participates in civil disobedience demonstrations organized by the Extinction Rebellion, a sociopolitical movement that uses nonviolent resistance to protest climate breakdown, biodiversity loss, and ecological collapse. At several demonstrations held in the financial district, Hanse is arrested on charges of unlawful assembly, obstructing public transit, and disorderly conduct. She is ultimately convicted of several minor criminal offenses. Hanse has signed a standard employment contract with the bank that allows it to terminate any employee who is convicted of a criminal offense.

Under the CFA Institute Code and Standards, Hanse's actions

A do not violate the Code and Standards.

B violate the Code and Standards because she has violated her employment contract with the bank.

C violate the Code and Standards because she is arrested for misconduct in the financial district.

Analysis A is correct. This case relates to what constitutes professional misconduct. Standard I(D): Professionalism, Misconduct prohibits CFA Institute members and candidates from engaging "in any professional conduct involving dishonesty, fraud, or deceit" or from committing "any act that reflects adversely on their professional reputation, integrity, or competence." Generally, Standard I(D) is not meant to cover legal transgressions resulting from acts of civil disobedience in support of personal beliefs because such conduct does not reflect adversely on the member's professional reputation, integrity, or competence.

B is incorrect. Hanse's arrest is not likely an automatic violation of bank policies. Although Hanse's employment contract gives the bank the option to fire her for criminal activity, the bank can apparently use discretion in retaining employees depending on the nature and circumstances of the conviction. Further, the Code and Standards are not automatically violated with every violation of an employer's employment contract unless the violations relate to conduct addressed in the Code and Standards.

C is incorrect. The fact that Hanse is arrested in the financial district is not a sufficient link to her professional activities to make her actions a violation of the Code and Standards. The acts of civil disobedience by Hanse reflect her personal beliefs, which would generally not be considered an act involving fraud, dishonesty, or deceit or an act that reflects poorly on her professional reputation, integrity, or competence.

Mang

Mang, CFA, is an investment adviser for a regional bank that has a number of discretionary, fee-based accounts for high-net-worth individuals. The bank's policies permit trade error corrections up to 30 days following a failure to place a trade when an adviser forgets to send an order to the trading desk. To rectify the error, the adviser is permitted to buy or sell a security at the current market price, with the price differential charged to the adviser personally through an internal error account.

On many occasions, Mang uses the trade error correction policy to benefit clients who are unhappy with their account's performance. Mang identifies a security whose price has increased in the past 30 days. He then tells the trade desk he mistakenly failed to buy that particular security some days before, when the price was substantially lower than the current market price. Once the request is approved, the trade desk purchases the security and charges the price differential to Mang personally through the error account. Shortly after that trade, Mang sends an order to sell the security and net a profit for the client. Mang then tells the client that he has given them a "gift" or "no-risk" trade.

Mang's actions

A violate the CFA Institute Code and Standards.

B are appropriate because Mang is acting for the benefit of the client.

C are appropriate because the bank is not harmed by Mang's actions.

Analysis A is correct. Mang is violating Standard I(D): Professionalism, Misconduct, which prohibits CFA Institute members and candidates from engaging in professional conduct involving dishonesty, fraud, or deceit. He is engaging in a misleading, fraudulent, and deceptive practice and misusing the bank's policies to enhance the performance of his clients' accounts when no legitimate error was made. He personally compensates clients as a means to mislead them about his ability as an investment adviser as well as the actual performance of their accounts.

B is incorrect. Although the trades benefit his clients, Mang is engaging in professional misconduct by misleading his clients with fraudulent and deceptive conduct.

C is incorrect. The fact that these practices did not cause financial losses for the bank does not make Mang's conduct any less deceptive, fraudulent, or misleading.

INTEGRITY OF CAPITAL MARKETS

II.

A. Material Nonpublic Information

> Members and Candidates who possess material nonpublic information that could affect the value of an investment must not act or cause others to act on the information.

Khatri

Khatri, a candidate in the CFA Program, plays on a local cricket team with his friends, including Patel, his brother-in-law and an attorney, and Ahuja, owner of a software development company called ZeroPower (ZP). Patel handles the legal work for ZP. Khatri, Patel, and many other friends in their circle have invested in ZP. Recently, a large global information technology company (GIT) made an offer to buy ZP at a substantial premium over the company's current share price. Patel is working with lawyers from GIT to assist in their due diligence.

One weekend, during a particularly intense period of negotiations, Khatri, Patel, and Ahuja are playing in a cricket tournament. Between matches, Khatri overhears Patel speaking with representatives of GIT on his cell phone. Although mention of a ZP acquisition is not made, Khatri hears Patel repeatedly reference the name "GIT". Khatri guesses that the acquisition of ZP is happening soon after seeing Patel and Ahuja huddled in private conversation several times over the course of the weekend. On Monday, Khatri calls his broker and increases his investment in ZP by 5,000 shares. One week later, ZP announces its acquisition by GIT and its share price increases 30%.

Khatri's actions are

- **A** a violation of the CFA Institute Code and Standards.
- **B** acceptable because Khatri's investment was based on his own speculation.
- **C** acceptable because Khatri received the information in a public environment.

Analysis A is correct. Khatri comes into possession of material nonpublic information by overhearing confidential information from Patel's phone calls and witnessing Patel's interactions with Ahuja that were prompted by those calls. Standard II(A): Integrity of Capital Markets, Material Nonpublic Information prohibits CFA Institute members and candidates from taking investment action based on material nonpublic information in their possession. Patel should have been more careful to keep his phone conversations private, but his failure to do so does not allow Khatri to use material nonpublic information received from overhearing those conversations.

B is incorrect. Although information about the acquisition of ZP was not explicitly stated, and Khatri was required to make some degree of deduction, he knew or should have known that this information was confidential, nonpublic, and material. His actions are different from mosaic theory, in which investment analysts combine nonmaterial and nonpublic information with public information in their analysis to derive insights.

C is incorrect. Although Patel's phone calls and interactions with Ahuja take place in a public environment, that does not mean the information is public.

Kwame

Kwame, CFA, is the chief financial officer of PH3D, a biotech firm that researches, develops, and commercializes pharmaceutical drugs for women's health. PH3D submits a drug application to the government regulator for a promising new drug. The regulator postpones two scheduled meetings with PH3D when unspecified deficiencies with the drug arise. The public announcement of the postponements results in a 10% drop in the company's stock price.

When the meeting between the regulator and PH3D finally takes place, PH3D presents unpublished preliminary test data with favorable indicators for the drug. The regulator reacts positively and gives PH3D provisional regulatory approval, contingent on further studies. Afterward, some of the sell-side research analysts covering PH3D ask Kwame about the meeting. Kwame responds by email, indicating that he believes the meeting with the regulator was "very positive and productive" and that the company was "pleasantly surprised" by the regulator's reaction. Kwame does not share the favorable preliminary test data. After Kwame's emails to the analysts, the company's stock price increases 19%.

Kwame's actions

- **A** violate the CFA Institute Code and Standards.
- **B** are appropriate because he does not share the unpublished preliminary test data with the analysts and restricts his comments to the general tenor of the meeting.
- **C** are appropriate because he responds to questions from research analysts covering the company.

Analysis A is correct. This case relates to selective disclosure of material nonpublic information. Standard II(A): Integrity of Capital Markets, Material Nonpublic Information prohibits CFA Institute members and candidates from causing others to act on material nonpublic information. Information is material if its disclosure would affect the price of a security or if reasonable investors would want to know the information before making an investment decision. Information is "nonpublic" until it has been disseminated or made available to the marketplace in general, and Kwame's response to select sell-side analysts inquiring about PH3D's meeting with the regulator would not be considered public dissemination of the information. In this case, both the drug's positive preliminary test results and the regulator's positive reaction to the new drug is information that would have an impact on the price of the security in addition to information that reasonable investors would want to know.

B is incorrect. Although Kwame does not share the test results with the analysts, he shares that the meeting with the regulator was "very positive and productive". This is information that reasonable investors would want to know.

C is incorrect. Kwame shares this information with a select group of analysts who regularly cover PH3D and not the general investing public. The fact that Kwame is only responding to questions is irrelevant. Kwame and PH3D should have issued a public announcement about the meeting with the regulator before, or at the same time, information was disclosed to analysts.

B. Market Manipulation

> Members and Candidates must not engage in practices that distort prices or artificially inflate trading volume with the intent to mislead market participants.

Abbha

Abbha, CFA, is a securities contractor working to assist Superior Energy (SE) with listing its shares on the Regional Security Exchange (RSX). SE files a prospectus for an offer of up to five million shares at $2 each to raise $10 million. The RSX current applicable listing rule requires that entities seeking admission to the RSX must meet a "minimum spread requirement" of at least 300 shareholders with a minimum value holding to qualify for listing. In their listing application, representatives of SE inform RSX that the minimum spread requirement of 300 shareholders has been met. These disclosures include as shareholders 31 people or companies arranged by Abbha, though none are actual buyers of SE securities. Abbha provides false names and addresses for the 31 shareholders. The SE share offer raises more than $3.5 million, with more than 1.75 million shares issued. SE is admitted to the official list of the RSX, with its shares quoted on the exchange. Over time, the price of SE shares steadily increases, the company attracts hundreds of investors and shareholders, and early investors achieve an excellent investment return.

Abbha's actions

A violate the CFA Institute Code and Standards.

B are acceptable because SE proved to be a strong company with excellent performance.

C are acceptable because no investors were harmed by the technical violation of RSX rules.

Analysis A is correct. Abbha engaged in information-based manipulation by inflating the number of initial shareholders in SE securities to create the impression that its shares would have sufficient liquidity, and giving new investors comfort in buying them. Standard II(B): Integrity of Capital Markets, Market Manipulation prohibits CFA Institute members and candidates from engaging in practices that artificially inflate trading volume to mislead market participants. To circumvent regulatory requirements and drive investor interest, Abbha's misrepresentations overstated initial interest in SE securities, thereby creating false pretenses under which the company was listed.

B is incorrect. The fact that SE ultimately proved to be a bona fide and well performing investment does not mitigate Abbha's unethical, misleading, and deceitful conduct.

C is incorrect. The purpose of the minimum spread requirement for RSX security listings is to demonstrate that investor interest in a company is sufficient to justify the company's listing. Th minimum spread requirement ensures some level of liquidity at the time a company is listed and prevents lower quality applicants, that are unable to attract sufficient investor interest, from being listed. Falsifying the number of initial investors goes beyond a technical violation of RSX rules and has substantial consequences.

III. DUTIES TO CLIENTS

A. Loyalty, Prudence, and Care

> Members and Candidates have a duty of loyalty to their clients and must act with reasonable care and exercise prudent judgment. Members and Candidates must act for the benefit of their clients and place their clients' interests before their employer's or their own interests.

Maste

Maste, CFA, is the sole director of Dov Services (Dov), a firm that sells financial products and advice. Maste directs Dov's authorized representatives to incorporate the Dov Client Protection Policy into their contracts with clients. The protection policy, which sets forth the terms for providing financial advice, states that it "contains a number of client protections designed to ensure that you (the client) receive the best possible advice and the maximum protection available under the law." The protection policy's terms (1) excuse Dov and its authorized representatives from various liabilities arising from their failure to act in a client's best interest, (2) relieve Dov and its authorized representatives of their duty to conduct suitability analyses of clients and investments, and (3) lead clients to believe that they cannot make claims against Dov or its representatives for violating securities law.

Maste's actions

A violate the CFA Institute Code and Standards.

B are appropriate because Dov and Maste fully disclose the terms of the Dov Client Protection Policy to clients.

C are appropriate because Dov and Maste are free to negotiate the terms of advisory agreements with clients.

Analysis A is correct. The terms of the Dov Client Protection Policy improperly attempt to use disclosure to relieve Dov and Maste of their fundamental ethical (and very likely legal) obligations to clients by limiting liability for failures to act in the client's best interests or to provide appropriate advice. Standard III(A): Duties to Clients, Loyalty, Prudence, and Care sets forth a duty of loyalty on the part of CFA Institute members and candidates to their clients and requires them to act for the benefit of their clients and to place their clients' interests before their own. In addition, Standard V(A): Investment Analysis, Recommendations, and Actions, Diligence and

Reasonable Basis requires members and candidates to provide diligent, independent, and thorough advice and to have a reasonable and adequate basis for investment action; Standard III(B): Duties to Clients, Fair Dealing requires members and candidates to conduct a suitability analysis for any investment recommendation to their clients; and Standard I(C): Professionalism, Misrepresentation says that members and candidates cannot make any misrepresentations relating to investment services. These components of the CFA Institute Code and Standards define the fundamental principles applicable to investment professionals and detail the conduct investors should expect from their financial advisers.

B is incorrect. Disclosure cannot be used to relieve advisers of their fundamental ethical obligations to clients. The Dov Client Protection Policy is also deceptive because it misrepresents the client's right to bring legal action for ethical and regulatory violations and appears to give the false impression that a client would benefit from its terms.

C is incorrect. Although in general, clients and advisers are free to negotiate the terms of advisory agreements, for advisers to use the client agreement to create a significant imbalance between the rights and obligations of the adviser and their client or to limit the fundamental ethical obligations of loyalty, prudence, and care to their client is improper.

Gaini

Gaini, CFA, is a commodities trader with a number of retail clients, including Laube, who opens a self-directed foreign exchange retail account for which Gaini does not advise on any trades. Laube signs Gaini's standard customer agreement, which contains provisions relating to margin requirements and liquidation in the event of margin deficits. The agreement authorizes Gaini, at his discretion, "to liquidate, without notice, any or all open positions in an account with insufficient margin."

Laube initially purchases two 100,000 US dollar/Swiss franc (USD/CHF) contracts and sets her margin requirement to $4,000. Laube then places two additional 100,000 USD/CHF "pending limit" orders that will execute if the contract trading prices reach a specified level. Each order has a different limit price, so one will execute before the other. Shortly after placing the orders, Laube goes on an extended vacation. While she is away, execution on the first limit order occurs, and the margin requirement increases to $6,000. Then, the next limit price is reached, and the second order executes, increasing the margin requirement to $8,000. Post trades, while Laube is still on vacation, her account balance drops to $6,900. Without notice, Gaini liquidates all positions in Laube's account, realizing a loss of $37,000 from the liquidation.

Gaini's actions

A are appropriate because Gaini followed the policy and procedures set forth in Laube's client agreement.

B violate the CFA Institute Code and Standards because Gaini has a duty to act in the best interest of Laube by protecting her financial position.

C are inappropriate because Gaini could have met the margin requirement by liquidating only one foreign exchange contract position.

Analysis A is correct. This case relates to Standard III(A): Duties to Clients, Loyalty, Prudence, and Care, which imposes on CFA Institute members and candidates a duty to act with loyalty, prudence, and care in their clients' best interests as well as in relation to other clients and the markets. However the conduct that fulfills this responsibility depends on the nature of the relationship with the client. Laube self-directs her trades which are executed by Gaini, who does not provide investment advice. Additionally, Laube freely contracts with Gaini regarding the consequences of insufficient margins in her account, and these policies and procedures are clearly set forth in her client agreement. Nothing indicates that Gaini misled Laube about the margin requirements or the liquidation procedures. Although the contractual provision authorizing liquidation without notice does not waive Laube's right to be dealt with in good faith, the margin and liquidation provisions are not an attempt by Gaini to waive his duty of loyalty, prudence, and care. They simply set forth the parameters of the relationship, which Gaini and Laube are free to negotiate. Although Laube suffered significant trading losses, Gaini liquidated Laube's open positions in accordance with the terms of the customer agreement, serving to protect the integrity of the market.

B is incorrect. Given the facts of this case, Gaini's responsibility to protect his own financial position and that of his other customers supersedes any duties he owes to Laube when she defaults on the margin requirements. Because of the limited nature of the relationship between Laube and Gaini, Gaini is not obligated to actively monitor Laube's account to protect her financial interests. Doing so would require Gaini to continuously monitor a potentially deteriorating market or to take on additional risk management measures that he has not agreed to and that Laube has not contracted for.

C is incorrect. Gaini is not obligated to use a less drastic alternative by closing only one foreign exchange contract position and does not breach a duty to Laube by liquidating all open positions when the parties have contractually agreed that total liquidation to meet a margin deficit may be done at Gaini's discretion without notice.

Braung

A regional government hires Braung, CFA, and his firm to serve as its financial adviser for issuing general obligation bonds. The municipality conducts several bond offerings over the years for constructing municipal facilities, including a maximum-security detention facility and two school buildings. In connection with the bond issues, Braung makes a number of trips to New York City to meet with ratings agencies. The trips are typically planned for a Monday or Friday because the costs for weekend travel are less. Braung's wife accompanies him on the trips, and they typically spend the weekend either before or after the meetings in New York City to enjoy sporting events, theater performances, and museums. Braung often makes a number of train and hotel changes after a trip is booked to accommodate meetings with other clients. Braung submits his travel expenses to his supervisor, who deducts costs she believes are unrelated to the business purpose of the trip and submits the bills to the municipality for reimbursement.

Which of the following expenses can most likely be billed to the government entity issuing the bonds?

A Braung's accommodation and meal expenses for the weekend days because the travel rates are less expensive over a weekend

B Tickets to the sporting and theater events, as long as they do not exceed a reasonable amount for business entertainment

C Flight and hotel change fees that result from the regular course of Braung's business activities

Analysis A is correct. This case relates to Standard III(A): Duties to Clients, Loyalty, Prudence, and Care, which states that CFA Institute members and candidates have a duty of loyalty to their clients, must act for their clients' benefit, and must place client interests before their own. Under this standard, investment professionals, including municipal security dealers, must not engage in any deceptive, dishonest, or unfair practice when handling client accounts. The savings in travel fees for booking a weekend travel schedule could be greater than the additional accommodation and meal expenses for Braung to stay in New York City the extra days, making the cost to the client lower. Under these circumstances, Braung would be meeting his duty of loyalty to the client by choosing the most inexpensive travel schedule overall, thus limiting costs to the client. Charging personal theater and sporting event tickets or travel change fees associated with seeing unrelated clients would not qualify as chargeable expenses to the client. The best scenario would be to detail in writing at the outset of the engagement which expenses would be considered reimbursable.

B is incorrect. In the context of conflicts of interest, the CFA Institute Code and Standards allow members and candidates to accept or provide modest gifts and entertainment in the ordinary course of business (a gift basket at the holidays from a vendor or to a client, for example). But that “ordinary course of business” does not allow investment professionals to charge clients for obviously extraneous entertainment expenses not connected to a business meeting. Charging excessive or lavish expenses for the personal benefit of the investment professional at the expense of the client can constitute a deceptive, dishonest, or unfair practice that violates Standard III(A).

C is incorrect. Although busy investment professionals might be forced by other priorities to change travel arrangements when a client trip has already been scheduled, additional expenses resulting from the change most likely must be borne by the investment professional (Braung) as an overhead cost, not charged to the client. (Under some limited circumstances, those expenses might be charged to the client who necessitates the travel changes.)

B. Fair Dealing

> Members and Candidates must deal fairly and objectively with all clients when providing investment analysis, making investment recommendations, taking investment action, or engaging in other professional activities.

Scherzer

Scherzer, CFA, is the head of research at a large investment management firm. She publishes monthly “Recommendation Update Reports” to communicate to clients any investment recommendation changes her firm has made. The reports are sent to clients via email on the first Friday of the month and posted on the firm’s website the following Monday. The firm’s internal policy is that a change in recommendation can only be made once a month through this report. Scherzer also publishes weekly reports with information gathered by analysts that might implicitly signal a future change in recommendation or lead a reader to infer that a recommendation will be changing. Although the monthly “Recommendation Update Reports” are sent to all clients, clients who wish to receive the weekly publication must pay an annual fee of $1,000. This option is available to any client and is fully disclosed as part of every client agreement.

To comply with the CFA Institute Code and Standards, Scherzer is required to

A do nothing because her actions comply with the Code and Standards.

B publish the monthly report on the firm’s website at the same time it is sent to clients.

C send the weekly reports to all clients at no additional charge.

Analysis A is correct. Standard III(B): Duties to Clients, Fair Dealing requires CFA Institute members and candidates to deal fairly and objectively with all clients when providing investment analysis and making investment recommendations. It also requires that information about investment recommendations be disseminated in such a manner that all clients have a fair opportunity to act on the information. Scherzer distributes the recommendation update reports to all clients simultaneously via email, giving each an opportunity to act on the information. Scherzer is also not limited by the Code and Standards in communicating with clients about investments only when an investment recommendation is provided and can provide updated or clarifying information to clients as appropriate. In addition, Scherzer does not violate the Code and Standards by providing weekly reports to clients who pay a subscription fee for the service as long as (1) all clients are eligible to purchase the service, (2) the different levels of service are disclosed to clients, and (3) no change in recommendations is made that would disadvantage existing clients who receive only the monthly report.

B is incorrect. The Code and Standards do not require Scherzer to publish the reports to the public on the firm's website.

C is incorrect. The Code and Standards allow CFA Institute members and candidates to provide more personal, specialized, or in-depth service to clients who are willing to pay for premium services through additional fees as long as (1) those services do not disadvantage other clients who do not pay additional fees, (2) the differing levels of service are disclosed to clients, and (3) the services are made available to all.

C. Suitability

1 When Members and Candidates are in an advisory relationship with a client, they must:
   a Make a reasonable inquiry into a client's or prospective client's investment experience, risk and return objectives, and financial constraints prior to making any investment recommendation or taking investment action and must reassess and update this information regularly.
   b Determine that an investment is suitable to the client's financial situation and consistent with the client's written objectives, mandates, and constraints before making an investment recommendation or taking investment action.
   c Judge the suitability of investments in the context of the client's total portfolio.
2 When Members and Candidates are responsible for managing a portfolio to a specific mandate, strategy, or style, they must make only investment recommendations or take only investment actions that are consistent with the stated objectives and constraints of the portfolio.

Marte

Marte, CFA, is an asset manager in Puerto Rico, a US territory. Residents of Puerto Rico receive significant tax advantages when they invest in local securities. To capitalize on this advantage, Marte's firm offers its clients shares in a closed-end investment fund, organized under Puerto Rico's financial laws and regulations, that holds at least 67% local securities and is permitted to borrow against up to 50% of its assets. The fund is usually leveraged to the extent legally permitted. Many of Marte's clients have a modest net worth and conservative or moderate investment objectives. Marte convinces them to invest 85% or more of their assets in shares of the closed-end fund.

Marte's actions

A violate the CFA Institute Code and Standards.

B are appropriate because they take advantage of the fund's unique tax benefits for his clients.

C are appropriate as long as Marte fully discloses the risks and benefits of the fund to his clients.

Analysis A is correct. Standard III(C): Duties to Clients, Suitability states that CFA Institute members and candidates in an advisory relationship with clients must "determine that an investment is suitable to the client's financial situation and consistent with the client's written objectives, mandates, and constraints before making an investment recommendation or taking investment action." Given the financial circumstances and conservative investment objectives of his clients, the high concentration of the fund's shares in his clients' accounts, combined with the leverage, makes the large investment in the fund unsuitable because greater portfolio diversification is needed. Despite the favorable tax advantages, highly concentrated investments carry an increased risk that a single market event affecting the value of the fund's shares would significantly decrease the total account value. This risk is exacerbated by the fact that the closed-end fund is internally leveraged, which could magnify the fund's loss during a market event that causes share values to drop steeply.

B is incorrect. Given the favorable tax advantages of the investment vehicle, investment in shares of the closed-end fund might be suitable and appropriate for his clients at some level. But given the financial circumstances and investment objectives of his clients, the high concentration, combined with the leverage, makes the large investment in the fund unsuitable.

C is incorrect. Although Marte should always fully disclose the risks and benefits of his recommendations to his clients, doing so does not take away the need for the investment to be suitable for the clients' financial circumstances, risk tolerance, and investment goals.

Duri

Duri, CFA, is a registered account representative providing financial advice to retail clients. She is also a principal partner of Tabak Accountants. Duri assists a number of advisory clients who want to move their retirement assets from existing superannuation (pension) accounts to establish self-managed superannuation funds (SMSFs) that have the goal of investing in direct residential property. When clients express interest in these types of SMSFs, Duri accepts their reasons for wanting to invest in direct property and assumes that they have the time and expertise to manage their superannuation affairs. She reclassifies their investment objectives as "growth" to match the new investment strategy. Duri charges her clients for establishing the SMSFs and recommends that her firm, Tabak Accountants, prepare the clients' annual accounts and tax returns.

Duri's actions

A violate the CFA Institute Code and Standards.

B are acceptable because she is following the directives of her clients.

C are acceptable if the services provided by Tabak Accountants are reasonable and the costs of services are competitive.

Analysis A is correct. Standard III(C): Duties to Clients, Suitability requires CFA Institute members and candidates in an advisory relationship to make a reasonable inquiry into a client's investment experience, risk and return objectives, and financial constraints to determine whether an investment is suitable for the client's financial situation and consistent with the client's objectives and mandates before taking investment action. When Duri's clients want to switch their retirement assets into an SMSF that invests in direct residential property, Duri appears to accept their reasoning without doing a suitability analysis. The case does not indicate whether Duri did any of the following:

- Assess the reasons a client wants to invest in direct property
- Consider the client's financial goals

- Compare the benefits, risks, and costs associated with establishing an SMSF or owning rental property to existing superannuation account investment alternatives
- Consider asset diversification
- Determine whether the client intends to draw a pension from the SMSF once the individual reaches retirement, or how the client will do that with an illiquid asset such as investment property
- Evaluate whether the client has the time and expertise to manage his or her superannuation affairs, including managing a rental property
- Consider whether the SMSF investment strategy will remain viable if the client's income is reduced or the property is unoccupied for a period of time

Without analyzing these and other factors, for Duri to move her clients to the SMSFs would be inappropriate.

B is incorrect. If her clients request a change in their investment strategy, Duri has a responsibility as their investment adviser to conduct an analysis to ensure that the new direction is suitable and appropriate. Duri appears to have classified all clients as growth investors to superficially justify their investment in an SMSF vehicle without doing an analysis of whether this classification was an accurate assessment.

C is incorrect. Duri appears to be recommending that her firm provide clients' accounting services for the SMSF investments to create additional income for herself. But Standard III(A): Duties to Clients, Loyalty, Prudence, and Care requires members and candidates to act for the benefit of their clients and to place their clients' interests before their own. At a minimum, Duri's relationship with Tabak must be fully disclosed to clients as a potential conflict of interest.

D. Performance Presentation

> When communicating investment performance information, Members and Candidates must make reasonable efforts to ensure that it is fair, accurate, and complete.

Jergenn

Jergenn, CFA, is the portfolio manager for the Volare Investment Management (VIM) fund, a registered collective investment scheme (CIS) organized under the laws of South Africa. VIM's 2022 regulatory disclosure and marketing material for the fund, produced by Jergenn, presents annual investment performance data for the 2014–2020 period that is accurate and calculated correctly. The performance history is that of a composite of separate accounts that followed the strategy used by the VIM fund before the assets were moved over to the CIS environment in 2021.

In presenting the fund's performance history, Jergenn's actions

A violate the CFA Institute Code and Standards.

B are appropriate because the investment performance is accurate.

C are appropriate as long as the performance calculations are net of fees.

Analysis A is correct. Standard III(D): Duties to Clients, Performance Presentation states that CFA Institute members and candidates must make a reasonable effort to ensure that investment performance information is fair, accurate, and complete. Although the performance information presented by Jergenn is calculated correctly and includes technically accurate data, Jergenn's failure to indicate clearly that the performance data applied to a period before the VIM fund was registered as a CIS has the potential to mislead investors into believing that the CIS fund had a long track record. To meet the "fair, accurate, and complete" requirement of the standard, Jergenn should disclose that the 2014–2020 performance history was for a prior but similar entity with the same investment strategy and objectives and that the VIM fund, as a CIS, has been in existence since only 2021.

B is incorrect. Although the performance is accurate and correctly calculated, the lack of disclosure about the nature of the accounts makes the presentation misleading.

C is incorrect. Performance can be presented either net or gross of fees as long as investors are sufficiently informed about how the performance is calculated and what effect fees might have on the returns.

E. Preservation of Confidentiality

> Members and Candidates must keep information about current, former, and prospective clients confidential unless:
>
> 1. The information concerns illegal activities on the part of the client or prospective client,
> 2. Disclosure is required by law, or
> 3. The client or prospective client permits disclosure of the information.

Giddings and Marsh

Giddings, CFA, is responsible for compliance at GWH, a large broker/dealer and investment adviser. In connection with GWH's wealth management business, the company maintains personally identifiable information (names, addresses, phone numbers, account numbers, balances, and holdings) of its clients. Giddings adopts a number of policies and restrictions, including a Code of Conduct, that address employees' access to and handling of this confidential information. Marsh, CFA, who works for GWH as a client services associate, decides to download client data to his personal server located at his residence to facilitate his telecommuting. Marsh's server is hacked, and portions of the personal client information Marsh downloaded are posted for sale on the internet.

Did either Marsh or Giddings violate the CFA Institute Code and Standards?

A Marsh violated the Code and Standards.

B Giddings violated the Code and Standards.

C Both Marsh and Giddings violated the Code and Standards.

Analysis C is correct. Standard III(E): Duties to Clients, Preservation of Confidentiality requires that CFA Institute members and candidates keep information about current, former, and prospective clients confidential unless the information concerns illegal activities, disclosure is required by law, or the client permits disclosure. Although the standard does not require investment professionals to be information security experts, they must make reasonable efforts to prevent the accidental distribution of confidential information by following compliance procedures. Assuming Marsh did not have permission to download client data to his personal server, his misappropriation of client information for his own purposes is a violation of Standard III(E). And although he was not responsible for the distribution of the information, his misconduct facilitated the publication of the information. Marsh needed to understand and follow the firm's electronic information communication and storage procedures.

As the compliance officer, Giddings is responsible for implementing sufficient compliance policies and procedures reasonably designed to protect client information. He is responsible for ensuring the confidentiality of customer information, protecting against any potential threats to the security or integrity of the records, and preventing unauthorized access or use of client information that could harm GWH clients. Although GWH has policies restricting access to and handling of client information, the extent of those safeguards was apparently insufficient because Marsh was able to download company files to his personal computer. The fact that client information was accessed and published makes the effectiveness of Giddings's compliance efforts questionable. Even if the policies were sufficient, nothing indicates that Giddings audited and/or tested the effectiveness of the safeguards meant to keep client information confidential.

A is incorrect. Both Marsh and Giddings violated the Code and Standards.

B is incorrect. Both Marsh and Giddings violated the Code and Standards.

IV. DUTIES TO EMPLOYERS

A. Loyalty

> In matters related to their employment, Members and Candidates must act for the benefit of their employer and not deprive their employer of the advantage of their skills and abilities, divulge confidential information, or otherwise cause harm to their employer.

Nickoli

Nickoli, CFA, is an investment counselor with HHI Capital Management (HHI). A colleague at her local society encourages Nickoli to leave HHI and join her at Vesuvius Asset Advisers. Nickoli eventually agrees and decides to leave at the beginning of the new year. In the weeks before submitting her resignation, she tells her clients that they will likely be working with a new investment counselor because she is leaving HHI. Her clients express surprise, and when asked for details about why she is leaving, Nickoli shares that she is frustrated by the firm's structure, disagrees with the direction the firm is going, lacks confidence in the current leadership, doubts the firm will be able to attract and retain good people, and believes other HHI employees have been mistreated and will also be leaving soon. Several of Nickoli's HHI clients indicate that they would like information about Vesuvius and might be interested in switching their accounts. After submitting her resignation, Nickoli immediately shares the names of the interested clients with Vesuvius, and after the first of the year, she begins soliciting them to transfer their accounts from HHI to her new firm.

Nickoli's conduct is

A a violation of the Code and Standards.

B acceptable because she did not solicit clients until after she left HHI.

C acceptable because she is looking out for her clients' best interests and believes Vesuvius provides better service.

Analysis A is correct. Standard IV(A): Duties to Employers, Loyalty states that CFA Institute members and candidates "must act for the benefit of their employer and not . . . otherwise cause harm to their employer." Although a departing employee is generally free to make arrangements or preparations to change firms before terminating the relationship, those preparations must not conflict with the employee's continued duty to act in the best interests of the current employer and not otherwise undermine, disparage, or cause harm to that employer. In this case, Nickoli decides to leave HHI and join Vesuvius weeks before she notifies the firm and submits her resignation. During that time, Standard IV(A) obligates her to continue to act in her current employer's best interests and not engage in any activities that would conflict with this duty until her resignation becomes effective. Nickoli violates her duty of loyalty to HHI by making disparaging and harmful statements about the firm to its clients in the weeks before she submits her resignation and by promoting Vesuvius to HHI clients while she is still employed by HHI.

B is incorrect. Although she did not make actual solicitations until after she left HHI, Nickoli used the final weeks of her employment with HHI to determine which of the firm's clients might be interested in receiving information about Vesuvius and possibly transferring their accounts from HHI and then contacted them after her move to Vesuvius.

C is incorrect. Although investment professionals should protect their clients' best interests, even if Nickoli believes the clients will be better off with her at Vesuvius, the clients' relationship is with HHI. She is a representative of HHI, so she must not act in a manner harmful to the firm while still employed there.

Kuznetsov

Kuznetsov, CFA, is a portfolio manager for a medium-sized investment firm that encourages its employees to sell proprietary investment products to their clients. Kuznetsov follows the encouragement, and within a year, he becomes the firm's top seller of these investment products. He receives glowing performance reviews and a large bonus. However, Kuznetsov begins to realize that the firm's investment products are underperforming, and more expensive than other, external investment options that are suitable for his clients and present stronger growth opportunities. He stops selling the firm's investment products to his clients. Although his supervisor puts increasing pressure on him to resume selling the firm's products, Kuznetsov refuses. He complains several times to management that he is being pressured to place the firm's interests above his clients. He secretly records several conversations with his supervisor and makes copies of client records that document what he considers inappropriate conduct by his supervisor. When management ignores his complaints and Kuznetsov's supervisor begins giving him poor performance reviews, he files a complaint with the local regulator against his supervisor and his firm, providing the recordings and copies of client files as evidence. After the firm becomes aware of Kuznetsov's actions, he is let go.

Kuznetsov's actions are

A appropriate because he is protecting his clients' interests.

B a violation of the CFA Institute Code and Standards because he fails to keep client information confidential.

C a violation of the CFA Institute Code and Standards because he violates his duty of loyalty to his employer by taking his dispute with his supervisor to the regulator and exposing the employer to financial and reputational harm.

Analysis A is correct. Standard IV(A): Duties to Employers, Loyalty states that CFA Institute members and candidates "must act for the benefit of their employer and not . . . divulge confidential information, or otherwise cause harm to their employer." But occasionally, circumstances might arise in which investment professionals need to engage in conduct contrary to their employer's interests to protect clients' interests. In pressuring Kuznetsov to sell more expensive, suboptimal investment products to his clients, the employer is acting in its own interests and not in the clients' best interests. Kuznetsov is clearly acting in the best interests of the clients by notifying the regulators of his employer's unethical practices.

B is incorrect. Copying client records to give to the regulator is a justifiable step for Kuznetsov to take to protect his clients' interests.

C is incorrect. In general, Kuznetsov's decision to record conversations with his supervisor and report the employer to the regulator is justified because he is attempting to protect his clients' interests by calling out his employer's unethical (and possibly illegal) conduct. However, certain jurisdictions might have laws against recording conversations without the other party's consent. Under these circumstances, his "whistle-blowing" activity is not a violation of the Code and Standards.

Clemence

Clemence, CFA, is a wealth management adviser for DeLaurier Strategic Advisers, where she is responsible for financial planning and wealth management for more than 400 retail clients. She met many of her clients through her spouse, who is a well-known attorney, and her sister, who is a physician. Clemence decides to leave DeLaurier to take a position at another firm, where she will not be expected to generate new advisory clients but will take on more research and investment management responsibilities. She leaves DeLaurier on good terms, providing her supervisor with all the background and information to seamlessly transition her clients to a new account manager. Not all of Clemence's clients have sufficient assets under management to become clients at her new firm. On the day Clemence leaves DeLaurier, she downloads a spreadsheet of DeLaurier's clients, prospects, and former clients and sends it to her personal email. The list includes names, assets under management, addresses, and phone numbers. Clemence intends to contact her clients as a courtesy to inform them of her new position, thank them for being clients, and express her confidence that DeLaurier will continue to provide them with competent and professional service even though she has left the firm.

Clemence's actions are

A a violation of the CFA Institute Code and Standards.

B appropriate because she is protecting the interests of her clients.

C appropriate as long as she contacts only those clients who are personal friends to inform them of her new position.

Analysis A is correct. Standard IV(A): Duties to Employers, Loyalty requires that CFA Institute members and candidates "act for the benefit of their employer and not . . . divulge confidential information, or otherwise cause harm to their employer." Clemence has violated her duty of loyalty to her employer by downloading the client list and taking it with her for use after she leaves DeLaurier. The client list is DeLaurier's property. The list also contains proprietary confidential information about DeLaurier clients that Clemence is improperly using for her own purposes, no matter how well intentioned those purposes might be. Clemence could have asked her employer's permission to send a "thank you" note to her clients immediately before leaving DeLaurier, thus eliminating the need to take client information with her.

B is incorrect. Clemence is clearly not motivated to use the client list and information to benefit her new firm but appears rather to be working for DeLaurier to protect the interests of her former clients and make them feel comfortable in continuing to use DeLaurier as their financial adviser. But in obtaining information about her clients, Clemence goes too far and also takes more information than perhaps intended. In addition to information about her clients she has also obtained information about the firm's former, current, and prospective clients.

C is incorrect. Clemence may contact her former clients who are friends through personal channels, such as social media or a personal contact, but she cannot use DeLaurier's property to facilitate this communication. As an alternative, she could ask DeLaurier's permission to take her clients' contact information so she can send them a final "thank you" correspondence.

B. Additional Compensation Arrangements

> Members and Candidates must not accept gifts, benefits, compensation, or consideration that competes with or might reasonably be expected to create a conflict of interest with their employer's interest unless they obtain written consent from all parties involved.

Estevez

Estevez, CFA, is a senior research analyst with BIR, a boutique investment research firm covering micro- and small-cap companies that hire BIR to provide research coverage (called issuer-paid research) to promote their stock to investors. Because of BIR's stellar reputation, its research services are in high demand among both investors and companies seeking potential investors. Estevez helps BIR select the companies it covers and oversees a team of junior research analysts. Some companies encourage Estevez to select their company for BIR research coverage by giving her a separate bonus if they are included in the BIR research universe. Estevez's actions are

A a violation of the CFA Institute Code and Standards because her independence and objectivity in conducting the research are compromised.

B acceptable as long as Estevez does not use material nonpublic information from the company.

C acceptable as long as her company approves in writing the payments offered by covered companies.

Analysis C is correct. Standard IV(B): Duties to Employers, Additional Compensation Arrangements requires members and candidates to obtain written consent from their employer before they accept any gift, benefit, or compensation that "might reasonably be expected to create a conflict of interest with their employer's interest." Estevez receives a benefit from companies that BIR selects to receive research coverage. Estevez's involvement in the process of choosing the companies that BIR agrees to research presents a conflict of interest because she might favor those companies that pay her the bonus. Estevez would have to disclose to her company that she is receiving the benefit from the selected company, and BIR would have to disclose to the users of the report that the company paid Estevez the bonus and is paying for the research coverage.

A is incorrect. Issuer-paid research itself is not automatically unethical nor does it automatically compromise an analyst's independence and objectivity. However this area is filled with potential conflicts, so important safeguards must be observed. BIR must adopt strict procedures to protect the analyst's objectivity from being influenced by the company. Such safeguards must also include full disclosure of any conflicts of interest on the part of the analyst conducting the research; full disclosure of any compensation arrangements, including the source of the payment for the research; and policies that separate the research recommendations from the level or nature of the payment.

B is incorrect. Even if she is not using material nonpublic information from the company, this does not alleviate Estevez's need for full disclosure and employer consent regarding her bonuses from companies.

C. Responsibilities of Supervisors

> Members and Candidates must make reasonable efforts to ensure that anyone subject to their supervision or authority complies with applicable laws, rules, regulations, and the Code and Standards.

Duhih

RC Group (RCG) is a registered futures commission merchant with several branch offices, including one in Memphis, Tennessee (USA). Duhih, CFA, is hired to be the branch manager of the Memphis office, supervising several employees, including Lewes, CFA. Duhih allows Lewes to work from home, so Lewes does not have a physical workspace in the Memphis office or even access to the building, and Duhih only occasionally checks in with Lewes regarding work. Unknown to either Duhih or RCG, Lewes also works for another futures commission merchant, called AFCM. Lewes arranges swap agreements for AFCM, including orders with several cattle feed yard clients. Working with another employee at RCG, Lewes opens new futures accounts for the feed yard clients RCG represents. Although the other RCG employee receives all the commissions for the accounts, she secretly splits them with Lewes. Duhih is unaware of Lewes' participation in this commission sharing arrangement.

Duhih's actions as a supervisor are

A a violation of the CFA Institute Code and Standards.

B acceptable if the RCG headquarters conducts regular audits of the Memphis branch.

C acceptable if RCG did not develop adequate policies and procedures for the detection and deterrence of possible employee misconduct.

Analysis A is correct. Standard IV(C): Duties to Employers, Responsibilities of Supervisors states that members and candidates "must make reasonable efforts to ensure that anyone subject to their supervision or authority complies with applicable laws, rules, regulations, and the Code and Standards." At a minimum, supervisors must make reasonable efforts to detect and prevent legal, regulatory, and policy violations by ensuring that effective compliance systems have been established. Although the case does not clearly present what steps Duhih took to diligently exercise supervisory responsibility, the fact that Lewes works from home and did not have access to the branch office suggests that Duhih's supervision is minimal at best and noticeably ineffective. In addition, Duhih should support his company by encouraging it to adopt a code of ethics, have clearly written compliance policies and procedures, and employee training on appropriate ethical behavior.

B is not correct. A regular audit of the Memphis branch by compliance personnel from RCG headquarters could be an excellent way to ensure that branch employees are complying with applicable law, regulations, and RCG policies. However, it would not be a substitute for effective and regular supervision by Duhih, the onsite branch manager.

C is incorrect. Supervisors must understand what constitutes an adequate compliance system and make reasonable efforts to ensure that appropriate compliance procedures are established, documented, communicated to covered personnel, and followed. If Duhih knew that RCG did not have adequate policies and procedures for detecting and deterring potential employee misconduct, he was obligated to tell management about the shortcoming, help develop adequate compliance policies, or decline supervisory responsibility. A lack of adequate policies is not an excuse for failing to detect potential misconduct by RCG employees, including Lewes. Duhih should not have accepted supervisory responsibility until he understood not only RCG's policies and procedures but also the company's expectations of him with regard to maintaining his subordinates' compliance with them.

Denikin & Denikin

Sasha Denikin, CFA, began his investment career as a research analyst for Galak Investment Partners, a company founded and controlled by his father, Franz Denikin, CFA. After several years, Franz transfers ownership of Galak to Sasha, who becomes a director of the company. Franz tells the firm's clients that he retains management of all client accounts. When the longtime CCO announces her retirement, Sasha is promoted to chief compliance officer (CCO). While Sasha has no previous compliance experience, the plan is for the prior CCO to retain compliance responsibilities as a consultant while mentoring and training Sasha. Despite his title, Sasha has no actual authority to supervise his father's conduct and Franz continues to exert absolute control over Galak. Sasha also does not have permission to contact clients or review Galak communications with clients because his father insists that all client contact go through him. During his tenure as CCO, Sasha raises multiple compliance issues to his father regarding his father's actions, but Sasha is powerless to enforce company policies and procedures concerning his father's conduct. After continuing to serve as Galak CCO for a time, Sasha finally resigns in frustration.

Sasha Denikin's actions are

A a violation of the CFA Institute Code and Standards.

B acceptable because Franz holds the actual power and client responsibilities at the firm and thus cannot be under the supervision of a subordinate.

C acceptable because Sasha resigns as CCO when he is frustrated by his inability to exercise his compliance responsibilities.

Analysis A is correct. Standard IV(C): Duties to Employers, Responsibilities of Supervisors requires CFA Institute members and candidates to "make reasonable efforts to ensure that anyone subject to their supervision or authority complies with applicable laws, rules, regulations, and the Code and Standards." Members and candidates with supervisory responsibility must make reasonable efforts to prevent and detect violations by ensuring that effective compliance systems are established. If they cannot carry out their supervisory responsibilities because of an inadequate compliance system, they should decline in writing to accept the supervisory responsibility until the firm adopts reasonable procedures to allow for the adequate exercise of that supervisory responsibility. Sasha knows that he is not qualified to serve as CCO of Galak, but still accepts the role. As CCO, Sasha has supervisory responsibility for all firm employees in compliance matters. However Sasha has difficulty performing his CCO duties because the firm's structure gives him no real authority over his father to ensure that his father's conduct complies with internal policies and regulatory rules. Sasha knows that his father is ignoring firm rules, but despite being CCO and a director of the company, Sasha has no real authority or ability to supervise and influence Franz's conduct.

B is incorrect. Franz's status as the founder and the one truly controlling the company does not exempt him from oversight by the CCO.

C is incorrect. If Sasha is unable to fulfill his responsibilities because of an inadequate compliance system or a lack of experience, he should refuse the CCO role until he has been fully trained and Galak adopts procedures allowing him to effectively exercise supervisory responsibility over his father. For example, Sasha could work with the outgoing CCO to establish a compliance program that includes a code of ethics, compliance policies and procedures, education and training programs, an incentive structure rewarding ethical behavior, and firmwide best practice standards to help meet his supervisory obligations. However, Sasha chooses to quit the CCO job and resign from the firm only after a significant period during which he fails to adequately meet his supervisory responsibilities, and, likely, the best interests of the firm's clients.

V. INVESTMENT ANALYSIS, RECOMMENDATIONS, AND ACTIONS

A. Diligence and Reasonable Basis

> Members and Candidates must:
>
> 1. Exercise diligence, independence, and thoroughness in analyzing investments, making investment recommendations, and taking investment actions.
>
> 2. Have a reasonable and adequate basis, supported by appropriate research and investigation, for any investment analysis, recommendation, or action.

Harrel and Chong

Corix Bioscience is a startup company in the manufacturing and distribution of cannabidiol (CBD) products. To promote the company, the Corix CEO hires Harrel, an independent research analyst and CFA Program candidate, to write and distribute a research report on the company. The CEO tells Harrel the following:

- Corix has an agreement with indigenous tribes that allows it to access tribal lands for commercial hemp and cannabis farming and to sell hemp and cannabis products in retail outlets on tribal lands.
- Corix has a certificate of compliance from the national regulator that permits the company to transport, process, and export industrial hemp products.
- The prior year's harvest of hemp exceeded expectations in both quality and quantity, resulting in a substantial inventory of product.

Harrelson includes all this information in a research report and provides a positive analysis of the company. However, Corix does not have agreements with indigenous tribes; nor does it have regulatory approval, the certificate of compliance is a forgery; and Corix never cultivated, or harvested significant quantities of commercial hemp.

Chong, a CFA charterholder and research analyst at Nature's Harvest Investment Management (NHIM), incorporates the information and conclusions from Harrel's research report into his own research on Corix and includes a "buy" recommendation on the company. Chong's report is distributed to portfolio managers at NHIM. Corix is ultimately shown to be a fake operation, leading to substantial losses for NHIM's clients.

Which individual(s) most likely violated the CFA Institute Code and Standards?

A Harrel only

B Chong only

C Both Harrel and Chong

Analysis C is correct. Standard V(A): Investment Analysis, Recommendations, and Actions, Diligence and Reasonable Basis states that CFA Institute members and candidates must exercise diligence and thoroughness in analyzing investments and must have a reasonable and adequate basis that is supported by appropriate research and investigation for any investment recommendation. Harrel and Chong do not meet the requirements of this standard. By relying on the statements given by Corix and not conducting an independent investigation into the accuracy of the information, Harrel did not exercise diligence and thoroughness in analyzing the company. Chong seemingly relied on Harrel's research to formulate his "buy" recommendation without conducting his own independent research. Investment professionals who rely on third-party research must make reasonable and diligent efforts to determine whether that research is correct. The facts do not indicate that Chong independently verified the information, critically assessed Harrel's research, or had reason to rely on Harrel's report based on past experience and familiarity with the quality of Harrel's work. In addition, Harrel works as an independent researcher and thus has no management structure that could enforce due diligence standards, which should raise concerns about the quality of Harrel's due diligence. Chong also did not recognize that Harrel's work was issuer-paid research and thus subject to increased scrutiny. NHIM should have had clear due diligence policies and procedures in place that Chong would need to follow before completing and disseminating research on behalf of the firm.

A is incorrect. Both Harrel and Chong violated the Code and Standards.

B is incorrect. Both Harrel and Chong violated the Code and Standards.

B. Communication with Clients and Prospective Clients

> Members and Candidates must:
>
> 1. Disclose to clients and prospective clients the basic format and general principles of the investment processes they use to analyze investments, select securities, and construct portfolios and must promptly disclose any changes that might materially affect those processes.
> 2. Disclose to clients and prospective clients significant limitations and risks associated with the investment process.
> 3. Use reasonable judgment in identifying which factors are important to their investment analyses, recommendations, or actions and include those factors in communications with clients and prospective clients.
> 4. Distinguish between fact and opinion in the presentation of investment analysis and recommendations.

Maalouf

Maalouf, CFA, works for a large wealth management firm. The firm's fees are calculated as a percentage of the asset value managed for each client account. The firm has a standard method for valuing assets and calculating fees, which is disclosed to clients when they open their account. Over time, the firm transitions to (1) using the market value of client assets at the end of the billing cycle instead of the average daily balance of the account; (2) including assets that were previously excluded, such as cash or cash equivalents, in the fee calculation; and (3) charging clients for a full billing period rather than prorating fees for clients that start or terminate accounts mid-billing period.

Under the CFA Institute Code and Standards, Maalouf

A must notify clients of the changes in the valuation and fee calculation method.

B cannot use end-of-cycle valuations, include cash equivalents, or charge fees for a full billing cycle for partial cycle accounts.

C can change the valuation and fee calculation methodology as long as actual fees charged to clients are lower.

Analysis A is correct. Standard V(B): Investment Analysis, Recommendations, and Actions, Communication with Clients and Prospective Clients requires CFA Institute members and candidates to disclose to clients the basic format and general principles of the investment process. Advisory fees are a critical part of that process. Developing and maintaining clear, frequent, and thorough communication with clients allows those clients to make well-informed decisions about their investments, including whether to engage or retain an investment adviser. Any changes to the method for valuing assets or calculating fees that are different from the process the client originally agreed to must be disclosed. Maalouf must notify his clients ahead of the changes.

B is incorrect. Using end-of-cycle valuations, including cash equivalents, or not prorating fees for new or terminated clients are acceptable methods for calculating fees as long as those policies are disclosed and agreed to by the client.

C is incorrect. Changing valuation and fee calculation policies over time for existing accounts in a way that might result in higher fees is permissible, however Maalouf and his firm should negotiate with their clients on fees and changing the calculation methodology that were disclosed when the clients opened their accounts. Changing the methodology for the fee calculation without disclosing the change is improper, even if it results in lower fees.

Dukis

Dukis, CFA, is a managing director at a global credit ratings service. She is responsible for the group that assigns new issue and surveillance credit ratings to commercial mortgage-backed securities (CMBSs). To determine the ratings, Dukis and her group calculate the debt service coverage ratio (DSCR) of each security, a key quantitative metric used to rate CMBSs. Shortly after the global financial crisis, the ratings agency changes its methodology for calculating the DSCR for certain securities so that it more accurately reflects risk. Dukis's group publishes subsequent credit ratings without disclosing the change. When the new methodology is used, the securities receive higher credit ratings than they would have received if the original methodology had been used.

Dukis's actions are

A appropriate because the new methodology more accurately reflects risk.

B a violation of the CFA Institute Code and Standards because she did not disclose the change in methodology to the investing public.

C appropriate because no disclosure is necessary, given that calculating DSCR is only one element in determining the overall rating of the security.

Analysis B is correct. Standard V(B): Investment Analysis, Recommendations, and Actions, Communication with Clients and Prospective Clients requires CFA Institute members and candidates to disclose to investors the basic format and general principles of the investment process they use to analyze investments as well as any changes that might materially affect those processes. Ratings agencies' consistency and transparency are important to investors. If rating methodologies are not applied consistently, ratings might not be easily comparable. Similarly, without transparency, investors cannot assess the methodologies used by the credit ratings agency or the application of those methodologies, which means they cannot determine how much weight to give the rating. Dukis should have disclosed to investors the change in methodology for calculating the DSCR.

A is incorrect. Even if the new methodology does more accurately reflect risk, the change must still be communicated to investors.

C is incorrect. The DSCR is clearly a key quantitative metric used to rate the securities because the change in methodology materially affected the credit ratings by rating them higher than the original method. A change in such a key factor must be disclosed.

C. Record Retention

> Members and Candidates must develop and maintain appropriate records to support their investment analyses, recommendations, actions, and other investment-related communications with clients and prospective clients.

Duermott

Duermott, CFA, is president of Enhanced Investment Strategies (EIS), a small investment firm. Most clients of EIS are longtime associates of Duermott who have had their investment portfolios with the firm for decades. Because of his close personal relationship with his clients, Duermott is very familiar with their investment profile, income and retirement requirements, and risk tolerance. He keeps up with all his clients' life-changing events—such as health issues, real estate purchases, children's university expenses, and retirement—and adjusts the clients' portfolios accordingly. Duermott regularly meets with his clients at EIS's offices, and he also sees them on numerous occasions outside the office, which give him additional opportunities to update them on their investments. EIS clients complete a client agreement and risk profile when opening their account, and those profiles are updated whenever Duermott finds the time to do so.

Duermott's business practices are

A a violation of the CFA Institute Code and Standards.

B acceptable because he regularly communicates with clients about their investments.

C acceptable because he adjusts clients' investments to ensure that they are suitable for the clients' needs given their changing income and risk profile.

Analysis A is correct. Standard V(C): Investment Analysis, Recommendations, and Actions, Record Retention states that CFA Institute members and candidates must "develop and maintain appropriate records to support their investment analyses, recommendations, actions, and other investment-related communications with clients and prospective clients." In this case, Duermott is personally close to his clients, but he updates client records only when he "finds the time to do so," which does not appear to be promptly or regularly. Duermott has a responsibility as his clients' adviser and as the president of his company to maintain appropriate records when client circumstances change. Without necessary, relevant, and up-to-date know-your-client information, Duermott would have difficulty establishing and proving that EIS has identified the needs and circumstances of its clients and has taken them into account in recommending investments. When client circumstances, investment goals, risk tolerances, or income needs change, records should be promptly updated and reviewed regularly to document these changes.

B is incorrect. Although Duermott is fulfilling his ethical obligations as an investment manager by communicating regularly with his clients, he is not keeping regular, up-to-date client records.

C is incorrect. Although Duermott is reviewing and adjusting client portfolios on a timely basis to meet clients' changing financial circumstances, he is not keeping regular, up-to-date client records.

VI. CONFLICTS OF INTEREST

A. Disclosure of Conflicts

> Members and Candidates must make full and fair disclosure of all matters that could reasonably be expected to impair their independence and objectivity or interfere with respective duties to their clients, prospective clients, and employer. Members and Candidates must ensure that such disclosures are prominent, are delivered in plain language, and communicate the relevant information effectively.

Reebh

Reebh, CFA, is the CEO and founding partner of Lux Asset Management (Lux). Reebh provides asset management and allocation services for high-net-worth individuals and several small institutional clients. His services include investing client funds with third-party subadvisers who have a specialty in a particular asset class. Reebh's clients are aware, and approve, of Lux's allocation of their assets to subadvisers. The third-party subadvisers make payments to Lux based on the total value of a client's assets placed or invested in the subadvisers' funds.

Reebh's actions are

A appropriate because Reebh has disclosed the use of subadvisers.

B inappropriate because the payments are an improper referral fee.

C inappropriate unless Reebh discloses the financial arrangement he has with the subadvisers to his clients.

Analysis C is correct. Standard VI(A): Conflicts of Interest, Disclosure of Conflicts requires CFA Institute members and candidates to "make full and fair disclosure of all matters that could reasonably be expected to impair their independence and objectivity or interfere with respective duties to their clients." The payments subadvisers make to Lux based on the value of the client assets Lux places with the subadvisers creates a potential conflict of interest because Reebh is thereby incentivized to hire subadvisers who pay the fee but who might not necessarily be the best subadvisers for his clients. To mitigate this conflict, Reebh must disclose the financial incentive to clients.

A is incorrect. Reebh has disclosed Lux's use of subadvisers, but he has not disclosed the financial incentive for Lux to use those subadvisers.

B is incorrect. Although referral arrangements might be acceptable with full disclosure to clients, Reebh is not referring clients to the subadvisers but hiring them directly on his clients' behalf.

B. Priority of Transactions

> Investment transactions for clients and employers must have priority over investment transactions in which a Member or Candidate is the beneficial owner.

Yang

Yang, CFA, is a research analyst at Dacco, a registered broker/dealer and investment adviser. While employed with Dacco, Yang establishes Prestige Trade Investments Limited (Prestige) and acts as investment adviser for the firm's clients. Yang is responsible for formulating Prestige's investment strategy and directs all trades on behalf of Prestige. Over several days, Yang purchases 50,000 shares of Zhongpin stock and 1,978 Zhongpin call options for his personal account at Dacco. Shortly after, Yang uses $29.8 million of Prestige's funds to purchase more than 3 million shares of Zhongpin stock.

Yang's actions are

A a violation of the CFA Institute Code and Standards.

B acceptable because Yang's personal investments are not in conflict with the investment advice being given to his clients at Prestige.

C acceptable as long as Prestige clients are not negatively affected by Yang's prior purchase of Zhongpin securities through his account at Dacco.

Analysis A is correct. Standard VI(B): Conflicts of Interest, Priority of Transactions states that "investment transactions for clients . . . must have priority over investment transactions in which a [CFA Institute] Member or Candidate is the beneficial owner." Yang is "front-running" his Prestige clients' trades. Front-running involves trading for one's personal account before trading for client accounts. Yang purchases Zhongpin stock and call options in his personal account at Dacco before directing the Zhongpin trades for clients at Prestige. The $29.8 million of Prestige funds invested in Zhongpin stock could have a material upward effect on the price of the stock and options Yang holds. Even the perception that Yang could profit by using Prestige's funds will diminish investors' trust in Yang and the capital markets.

B is incorrect. Yang's personal investments are tracking with his client investments, so no conflict exists between his personal trading and the investment actions/advice for clients. However the timing of the trades is the issue in this case because Yang is "front-running" his clients' trades.

C is incorrect. The fact that Prestige clients are not harmed by Yang's earlier trades for his personal accounts does not make his actions acceptable.

Kapadia

Kapadia, CFA, is a trader for an asset management company that manages several large global mutual funds. Kapadia executes the equity buy-and-sell orders for the portfolio managers of one of the company's mutual funds. He has the discretion to execute the orders at any time during the day, depending on market conditions. Before executing the orders, Kapadia contacts several close friends and relatives to give them information on which securities the mutual fund will be trading. In turn, these friends and relatives make trades that mirror the imminent trades to be executed by Kapadia on behalf of the mutual fund.

Kapadia's actions are

A a violation of the CFA Institute Code and Standards.

B inappropriate only if the client is harmed financially by the conduct.

C appropriate because he does not share confidential information about individual clients.

Analysis A is correct. Standard VI(B): Conflicts of Interest, Priority of Transactions states that "investment transactions for clients must have priority over investment transactions" for a member or candidate's personal benefit. Kapadia is facilitating front-running by his friends and relatives on the trades of his employer's mutual fund. Front-running is the unethical and often illegal practice of trading on advance information for one's personal account before trading for client accounts to gain an economic advantage. Although Kapadia might not directly benefit financially, he benefits personally by providing the information to those with whom he has close relationships.

B is incorrect. Kapadia's engagement in the practice of front-running, which involves trading in personal accounts before trading for client accounts, is unethical and inappropriate even if the trades of Kapadia's friends and relatives do not disadvantage the mutual fund by moving the price of the security or causing the fund to lose the price advantage or any profit from its own trades.

C is incorrect. Although Kapadia does not share the confidential information of individual clients or individual investors in the fund, he does share confidential information about the fund itself.

Perrkins

Perrkins, CFA, is the chief investment officer of GT Financial (GTF). Perrkins's wife is GTF's compliance officer. GTF has several dozen retail clients and total assets under management of $70 million. All client assets are managed on a discretionary basis. Perrkins frequently makes trades for his clients using an omnibus trading account through a broker/dealer, which allows Perrkins to buy and sell securities in a block trade on behalf of multiple clients simultaneously. Perrkins regularly allocates the securities purchases to individual client accounts after the market closes. Over one six-month period, Perrkins allocates 75% of the profitable trades to nine accounts that Perrkins and his wife own or control. At the same time, he allocates 82% of the unprofitable trades to the account of the three largest GTF clients.

Perrkins's actions are

A a violation of the CFA Institute Code and Standards.

B acceptable as long as he discloses the trade allocation practices to his clients.

C acceptable as long as he reverses his trade allocation practices to favor the larger clients so that they are not harmed over the long term.

Analysis A is correct. Standard VI(B): Conflicts of Interest, Priority of Transactions states that investment transactions for clients have priority over personal transactions. By trading in the firm's omnibus account and then delaying the allocation of trades to a specific account until he has an opportunity to observe the security's intraday performance, Perrkins can pick the winning trades for accounts in which he has a beneficial interest. This practice is a violation of Standard VI(B). He then allocates the losing trades to clients' accounts that are large enough to absorb incremental trading losses without arousing suspicion that the losses are due to fraud. Perrkins's actions are also likely a violation of Standard III(B): Duties to Clients, Fair Dealing, which states that members and candidates "must deal fairly and objectively with all clients when . . . taking investment action, or engaging in other professional activities."

B is incorrect. Disclosure to his clients as to how he is allocating the trades does not make Perrkins unethical, fraudulent behavior acceptable.

C is incorrect. Perrkins cannot temporarily favor his personal interests over his clients' interests with the intent of rectifying the situation for his clients in the future. Ethical conduct is not subject to a ledger-keeping exercise. CFA Institute members and candidates must comply with the ethical principles and requirements of the Code and Standards at all times.

C. Referral Fees

> Members and Candidates must disclose to their employer, clients, and prospective clients, as appropriate, any compensation, consideration, or benefit received from or paid to others for the recommendation of products or services.

Kiang

Kiang, CFA, is a successful investment adviser with several high-net-worth clients who are very happy with his services. Many of Kiang's clients recommend his advisory services to their friends and family, and Kiang encourages these recommendations to build his business. Each year, Kiang hosts an elaborate party for clients who have referred new clients to his advisory firm. At the party, Kiang distributes nominal gift cards to attendees. In some cases, Kiang offers discounts on advisory fees to clients who sent him referrals that proved particularly lucrative. Many of the clients attending these celebrations were referred to Kiang by other clients, and they, in turn, continue the cycle of recommending Kiang to a wider circle of friends and family.

Kiang's actions most likely are

A acceptable as a reward for client loyalty.

B acceptable because he treats all clients fairly.

C a violation of the CFA Institute Code and Standards.

Analysis C is correct. Standard VI(C): Conflicts of Interest, Referral Fees states that members and candidates "must disclose to their employer, clients, and prospective clients, as appropriate, any compensation, consideration, or benefit . . . paid to others for the recommendation of . . . services." In this case, Kiang hosts an elaborate party, distributes gift cards, and, in some cases, offers discounted advisory fees to clients who have referred particularly lucrative clients. Under the standard, these benefits would be considered referral fees that must be disclosed. The facts do not indicate that Kiang makes any disclosure to potential clients. The fact that some prospects, upon becoming clients, become aware that he pays for referrals when they receive discounted fees is insufficient disclosure. It would be acceptable if Kiang were to host a party or give gift cards to all his clients to reward their loyalty, regardless as to whether or not they provided referrals.

A is incorrect. The party, gift cards, and discounted advisory fees are more than a reward for client loyalty and would be considered referral fees that are provided only to existing clients for recommending Kiang's services.

B is incorrect. Arguably, Kiang treats his clients fairly because he offers all of his clients the opportunity to receive these benefits and fee discounts as long as they refer others to his business and whether the clients access these benefits by making referrals is up to them. However regardless as to whether Kiang is treating all clients fairly, he is violating Standard VI(C) by not disclosing the benefits and compensation he awards for referrals.

RESPONSIBILITIES AS A CFA INSTITUTE MEMBER OR CFA CANDIDATE

VII.

A. Conduct as Participants in CFA Institute Programs

> Members and Candidates must not engage in any conduct that compromises the reputation or integrity of CFA Institute or the CFA designation or the integrity, validity, or security of the CFA Institute programs.

Taveras

Taveras, CFA, leads an exam preparation course sponsored by his local society. The society hosts a celebration for the students after the exam is over. During the celebration, several of Taveras's students describe their experience of taking the exam. Most give their opinion on the relative difficulty of the exam compared to their expectations, and some describe their surprise about areas of the curriculum that were not tested. Taveras asks his students for their opinions on the most difficult exam questions.

Under the CFA Institute Code and Standards, Taveras is most likely

A prohibited from discussing the exam with students after it is over.

B free to pass along information about the exam to candidates in future prep courses to help prepare them for the exam.

C allowed to share the opinions of his students about the difficulty of the exam with candidates in future prep courses to emphasize the need to thoroughly prepare.

Analysis B is correct. Standard VII(A): Responsibilities as a CFA Institute Member or CFA Candidate, Conduct as Participants in CFA Institute Programs states that candidates "must not engage in any conduct that compromises . . . the integrity, validity, or security of CFA Institute programs. For Taveras to share with future prep course participants the opinions of his previous students who found the CFA exam more difficult than expected is acceptable as a way to encourage future students to study the curriculum thoroughly and prepare as much as possible. However, Tavares should not solicit or pass on information regarding the specifics of the exam.

A is incorrect. A group of candidates who collectively completed the rigorous process of studying for and taking the CFA exam will naturally want to celebrate the accomplishment and discuss the exam after it is over. Candidates are allowed to discuss their exam experience with Taveras in general terms, but they cannot provide specific information about the exam regarding the questions or the general areas tested.

C is incorrect. Taveras cannot pass along specific information to future candidates and should not be soliciting information about specific questions; doing so would be a violation of Standard VII(A), which is designed to protect the integrity and security of future exams.

B. Reference to CFA Institute, the CFA Designation, and the CFA Program

> When referring to CFA Institute, CFA Institute membership, the CFA designation, or candidacy in the CFA Program, Members and Candidates must not misrepresent or exaggerate the meaning or implications of membership in CFA Institute, holding the CFA designation, or candidacy in the CFA program.

Ahmed

Ahmed recently earned his CFA designation and joined a medium-sized hedge fund as a senior analyst. His supervisor, Bennett, the firm's founder, earned her CFA designation 10 years ago and proudly uses the CFA designation on her business card and on all marketing materials for the fund. Bennett shares with Ahmed that she has not paid her CFA Institute membership dues for the past four years and no longer participates in the organization's continuing education program. When Ahmed asks Bennett about her use of the designation, she states that by passing the CFA exam she earned the CFA charter, and that the credential is like a university degree that cannot be taken away. Later, during a marketing meeting by the two to a potential investor, the investor notes that he narrowed his manager search to only firms that employ CFA charterholders in senior positions. When he asks Bennett if everyone in the firm on the investment side is a CFA charterholder, she responds, "Yes, that is correct." Ahmed does not respond.

Did either Ahmed or Bennett violate the CFA Institute Code and Standards?

A Ahmed violated the Code and Standards, but Bennett did not.

B Bennett violated the Code and Standards, but Ahmed did not.

C Both Ahmed and Bennett violated the Code and Standards.

Analysis C is correct. Standard VII(B): Responsibilities as a CFA Institute Member or CFA Candidate, Reference to CFA Institute, the CFA Designation, and the CFA Program states that when referring to the CFA designation, members and candidates "must not misrepresent . . . holding the CFA designation." The CFA designation is not like a degree from a university because once individuals have been granted the right to use the designation, they must also satisfy CFA Institute membership requirements, which include paying dues, to maintain the right to refer to themselves as CFA charterholders. Participation in the CFA Institute Professional Learning program is not mandatory for members to maintain their designation, but it is encouraged as a way to meet the CFA Institute Code of Ethics provision that members maintain and improve their professional competence. Bennett's membership is considered lapsed because she has not been paying dues to CFA Institute. Until she reactivates her membership, Bennett is violating Standard VII(B) by continuing to use the CFA designation and representing herself as a charterholder to a potential client. Ahmed knows that Bennett's CFA Institute membership has lapsed. Standard I(A): Professionalism, Knowledge of the Law prohibits members and candidates from knowingly participating or assisting in the violations of others and requires members and candidates to dissociate from any unethical or illegal conduct. By staying silent in a sales meeting in which he knows false information is being given to a potential investor that could cause harm to that investor, Ahmed would be seen as assisting Bennett in providing that false information, even though Ahmed is not actively engaging in the misconduct himself. Best practice would be for Ahmed to address Bennett directly about her conduct and ask her to reinstate her membership or correct the statement she made to the potential investor. If Bennett refuses to take corrective action, Ahmed should report her conduct to the fund's compliance department for it to address and should dissociate himself from the activity by not participating in any additional sales meetings with Bennett.

A is not correct because both Ahmed and Bennett violated the Code and Standards

B is not correct because both Ahmed and Bennett violated the Code and Standards.

Glossary

A priori probability A probability based on logical analysis rather than on observation or personal judgment.

Abandonment option The ability to terminate an investment at some future time if the financial results are disappointing.

Abnormal return The amount by which a security's actual return differs from its expected return, given the security's risk and the market's return.

Absolute advantage A country's ability to produce a good or service at a lower absolute cost than its trading partner.

Absolute dispersion The amount of variability present without comparison to any reference point or benchmark.

Absolute frequency The actual number of observations counted for each unique value of the variable (also called raw frequency).

Accelerated book build An offering of securities by an investment bank acting as principal that is accomplished in only one or two days.

Accelerated methods Depreciation methods that allocate a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Accounting costs Monetary value of economic resources used in performing an activity. These can be explicit, out-of-pocket, current payments, or an allocation of historical payments (depreciation) for resources. They do not include implicit opportunity costs.

Accounting profit Income as reported on the income statement, in accordance with prevailing accounting standards, before the provisions for income tax expense. Also called *income before taxes* or *pretax income*.

Accounts payable Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Accrued expenses Liabilities related to expenses that have been incurred but not yet paid as of the end of an accounting period—an example of an accrued expense is rent that has been incurred but not yet paid, resulting in a liability "rent payable." Also called *accrued liabilities*.

Accrued interest Interest earned but not yet paid.

Acquisition method A method of accounting for a business combination where the acquirer is required to measure each identifiable asset and liability at fair value. This method was the result of a joint project of the IASB and FASB aiming at convergence in standards for the accounting of business combinations.

Action lag Delay from policy decisions to implementation.

Active investment An approach to investing in which the investor seeks to outperform a given benchmark.

Active return The return on a portfolio minus the return on the portfolio's benchmark.

Activity ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory. Also called *asset utilization ratios* or *operating efficiency ratios*.

Add-on rates Bank certificates of deposit, repos, and indexes such as Libor and Euribor are quoted on an add-on rate basis (bond equivalent yield basis).

Addition rule for probabilities A principle stating that the probability that A or B occurs (both occur) equals the probability that A occurs, plus the probability that B occurs, minus the probability that both A and B occur.

Agency bond See *quasi-government bond*.

Agency costs Costs associated with the conflict of interest present between principals and agents when a company is managed by non-owners. Agency costs result from the inherent conflicts of interest between managers, bondholders, and equity owners.

Agency costs of debt Costs arising from conflicts of interest between managers and debtholders.

Agency costs of equity The smaller the stake managers have in the company, the less their share in bearing the cost of excessive perquisite consumption—consequently, the less their desire to give their best efforts in running the company.

Agency RMBS In the United States, securities backed by residential mortgage loans and guaranteed by a federal agency or guaranteed by either of the two GSEs (Fannie Mae and Freddie Mac).

Aggregate demand The quantity of goods and services that households, businesses, government, and non-domestic customers want to buy at any given level of prices.

Aggregate demand curve Inverse relationship between the price level and real output.

Aggregate income The value of all the payments earned by the suppliers of factors used in the production of goods and services.

Aggregate output The value of all the goods and services produced during a specified period.

Aggregate supply The quantity of goods and services producers are willing to supply at any given level of price.

Aggregate supply curve The level of domestic output that companies will produce at each price level.

All-or-nothing (AON) orders An order that includes the instruction to trade only if the trade fills the entire quantity (size) specified.

Allocationally efficient A characteristic of a market, a financial system, or an economy that promotes the allocation of resources to their highest value uses.

Alternative data Non-traditional data types generated by the use of electronic devices, social media, satellite and sensor networks, and company exhaust.

Alternative hypothesis The hypothesis that is accepted if the null hypothesis is rejected.

Alternative investment markets Market for investments other than traditional securities investments (i.e., traditional common and preferred shares and traditional fixed income instruments). The term usually encompasses direct and indirect investment in real estate (including timberland and farmland) and commodities (including precious metals); hedge funds, private equity, and other investments requiring specialized due diligence.

Alternative trading systems Trading venues that function like exchanges but that do not exercise regulatory authority over their subscribers except with respect to the conduct

of the subscribers' trading in their trading systems. Also called *electronic communications networks* or *multilateral trading facilities*.

American depository receipt A US dollar-denominated security that trades like a common share on US exchanges.

American depository share The underlying shares on which American depository receipts are based. They trade in the issuing company's domestic market.

American-style Type of option contract that can be exercised at any time up to the option's expiration date.

Amortisation The process of allocating the cost of intangible long-term assets having a finite useful life to accounting periods; the allocation of the amount of a bond premium or discount to the periods remaining until bond maturity.

Amortised cost The historical cost (initially recognised cost) of an asset, adjusted for amortisation and impairment.

Amortizing bond Bond with a payment schedule that calls for periodic payments of interest and repayments of principal.

Amortizing loans Loans with a payment schedule that calls for periodic payments of interest and repayments of principal.

Analysis of variance (ANOVA) The analysis that breaks the total variability of a dataset (such as observations on the dependent variable in a regression) into components representing different sources of variation. With reference to regression, ANOVA provides the inputs for an F-test of the significance of the regression as a whole, as well as the inputs for the coefficient of determination and the standard error of the estimate.

Analytical duration The use of mathematical formulas to estimate the impact of benchmark yield-to-maturity changes on bond prices.

Anchoring and adjustment bias An information-processing bias in which the use of a psychological heuristic influences the way people estimate probabilities.

Annual percentage rate The cost of borrowing expressed as a yearly rate.

Annuity A finite set of level sequential cash flows.

Annuity due An annuity having a first cash flow that is paid immediately.

Anomalies Apparent deviations from market efficiency.

Antidilutive With reference to a transaction or a security, one that would increase earnings per share (EPS) or result in EPS higher than the company's basic EPS—antidilutive securities are not included in the calculation of diluted EPS.

Arbitrage 1) The simultaneous purchase of an undervalued asset or portfolio and sale of an overvalued but equivalent asset or portfolio, in order to obtain a riskless profit on the price differential. Taking advantage of a market inefficiency in a risk-free manner. 2) The condition in a financial market in which equivalent assets or combinations of assets sell for two different prices, creating an opportunity to profit at no risk with no commitment of money. In a well-functioning financial market, few arbitrage opportunities are possible. 3) A risk-free operation that earns an expected positive net profit but requires no net investment of money.

Arbitrage-free pricing The overall process of pricing derivatives by arbitrage and risk neutrality. Also called the *principle of no arbitrage*.

Arbitrageurs Traders who engage in arbitrage. See *arbitrage*.

Arithmetic mean The sum of the observations divided by the number of observations.

Artificial intelligence Computer systems that exhibit cognitive and decision-making ability comparable (or superior) to that of humans.

Ask The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size). See *offer*.

Ask size The maximum quantity of an asset that pertains to a specific ask price from a trader. For example, if the ask for a share issue is $30 for a size of 1,000 shares, the trader is offering to sell at $30 up to 1,000 shares.

Asset allocation The process of determining how investment funds should be distributed among asset classes.

Asset-backed securities A type of bond issued by a legal entity called a *special purpose entity* (SPE) on a collection of assets that the SPE owns. Also, securities backed by receivables and loans other than mortgages.

Asset-based valuation models Valuation based on estimates of the market value of a company's assets.

Asset beta The unlevered beta; which reflects the business risk of the assets; the assets' systematic risk.

Asset class A group of assets that have similar characteristics, attributes, and risk–return relationships.

Asset swap Converts the periodic fixed coupon of a specific bond to an MRR plus or minus a spread.

Asset utilization ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Assets Resources controlled by an enterprise as a result of past events and from which future economic benefits to the enterprise are expected to flow.

Assignment of accounts receivable The use of accounts receivable as collateral for a loan.

Asymmetric information The differential of information between corporate insiders and outsiders regarding the company's performance and prospects. Managers typically have more information about the company's performance and prospects than owners and creditors.

At the money An option in which the underlying's price equals the exercise price.

Auction A type of bond issuing mechanism often used for sovereign bonds that involves bidding.

Autarkic price The price of a good or service in an autarkic economy.

Autarky A state in which a country does not trade with other countries.

Automatic stabilizer A countercyclical factor that automatically comes into play as an economy slows and unemployment rises.

Availability bias An information-processing bias in which people take a heuristic approach to estimating the probability of an outcome based on how easily the outcome comes to mind.

Available-for-sale Under US GAAP, debt securities not classified as either held-to-maturity or held-for-trading securities. The investor is willing to sell but not actively planning to sell. In general, available-for-sale debt securities are reported at fair value on the balance sheet, with unrealized gains included as a component of other comprehensive income.

Average fixed cost Total fixed cost divided by quantity produced.

Average life See *weighted average life*.

Average product Measures the productivity of inputs on average and is calculated by dividing total product by the total number of units for a given input that is used to generate that output.

Average revenue Total revenue divided by quantity sold.

Average total cost Total cost divided by quantity produced.

Average variable cost Total variable cost divided by quantity produced.

Back-testing With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backfill bias A bias introduced if a fund backfills its past performance data to the database when it first starts reporting actual performance.

Backtesting With reference to portfolio strategies, the application of a strategy's portfolio selection rules to historical data to assess what would have been the strategy's historical performance.

Backup line of credit A type of credit enhancement provided by a bank to an issuer of commercial paper to ensure that the issuer will have access to sufficient liquidity to repay maturing commercial paper if issuing new paper is not a viable option.

Backwardation A condition in the futures markets in which the spot price exceeds the futures price, the forward curve is downward sloping, and the convenience yield is high.

Balance of payments A double-entry bookkeeping system that summarizes a country's economic transactions with the rest of the world for a particular period of time, typically a calendar quarter or year.

Balance of trade deficit When the domestic economy is spending more on non-domestic goods and services than non-domestic economies are spending on domestic goods and services.

Balance sheet The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet). Also called *statement of financial position* or *statement of financial condition.*

Balance sheet ratios Financial ratios involving balance sheet items only.

Balanced With respect to a government budget, one in which spending and revenues (taxes) are equal.

Balloon payment Large payment required at maturity to retire a bond's outstanding principal amount.

Bar chart A chart for plotting the frequency distribution of categorical data, where each bar represents a distinct category and each bar's height is proportional to the frequency of the corresponding category. In technical analysis, a bar chart that plots four bits of data for each time interval—the high, low, opening, and closing prices. A vertical line connects the high and low prices. A cross-hatch left indicates the opening price and a cross-hatch right indicates the closing price.

Barter economy An economy where economic agents as house-holds, corporations, and governments "pay" for goods and services with another good or service.

Base-rate neglect A type of representativeness bias in which the base rate or probability of the categorization is not adequately considered.

Base rates The reference rate on which a bank bases lending rates to all other customers.

Basic EPS Net earnings available to common shareholders (i.e., net income minus preferred dividends) divided by the weighted average number of common shares outstanding.

Basis point Used in stating yield spreads, one basis point equals one-hundredth of a percentage point, or 0.01%.

Basket of listed depository receipts An exchange-traded fund (ETF) that represents a portfolio of depository receipts.

Bayes' formula The rule for updating the probability of an event of interest—given a set of prior probabilities for the event, information, and information given the event—if you receive new information.

Bearer bonds Bonds for which ownership is not recorded; only the clearing system knows who the bond owner is.

Bearish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from above; this movement is considered bearish. A **death cross** is a bearish crossover based on 50-day and 200-day moving averages.

Behavioral finance A field of finance that examines the psychological variables that affect and often distort the investment decision making of investors, analysts, and portfolio managers.

Behind the market Said of prices specified in orders that are worse than the best current price; e.g., for a limit buy order, a limit price below the best bid.

Benchmark A comparison portfolio; a point of reference or comparison.

Benchmark issue The latest sovereign bond issue for a given maturity. It serves as a benchmark against which to compare bonds that have the same features but that are issued by another type of issuer.

Benchmark rate Typically the yield-to-maturity on a government bond having the same or close to the same time-to-maturity.

Benchmark spread The yield spread over a specific benchmark, usually measured in basis points.

Bernoulli random variable A random variable having the outcomes 0 and 1.

Bernoulli trial An experiment that can produce one of two outcomes.

Best bid The highest bid in the market.

Best effort offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-efforts offering An offering of a security using an investment bank in which the investment bank, as agent for the issuer, promises to use its best efforts to sell the offering but does not guarantee that a specific amount will be sold.

Best-in-class An ESG implementation approach that seeks to identify the most favorable companies in an industry based on ESG considerations.

Best offer The lowest offer (ask price) in the market.

Beta A measure of the sensitivity of a given investment or portfolio to movements in the overall market.

Bid The price at which a dealer or trader is willing to buy an asset, typically qualified by a maximum quantity.

Bid–ask spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid–offer spread The difference between the prices at which dealers will buy from a customer (bid) and sell to a customer (offer or ask). It is often used as an indicator of liquidity.

Bid size The maximum quantity of an asset that pertains to a specific bid price from a trader.

Big Data The vast amount of data being generated by industry, governments, individuals, and electronic devices that arises from both traditional and non-traditional data sources.

Bilateral loan A loan from a single lender to a single borrower.

Bimodal A distribution that has two most frequently occurring values.

Binomial model A model for pricing options in which the underlying price can move to only one of two possible new prices.

Binomial random variable The number of successes in n Bernoulli trials for which the probability of success is constant for all trials and the trials are independent.

Binomial tree The graphical representation of a model of asset price dynamics in which, at each period, the asset moves up with probability p or down with probability $(1 - p)$.

Bitcoin A cryptocurrency using blockchain technology that was created in 2009.

Bivariate correlation See *Pearson correlation*.

Block brokers A broker (agent) that provides brokerage services for large-size trades.

Blockchain A type of digital ledger in which information is recorded sequentially and then linked together and secured using cryptographic methods.

Blue chip Widely held large market capitalization companies that are considered financially sound and are leaders in their respective industry or local stock market.

Bollinger Bands A price-based technical analysis indicator consisting of a line representing the moving average, a higher line representing the moving average plus a set number of standard deviations from the average (for the same number of periods as was used to calculate the moving average), and a lower line representing the moving average minus the same number of standard deviations.

Bond Contractual agreement between the issuer and the bondholders.

Bond equivalent yield A calculation of yield that is annualized using the ratio of 365 to the number of days to maturity. Bond equivalent yield allows for the restatement and comparison of securities with different compounding periods.

Bond indenture The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *trust deed*.

Bond market vigilantes Bond market participants who might reduce their demand for long-term bonds, thus pushing up their yields.

Bond yield plus risk premium approach An estimate of the cost of common equity that is produced by summing the before-tax cost of debt and a risk premium that captures the additional yield on a company's stock relative to its bonds. The additional yield is often estimated using historical spreads between bond yields and stock yields.

Bonus issue of shares A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Book building Investment bankers' process of compiling a "book" or list of indications of interest to buy part of an offering.

Book value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. Also called *carrying value*.

Boom An expansionary phase characterized by economic growth "testing the limits" of the economy.

Bootstrap A resampling method that repeatedly draws samples with replacement of the selected elements from the original observed sample. Bootstrap is usually conducted by using computer simulation and is often used to find standard error or construct confidence intervals of population parameters.

Bottom-up analysis An investment selection approach that focuses on company-specific circumstances rather than emphasizing economic cycles or industry analysis.

Box and whisker plot A graphic for visualizing the dispersion of data across quartiles. It consists of a "box" with "whiskers" connected to the box.

Breakdown A breakdown occurs when the price of an asset moves below a support level.

Breakeven point The number of units produced and sold at which the company's net income is zero (Revenues = Total cost); in the case of perfect competition, the quantity at which price, average revenue, and marginal revenue equal average total cost.

Breakout A breakout occurs when the price of an asset moves above a resistance level.

Bridge financing Interim financing that provides funds until permanent financing can be arranged.

Broad money Encompasses narrow money plus the entire range of liquid assets that can be used to make purchases.

Broker 1) An agent who executes orders to buy or sell securities on behalf of a client in exchange for a commission. 2) See *futures commission merchants*.

Broker–dealer A financial intermediary (often a company) that may function as a principal (dealer) or as an agent (broker) depending on the type of trade.

Brokered market A market in which brokers arrange trades among their clients.

Brownfield investment Investing in *existing* infrastructure assets.

Bubble line chart A line chart that uses varying-sized bubbles to represent a third dimension of the data. The bubbles are sometimes color-coded to present additional information.

Budget surplus/deficit The difference between government revenue and expenditure for a stated fixed period of time.

Bullet bond Bond in which the principal repayment is made entirely at maturity.

Bullish crossover A technical analysis term that describes a situation where a short-term moving average crosses a longer-term moving average from below; this movement is considered bullish. A **golden cross** is a bullish crossover based on 50-day and 200-day moving averages.

Business cycles Are recurrent expansions and contractions in economic activity affecting broad segments of the economy.

Business risk The risk associated with operating earnings. Operating earnings are uncertain because total revenues and many of the expenditures contributed to produce those revenues are uncertain.

Buy-side firm An investment management company or other investor that uses the services of brokers or dealers (i.e., the client of the sell side firms).

Buyback A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Call An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call market A market in which trades occur only at a particular time and place (i.e., when the market is called).

Call money rate The interest rate that buyers pay for their margin loan.

Call option An option that gives the holder the right to buy an underlying asset from another party at a fixed price over a specific period of time.

Call protection The time during which the issuer of the bond is not allowed to exercise the call option.

Callable bond A bond containing an embedded call option that gives the issuer the right to buy the bond back from the investor at specified prices on pre-determined dates.

Calmar ratio A comparison of the average annual compounded return to the maximum drawdown risk.

Candlestick chart A price chart with four bits of data for each time interval. A candle indicates the opening and closing price for the interval. The body of the candle is shaded if the opening price was higher than the closing price, and the body is white (or clear) if the opening price was lower than the closing price. Vertical lines known as wicks or shadows extend from the top and bottom of the candle to indicate, respectively, the high and low prices for the interval.

Cannibalization Cannibalization occurs when an investment takes customers and sales away from another part of the company.

Cap rate A metric by which real estate managers are often judged; the annual rent actually earned (net of any vacancies) divided by the price originally paid for the property.

Capacity The ability of the borrower to make its debt payments on time.

Capital account A component of the balance of payments account that measures transfers of capital.

Capital allocation line (CAL) A graph line that describes the combinations of expected return and standard deviation of return available to an investor from combining the optimal portfolio of risky assets with the risk-free asset.

Capital allocation process The process that companies use for decision making on capital investments—those projects with a life of one year or longer.

Capital asset pricing model (CAPM) An equation describing the expected return on any asset (or portfolio) as a linear function of its beta relative to the market portfolio.

Capital consumption allowance A measure of the wear and tear (depreciation) of the capital stock that occurs in the production of goods and services.

Capital deepening investment Increases the stock of capital relative to labor.

Capital expenditure Expenditure on physical capital (fixed assets).

Capital light Capital light businesses require little incremental investment in fixed assets or working capital to enable revenue growth.

Capital loss ratio The percentage of capital in deals that have been realized below cost, net of any recovered proceeds, divided by total invested capital.

Capital market expectations An investor's expectations concerning the risk and return prospects of asset classes.

Capital market line (CML) The line with an intercept point equal to the risk-free rate that is tangent to the efficient frontier of risky assets; represents the efficient frontier when a risk-free asset is available for investment.

Capital market securities Securities with maturities at issuance longer than one year.

Capital markets Financial markets that trade securities of longer duration, such as bonds and equities.

Capital rationing A capital rationing environment assumes that the company has a fixed amount of funds to invest.

Capital restrictions Controls placed on foreigners' ability to own domestic assets and/or domestic residents' ability to own foreign assets.

Capital stock The accumulated amount of buildings, machinery, and equipment used to produce goods and services.

Capital structure The mix of debt and equity that a company uses to finance its business; a company's specific mixture of long-term financing.

Carry The net of the costs and benefits of holding, storing, or "carrying" an asset.

Carrying amount The amount at which an asset or liability is valued according to accounting principles.

Carrying value The net amount shown for an asset or liability on the balance sheet; book value may also refer to the company's excess of total assets over total liabilities. For a bond, the purchase price plus (or minus) the amortized amount of the discount (or premium).

Cartel Participants in collusive agreements that are made openly and formally.

Cash collateral account Form of external credit enhancement whereby the issuer immediately borrows the credit-enhancement amount and then invests that amount, usually in highly rated short-term commercial paper.

Cash conversion cycle A financial metric that measures the length of time required for a company to convert cash invested in its operations to cash received as a result of its operations; equal to days of inventory on hand + days of sales outstanding − number of days of payables. Also called *net operating cycle*.

Cash flow additivity principle The principle that dollar amounts indexed at the same point in time are additive.

Cash flow from operating activities The net amount of cash provided from operating activities.

Cash flow from operations The net amount of cash provided from operating activities.

Cash flow yield The internal rate of return on a series of cash flows.

Cash market securities Money market securities settled on a "same day" or "cash settlement" basis.

Cash markets See *spot markets*.

Cash prices See *spot prices*.

Cash-settled forwards See *non-deliverable forwards*.

Catch-up clause A clause in an agreement that favors the GP. For a GP who earns a 20% performance fee, a catch-up clause allows the GP to receive 100% of the distributions above the hurdle rate *until* she receives 20% of the profits generated, and then every excess dollar is split 80/20 between the LPs and GP.

Categorical data Values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize (also called **qualitative data**).

CBOE Volatility Index (VIX) A measure of near-term market volatility as conveyed by S&P 500 stock index option prices.

Central bank funds market The market in which deposit-taking banks that have an excess reserve with their national central bank can lend money to banks that need funds for maturities ranging from overnight to one year. Called the Federal or Fed funds market in the United States.

Central bank funds rate Interest rate at which central bank funds are bought (borrowed) and sold (lent) for maturities ranging from overnight to one year. Called Federal or Fed funds rate in the United States.

Central banks The dominant bank in a country, usually with official or semi-official governmental status.

Central limit theorem The theorem that states the sum (and the mean) of a set of independent, identically distributed random variables with finite variances is normally distributed, whatever distribution the random variables follow.

Certificate of deposit An instrument that represents a specified amount of funds on deposit with a bank for a specified maturity and interest rate. CDs are issued in various denominations and can be negotiable or non-negotiable.

Change in polarity principle A tenet of technical analysis that states that once a support level is breached, it becomes a resistance level. The same holds true for resistance levels: Once breached, they become support levels.

Change of control put A covenant giving bondholders the right to require the issuer to buy back their debt, often at par or at some small premium to par value, in the event that the borrower is acquired.

Character The quality of a debt issuer's management.

Chartist An individual who uses charts or graphs of a security's historical prices or levels to forecast its future trends.

Chi-square test of independence A statistical test for detecting a potential association between categorical variables.

Classical cycle Refers to fluctuations in the level of economic activity when measured by GDP in volume terms.

Classified balance sheet A balance sheet organized so as to group together the various assets and liabilities into subcategories (e.g., current and noncurrent).

Clawback A requirement that the general partner return any funds distributed as incentive fees until the limited partners have received back their initial investment and a percentage of the total profit.

Clearing The process by which the exchange verifies the execution of a transaction and records the participants' identities.

Clearing instructions Instructions that indicate how to arrange the final settlement ("clearing") of a trade.

Clearinghouse An entity associated with a futures market that acts as middleman between the contracting parties and guarantees to each party the performance of the other.

Closed economy An economy that does not trade with other countries; an *autarkic economy*.

Closed-end fund A mutual fund in which no new investment money is accepted. New investors invest by buying existing shares, and investors in the fund liquidate by selling their shares to other investors.

Cluster sampling A procedure that divides a population into subpopulation groups (clusters) representative of the population and then randomly draws certain clusters to form a sample.

Clustered bar chart See *grouped bar chart*.

Co-investing In co-investing, the investor invests in assets *indirectly* through the fund but also possesses rights (known as co-investment rights) to invest *directly* in the same assets. Through co-investing, an investor is able to make an investment *alongside* a fund when the fund identifies deals.

Code of ethics An established guide that communicates an organization's values and overall expectations regarding member behavior. A code of ethics serves as a general guide for how community members should act.

Coefficient of determination The percentage of the variation of the dependent variable that is explained by the independent variable. Also referred to as the "R-squared" or "R^2."

Coefficient of variation The ratio of a set of observations' standard deviation to the observations' mean value.

Cognitive cost The effort involved in processing new information and updating beliefs.

Cognitive dissonance The mental discomfort that occurs when new information conflicts with previously held beliefs or cognitions.

Cognitive errors Behavioral biases resulting from faulty reasoning; cognitive errors stem from basic statistical, information-processing, or memory errors.

Coincident economic indicators Turning points that are usually close to those of the overall economy; they are believed to have value for identifying the economy's present state.

Collateral Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Collateral manager Buys and sells debt obligations for and from the CDO's portfolio of assets (i.e., the collateral) to generate sufficient cash flows to meet the obligations to the CDO bondholders.

Collateral trust bonds Bonds secured by securities, such as common shares, other bonds, or other financial assets.

Collateralized bond obligations A structured asset-backed security that is collateralized by a pool of bonds.

Collateralized debt obligation Generic term used to describe a security backed by a diversified pool of one or more debt obligations.

Collateralized loan obligations A structured asset-backed security that is collateralized by a pool of loans.

Collateralized mortgage obligation A security created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collateralized mortgage obligations Securities created through the securitization of a pool of mortgage-related products (mortgage pass-through securities or pools of loans).

Collaterals Assets or financial guarantees underlying a debt obligation that are above and beyond the issuer's promise to pay.

Combination A listing in which the order of the listed items does not matter.

Combination formula (binomial formula) The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is ${}_nC_r = \binom{n}{r} = \dfrac{n!}{(n-r)!r!}$.

Commercial paper A short-term, negotiable, unsecured promissory note that represents a debt obligation of the issuer.

Committed (Regular) lines of credit A bank commitment to extend credit up to a pre-specified amount; the commitment is considered a short-term liability and is usually in effect for 364 days (one day short of a full year).

Committed capital The amount that the limited partners have agreed to provide to the private equity fund.

Commodity swap A swap in which the underlying is a commodity such as oil, gold, or an agricultural product.

Common market Level of economic integration that incorporates all aspects of the customs union and extends it by allowing free movement of factors of production among members.

Common shares A type of security that represent an ownership interest in a company.

Common-size analysis The restatement of financial statement items using a common denominator or reference item that allows one to identify trends and major differences; an example is an income statement in which all items are expressed as a percent of revenue.

Common stock See *common shares.*

Company analysis Analysis of an individual company.

Comparable company A company that has similar business risk, usually in the same industry and preferably with a single line of business.

Comparative advantage A country's ability to produce a good or service at a lower relative cost, or opportunity cost, than its trading partner.

Competitive strategy A company's plans for responding to the threats and opportunities presented by the external environment.

Complement The event not-S, written S^{C}, given the event S. Note that $P(S) + P(S^{C}) = 1$.

Complements Goods that tend to be used together; technically, two goods whose cross-price elasticity of demand is negative.

Complete markets Informally, markets in which the variety of distinct securities traded is so broad that any desired payoff in a future state-of-the-world is achievable.

Component cost of capital The rate of return required by suppliers of capital for an individual source of a company's funding, such as debt or equity.

Compounding The process of accumulating interest on interest.

Comprehensive income The change in equity of a business enterprise during a period from nonowner sources; includes all changes in equity during a period except those resulting from investments by owners and distributions to owners; comprehensive income equals net income plus other comprehensive income.

Conditional expected value The expected value of a stated event given that another event has occurred.

Conditional pass-through covered bonds Covered bonds that convert to pass-through securities after the original maturity date if all bond payments have not yet been made and the sponsor is in default.

Conditional probability The probability of an event given (conditioned on) another event.

Conditional variances The variance of one variable, given the outcome of another.

Confidence level The complement of the level of significance.

Confirmation bias A belief perseverance bias in which people tend to look for and notice what confirms their beliefs, to ignore or undervalue what contradicts their beliefs, and to misinterpret information as support for their beliefs.

Confusion matrix A type of contingency table used for evaluating the performance of a classification model.

Conservatism bias A belief perseverance bias in which people maintain their prior views or forecasts by inadequately incorporating new information.

Consolidation The movement of a stock's price within a well-defined range of trading levels for a period of time. The price consolidates between a support level and a resistance level.

Constant-yield price trajectory A graph that illustrates the change in the price of a fixed-income bond over time assuming no change in yield-to-maturity. The trajectory shows the "pull to par" effect on the price of a bond trading at a premium or a discount to par value.

Constituent securities With respect to an index, the individual securities within an index.

Consumer surplus The difference between the value that a consumer places on units purchased and the amount of money that was required to pay for them.

Contango A condition in the futures markets in which the spot price is lower than the futures price, the forward curve is upward sloping, and there is little or no convenience yield.

Contingency provision Clause in a legal document that allows for some action if a specific event or circumstance occurs.

Contingency table A tabular format that displays the frequency distributions of two or more categorical variables simultaneously and is used for finding patterns between the variables. A contingency table for two categorical variables is also known as a two-way table.

Contingent claim A security whose value is based on some other underlying security.

Contingent claims Derivatives in which the payoffs occur if a specific event occurs; generally referred to as options.

Contingent convertible bonds Bonds that automatically convert into equity if a specific event or circumstance occurs, such as the issuer's equity capital falling below the minimum requirement set by the regulators. Also called *CoCos.*

Continuation pattern A type of pattern used in technical analysis to predict the resumption of a market trend that was in place prior to the formation of a pattern.

Continuous data Data that can be measured and can take on any numerical value in a specified range of values.

Continuous random variable A random variable for which the range of possible outcomes is the real line (all real numbers between $-\infty$ and $+\infty$) or some subset of the real line.

Continuous trading market A market in which trades can be arranged and executed any time the market is open.

Continuously compounded return The natural logarithm of 1 plus the holding period return, or equivalently, the natural logarithm of the ending price over the beginning price.

Contra account An account that offsets another account.

Contract rate See *mortgage rate.*

Contraction The period of a business cycle after the peak and before the trough; often called a *recession* or, if exceptionally severe, called a *depression*.

Contraction risk The risk that when interest rates decline, the security will have a shorter maturity than was anticipated at the time of purchase because borrowers refinance at the new, lower interest rates.

Contractionary Tending to cause the real economy to contract.

Contractionary fiscal policy A fiscal policy that has the objective to make the real economy contract.

Contracts for differences See *non-deliverable forwards.*

Contribution margin The amount available for fixed costs and profit after paying variable costs; revenue minus variable costs.

Controlling shareholders A particular shareholder or block of shareholders holding a percentage of shares that gives them significant voting power.

Convenience sampling A procedure of selecting an element from a population on the basis of whether or not it is accessible to a researcher or how easy it is for a researcher to access the element.

Convenience yield A non-monetary advantage of holding an asset.

Conventional bond See *plain vanilla bond*.

Conventional cash flows A conventional cash flow pattern is one with an initial outflow followed by a series of inflows.

Convergence The tendency for differences in output per capita across countries to diminish over time. In technical analysis, the term describes the case when an indicator moves in the same manner as the security being analyzed.

Conversion premium The difference between the convertible bond's price and its conversion value.

Conversion price For a convertible bond, the price per share at which the bond can be converted into shares.

Conversion ratio For a convertible bond, the number of common shares that each bond can be converted into.

Conversion value For a convertible bond, the current share price multiplied by the conversion ratio.

Convertible bond Bond that gives the bondholder the right to exchange the bond for a specified number of common shares in the issuing company.

Convertible preference shares A type of equity security that entitles shareholders to convert their shares into a specified number of common shares.

Convexity adjustment For a bond, one half of the annual or approximate convexity statistic multiplied by the change in the yield-to-maturity squared.

Core inflation Refers to the inflation rate calculated based on a price index of goods and services except food and energy.

Corporate governance The system of internal controls and procedures by which individual companies are managed.

Correlation A measure of the linear relationship between two random variables.

Correlation coefficient A number between −1 and +1 that measures the consistency or tendency for two investments to act in a similar way. It is used to determine the effect on portfolio risk when two assets are combined.

Cost averaging The periodic investment of a fixed amount of money.

Cost of capital The cost of financing to a company; the rate of return that suppliers of capital require as compensation for their contribution of capital.

Cost of carry See *carry*.

Cost of debt The cost of debt financing to a company, such as when it issues a bond or takes out a bank loan; the rate of return required by a company's bondholders.

Cost of equity The rate of return required by a company's common stockholders.

Cost of preferred stock The cost to a company of issuing preferred stock; the dividend yield that a company must commit to pay preferred stockholders.

Cost structure The mix of a company's variable costs and fixed costs.

Counterparty risk The risk that the other party to a contract will fail to honor the terms of the contract.

Coupon rate The interest rate promised in a contract; this is the rate used to calculate the periodic interest payments.

Cournot assumption Assumption in which each firm determines its profit-maximizing production level assuming that the other firms' output will not change.

Covariance A measure of the co-movement (linear association) between two random variables.

Covariance matrix A matrix or square array whose entries are covariances; also known as a variance–covariance matrix.

Covenants The terms and conditions of lending agreements that the issuer must comply with; they specify the actions that an issuer is obligated to perform (affirmative covenant) or prohibited from performing (negative covenant).

Covered bond Debt obligation secured by a segregated pool of assets called the cover pool. The issuer must maintain the value of the cover pool. In the event of default, bondholders have recourse against both the issuer and the cover pool.

Covered bonds Senior debt obligations issued by a financial institution and backed by a segregated pool of assets. Covered bond investors have recourse to both the underlying assets and the financial institution.

Credit analysis The evaluation of credit risk; the evaluation of the creditworthiness of a borrower or counterparty.

Credit default swap (CDS) A type of credit derivative in which one party, the credit protection buyer who is seeking credit protection against a third party, makes a series of regularly scheduled payments to the other party, the credit protection seller. The seller makes no payments until a credit event occurs.

Credit derivatives A contract in which one party has the right to claim a payment from another party in the event that a specific credit event occurs over the life of the contract.

Credit enhancements Provisions that may be used to reduce the credit risk of a bond issue.

Credit-linked coupon bond Bond for which the coupon changes when the bond's credit rating changes.

Credit-linked note (CLN) Fixed-income security in which the holder of the security has the right to withhold payment of the full amount due at maturity if a credit event occurs.

Credit migration risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *downgrade risk*.

Credit risk The risk of loss caused by a counterparty's or debtor's failure to make a promised payment. Also called *default risk*.

Credit spread option An option on the yield spread on a bond.

Credit tranching A structure used to redistribute the credit risk associated with the collateral; a set of bond classes created to allow investors a choice in the amount of credit risk that they prefer to bear.

Creditworthiness The perceived ability of the borrower to pay its debt obligations in a timely manner; it represents the ability of a company to withstand adverse impacts on its cash flows.

Critical values Values of the test statistic at which the decision changes from fail to reject the null hypothesis to reject the null hypothesis.

Cross-default Covenant or contract clause that specifies a borrower is considered in default if they default on another debt obligation.

Cross-default provisions Provisions whereby events of default, such as non-payment of interest on one bond, trigger default on all outstanding debt; implies the same default probability for all issues.

Cross-price elasticity of demand The percentage change in quantity demanded for a given percentage change in the price of another good; the responsiveness of the demand for Product A that is associated with the change in price of Product B.

Cross-sectional analysis Analysis that involves comparisons across individuals in a group over a given time period or at a given point in time.

Cross-sectional data A list of the observations of a specific variable from multiple observational units at a given point in time. The observational units can be individuals, groups, companies, trading markets, regions, etc.

Crossing networks Trading systems that match buyers and sellers who are willing to trade at prices obtained from other markets.

Crowding out The thesis that government borrowing may divert private sector investment from taking place.

Cryptocurrency An electronic medium of exchange that lacks physical form.

Cryptography An algorithmic process to encrypt data, making the data unusable if received by unauthorized parties.

Cumulative absolute frequency Cumulates (i.e., adds up) in a frequency distribution the absolute frequencies as one moves from the first bin to the last bin.

Cumulative distribution function A function giving the probability that a random variable is less than or equal to a specified value.

Cumulative frequency distribution chart A chart that plots either the cumulative absolute frequency or the cumulative relative frequency on the y-axis against the upper limit of the interval and allows one to see the number or the percentage of the observations that lie below a certain value.

Cumulative preference shares Preference shares for which any dividends that are not paid accrue and must be paid in full before dividends on common shares can be paid.

Cumulative relative frequency A sequence of partial sums of the relative frequencies in a frequency distribution.

Cumulative voting A voting process whereby shareholders can accumulate and vote all their shares for a single candidate in an election, as opposed to having to allocate their voting rights evenly among all candidates.

Currencies Monies issued by national monetary authorities.

Currency option bonds Bonds that give bondholders the right to choose the currency in which they want to receive interest payments and principal repayments.

Currency swap A swap in which each party makes interest payments to the other in different currencies.

Current account A component of the balance of payments account that measures the flow of goods and services.

Current assets Assets that are expected to be consumed or converted into cash in the near future, typically one year or less. *Also called liquid assets.*

Current cost With reference to assets, the amount of cash or cash equivalents that would have to be paid to buy the same or an equivalent asset today; with reference to liabilities, the undiscounted amount of cash or cash equivalents that would be required to settle the obligation today.

Current government spending With respect to government expenditures, spending on goods and services that are provided on a regular, recurring basis including health, education, and defense.

Current liabilities Short-term obligations, such as accounts payable, wages payable, or accrued liabilities, that are expected to be settled in the near future, typically one year or less.

Current yield The sum of the coupon payments received over the year divided by the flat price; also called the *income* or *interest yield* or *running yield*.

Curve duration The sensitivity of the bond price (or the market value of a financial asset or liability) with respect to a benchmark yield curve.

Customs union Extends the free trade area (FTA) by not only allowing free movement of goods and services among members, but also creating a common trade policy against nonmembers.

CVaR Conditional VaR, a tail loss measure. The weighted average of all loss outcomes in the statistical distribution that exceed the VaR loss.

Cyclical A cyclical company is one whose profits are strongly correlated with the strength of the overall economy.

Cyclical companies Companies with sales and profits that regularly expand and contract with the business cycle or state of economy.

Daily settlement See *mark to market* and *marking to market.*

Dark pools Alternative trading systems that do not display the orders that their clients send to them.

Data A collection of numbers, characters, words, and text—as well as images, audio, and video—in a raw or organized format to represent facts or information.

Data mining The practice of determining a model by extensive searching through a dataset for statistically significant patterns. Also called *data snooping.*

Data science An interdisciplinary field that brings computer science, statistics, and other disciplines together to analyze and produce insights from Big Data.

Data snooping The practice of determining a model by extensive searching through a dataset for statistically significant patterns.

Data table see **two-dimensional rectangular array**.

Day order An order that is good for the day on which it is submitted. If it has not been filled by the close of business, the order expires unfilled.

Days of inventory on hand An activity ratio equal to the number of days in the period divided by inventory turnover over the period.

Dealers A financial intermediary that acts as a principal in trades.

Dealing securities Securities held by banks or other financial intermediaries for trading purposes.

Death cross See **Bearish crossover**.

Debentures Type of bond that can be secured or unsecured.

Debt rating An independent measure of the quality and safety of a company's debt based upon an analysis of the company's ability to pay the promised cash flows in full and on time. It includes an analysis of any indentures.

Debt-rating approach A method for estimating a company's before-tax cost of debt based on the yield on comparably rated bonds for maturities that closely match that of the company's existing debt.

Debt tax shield The tax benefit from interest paid on debt being tax deductible from income, equal to the marginal tax rate multiplied by the value of the debt.

Debt-to-assets ratio A solvency ratio calculated as total debt divided by total assets.

Debt-to-capital ratio A solvency ratio calculated as total debt divided by total debt plus total shareholders' equity.

Debt-to-equity ratio A solvency ratio calculated as total debt divided by total shareholders' equity.

Deciles Quantiles that divide a distribution into 10 equal parts.

Declaration date The day that the corporation issues a statement declaring a specific dividend.

Decreasing returns to scale When a production process leads to increases in output that are proportionately smaller than the increase in inputs.

Deductible temporary differences Temporary differences that result in a reduction of or deduction from taxable income in a future period when the balance sheet item is recovered or settled.

Deep learning Machine learning using neural networks with many hidden layers.

Deep learning nets Machine learning using neural networks with many hidden layers.

Default probability The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default risk*.

Default risk The probability that a borrower defaults or fails to meet its obligation to make full and timely payments of principal and interest, according to the terms of the debt security. Also called *default probability*.

Default risk premium An extra return that compensates investors for the possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount.

Defensive companies Companies with sales and profits that have little sensitivity to the business cycle or state of the economy.

Defensive interval ratio A liquidity ratio that estimates the number of days that an entity could meet cash needs from liquid assets; calculated as (cash + short-term marketable investments + receivables) divided by daily cash expenditures.

Deferred coupon bond Bond that pays no coupons for its first few years but then pays a higher coupon than it otherwise normally would for the remainder of its life. Also called *split coupon bond*.

Deferred income A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service.

Deferred tax assets A balance sheet asset that arises when an excess amount is paid for income taxes relative to accounting profit. The taxable income is higher than accounting profit and income tax payable exceeds tax expense. The company expects to recover the difference during the course of future operations when tax expense exceeds income tax payable.

Deferred tax liabilities A balance sheet liability that arises when a deficit amount is paid for income taxes relative to accounting profit. The taxable income is less than the accounting profit and income tax payable is less than tax expense. The company expects to eliminate the liability over the course of future operations when income tax payable exceeds tax expense.

Defined benefit pension plans Plans in which the company promises to pay a certain annual amount (defined benefit) to the employee after retirement. The company bears the investment risk of the plan assets.

Defined contribution pension plans Individual accounts to which an employee and typically the employer makes contributions during their working years and expect to draw on the accumulated funds at retirement. The employee bears the investment and inflation risk of the plan assets.

Deflation Negative inflation.

Degree of confidence The probability that a confidence interval includes the unknown population parameter.

Degree of financial leverage (DFL) The ratio of the percentage change in net income to the percentage change in operating income; the sensitivity of the cash flows available to owners when operating income changes.

Degree of operating leverage (DOL) The ratio of the percentage change in operating income to the percentage change in units sold; the sensitivity of operating income to changes in units sold.

Degree of total leverage The ratio of the percentage change in net income to the percentage change in units sold; the sensitivity of the cash flows to owners to changes in the number of units produced and sold.

Degrees of freedom The number of independent variables used in defining sample statistics, such as variance, and the probability distributions they measure.

Delta The sensitivity of the derivative price to a small change in the value of the underlying asset.

Demand curve Graph of the inverse demand function. A graph showing the demand relation, either the highest quantity willingly purchased at each price or the highest price willingly paid for each quantity.

Demand function A relationship that expresses the quantity demanded of a good or service as a function of own-price and possibly other variables.

Demand shock A typically unexpected disturbance to demand, such as an unexpected interruption in trade or transportation.

Dependent With reference to events, the property that the probability of one event occurring depends on (is related to) the occurrence of another event.

Dependent variable The variable whose variation about its mean is to be explained by the regression; the left-side variable in a regression equation. Also referred to as the *explained variable*.

Depository bank A bank that raises funds from depositors and other investors and lends it to borrowers.

Depository institutions Commercial banks, savings and loan banks, credit unions, and similar institutions that raise funds from depositors and other investors and lend it to borrowers.

Depository receipt A security that trades like an ordinary share on a local exchange and represents an economic interest in a foreign company.

Depreciation The process of systematically allocating the cost of long-lived (tangible) assets to the periods during which the assets are expected to provide economic benefits.

Derivative pricing rule A pricing rule used by crossing networks in which a price is taken (derived) from the price that is current in the asset's primary market.

Derivatives A financial instrument whose value depends on the value of some underlying asset or factor (e.g., a stock price, an interest rate, or exchange rate).

Descriptive statistics Measures that summarize central tendency and spread variation in the data's distribution.

Diffuse prior The assumption of equal prior probabilities.

Diffusion index Reflects the proportion of the index's components that are moving in a pattern consistent with the overall index.

Diluted EPS The EPS that would result if all dilutive securities were converted into common shares.

Diluted shares The number of shares that would be outstanding if all potentially dilutive claims on common shares (e.g., convertible debt, convertible preferred stock, and employee stock options) were exercised.

Diminishing balance method An accelerated depreciation method, i.e., one that allocates a relatively large proportion of the cost of an asset to the early years of the asset's useful life.

Diminishing marginal productivity Describes a state in which each additional unit of input produces less output than previously.

Direct format With reference to the cash flow statement, a format for the presentation of the statement in which cash flow from operating activities is shown as operating cash receipts less operating cash disbursements. Also called *direct method*.

Direct investing Occurs when an investor makes a direct investment in an asset without the use of an intermediary.

Direct method See *direct format*.

Direct taxes Taxes levied directly on income, wealth, and corporate profits.

Direct write-off method An approach to recognizing credit losses on customer receivables in which the company waits until such time as a customer has defaulted and only then recognizes the loss.

Discount To reduce the value of a future payment in allowance for how far away it is in time; to calculate the present value of some future amount. Also, the amount by which an instrument is priced below its face (par) value.

Discount margin See *required margin.*

Discount rates In general, the interest rates used to calculate present values. In the money market, however, a discount rate is a specific type of quoted rate.

Discounted cash flow models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security.

Discouraged worker A person who has stopped looking for a job or has given up seeking employment.

Discrete data Numerical values that result from a counting process; therefore, practically speaking, the data are limited to a finite number of values.

Discrete random variable A random variable that can take on at most a countable number of possible values.

Discriminatory pricing rule A pricing rule used in continuous markets in which the limit price of the order or quote that first arrived determines the trade price.

Diseconomies of scale Increase in cost per unit resulting from increased production.

Dispersion The variability of a population or sample of observations around the central tendency.

Display size The size of an order displayed to public view.

Disposition effect As a result of loss aversion, an emotional bias whereby investors are reluctant to dispose of losers. This reluctance results in an inefficient and gradual adjustment to deterioration in fundamental value.

Distributed ledger A type of database that may be shared among entities in a network.

Distributed ledger technology Technology based on a distributed ledger.

Divergence In technical analysis, a term that describes the case when an indicator moves differently from the security being analyzed.

Diversification ratio The ratio of the standard deviation of an equally weighted portfolio to the standard deviation of a randomly selected security.

Dividend A distribution paid to shareholders based on the number of shares owned.

Dividend discount model (DDM) A present value model that estimates the intrinsic value of an equity share based on the present value of its expected future dividends.

Dividend payout ratio The ratio of cash dividends paid to earnings for a period.

Divisor A number (denominator) used to determine the value of a price return index. It is initially chosen at the inception of an index and subsequently adjusted by the index provider, as necessary, to avoid changes in the index value that are unrelated to changes in the prices of its constituent securities.

Doji In the Japanese terminology used in candlestick charting, the doji signifies that after a full day of trading, the positive price influence of buyers and the negative price influence of sellers exactly counteracted each other—with opening and closing prices that are virtually equal—which suggests that the market under analysis is in balance.

Domestic content provisions Stipulate that some percentage of the value added or components used in production should be of domestic origin.

Double bottom In technical analysis, a reversal pattern that is formed when the price reaches a low, rebounds, and then declines back to the first low level. A double bottom is used to predict a change from a downtrend to an uptrend.

Double coincidence of wants A prerequisite to barter trades, in particular that both economic agents in the transaction want what the other is selling.

Double declining balance depreciation An accelerated depreciation method that involves depreciating the asset at double the straight-line rate. This rate is multiplied by the book value of the asset at the beginning of the period (a declining balance) to calculate depreciation expense.

Double top In technical analysis, a reversal pattern that is formed when an uptrend reverses twice at roughly the same high price level. A double top is used to predict a change from an uptrend to a downtrend.

Down transition probability The probability that an asset's value moves down in a model of asset price dynamics.

Downgrade risk The risk that a bond issuer's creditworthiness deteriorates, or migrates lower, leading investors to believe the risk of default is higher. Also called *credit migration risk.*

Downside risk Risk of incurring returns below a specified value.

Downtrend A pattern that occurs when the price of an asset moves lower over a period of time.

Drag on liquidity When receipts lag, creating pressure from the decreased available funds.

Drawdown A percentage peak-to-trough reduction in net asset value.

Dual-currency bonds Bonds that make coupon payments in one currency and pay the par value at maturity in another currency.

DuPont analysis An approach to decomposing return on investment, e.g., return on equity, as the product of other financial ratios.

Duration A measure of the approximate sensitivity of a security to a change in interest rates (i.e., a measure of interest rate risk).

Duration gap A bond's Macaulay duration minus the investment horizon.

Dutch Book Theorem A result in probability theory stating that inconsistent probabilities create profit opportunities.

Early repayment option See *prepayment option.*

Earnings per share The amount of income earned during a period per share of common stock.

Earnings surprise The portion of a company's earnings that is unanticipated by investors and, according to the efficient market hypothesis, merits a price adjustment.

Economic costs All the remuneration needed to keep a productive resource in its current employment or to acquire the resource for productive use; the sum of total accounting costs and implicit opportunity costs.

Economic indicator A variable that provides information on the state of the overall economy.

Economic loss The amount by which accounting profit is less than normal profit.

Economic profit Equal to accounting profit less the implicit opportunity costs not included in total accounting costs; the difference between total revenue (TR) and total cost (TC). Also called *abnormal profit* or *supernormal profit.*

Economic stabilization Reduction of the magnitude of economic fluctuations.

Economic union Incorporates all aspects of a common market and in addition requires common economic institutions and coordination of economic policies among members.

Economies of scale Reduction in cost per unit resulting from increased production.

Effective annual rate The amount by which a unit of currency will grow in a year with interest on interest included.

Effective convexity A *curve convexity* statistic that measures the secondary effect of a change in a benchmark yield curve on a bond's price.

Effective duration The sensitivity of a bond's price to a change in a benchmark yield curve.

Effective interest rate The borrowing rate or market rate that a company incurs at the time of issuance of a bond.

Efficient market A market in which asset prices reflect new information quickly and rationally.

Elastic Said of a good or service when the magnitude of elasticity is greater than one.

Elasticity The percentage change in one variable for a percentage change in another variable; a general measure of how sensitive one variable is to a change in the value of another variable.

Elasticity of demand A measure of the sensitivity of quantity demanded to a change in a product's own price: $\%\Delta Q^D/\%\Delta P$.

Elasticity of supply A measure of the sensitivity of quantity supplied to a change in price: $\%\Delta Q^S/\%\Delta P$.

Electronic communications networks See *alternative trading systems.*

Embedded option Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Embedded options Contingency provisions that provide the issuer or the bondholders the right, but not the obligation, to take action. These options are not part of the security and cannot be traded separately.

Emotional biases Behavioral biases resulting from reasoning influenced by feelings; emotional biases stem from impulse or intuition.

Empirical duration The use of statistical methods and historical bond prices to estimate the price–yield relationship for a specific bond or portfolio of bonds.

Empirical probability The probability of an event estimated as a relative frequency of occurrence.

Employed The number of people with a job.

Endowment bias An emotional bias in which people value an asset more when they hold rights to it than when they do not.

Engagement/active ownership An ESG investment approach that uses shareholder power to influence corporate behavior through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is directed by ESG guidelines.

Enterprise risk management An overall assessment of a company's risk position. A centralized approach to risk management sometimes called firmwide risk management.

Enterprise value A measure of a company's total market value from which the value of cash and short-term investments have been subtracted.

Equal weighting An index weighting method in which an equal weight is assigned to each constituent security at inception.

Equipment trust certificates Bonds secured by specific types of equipment or physical assets.

Equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Equity risk premium The expected return on equities minus the risk-free rate; the premium that investors demand for investing in equities.

Equity swap A swap transaction in which at least one cash flow is tied to the return to an equity portfolio position, often an equity index.

Error term The difference between an observation and its expected value, where the expected value is based on the true underlying population relation between the dependent and independent variables. Also known simply as the *error*.

ESG An acronym that encompasses environmental, social, and governance.

ESG integration An ESG investment approach that focuses on systematic consideration of material ESG factors in asset allocation, security selection, and portfolio construction decisions for the purpose of achieving the product's stated investment objectives. Used interchangeably with **ESG investing**.

ESG investing The consideration of environmental, social, and governance factors in the investment process. Used interchangeably with **ESG integration**.

Estimate The particular value calculated from sample observations using an estimator.

Estimated parameters With reference to a regression analysis, the estimated values of the population intercept and population slope coefficients in a regression.

Estimator An estimation formula; the formula used to compute the sample mean and other sample statistics are examples of estimators.

Ethical principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Ethics The study of moral principles or of making good choices. Ethics encompasses a set of moral principles and rules of conduct that provide guidance for our behavior.

Eurobonds Type of bond issued internationally, outside the jurisdiction of the country in whose currency the bond is denominated.

European-style Said of an option contract that can only be exercised on the option's expiration date.

Event Any outcome or specified set of outcomes of a random variable.

Ex-dividend date The first date that a share trades without (i.e., "ex") the dividend.

Excess kurtosis Degree of kurtosis (fatness of tails) relative to the kurtosis of the normal distribution.

Exchanges Places where traders can meet to arrange their trades.

Execution instructions Instructions that indicate how to fill an order.

Exercise The process of using an option to buy or sell the underlying.

Exercise price The fixed price at which an option holder can buy or sell the underlying. Also called *strike price*, *striking price*, or *strike*.

Exercise value The value obtained if an option is exercised based on current conditions. Also known as *intrinsic value*.

Exhaustive Covering or containing all possible outcomes.

Expansion The period of a business cycle after its lowest point and before its highest point.

Expansionary Tending to cause the real economy to grow.

Expansionary fiscal policy Fiscal policy aimed at achieving real economic growth.

Expected inflation The level of inflation that economic agents expect in the future.

Expected loss Default probability times loss severity given default.

Expected return on the portfolio ($E(R_p)$) The weighted average of the expected returns (R_1 to R_n) on the component securities using their respective weights (w_1 to w_n).

Expected value The probability-weighted average of the possible outcomes of a random variable.

Expenses Outflows of economic resources or increases in liabilities that result in decreases in equity (other than decreases because of distributions to owners); reductions in net assets associated with the creation of revenues.

Export subsidy Paid by the government to the firm when it exports a unit of a good that is being subsidized.

Exports Goods and services that an economy sells to other countries.

Extension risk The risk that when interest rates rise, fewer prepayments will occur because homeowners are reluctant to give up the benefits of a contractual interest rate that now looks low. As a result, the security becomes longer in maturity than anticipated at the time of purchase.

Externalities Spillover effects of production and consumption activities onto others who are not directly involved in a particular transaction, activity, or decision.

Externality An effect of a market transaction that is borne by parties other than those who transacted.

Extra dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Extreme value theory A branch of statistics that focuses primarily on extreme outcomes.

Face value The amount of cash payable by a company to the bondholders when the bonds mature; the promised payment at maturity separate from any coupon payment.

Factoring arrangement When a company sells its accounts receivable to a lender (known as a factor) who assumes responsibility for the credit granting and collection process.

Fair value The amount at which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's-length transaction; the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants.

False discovery approach An adjustment in the *p*-values for tests performed multiple times.

False discovery rate The rate of Type I errors in testing a null hypothesis multiple times for a given level of significance.

Fat-Tailed Describes a distribution that has fatter tails than a normal distribution (also called leptokurtic).

Fed funds rate The US interbank lending rate on overnight borrowings of reserves.

Federal funds rate The US interbank lending rate on overnight borrowings of reserves.

Fiat money Money that is not convertible into any other commodity.

Fiduciary An entity designated to represent the rights and responsibilities of a beneficiary whose assets they are managing, such as a bond trustee acting on behalf of fixed-income investors.

Fiduciary call A combination of a European call and a risk-free bond that matures on the option expiration day and has a face value equal to the exercise price of the call.

FIFO method The first in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the order in which they were placed in inventory.

Fill or kill See *immediate or cancel order*.

Finance lease From the lessee perspective, under US GAAP, a type of lease which is more akin to the purchase of an asset by the lessee. From the lessor perspective, under IFRS, a lease which "transfers substantially all the risks and rewards incidental to ownership of an underlying asset."

Financial account A component of the balance of payments account that records investment flows.

Financial distress Heightened uncertainty regarding a company's ability to meet its various obligations because of diminished earnings power or actual current losses.

Financial flexibility The ability to react and adapt to financial adversity and opportunities.

Financial leverage The extent to which a company can effect, through the use of debt, a proportional change in the return on common equity that is greater than a given proportional change in operating income; also, short for the financial leverage ratio.

Financial leverage ratio A measure of financial leverage calculated as average total assets divided by average total equity.

Financial risk The risk that environmental, social, or governance risk factors will result in significant costs or other losses to a company and its shareholders; the risk arising from a company's obligation to meet required payments under its financing agreements.

Financing activities Activities related to obtaining or repaying capital to be used in the business (e.g., equity and long-term debt).

Fintech Technological innovation in the design and delivery of financial services and products in the financial industry.

Firm commitment offering See *underwritten offering*.

First-degree price discrimination Where a monopolist is able to charge each customer the highest price the customer is willing to pay.

First lien debt Debt secured by a pledge of certain assets that could include buildings, but it may also include property and equipment, licenses, patents, brands, etc.

First mortgage debt Debt secured by a pledge of a specific property.

Fiscal multiplier The ratio of a change in national income to a change in government spending.

Fiscal policy The use of taxes and government spending to affect the level of aggregate expenditures.

Fisher effect The thesis that the real rate of interest in an economy is stable over time so that changes in nominal interest rates are the result of changes in expected inflation.

Fixed charge coverage A solvency ratio measuring the number of times interest and lease payments are covered by operating income, calculated as (EBIT + lease payments) divided by (interest payments + lease payments).

Fixed costs Costs that remain at the same level regardless of a company's level of production and sales.

Fixed-for-floating interest rate swap An interest rate swap in which one party pays a fixed rate and the other pays a floating rate, with both sets of payments in the same currency. Also called *plain vanilla swap* or *vanilla swap*.

Fixed-rate perpetual preferred stock Nonconvertible, noncallable preferred stock that has a fixed dividend rate and no maturity date.

Flag A technical analysis continuation pattern formed by parallel trendlines, typically over a short period.

Flat price The full price of a bond minus the accrued interest; also called the *quoted* or *clean* price.

Float-adjusted market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by adjusting its market capitalization for its market float.

Floaters See *floating-rate notes*.

Floating-rate notes Notes on which interest payments are not fixed but instead vary from period to period depending on the current level of a reference interest rate.

Flotation cost Fees charged to companies by investment bankers and other costs associated with raising new capital.

Foreclosure Allows the lender to take possession of a mortgaged property if the borrower defaults and then sell it to recover funds.

Foreign currency reserves Holding by the central bank of non-domestic currency deposits and non-domestic bonds.

Foreign direct investment Direct investment by a firm in one country (the source country) in productive assets in a foreign country (the host country).

Foreign exchange gains (or losses) Gains (or losses) that occur when the exchange rate changes between the investor's currency and the currency that foreign securities are denominated in.

Foreign portfolio investment Shorter-term investment by individuals, firms, and institutional investors (e.g., pension funds) in foreign financial instruments such as foreign stocks and foreign government bonds.

Forward commitments Class of derivatives that provides the ability to lock in a price to transact in the future at a previously agreed-upon price.

Forward contract An agreement between two parties in which one party, the buyer, agrees to buy from the other party, the seller, an underlying asset at a later date for a price established at the start of the contract.

Forward curve A series of forward rates, each having the same time frame.

Forward market For future delivery, beyond the usual settlement time period in the cash market.

Forward price The fixed price or rate at which the transaction scheduled to occur at the expiration of a forward contract will take place. This price is agreed on at the initiation date of the contract.

Forward rate The interest rate on a bond or money market instrument traded in a forward market. A forward rate can be interpreted as an incremental, or marginal, return for extending the time-to-maturity for an additional time period.

Forward rate agreements A forward contract calling for one party to make a fixed interest payment and the other to make an interest payment at a rate to be determined at the contract expiration.

Fractile A value at or below which a stated fraction of the data lies. Also called quantile.

Fractional reserve banking Banking in which reserves constitute a fraction of deposits.

Framing bias An information-processing bias in which a person answers a question differently based on the way in which it is asked (framed).

Free cash flow The actual cash that would be available to the company's investors after making all investments necessary to maintain the company as an ongoing enterprise (also referred to as free cash flow to the firm); the internally generated funds that can be distributed to the company's investors (e.g., shareholders and bondholders) without impairing the value of the company.

Free cash flow hypothesis The hypothesis that higher debt levels discipline managers by forcing them to make fixed debt service payments and by reducing the company's free cash flow.

Free-cash-flow-to-equity models Valuation models based on discounting expected future free cash flow to equity.

Free float The number of shares that are readily and freely tradable in the secondary market.

Free trade When there are no government restrictions on a country's ability to trade.

Free trade areas One of the most prevalent forms of regional integration, in which all barriers to the flow of goods and services among members have been eliminated.

Frequency distribution A tabular display of data constructed either by counting the observations of a variable by distinct values or groups or by tallying the values of a numerical variable into a set of numerically ordered bins (also called a one-way table).

Frequency polygon A graph of a frequency distribution obtained by drawing straight lines joining successive points representing the class frequencies.

Frequency table A representation of the frequency of occurrence of two discrete variables.

Full price The price of a security with accrued interest; also called the *invoice* or *dirty* price.

Fund investing In fund investing, the investor invests in assets indirectly by contributing capital to a fund as part of a group of investors. Fund investing is available for all major alternative investment types.

Fundamental analysis The examination of publicly available information and the formulation of forecasts to estimate the intrinsic value of assets.

Fundamental value The underlying or true value of an asset based on an analysis of its qualitative and quantitative characteristics. Also called *intrinsic value*.

Fundamental weighting An index weighting method in which the weight assigned to each constituent security is based on its underlying company's size. It attempts to address the disadvantages of market-capitalization weighting by using measures that are independent of the constituent security's price.

Funds of hedge funds Funds that hold a portfolio of hedge funds, more commonly shortened to *funds of funds*.

Fungible Freely exchangeable, interchangeable, or substitutable with other things of the same type. Money and commodities are the most common examples.

Future value (FV) The amount to which a payment or series of payments will grow by a stated future date.

Futures contract A variation of a forward contract that has essentially the same basic definition but with some additional features, such as a clearinghouse guarantee against credit losses, a daily settlement of gains and losses, and an organized electronic or floor trading facility.

Futures price The agreed-upon price of a futures contract.

FX swap The combination of a spot and a forward FX transaction.

G-spread The yield spread in basis points over an actual or interpolated government bond.

Gains Asset inflows not directly related to the ordinary activities of the business.

Game theory The set of tools decision makers use to incorporate responses by rival decision makers into their strategies.

Gamma A numerical measure of how sensitive an option's delta (the sensitivity of the derivative's price) is to a change in the value of the underlying.

Gap opening A gap is an area of a chart where a security's price either rises or falls from the previous day's close with no trading occurring in between. A gap opening is the start of a new trading session with a gap.

GDP deflator A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

General partner (GP) The partner that runs the business and theoretically bears unlimited liability for the business's debts and obligations.

Geometric mean A measure of central tendency computed by taking the nth root of the product of n non-negative values.

Giffen goods Goods that are consumed more as the price of the good rises because it is a very inferior good whose income effect overwhelms its substitution effect when price changes.

Gilts Bonds issued by the UK government.

Global depository receipt A depository receipt that is issued outside of the company's home country and outside of the United States.

Global minimum-variance portfolio The portfolio on the minimum-variance frontier with the smallest variance of return.

Global registered share A common share that is traded on different stock exchanges around the world in different currencies.

Gold standard With respect to a currency, if a currency is on the gold standard a given amount can be converted into a prespecified amount of gold.

Golden Cross See **Bullish crossover**.

Good-on-close An execution instruction specifying that an order can only be filled at the close of trading. Also called *market on close*.

Good-on-open An execution instruction specifying that an order can only be filled at the opening of trading.

Good-till-cancelled order An order specifying that it is valid until the entity placing the order has cancelled it (or, commonly, until some specified amount of time such as 60 days has elapsed, whichever comes sooner).

Goodwill An intangible asset that represents the excess of the purchase price of an acquired company over the value of the net assets acquired.

Government equivalent yield A yield that restates a yield-to-maturity based on a 30/360 day count to one based on actual/actual.

Green bonds Bonds used in green finance whereby the proceeds are earmarked toward environmental-related products.

Green finance A type of finance that addresses environmental concerns while achieving economic growth.

Green loans Any loan instruments made available exclusively to finance or re-finance, in whole or in part, new and/or existing eligible green projects. Green loans are commonly aligned in the market with the Green Loan Principles.

Greenfield investment Investing in infrastructure assets *that are to be constructed*.

Grey market The forward market for bonds about to be issued. Also called "when issued" market.

Gross domestic product The market value of all final goods and services produced within the economy during a given period (output definition) or, equivalently, the aggregate income earned by all households, all companies, and the government within the economy during a given period (income definition).

Gross margin Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit Sales minus the cost of sales (i.e., the cost of goods sold for a manufacturing company).

Gross profit margin The ratio of gross profit to revenues.

Grouped bar chart A bar chart for showing joint frequencies for two categorical variables (also known as a **clustered bar chart**).

Grouping by function With reference to the presentation of expenses in an income statement, the grouping together of expenses serving the same function, e.g. all items that are costs of goods sold.

Grouping by nature With reference to the presentation of expenses in an income statement, the grouping together of expenses by similar nature, e.g., all depreciation expenses.

Growth cycle Refers to fluctuations in economic activity around the long-term potential trend growth level, focusing on how much actual economic activity is below or above trend growth in economic activity.

Growth cyclical A term sometimes used to describe companies that are growing rapidly on a long-term basis but that still experience above-average fluctuation in their revenues and profits over the course of a business cycle.

Growth investors With reference to equity investors, investors who seek to invest in high-earnings-growth companies.

Growth option The ability to make additional investments in a project at some future time if the financial results are strong. Also called *expansion option*.

Growth rate cycle Refers to fluctuations in the growth rate of economic activity.

Guarantee certificate A type of structured financial instrument that provides investors with capital protection. It combines a zero-coupon bond and a call option on some underlying asset.

Haircut See *repo margin*.

Halo effect An emotional bias that extends a favorable evaluation of some characteristics to other characteristics.

Hard-bullet covered bonds Covered bonds for which a bond default is triggered and bond payments are accelerated in the event of sponsor default if payments do not occur according to the original maturity schedule.

Harmonic mean A type of weighted mean computed as the reciprocal of the arithmetic average of the reciprocals.

Head and shoulders pattern In technical analysis, a reversal pattern that is formed in three parts: a left shoulder, a head, and a right shoulder. A head and shoulders pattern is used to predict a change from an uptrend to a downtrend.

Headline inflation Refers to the inflation rate calculated based on the price index that includes all goods and services in an economy.

Heat map A type of graphic that organizes and summarizes data in a tabular format and represents it using a color spectrum.

Hedge funds Private investment vehicles that typically use leverage, derivatives, and long and short investment strategies.

Hedge portfolio A hypothetical combination of the derivative and its underlying that eliminates risk.

Held-to-maturity Debt (fixed-income) securities that a company intends to hold to maturity; these are presented at their original cost, updated for any amortisation of discounts or premiums.

Herding Clustered trading that may or may not be based on information.

Heteroskedasticity A violation of the simple linear regression assumption of homoskedasticity; with heteroskedasticity, the variability of the residuals is not equal across all observations.

Hidden order An order that is exposed not to the public but only to the brokers or exchanges that receive it.

High-frequency trading A form of algorithmic trading that makes use of vast quantities of data to execute trades on ultra-high-speed networks in fractions of a second.

High-water mark The highest value, net of fees, which a fund has reached in history. It reflects the highest cumulative return used to calculate an incentive fee.

Hindsight bias A bias with selective perception and retention aspects in which people may see past events as having been predictable and reasonable to expect.

Histogram A chart that presents the distribution of numerical data by using the height of a bar or column to represent the absolute frequency of each bin or interval in the distribution.

Historical cost In reference to assets, the amount paid to purchase an asset, including any costs of acquisition and/or preparation; with reference to liabilities, the amount of proceeds received in exchange in issuing the liability.

Historical equity risk premium approach An estimate of a country's equity risk premium that is based on the historical averages of the risk-free rate and the rate of return on the market portfolio.

Holder-of-record date The date that a shareholder listed on the corporation's books will be deemed to have ownership of the shares for purposes of receiving an upcoming dividend.

Holding period return The return that an investor earns during a specified holding period; a synonym for total return.

Home bias A preference for securities listed on the exchanges of one's home country.

Homogeneity of expectations The assumption that all investors have the same economic expectations and thus have the same expectations of prices, cash flows, and other investment characteristics.

Homoskedasticity An assumption of simple linear regression such that the variability of the residuals is equal across all observations.

Horizon yield The internal rate of return between the total return (the sum of reinvested coupon payments and the sale price or redemption amount) and the purchase price of the bond.

Horizontal analysis Common-size analysis that involves comparing a specific financial statement with that statement in prior or future time periods; also, cross-sectional analysis of one company with another.

Horizontal demand schedule Implies that at a given price, the response in the quantity demanded is infinite.

Hostile takeover An attempt by one entity to acquire a company without the consent of the company's management.

Household A person or a group of people living in the same residence, taken as a basic unit in economic analysis.

Human capital The accumulated knowledge and skill that workers acquire from education, training, or life experience and the corresponding present value of future earnings to be generated by said skilled individual.

Hurdle rate The minimum rate of return on investment that a fund must reach before a GP receives carried interest.

Hypothesis A proposed explanation or theory that can be tested.

Hypothesis testing The process of testing of hypotheses about one or more populations using statistical inference.

I-spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Iceberg order An order in which the display size is less than the order's full size.

If-converted method A method for accounting for the effect of convertible securities on earnings per share (EPS) that specifies what EPS would have been if the convertible securities had been converted at the beginning of the period, taking account of the effects of conversion on net income and the weighted average number of shares outstanding.

Illusion of control bias A bias in which people tend to believe that they can control or influence outcomes when, in fact, they cannot.

Immediate or cancel order An order that is valid only upon receipt by the broker or exchange. If such an order cannot be filled in part or in whole upon receipt, it cancels immediately. Also called *fill or kill*.

Impact investing Investments made with the intention to generate positive, measurable social and environmental impact alongside a financial return.

Impact lag The lag associated with the result of actions affecting the economy with delay.

Implicit price deflator for GDP A gauge of prices and inflation that measures the aggregate changes in prices across the overall economy.

Implicit selection bias One type of selection bias introduced through the presence of a threshold that filters out some unqualified members.

Implied forward rates Calculated from spot rates, an implied forward rate is a breakeven reinvestment rate that links the return on an investment in a shorter-term zero-coupon bond to the return on an investment in a longer-term zero-coupon bond.

Implied volatility The volatility that option traders use to price an option, implied by the price of the option and a particular option-pricing model.

Import license Specifies the quantity of a good that can be imported into a country.

Imports Goods and services that a domestic economy (i.e., house-holds, firms, and government) purchases from other countries.

In the money Options that, if exercised, would result in the value received being worth more than the payment required to exercise.

Income Increases in economic benefits in the form of inflows or enhancements of assets, or decreases of liabilities that result in an increase in equity (other than increases resulting from contributions by owners).

Income elasticity of demand A measure of the responsiveness of demand to changes in income, defined as the percentage change in quantity demanded divided by the percentage change in income.

Income tax paid The actual amount paid for income taxes in the period; not a provision, but the actual cash outflow.

Income tax payable The income tax owed by the company on the basis of taxable income.

Increasing marginal returns When the marginal product of a resource increases as additional units of that input are employed.

Increasing returns to scale When a production process leads to increases in output that are proportionately larger than the increase in inputs.

Incremental cash flow The cash flow that is realized because of a decision; the changes or increments to cash flows resulting from a decision or action.

Indenture Legal contract that describes the form of a bond, the obligations of the issuer, and the rights of the bondholders. Also called the *trust deed*.

Independent With reference to events, the property that the occurrence of one event does not affect the probability of another event occurring. With reference to two random variables X and Y, they are independent if and only if $P(X,Y) = P(X)P(Y)$.

Independent projects Independent projects are capital investments whose cash flows are independent of each other.

Independent variable A variable used to explain the dependent variable in a regression; a right-side variable in a regression equation. Also referred to as the *explanatory variable*.

Independently and identically distributed With respect to random variables, the property of random variables that are independent of each other but follow the identical probability distribution.

Index-linked bond Bond for which coupon payments and/or principal repayment are linked to a specified index.

Indexing An investment strategy in which an investor constructs a portfolio to mirror the performance of a specified index.

Indicator variable A variable that takes on only one of two values, 0 or 1, based on a condition. In simple linear regression, the slope is the difference in the dependent variable for the two conditions. Also referred to as a *dummy variable*.

Indifference curve A curve representing all the combinations of two goods or attributes such that the consumer is entirely indifferent among them.

Indirect format With reference to cash flow statements, a format for the presentation of the statement which, in the operating cash flow section, begins with net income then shows additions and subtractions to arrive at operating cash flow. Also called *indirect method*.

Indirect method See *indirect format*.

Indirect taxes Taxes such as taxes on spending, as opposed to direct taxes.

Industry A group of companies offering similar products and/or services.

Industry analysis The analysis of a specific branch of manufacturing, service, or trade.

Inelastic Said of a good or service when the magnitude of elasticity is less than one. Insensitive to price changes.

Inferior goods A good whose consumption decreases as income increases.

Inflation The percentage increase in the general price level from one period to the next; a sustained rise in the overall level of prices in an economy.

Inflation-linked bond Type of index-linked bond that offers investors protection against inflation by linking the bond's coupon payments and/or the principal repayment to an index of consumer prices. Also called *linkers*.

Inflation premium An extra return that compensates investors for expected inflation.

Inflation reports A type of economic publication put out by many central banks.

Inflation uncertainty The degree to which economic agents view future rates of inflation as difficult to forecast.

Information cascade The transmission of information from those participants who act first and whose decisions influence the decisions of others.

Information-motivated traders Traders that trade to profit from information that they believe allows them to predict future prices.

Informationally efficient market A market in which asset prices reflect new information quickly and rationally.

Initial coin offering An unregulated process whereby companies raise capital by selling crypto tokens to investors in exchange for fiat money or another agreed-upon cryptocurrency.

Initial margin The amount that must be deposited in a clearinghouse account when entering into a futures contract.

Initial margin requirement The margin requirement on the first day of a transaction as well as on any day in which additional margin funds must be deposited.

Initial public offering (IPO) The first issuance of common shares to the public by a formerly private corporation.

Input productivity The amount of output produced by workers in a given period of time—for example, output per hour worked; measures the efficiency of labor.

Intangible assets Assets lacking physical substance, such as patents and trademarks.

Interbank market The market of loans and deposits between banks for maturities ranging from overnight to one year.

Intercept The expected value of the dependent variable when the independent variable in a simple linear regression is equal to zero.

Interest Payment for lending funds.

Interest coverage A solvency ratio calculated as EBIT divided by interest payments.

Interest-only mortgage A loan in which no scheduled principal repayment is specified for a certain number of years.

Interest rate A rate of return that reflects the relationship between differently dated cash flows; a discount rate.

Interest rate effect The effect through which price level changes, through demand for money, impact interest rate, which in turn impacts investment and consumption.

Interest rate swap A swap in which the underlying is an interest rate. Can be viewed as a currency swap in which both currencies are the same and can be created as a combination of currency swaps.

Intermarket analysis A field within technical analysis that combines analysis of the major categories of securities—namely, equities, bonds, currencies, and commodities—to identify market trends and possible inflections in trends.

Internal rate of return The discount rate that makes net present value equal 0; the discount rate that makes the present value of an investment's costs (outflows) equal to the present value of the investment's benefits (inflows).

Internet of Things A network arrangement of structures and devices whereby the objects on the network are able to interact and share information.

Interpolated spread The yield spread of a specific bond over the standard swap rate in that currency of the same tenor.

Interquartile range The difference between the third and first quartiles of a dataset.

Interval With reference to grouped data, a set of values within which an observation falls.

Intrinsic value See *exercise value*.

Inventory investment Net change in business inventory.

Inventory turnover An activity ratio calculated as cost of goods sold divided by average inventory.

Inverse demand function A restatement of the demand function in which price is stated as a function of quantity.

Inverse floater A type of leveraged structured financial instrument. The cash flows are adjusted periodically and move in the opposite direction of changes in the reference rate.

Inverse transformation method A method using randomly generated numbers from the continuous uniform distribution to generate random observations from any distribution.

Investing activities Activities associated with the acquisition and disposal of property, plant, and equipment; intangible assets; other long-term assets; and both long-term and short-term investments in the equity and debt (bonds and loans) issued by other companies.

Investment banks Financial intermediaries that provide advice to their mostly corporate clients and help them arrange transactions such as initial and seasoned securities offerings.

Investment policy statement (IPS) A written planning document that describes a client's investment objectives and risk tolerance over a relevant time horizon, along with constraints that apply to the client's portfolio.

Investment property Property used to earn rental income or capital appreciation (or both).

Jackknife A resampling method that repeatedly draws samples by taking the original observed data sample and leaving out one observation at a time (without replacement) from the set.

January effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January. Also called *turn-of-the-year effect.*

Joint frequencies The entry in the cells of the contingency table that represent the joining of one variable from a row and the other variable from a column to count observations.

Joint probability The probability of the joint occurrence of stated events.

Joint probability function A function giving the probability of joint occurrences of values of stated random variables.

Judgmental Sampling A procedure of selectively handpicking elements from the population based on a researcher's knowledge and professional judgment.

Key rate duration A method of measuring the interest rate sensitivities of a fixed-income instrument or portfolio to shifts in key points along the yield curve.

Keynesians Economists who believe that fiscal policy can have powerful effects on aggregate demand, output, and employment when there is substantial spare capacity in an economy.

Kurtosis The statistical measure that indicates the combined weight of the tails of a distribution relative to the rest of the distribution.

Labor force The portion of the working age population (over the age of 16) that is employed or is available for work but not working (unemployed).

Labor productivity The quantity of goods and services (real GDP) that a worker can produce in one hour of work.

Lagging economic indicators Turning points that take place later than those of the overall economy; they are believed to have value in identifying the economy's past condition.

Law of demand The principle that as the price of a good rises, buyers will choose to buy less of it, and as its price falls, they will buy more.

Law of diminishing marginal returns The observation that a variable factor's marginal product must eventually fall as more of it is added to a fixed amount of the other factors.

Law of diminishing returns The smallest output that a firm can produce such that its long run average costs are minimized.

Law of one price The condition in a financial market in which two equivalent financial instruments or combinations of financial instruments can sell for only one price. Equivalent to the principle that no arbitrage opportunities are possible.

Lead underwriter The lead investment bank in a syndicate of investment banks and broker–dealers involved in a securities underwriting.

Leading economic indicators Turning points that usually precede those of the overall economy; they are believed to have value for predicting the economy's future state, usually near-term.

Lease A contract that conveys the right to use an asset for a period of time in exchange for consideration.

Legal tender Something that must be accepted when offered in exchange for goods and services.

Lender of last resort An entity willing to lend money when no other entity is ready to do so.

Leptokurtic Describes a distribution that has fatter tails than a normal distribution (also called fat-tailed).

Lessee The party obtaining the use of an asset through a lease.

Lessor The owner of an asset that grants the right to use the asset to another party.

Letter of credit Form of external credit enhancement whereby a financial institution provides the issuer with a credit line to reimburse any cash flow shortfalls from the assets backing the issue.

Level of significance The probability of a Type I error in testing a hypothesis.

Leverage In the context of corporate finance, leverage refers to the use of fixed costs within a company's cost structure. Fixed costs that are operating costs (such as depreciation or rent) create operating leverage. Fixed costs that are financial costs (such as interest expense) create financial leverage.

Leveraged buyout A transaction whereby the target company's management team converts the target to a privately held company by using heavy borrowing to finance the purchase of the target company's outstanding shares.

Liabilities Present obligations of an enterprise arising from past events, the settlement of which is expected to result in an outflow of resources embodying economic benefits; creditors' claims on the resources of a company.

LIFO layer liquidation With respect to the application of the LIFO inventory method, the liquidation of old, relatively low-priced inventory; happens when the volume of sales rises above the volume of recent purchases so that some sales are made from relatively old, low-priced inventory. Also called *LIFO liquidation*.

LIFO method The last in, first out, method of accounting for inventory, which matches sales against the costs of items of inventory in the reverse order the items were placed in inventory (i.e., inventory produced or acquired last are assumed to be sold first).

LIFO reserve The difference between the reported LIFO inventory carrying amount and the inventory amount that would have been reported if the FIFO method had been used (in other words, the FIFO inventory value less the LIFO inventory value).

Likelihood The probability of an observation, given a particular set of conditions.

Limit down A limit move in the futures market in which the price at which a transaction would be made is at or below the lower limit.

Limit order Instructions to a broker or exchange to obtain the best price immediately available when filling an order, but in no event accept a price higher than a specified (limit) price when buying or accept a price lower than a specified (limit) price when selling.

Limit order book The book or list of limit orders to buy and sell that pertains to a security.

Limit up A limit move in the futures market in which the price at which a transaction would be made is at or above the upper limit.

Limitations on liens Meant to put limits on how much secured debt an issuer can have.

Limited partners Partners with limited liability. Limited partnerships in hedge and private equity funds are typically restricted to investors who are expected to understand and to be able to assume the risks associated with the investments.

Limited partnership agreement (LPA) A legal document that outlines the rules of the partnership and establishes the framework that ultimately guides the fund's operations throughout its life.

Lin-log model A regression model in which the independent variable is in logarithmic form.

Line chart A type of graph used to visualize ordered observations. In technical analysis, a plot of price data, typically closing prices, with a line connecting the points.

Linear interpolation The estimation of an unknown value on the basis of two known values that bracket it, using a straight line between the two known values.

Linear regression Regression that models the straight-line relationship between the dependent and independent variables. Also known as *least squares regression* and *ordinary least squares regression*.

Linear scale A scale in which equal distances correspond to equal absolute amounts. Also called an *arithmetic scale*.

Linker See *inflation-linked bond*.

Liquid market Said of a market in which traders can buy or sell with low total transaction costs when they want to trade.

Liquidation To sell the assets of a company, division, or subsidiary piecemeal, typically because of bankruptcy; the form of bankruptcy that allows for the orderly satisfaction of creditors' claims after which the company ceases to exist.

Liquidity The extent to which a company is able to meet its short-term obligations using cash flows and those assets that can be readily transformed into cash.

Liquidity premium An extra return that compensates investors for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

Liquidity ratios Financial ratios measuring the company's ability to meet its short-term obligations to creditors.

Liquidity risk The risk that a financial instrument cannot be purchased or sold without a significant concession in price due to the size of the market.

Liquidity trap A condition in which the demand for money becomes infinitely elastic (horizontal demand curve) so that injections of money into the economy will not lower interest rates or affect real activity.

Load fund A mutual fund in which, in addition to the annual fee, a percentage fee is charged to invest in the fund and/or for redemptions from the fund.

Loan-to-value ratio The amount of the property's mortgage divided by the purchase price

Locked limit A condition in the futures markets in which a transaction cannot take place because the price would be beyond the limits.

Lockup period The minimum holding period before investors are allowed to make withdrawals or redeem shares from a fund.

Log-lin model A regression model in which the dependent variable is in logarithmic form.

Log-log model A regression model in which both the dependent and independent variables are in logarithmic form. Also known as the *double-log model*.

Logarithmic scale A scale in which equal distances represent equal proportional changes in the underlying quantity.

London interbank offered rate (Libor) Collective name for multiple rates at which a select set of banks believes they could borrow unsecured funds from other banks in the London interbank market for different currencies and different borrowing periods ranging from overnight to one year.

Long The buyer of a derivative contract. Also refers to the position of owning a derivative.

Long-lived assets Assets that are expected to provide economic benefits over a future period of time, typically greater than one year. Also called *long-term assets*.

Long position A position in an asset or contract in which one owns the asset or has an exercisable right under the contract.

Long-run average total cost The curve describing average total cost when no costs are considered fixed.

Look-ahead bias A bias caused by using information that was unavailable on the test date.

Loss aversion The tendency of people to dislike losses more than they like comparable gains.

Loss-aversion bias A bias in which people tend to strongly prefer avoiding losses as opposed to achieving gains.

Loss severity Portion of a bond's value (including unpaid interest) an investor loses in the event of default.

Losses Asset outflows not directly related to the ordinary activities of the business.

Lower bound The lowest possible value of an option.

M^2 A measure of what a portfolio would have returned if it had taken on the same total risk as the market index.

M^2 alpha Difference between the risk-adjusted performance of the portfolio and the performance of the benchmark.

Macaulay duration The approximate amount of time a bond would have to be held for the market discount rate at purchase to be realized if there is a single change in interest rate. It indicates the point in time when the coupon reinvestment and price effects of a change in yield-to-maturity offset each other.

Machine learning Computer based techniques that seek to extract knowledge from large amounts of data by "learning" from known examples and then generating structure or predictions. ML algorithms aim to "find the pattern, apply the pattern."

Macroeconomics The branch of economics that deals with aggregate economic quantities, such as national output and national income.

Maintenance covenants Covenants in bank loan agreements that require the borrower to satisfy certain financial ratio tests while the loan is outstanding.

Maintenance margin The minimum amount that is required by a futures clearinghouse to maintain a margin account and to protect against default. Participants whose margin balances drop below the required maintenance margin must replenish their accounts.

Maintenance margin requirement The margin requirement on any day other than the first day of a transaction.

Management buy-in Leveraged buyout in which the current management team is being replaced and the acquiring team will be involved in managing the company.

Management buyout A leveraged buyout event in which a group of investors consisting primarily of the company's existing management purchases at least controlling interest in its outstanding shares. At the extreme, they may purchase all shares and take the company private.

Management fee A fee based on assets under management or committed capital, as applicable—also called a *base fee*.

MAR ratio A variation of the Calmar ratio that uses a full investment history and the average drawdown.

Margin The amount of money that a trader deposits in a margin account. The term is derived from the stock market practice in which an investor borrows a portion of the money required to purchase a certain amount of stock. In futures markets, there is no borrowing so the margin is more of a down payment or performance bond.

Margin bond A cash deposit required by the clearinghouse from the participants to a contract to provide a credit guarantee. Also called a *performance bond*.

Margin call A request for the short to deposit additional funds to bring their balance up to the initial margin.

Margin loan Money borrowed from a broker to purchase securities.

Marginal cost The cost of producing an additional unit of a good.

Marginal frequencies The sums determined by adding joint frequencies across rows or across columns in a contingency table.

Marginal product Measures the productivity of each unit of input and is calculated by taking the difference in total product from adding another unit of input (assuming other resource quantities are held constant).

Marginal propensity to consume The proportion of an additional unit of disposable income that is consumed or spent; the change in consumption for a small change in income.

Marginal propensity to save The proportion of an additional unit of disposable income that is saved (not spent).

Marginal revenue The change in total revenue divided by the change in quantity sold; simply, the additional revenue from selling one more unit.

Marginal value curve A curve describing the highest price consumers are willing to pay for each additional unit of a good.

Mark to market The revaluation of a financial asset or liability to its current market value or fair value.

Mark-to-market Refers to the current expected fair market value for which a given security would likely be available for purchase or sale if traded in current market conditions.

Market anomaly Change in the price or return of a security that cannot directly be linked to current relevant information known in the market or to the release of new information into the market.

Market bid–ask spread The difference between the best bid and the best offer.

Market-capitalization weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its market capitalization by the total market capitalization (sum of the market capitalization) of all securities in the index. Also called *value weighting*.

Market discount rate The rate of return required by investors given the risk of the investment in a bond; also called the *required yield* or the *required rate of return*.

Market float The number of shares that are available to the investing public.

Market liquidity risk The risk that the price at which investors can actually transact—buying or selling—may differ from the price indicated in the market.

Market model A regression equation that specifies a linear relationship between the return on a security (or portfolio) and the return on a broad market index.

Market multiple models Valuation models based on share price multiples or enterprise value multiples.

Market-on-close An execution instruction specifying that an order can only be filled at the close of trading.

Market order Instructions to a broker or exchange to obtain the best price immediately available when filling an order.

Market-oriented investors With reference to equity investors, investors whose investment disciplines cannot be clearly categorized as value or growth.

Market rate of interest The rate demanded by purchasers of bonds, given the risks associated with future cash payment obligations of the particular bond issue.

Market reference rate (MRR) Collective name for a set of rates covering different currencies for different maturities, ranging from overnight to one year.

Market risk The risk that arises from movements in interest rates, stock prices, exchange rates, and commodity prices.

Market value The price at which an asset or security can currently be bought or sold in an open market.

Marketable limit order A buy limit order in which the limit price is placed above the best offer, or a sell limit order in which the limit price is placed below the best bid. Such orders generally will partially or completely fill right away.

Markowitz efficient frontier The graph of the set of portfolios offering the maximum expected return for their level of risk (standard deviation of return).

Matching principle The accounting principle that expenses should be recognized in the same period in which the associated revenue is recognized.

Matrix pricing Process of estimating the market discount rate and price of a bond based on the quoted or flat prices of more frequently traded comparable bonds.

Maturity premium An extra return that compensates investors for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended.

Maturity structure A factor explaining the differences in yields on similar bonds; also called *term structure*.

Mean absolute deviation With reference to a sample, the mean of the absolute values of deviations from the sample mean.

Mean square error (MSE) The sum of squares error divided by the degrees of freedom, $n - k - 1$; in a simple linear regression, $n - k - 1 = n - 2$.

Mean square regression (MSR) The sum of squares regression divided by the number of independent variables k; in a simple linear regression, $k = 1$.

Mean–variance analysis An approach to portfolio analysis using expected means, variances, and covariances of asset returns.

Measure of central tendency A quantitative measure that specifies where data are centered.

Measure of value A standard for measuring value; a function of money.

Measures of location Quantitative measures that describe the location or distribution of data. They include not only measures of central tendency but also other measures, such as percentiles.

Median The value of the middle item of a set of items that has been sorted into ascending or descending order (i.e., the 50th percentile).

Medium of exchange Any asset that can be used to purchase goods and services or to repay debts; a function of money.

Medium-term note A corporate bond offered continuously to investors by an agent of the issuer, designed to fill the funding gap between commercial paper and long-term bonds.

Mental accounting bias An information-processing bias in which people treat one sum of money differently from another equal-sized sum based on which mental account the money is assigned to.

Menu costs A cost of inflation in which businesses constantly have to incur the costs of changing the advertised prices of their goods and services.

Mesokurtic Describes a distribution with kurtosis equal to that of the normal distribution, namely, kurtosis equal to three.

Microeconomics The branch of economics that deals with markets and decision making of individual economic units, including consumers and businesses.

Minimum efficient scale The smallest output that a firm can produce such that its long-run average total cost is minimized.

Minimum-variance portfolio The portfolio with the minimum variance for each given level of expected return.

Minority shareholders Particular shareholders or a block of shareholders holding a small proportion of a company's outstanding shares, resulting in a limited ability to exercise control in voting activities.

Minsky moment Named for Hyman Minksy. A point in a business cycle when, after individuals become overextended in borrowing to finance speculative investments, people start realizing that something is likely to go wrong and a panic ensues, leading to asset sell-offs.

Modal interval With reference to grouped data, the interval containing the greatest number of observations (i.e., highest frequency).

Mode The most frequently occurring value in a distribution.

Modern portfolio theory (MPT) The analysis of rational portfolio choices based on the efficient use of risk.

Modified duration A measure of the percentage price change of a bond given a change in its yield-to-maturity.

Momentum oscillator A graphical representation of market sentiment that is constructed from price data and calculated so that it oscillates either between a low and a high or around some number.

Monetarists Economists who believe that the rate of growth of the money supply is the primary determinant of the rate of inflation.

Monetary policy Actions taken by a nation's central bank to affect aggregate output and prices through changes in bank reserves, reserve requirements, or its target interest rate.

Monetary transmission mechanism The process whereby a central bank's interest rate gets transmitted through the economy and ultimately affects the rate of increase of prices.

Monetary union An economic union in which the members adopt a common currency.

Money A generally accepted medium of exchange and unit of account.

Money convexity For a bond, the annual or approximate convexity multiplied by the full price.

Money creation The process by which changes in bank reserves translate into changes in the money supply.

Money duration A measure of the price change in units of the currency in which the bond is denominated given a change in its yield-to-maturity.

Money market The market for short-term debt instruments (one-year maturity or less).

Money market securities Fixed-income securities with maturities at issuance of one year or less.

Money multiplier Describes how a change in reserves is expected to affect the money supply; in its simplest form, 1 divided by the reserve requirement.

Money neutrality The thesis that an increase in the money supply leads in the long-run to an increase in the price level, while leaving real variables like output and employment unaffected.

Money-weighted return The internal rate of return on a portfolio, taking account of all cash flows.

Moneyness The relationship between the price of the underlying and an option's exercise price.

Monopolistic competition Highly competitive form of imperfect competition; the competitive characteristic is a notably large number of firms, while the monopoly aspect is the result of product differentiation.

Monopoly In pure monopoly markets, there are no substitutes for the given product or service. There is a single seller, which exercises considerable power over pricing and output decisions.

Monte Carlo simulation A technique involving the use of computer software to represent the operation of a complex financial system. A characteristic feature is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Moral principles Beliefs regarding what is good, acceptable, or obligatory behavior and what is bad, unacceptable, or forbidden behavior.

Mortgage-backed securities Debt obligations that represent claims to the cash flows from pools of mortgage loans, most commonly on residential property.

Mortgage loan A loan secured by the collateral of some specified real estate property that obliges the borrower to make a predetermined series of payments to the lender.

Mortgage pass-through security A security created when one or more holders of mortgages form a pool of mortgages and sell shares or participation certificates in the pool.

Mortgage rate The interest rate on a mortgage loan; also called *contract rate* or *note rate*.

Moving average The average of the closing price of a security over a specified number of periods. With each new period, the average is recalculated.

Moving-average convergence/divergence oscillator A momentum oscillator that is based on the difference between short-term and long-term moving averages of a security's price.

Multi-factor model A model that explains a variable in terms of the values of a set of factors.

Multi-market indexes Comprised of indexes from different countries, designed to represent multiple security markets.

Multi-step format With respect to the format of the income statement, a format that presents a subtotal for gross profit (revenue minus cost of goods sold).

Multilateral trading facilities See *alternative trading systems*.

Multinational corporation A company operating in more than one country or having subsidiary firms in more than one country.

Multinomial formula (general formula for labeling problems) The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by $\dfrac{n!}{n_1!n_2!\ldots n_k!}$.

Multiple of invested capital (MOIC) A simplified calculation that measures the total value of all distributions and residual asset values relative to an initial total investment; also known as a *money multiple*.

Multiple testing problem The risk of getting statistically significant test results when performing a test multiple times.

Multiplication rule for counting If one task can be done in n_1 ways, and a second task, given the first, can be done in n_2 ways, and a third task, given the first two tasks, can be done in n_3 ways, and so on for k tasks, then the number of ways the k tasks can be done is $(n_1)(n_2)(n_3) \ldots (n_k)$.

Multiplication rule for expected value of the product of uncorrelated random variables This rule is $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated.

Multiplication rule for independent events The rule that when two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B.

Multiplication rule for probability The rule that the joint probability of events A and B equals the probability of A given B times the probability of B.

Multiplier models Valuation models based on share price multiples or enterprise value multiples.

Multivariate distribution A probability distribution that specifies the probabilities for a group of related random variables.

Multivariate normal distribution A probability distribution for a group of random variables that is completely defined by the means and variances of the variables plus all the correlations between pairs of the variables.

Muni See *municipal bond*.

Municipal bond A type of non-sovereign bond issued by a state or local government in the United States. It very often (but not always) offers income tax exemptions.

Mutual fund A comingled investment pool in which investors in the fund each have a pro-rata claim on the income and value of the fund.

Mutually exclusive Indicates that only one event can occur at a time.

Mutually exclusive projects Mutually exclusive projects compete directly with each other. For example, if Projects A and B are mutually exclusive, you can choose A or B but you cannot choose both.

***n* Factorial** For a positive integer n, the product of the first n positive integers; 0 factorial equals 1 by definition. n factorial is written as $n!$.

Narrow money The notes and coins in circulation in an economy, plus other very highly liquid deposits.

Nash equilibrium When two or more participants in a non-coop-erative game have no incentive to deviate from their respective equilibrium strategies given their opponent's strategies.

National income The income received by all factors of production used in the generation of final output. National income equals gross domestic product (or, in some countries, gross national product) minus the capital consumption allowance and a statistical discrepancy.

Natural language processing Computer programs developed to analyze and interpret human language.

Natural rate of unemployment Effective unemployment rate, below which pressure emerges in labor markets.

Negative screening An ESG investment style that focuses on the exclusion of certain sectors, companies, or practices in a fund or portfolio on the basis of specific ESG criteria.

Net book value The remaining (undepreciated) balance of an asset's purchase cost. For liabilities, the face value of a bond minus any unamortized discount, or plus any unamortized premium.

Net exports The difference between the value of a country's exports and the value of its imports (i.e., value of exports minus imports).

Net income The difference between revenue and expenses; what remains after subtracting all expenses (including depreciation, interest, and taxes) from revenue.

Net present value The present value of an investment's cash inflows (benefits) minus the present value of its cash outflows (costs).

Net profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also called *profit margin* or *return on sales*.

Net realisable value Estimated selling price in the ordinary course of business less the estimated costs necessary to make the sale.

Net revenue Revenue after adjustments (e.g., for estimated returns or for amounts unlikely to be collected).

Net tax rate The tax rate net of transfer payments.

Neural networks Computer programs based on how our own brains learn and process information.

Neutral rate of interest The rate of interest that neither spurs on nor slows down the underlying economy.

No-load fund A mutual fund in which there is no fee for investing in the fund or for redeeming fund shares, although there is an annual fee based on a percentage of the fund's net asset value.

Node Each value on a binomial tree from which successive moves or outcomes branch.

Nominal data Categorical values that are not amenable to being organized in a logical order. An example of nominal data is the classification of publicly listed stocks into sectors.

Nominal GDP The value of goods and services measured at current prices.

Nominal risk-free interest rate The sum of the real risk-free interest rate and the inflation premium.

Non-accelerating inflation rate of unemployment Effective unemployment rate below which pressure emerges in labor markets.

Non-agency RMBS In the United States, securities issued by private entities that are not guaranteed by a federal agency or a GSE.

Non-Bank lender Unlike typical banks, which make loans and take deposits, these lenders only make loans. Although they lend money to companies, they usually specialize in issuing mortgages to first-time homebuyers and parties looking to refinance an existing mortgage.

Non-cumulative preference shares Preference shares for which dividends that are not paid in the current or subsequent periods are forfeited permanently (instead of being accrued and paid at a later date).

Non-current assets Assets that are expected to benefit the company over an extended period of time (usually more than one year).

Non-current liabilities Obligations that broadly represent a probable sacrifice of economic benefits in periods generally greater than one year in the future.

Non-cyclical A company whose performance is largely independent of the business cycle. Also known as defensive.

Non-deliverable forwards Cash-settled forward contracts, used predominately with respect to foreign exchange forwards. Also called *contracts for differences*.

Non-financial risks Risks that arise from sources other than changes in the external financial markets, such as changes in accounting rules, legal environment, or tax rates.

Non-participating preference shares Preference shares that do not entitle shareholders to share in the profits of the company. Instead, shareholders are only entitled to receive a fixed dividend payment and the par value of the shares in the event of liquidation.

Non-probability sampling A sampling plan dependent on factors other than probability considerations, such as a sampler's judgment or the convenience to access data.

Non-recourse loan A loan in which the lender does not have a shortfall claim against the borrower, so the lender can look only to the property to recover the outstanding mortgage balance.

Non-renewable resources Finite resources that are depleted once they are consumed, such as oil and coal.

Non-sovereign bond A bond issued by a government below the national level, such as a province, region, state, or city.

Nonconventional cash flows In a nonconventional cash flow pattern, the initial outflow is not followed by inflows only, but the cash flows can flip from positive (inflows) to negative (outflows) again or even change signs several times.

Nonparametric test A test that is not concerned with a parameter or that makes minimal assumptions about the population from which a sample comes.

Nonsystematic risk Unique risk that is local or limited to a particular asset or industry that need not affect assets outside of that asset class.

Normal distribution A continuous, symmetric probability distribution that is completely described by its mean and its variance.

Normal goods Goods that are consumed in greater quantities as income increases.

Normal profit The level of accounting profit needed to just cover the implicit opportunity costs ignored in accounting costs.

Notching Ratings adjustment methodology where specific issues from the same borrower may be assigned different credit ratings.

Note rate See *mortgage rate*.

Notice period The length of time (typically 30 to 90 days) in advance that investors may be required to notify a fund of their intent to redeem some or all of their investment.

Notional principal An imputed principal amount.

Null hypothesis The hypothesis that is tested.

Numerical data Values that represent measured or counted quantities as a number. Also called **quantitative data**.

Objective probabilities Probabilities that generally do not vary from person to person; includes a priori and empirical probabilities.

Observation The value of a specific variable collected at a point in time or over a specified period of time.

Odds against *E* The reciprocal of odds for *E*.

Odds for *E* The probability of *E* divided by 1 minus the probability of *E*.

Off-the-run Seasoned government bonds are off-the-run securities; they are not the most recently issued or the most actively traded.

Offer The price at which a dealer or trader is willing to sell an asset, typically qualified by a maximum quantity (ask size).

Official interest rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks. Also called *official policy rate* or *policy rate*.

Official policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Oligopoly Market structure with a relatively small number of firms supplying the market.

On-the-run The most recently issued and most actively traded sovereign securities.

One-dimensional array The simplest format for representing a collection of data of the same data type.

One-sided hypothesis test A test in which the null hypothesis is rejected only if the evidence indicates that the population parameter is greater than or less than the hypothesized parameter; occurs when the alternative hypothesis is stated either as greater than or less than the hypothesized population parameter.

Open economy An economy that trades with other countries.

Open-end fund A mutual fund that accepts new investment money and issues additional shares at a value equal to the net asset value of the fund at the time of investment.

Open interest The number of outstanding contracts in a clearinghouse at any given time. The open interest figure changes daily as some parties open up new positions, while other parties offset their old positions.

Open market operations The purchase or sale of bonds by the national central bank to implement monetary policy. The bonds traded are usually sovereign bonds issued by the national government.

Operating activities Activities that are part of the day-to-day business functioning of an entity, such as selling inventory and providing services.

Operating breakeven The number of units produced and sold at which the company's operating profit is zero (revenues = operating costs).

Operating cash flow The net amount of cash provided from operating activities.

Operating efficiency ratios Ratios that measure how efficiently a company performs day-to-day tasks, such as the collection of receivables and management of inventory.

Operating lease A type of lease which is more akin to the rental of the underlying asset.

Operating leverage The use of fixed costs in operations.

Operating profit A company's profits on its usual business activities before deducting taxes. Also called *operating income*.

Operating profit margin A profitability ratio calculated as operating income (i.e., income before interest and taxes) divided by revenue. Also called *operating margin*.

Operating risk The risk attributed to the operating cost structure, in particular the use of fixed costs in operations; the risk arising from the mix of fixed and variable costs; the risk that a company's operations may be severely affected by environmental, social, and governance risk factors.

Operational independence A bank's ability to execute monetary policy and set interest rates in the way it thought would best meet the inflation target.

Operational risk The risk that arises from inadequate or failed people, systems, and internal policies, procedures, and processes, as well as from external events that are beyond the control of the organization but that affect its operations.

Operationally efficient Said of a market, a financial system, or an economy that has relatively low transaction costs.

Opportunity cost The value that investors forgo by choosing a particular course of action; the value of something in its best alternative use.

Optimal capital structure The capital structure at which the value of the company is maximized.

Option A financial instrument that gives one party the right, but not the obligation, to buy or sell an underlying asset from or to another party at a fixed price over a specific period of time. Also referred to as *contingent claim* or *option contract*.

Option-adjusted price The value of the embedded option plus the flat price of the bond.

Option-adjusted spread OAS = Z-spread – Option value (in basis points per year).

Option-adjusted yield The required market discount rate whereby the price is adjusted for the value of the embedded option.

Option contract See *option*.

Option premium The amount of money a buyer pays and seller receives to engage in an option transaction.

Order A specification of what instrument to trade, how much to trade, and whether to buy or sell.

Order-driven markets A market (generally an auction market) that uses rules to arrange trades based on the orders that traders submit; in their pure form, such markets do not make use of dealers.

Order precedence hierarchy With respect to the execution of orders to trade, a set of rules that determines which orders execute before other orders.

Ordinal data Categorical values that can be logically ordered or ranked.

Ordinary annuity An annuity with a first cash flow that is paid one period from the present.

Ordinary shares Equity shares that are subordinate to all other types of equity (e.g., preferred equity). Also called *common stock* or *common shares*.

Organized exchange A securities marketplace where buyers and seller can meet to arrange their trades.

Other comprehensive income Items of comprehensive income that are not reported on the income statement; comprehensive income minus net income.

Out-of-sample test A test of a strategy or model using a sample outside the period on which the strategy or model was developed.

Out of the money Options that, if exercised, would require the payment of more money than the value received and therefore would not be currently exercised.

Outcome A possible value of a random variable.

Over-the-counter (OTC) market A decentralized market where buy and sell orders initiated from various locations are matched through a communications network.

Overbought A market condition in which market sentiment is thought to be unsustainably bullish.

Overcollateralization Form of internal credit enhancement that refers to the process of posting more collateral than needed to obtain or secure financing.

Overconfidence bias A bias in which people demonstrate unwarranted faith in their own intuitive reasoning, judgments, and/or cognitive abilities.

Overfitting An undesirable result from fitting a model so closely to a dataset that it does not perform well on new data.

Oversold A market condition in which market sentiment is thought to be unsustainably bearish.

Own price The price of a good or service itself (as opposed to the price of something else).

Own-price elasticity of demand The percentage change in quantity demanded for a percentage change in good's own price, holding all other things constant.

Owners' equity The excess of assets over liabilities; the residual interest of shareholders in the assets of an entity after deducting the entity's liabilities. Also called *shareholders' equity* or *shareholders' funds*.

***p*-Value** The smallest level of significance at which the null is rejected.

Paired comparisons test See *test of the mean of the differences.*

Panel data A mix of time-series and cross-sectional data that contains observations through time on characteristics of across multiple observational units.

Par curve A sequence of yields-to-maturity such that each bond is priced at par value. The bonds are assumed to have the same currency, credit risk, liquidity, tax status, and annual yields stated for the same periodicity.

Par value The amount of principal on a bond.

Parallel shift A parallel yield curve shift implies that all rates change by the same amount in the same direction.

Parameter A descriptive measure computed from or used to describe a population of data, conventionally represented by Greek letters.

Parametric test Any test (or procedure) concerned with parameters or whose validity depends on assumptions concerning the population generating the sample.

Pari passu Covenant or contract clause that ensures a debt obligation is treated the same as the borrower's other senior debt instruments and is not subordinated to similar obligations.

Partial duration See *key rate duration.*

Participating preference shares Preference shares that entitle shareholders to receive the standard preferred dividend plus the opportunity to receive an additional dividend if the company's profits exceed a pre-specified level.

Pass-through rate The coupon rate of a mortgage pass-through security.

Passive investment A buy and hold approach in which an investor does not make portfolio changes based on short-term expectations of changing market or security performance.

Payable date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payment date The day that the company actually mails out (or electronically transfers) a dividend payment.

Payments system The system for the transfer of money.

Pearson correlation A parametric measure of the relationship between two variables.

Pecking order theory The theory that managers consider how their actions might be interpreted by outsiders and thus order their preferences for various forms of corporate financing. Forms of financing that are least visible to outsiders (e.g., internally generated funds) are most preferable to managers, and those that are most visible (e.g., equity) are least preferable.

Peer company See *comparable company.*

Peer group A group of companies engaged in similar business activities whose economics and valuation are influenced by closely related factors.

Pennant A technical analysis continuation pattern formed by trendlines that converge to form a triangle, typically over a short period.

Per capita real GDP Real GDP divided by the size of the population, often used as a measure of a country's average standard of living.

Per unit contribution margin The amount that each unit sold contributes to covering fixed costs—that is, the difference between the price per unit and the variable cost per unit.

Percentiles Quantiles that divide a distribution into 100 equal parts that sum to 100.

Perfect capital markets Markets in which, by assumption, there are no taxes, transaction costs, or bankruptcy costs and in which all investors have equal ("symmetric") information.

Perfect competition A market structure in which the individual firm has virtually no impact on market price, because it is assumed to be a very small seller among a very large number of firms selling essentially identical products.

Perfectly elastic When the quantity demanded or supplied of a given good is infinitely sensitive to a change in the value of a specified variable (e.g., price).

Perfectly inelastic When the quantity demanded or supplied of a given good is completely insensitive to a change in the value of a specified variable (e.g., price).

Performance bond See *margin bond.*

Performance evaluation The measurement and assessment of the outcomes of investment management decisions.

Performance fee Fees paid to the general partner from the limited partner(s) based on realized net profits.

Period costs Costs (e.g., executives' salaries) that cannot be directly matched with the timing of revenues and which are thus expensed immediately.

Periodicity The assumed number of periods in the year; typically matches the frequency of coupon payments.

Permanent differences Differences between tax and financial reporting of revenue (expenses) that will not be reversed at some future date. These result in a difference between the company's effective tax rate and statutory tax rate and do not result in a deferred tax item.

Permissioned networks Networks that are fully open only to select participants on a DLT network.

Permissionless networks Networks that are fully open to any user on a DLT network.

Permutation An ordered listing.

Permutation formula The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is ${}_nP_r = \dfrac{n!}{(n-r)!}$.

Perpetual bonds Bonds with no stated maturity date.

Perpetuity A perpetual annuity, or a set of never-ending level sequential cash flows, with the first cash flow occurring one period from now. A bond that does not mature.

Personal income A broad measure of household income that includes all income received by households, whether earned or unearned; measures the ability of consumers to make purchases.

Pet projects Investments in which influential managers want the corporation to invest. Often, unfortunately, pet projects are selected without undergoing normal capital allocation analysis.

Plain vanilla bond Bond that makes periodic, fixed coupon payments during the bond's life and a lump-sum payment of principal at maturity. Also called *conventional bond*.

Platykurtic Describes a distribution that has relatively less weight in the tails than the normal distribution (also called thin-tailed).

Point estimate A single numerical estimate of an unknown quantity, such as a population parameter.

Policy rate An interest rate that a central bank sets and announces publicly; normally the rate at which it is willing to lend money to the commercial banks.

Population All members of a specified group.

Portfolio company In private equity, the company in which the private equity fund is investing.

Portfolio demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments.

Portfolio planning The process of creating a plan for building a portfolio that is expected to satisfy a client's investment objectives.

Position The quantity of an asset that an entity owns or owes.

Positive screening The process of including sectors or companies based on specific ESG criteria, typically ESG performance relative to industry peers

Posterior probability An updated probability that reflects or comes after new information.

Potential GDP The level of real GDP that can be produced at full employment; measures the productive capacity of the economy.

Power of a test The probability of correctly rejecting the null—that is, rejecting the null hypothesis when it is false.

Precautionary money balances Money held to provide a buffer against unforeseen events that might require money.

Preference shares A type of equity interest which ranks above common shares with respect to the payment of dividends and the distribution of the company's net assets upon liquidation. They have characteristics of both debt and equity securities. Also called *preferred stock*.

Preferred stock See *preference shares*.

Premium In the case of bonds, premium refers to the amount by which a bond is priced above its face (par) value. In the case of an option, the amount paid for the option contract.

Prepaid expense A normal operating expense that has been paid in advance of when it is due.

Prepayment option Contractual provision that entitles the borrower to prepay all or part of the outstanding mortgage principal prior to the scheduled due date when the principal must be repaid. Also called *early repayment option*.

Prepayment penalty mortgages Mortgages that stipulate a monetary penalty if a borrower prepays within a certain time period after the mortgage is originated.

Prepayment risk The uncertainty that the timing of the actual cash flows will be different from the scheduled cash flows as set forth in the loan agreement due to the borrowers' ability to alter payments, usually to take advantage of interest rate movements.

Present value (PV) The present discounted value of future cash flows: For assets, the present discounted value of the future net cash inflows that the asset is expected to generate; for liabilities, the present discounted value of the future net cash outflows that are expected to be required to settle the liabilities.

Present value models Valuation models that estimate the intrinsic value of a security as the present value of the future benefits expected to be received from the security. Also called *discounted cash flow models*.

Pretax margin A profitability ratio calculated as earnings before taxes divided by revenue.

Price elasticity of demand Measures the percentage change in the quantity demanded, given a percentage change in the price of a given product.

Price index Represents the average prices of a basket of goods and services.

Price limits Limits imposed by a futures exchange on the price change that can occur from one day to the next.

Price multiple A ratio that compares the share price with some sort of monetary flow or value to allow evaluation of the relative worth of a company's stock.

Price priority The principle that the highest priced buy orders and the lowest priced sell orders execute first.

Price relative A ratio of an ending price over a beginning price; it is equal to 1 plus the holding period return on the asset.

Price return Measures *only* the price appreciation or percentage change in price of the securities in an index or portfolio.

Price return index An index that reflects *only* the price appreciation or percentage change in price of the constituent securities. Also called *price index*.

Price-setting option The option to adjust prices when demand varies from what is forecast.

Price stability In economics, refers to an inflation rate that is low on average and not subject to wide fluctuation.

Price takers Producers that must accept whatever price the market dictates.

Price to book value A valuation ratio calculated as price per share divided by book value per share.

Price to cash flow A valuation ratio calculated as price per share divided by cash flow per share.

Price to earnings ratio (P/E ratio or P/E) The ratio of share price to earnings per share.

Price to sales A valuation ratio calculated as price per share divided by sales per share.

Price value of a basis point A version of money duration, it is an estimate of the change in the full price of a bond given a 1 basis point change in the yield-to-maturity.

Price weighting An index weighting method in which the weight assigned to each constituent security is determined by dividing its price by the sum of all the prices of the constituent securities.

Priced risk Risk for which investors demand compensation for bearing (e.g., equity risk, company-specific factors, macroeconomic factors).

Primary bond market A market in which issuers first sell bonds to investors to raise capital.

Primary capital markets (primary markets) The market where securities are first sold and the issuers receive the proceeds.

Primary dealer Financial institution that is authorized to deal in new issues of sovereign bonds and that serves primarily as a trading counterparty of the office responsible for issuing sovereign bonds.

Primary market The market where securities are first sold and the issuers receive the proceeds.

Prime brokers Brokers that provide services that commonly include custody, administration, lending, short borrowing, and trading.

Principal The amount of funds originally invested in a project or instrument; the face value to be paid at maturity.

Principal–agent relationship A relationship in which a principal hires an agent to perform a particular task or service; also known as an *agency relationship*.

Principal amount Amount that an issuer agrees to repay the debtholders on the maturity date.

Principal business activity The business activity from which a company derives a majority of its revenues and/or earnings.

Principal value Amount that an issuer agrees to repay the debtholders on the maturity date.

Principle of no arbitrage See *arbitrage-free pricing*.

Prior probabilities Probabilities reflecting beliefs prior to the arrival of new information.

Priority of claims Priority of payment, with the most senior or highest ranking debt having the first claim on the cash flows and assets of the issuer.

Private equity fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *venture capital fund*.

Private equity funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *venture capital funds*.

Private equity securities Securities that are not listed on public exchanges and have no active secondary market. They are issued primarily to institutional investors via non-public offerings, such as private placements.

Private investment in public equity (PIPE) An investment in the equity of a publicly traded firm that is made at a discount to the market value of the firm's shares.

Private placement Typically, a non-underwritten, unregistered offering of securities that are sold only to an investor or a small group of investors. It can be accomplished directly between the issuer and the investor(s) or through an investment bank.

Probability A number between 0 and 1 describing the chance that a stated event will occur.

Probability density function A function with non-negative values such that probability can be described by areas under the curve graphing the function.

Probability distribution A distribution that specifies the probabilities of a random variable's possible outcomes.

Probability function A function that specifies the probability that the random variable takes on a specific value.

Probability sampling A sampling plan that allows every member of the population to have an equal chance of being selected.

Probability tree diagram A diagram with branches emanating from nodes representing either mutually exclusive chance events or mutually exclusive decisions.

Production-flexibility options The options to alter production when demand varies from what is forecast.

Production function Provides the quantitative link between the levels of output that the economy can produce and the inputs used in the production process.

Productivity The amount of output produced by workers during a given period—for example, output per hour worked measures the efficiency of labor.

Profession An occupational group that has specific education, expert knowledge, and a framework of practice and behavior that underpins community trust, respect, and recognition.

Profit The return that owners of a company receive for the use of their capital and the assumption of financial risk when making their investments.

Profit and loss (P&L) statement A financial statement that provides information about a company's profitability over a stated period of time. Also called the *income statement*.

Profit margin An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses.

Profitability ratios Ratios that measure a company's ability to generate profitable sales from its resources (assets).

Project sequencing To defer the decision to invest in a future project until the outcome of some or all of a current investment is known. Investments are sequenced over time, so that making an investment creates the option to invest in future projects.

Promissory note A written promise to pay a certain amount of money on demand.

Property, plant, and equipment Tangible assets that are expected to be used for more than one period in either the production or supply of goods or services, or for administrative purposes.

Prospectus The document that describes the terms of a new bond issue and helps investors perform their analysis on the issue.

Protective put An option strategy in which a long position in an asset is combined with a long position in a put.

Proxy contest Corporate takeover mechanism in which shareholders are persuaded to vote for a group seeking a controlling position on a company's board of directors.

Proxy voting A process that enables shareholders who are unable to attend a meeting to authorize another individual to vote on their behalf.

Public offer See *public offering*.

Public offering An offering of securities in which any member of the public may buy the securities. Also called *public offer*.

Public–private partnership (PPP) An agreement between the public sector and the private sector to finance, build, and operate public infrastructure, such as hospitals and toll roads.

Pull on liquidity When disbursements are paid too quickly or trade credit availability is limited, requiring companies to expend funds before they receive funds from sales that could cover the liability.

Pure discount bonds See *zero-coupon bond*.

Put An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Put–call–forward parity The relationship among puts, calls, and forward contracts.

Put–call parity An equation expressing the equivalence (parity) of a portfolio of a call and a bond with a portfolio of a put and the underlying, which leads to the relationship between put and call prices.

Put/call ratio A technical analysis indicator that evaluates market sentiment based on the volume of put options traded divided by the volume of call options traded for a particular financial instrument.

Put option An option that gives the holder the right to sell an underlying asset to another party at a fixed price over a specific period of time.

Putable bonds Bonds that give the bondholder the right to sell the bond back to the issuer at a predetermined price on specified dates.

Qualitative data see **categorical data**.

Quantile A value at or below which a stated fraction of the data lies. Also referred to as a fractile.

Quantitative data see **numerical data**.

Quantitative easing An expansionary monetary policy based on aggressive open market purchase operations.

Quantity equation of exchange An expression that over a given period, the amount of money used to purchase all goods and services in an economy, $M \times V$, is equal to monetary value of this output, $P \times Y$.

Quantity theory of money Asserts that total spending (in money terms) is proportional to the quantity of money.

Quartiles Quantiles that divide a distribution into four equal parts.

Quasi-fixed cost A cost that stays the same over a range of production but can change to another constant level when production moves outside of that range.

Quasi-government bond A bond issued by an entity that is either owned or sponsored by a national government. Also called *agency bond*.

Quintiles Quantiles that divide a distribution into five equal parts.

Quota rents Profits that foreign producers can earn by raising the price of their goods higher than they would without a quota.

Quotas Government policies that restrict the quantity of a good that can be imported into a country, generally for a specified period of time.

Quote-driven market A market in which dealers acting as principals facilitate trading.

Quoted interest rate A quoted interest rate that does not account for compounding within the year. Also called *stated annual interest rate*.

Quoted margin The specified yield spread over the reference rate, used to compensate an investor for the difference in the credit risk of the issuer and that implied by the reference rate.

Random number An observation drawn from a uniform distribution.

Random number generator An algorithm that produces uniformly distributed random numbers between 0 and 1.

Random variable A quantity whose future outcomes are uncertain.

Range The difference between the maximum and minimum values in a dataset.

Raw data Data available in their original form as collected.

Real exchange rate effect The effect through which changing price level impacts real exchange rate which in turn impacts net exports and aggregate demand.

Real GDP The value of goods and services produced, measured at base year prices.

Real income Income adjusted for the effect of inflation on the purchasing power of money. Also known as the *purchasing power of income*. If income remains constant and a good's price falls, real income is said to rise, even though the number of monetary units (e.g., dollars) remains unchanged.

Real interest rate Nominal interest rate minus the expected rate of inflation.

Real risk-free interest rate The single-period interest rate for a completely risk-free security if no inflation were expected.

Realizable (settlement) value With reference to assets, the amount of cash or cash equivalents that could currently be obtained by selling the asset in an orderly disposal; with reference to liabilities, the undiscounted amount of cash or cash equivalents expected to be paid to satisfy the liabilities in the normal course of business.

Rebalancing Adjusting the weights of the constituent securities in an index.

Rebalancing policy The set of rules that guide the process of restoring a portfolio's asset class weights to those specified in the strategic asset allocation.

Recession A period during which real GDP decreases (i.e., negative growth) for at least two successive quarters, or a period of significant decline in total output, income, employment, and sales usually lasting from six months to a year.

Recognition lag The lag in government response to an economic problem resulting from the delay in confirming a change in the state of the economy.

Recourse loan A loan in which the lender has a claim against the borrower for any shortfall between the outstanding mortgage balance and the proceeds received from the sale of the property.

Redemption yield See *yield-to-maturity*.

Redemptions Withdrawals of funds by investors, as allowed by the notice period and other terms in the partnership agreement.

Refinancing rate A type of central bank policy rate.

Registered bonds Bonds for which ownership is recorded by either name or serial number.

Regression analysis A tool for examining whether a variable is useful for explaining another variable.

Regression coefficients The intercept and slope coefficient of a regression.

Regret The feeling that an opportunity has been missed; typically, an expression of *hindsight bias*.

Regret-aversion bias An emotional bias in which people tend to avoid making decisions that will result in action out of fear that the decision will turn out poorly.

Relative dispersion The amount of dispersion relative to a reference value or benchmark.

Relative frequency The absolute frequency of each unique value of the variable divided by the total number of observations of the variable.

Relative price The price of a specific good or service in comparison with those of other goods and services.

Relative strength analysis A comparison of the performance of one asset with the performance of another asset or a benchmark, based on changes in the ratio of the two assets' prices over time.

Relative strength index (RSI) A technical analysis momentum oscillator that compares a security's gains with its losses over a set period.

Renewable resources Resources that can be replenished, such as a forest.

Rent Payment for the use of property.

Reorganization Agreements made by a company in bankruptcy under which a company's capital structure is altered and/or alternative arrangements are made for debt repayment; US Chapter 11 bankruptcy. The company emerges from bankruptcy as a going concern.

Replication The creation of an asset or portfolio from another asset, portfolio, and/or derivative.

Repo A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy back the same security from the purchaser at an agreed-on price and future date. The party who sells the

security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repo margin The difference between the market value of the security used as collateral and the value of the loan. Also called *haircut*.

Repo rate The interest rate on a repurchase agreement.

Representativeness bias A belief perseverance bias in which people tend to classify new information based on past experiences and classifications.

Repurchase agreement A form of collateralized loan involving the sale of a security with a simultaneous agreement by the seller to buy the same security back from the purchaser at an agreed-on price and future date. The party who sells the security at the inception of the repurchase agreement and buys it back at maturity is borrowing money from the other party, and the security sold and subsequently repurchased represents the collateral.

Repurchase date The date when the party who sold the security at the inception of a repurchase agreement buys back the security from the cash lending counterparty.

Repurchase price The price at which the party who sold the security at the inception of the repurchase agreement buys back the security from the cash lending counterparty.

Required margin The yield spread over or under the reference rate such that an FRN is priced at par value on a rate reset date.

Required rate of return See *market discount rate*.

Required yield See *market discount rate*.

Required yield spread The difference between the yield-to-maturity on a new bond and the benchmark rate; additional compensation required by investors for the difference in risk and tax status of a bond relative to a government bond. Sometimes called the *spread over the benchmark*.

Resampling A statistical method that repeatedly draws samples from the original observed data sample for the statistical inference of population parameters.

Reserve accounts Form of internal credit enhancement that relies on creating accounts and depositing in these accounts cash that can be used to absorb losses. Also called *reserve funds*.

Reserve funds See *reserve accounts*.

Reserve requirement The requirement for banks to hold reserves in proportion to the size of deposits.

Residual The difference between an observation and its predicted value, where the predicted value is based on the estimated linear relation between the dependent and independent variables using sample data.

Resistance In technical analysis, a price range in which selling activity is sufficient to stop the rise in the price of a security.

Responsible investing An overall (umbrella) term for several investment strategies—ESG investing, SRI investing, thematic investing, and impact investing—that incorporate ESG analysis in their investment processes.

Restricted payments A bond covenant meant to protect creditors by limiting how much cash can be paid out to shareholders over time.

Retracement In technical analysis, a reversal in the movement of a security's price such that it is counter to the prevailing longer-term price trend.

Return-generating model A model that can provide an estimate of the expected return of a security given certain parameters and estimates of the values of the independent variables in the model.

Return on assets (ROA) A profitability ratio calculated as net income divided by average total assets; indicates a company's net profit generated per dollar invested in total assets.

Return on equity (ROE) A profitability ratio calculated as net income divided by average shareholders' equity.

Return on invested capital A measure of the profitability of a company relative to the amount of capital invested by the equity- and debtholders.

Return on sales An indicator of profitability, calculated as net income divided by revenue; indicates how much of each dollar of revenues is left after all costs and expenses. Also referred to as *net profit margin*.

Return on total capital A profitability ratio calculated as EBIT divided by the sum of short- and long-term debt and equity.

Revaluation model Under IFRS, the process of valuing long-lived assets at fair value, rather than at cost less accumulated depreciation. Any resulting profit or loss is either reported on the income statement and/or through equity under revaluation surplus.

Revenue The amount charged for the delivery of goods or services in the ordinary activities of a business over a stated period; the inflows of economic resources to a company over a stated period.

Reversal pattern A type of pattern used in technical analysis to predict the end of a trend and a change in the direction of a security's price.

Reverse repo A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse repurchase agreement A repurchase agreement viewed from the perspective of the cash lending counterparty.

Reverse stock split A reduction in the number of shares outstanding with a corresponding increase in share price, but no change to the company's underlying fundamentals.

Revolving credit agreements (also known as revolvers) The most reliable form of short-term bank borrowing facilities; they are in effect for multiple years (e.g., three to five years) and can have optional medium-term loan features.

Rho The sensitivity of the option price to the risk-free rate.

Ricardian equivalence An economic theory that implies that it makes no difference whether a government finances a deficit by increasing taxes or issuing debt.

Risk Exposure to uncertainty. The chance of a loss or adverse outcome as a result of an action, inaction, or external event.

Risk averse The assumption that an investor will choose the least risky alternative.

Risk aversion The degree of an investor's inability and unwillingness to take risk.

Risk budgeting The establishment of objectives for individuals, groups, or divisions of an organization that takes into account the allocation of an acceptable level of risk.

Risk exposure The state of being exposed or vulnerable to a risk. The extent to which an organization is sensitive to underlying risks.

Risk governance The top-down process and guidance that directs risk management activities to align with and support the overall enterprise.

Risk management The process of identifying the level of risk an organization wants, measuring the level of risk the organization currently has, taking actions that bring the actual

level of risk to the desired level of risk, and monitoring the new actual level of risk so that it continues to be aligned with the desired level of risk.

Risk management framework The infrastructure, process, and analytics needed to support effective risk management in an organization.

Risk-neutral pricing Sometimes said of derivatives pricing, uses the fact that arbitrage opportunities guarantee that a risk-free portfolio consisting of the underlying and the derivative must earn the risk-free rate.

Risk-neutral probabilities Weights that are used to compute a binomial option price. They are the probabilities that would apply if a risk-neutral investor valued an option.

Risk premium An extra return expected by investors for bearing some specified risk.

Risk shifting Actions to change the distribution of risk outcomes.

Risk tolerance The amount of risk an investor is willing and able to bear to achieve an investment goal.

Risk transfer Actions to pass on a risk to another party, often, but not always, in the form of an insurance policy.

Robo-adviser A machine-based analytical tool or service that provides technology-driven investment solutions through online platforms.

Rule of 72 The principle that the approximate number of years necessary for an investment to double is 72 divided by the stated interest rate.

Running yield See *current yield*.

Safety-first rules Rules for portfolio selection that focus on the risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Sales Generally, a synonym for revenue; "sales" is generally understood to refer to the sale of goods, whereas "revenue" is understood to include the sale of goods or services.

Sales risk Uncertainty with respect to the quantity of goods and services that a company is able to sell and the price it is able to achieve; the risk related to the uncertainty of revenues.

Sample A subset of a population.

Sample correlation coefficient A standardized measure of how two variables in a sample move together. It is the ratio of the sample covariance to the product of the two variables' standard deviations.

Sample covariance A measure of how two variables in a sample move together.

Sample excess kurtosis A sample measure of the degree of a distribution's kurtosis in excess of the normal distribution's kurtosis.

Sample mean The sum of the sample observations divided by the sample size.

Sample selection bias Bias introduced by systematically excluding some members of the population according to a particular attribute—for example, the bias introduced when data availability leads to certain observations being excluded from the analysis.

Sample-size neglect A type of representativeness bias in which financial market participants incorrectly assume that small sample sizes are representative of populations (or "real" data).

Sample skewness A sample measure of the degree of asymmetry of a distribution.

Sample standard deviation The positive square root of the sample variance.

Sample statistic A quantity computed from or used to describe a sample.

Sample variance The sum of squared deviations around the mean divided by the degrees of freedom.

Sampling The process of obtaining a sample.

Sampling distribution The distribution of all distinct possible values that a statistic can assume when computed from samples of the same size randomly drawn from the same population.

Sampling error The difference between the observed value of a statistic and the estimate resulting from using subsets of the population.

Sampling plan The set of rules used to select a sample.

Say on pay A process whereby shareholders may vote on executive remuneration (compensation) matters.

Scatter plot A chart in which two variables are plotted along the axis and points on the chart represent pairs of the two variables. In regression, the dependent variable is plotted on the vertical axis and the independent variable is plotted along the horizontal axis. Also known as a scattergram and a *scatter diagram.*

Scatter plot matrix A tool for organizing scatter plots between pairs of variables, making it easy to inspect all pairwise relationships in one combined visual.

Scenario analysis Analysis that shows the changes in key financial quantities that result from given (economic) events, such as the loss of customers, the loss of a supply source, or a catastrophic event; a risk management technique involving examination of the performance of a portfolio under specified situations. Closely related to stress testing.

Screening The application of a set of criteria to reduce a set of potential investments to a smaller set having certain desired characteristics.

Seasoned offering An offering in which an issuer sells additional units of a previously issued security.

Second-degree price discrimination When the monopolist charges different per-unit prices using the quantity purchased as an indicator of how highly the customer values the product.

Second lien A secured interest in the pledged assets that ranks below first lien debt in both collateral protection and priority of payment.

Secondary bond markets Markets in which existing bonds are traded among investors.

Secondary market The market where securities are traded among investors.

Secondary precedence rules Rules that determine how to rank orders placed at the same time.

Sector A group of related industries (GICS definition).

Sector indexes Indexes that represent and track different economic sectors—such as consumer goods, energy, finance, health care, and technology—on either a national, regional, or global basis.

Secured ("asset-based") loan A loan that is backed by specific, secured company assets.

Secured bonds Bonds secured by assets or financial guarantees pledged to ensure debt repayment in case of default.

Secured debt Debt in which the debtholder has a direct claim—a pledge from the issuer—on certain assets and their associated cash flows.

Securitization A process that involves moving assets into a special legal entity, which then uses the assets as guarantees to secure a bond issue.

Securitized assets Assets that are typically used to create asset-backed bonds; for example, when a bank securitizes a pool of loans, the loans are said to be securitized.

Security characteristic line A plot of the excess return of a security on the excess return of the market.

Security market index A portfolio of securities representing a given security market, market segment, or asset class.

Security market line (SML) The graph of the capital asset pricing model.

Security selection The process of selecting individual securities; typically, security selection has the objective of generating superior risk-adjusted returns relative to a portfolio's benchmark.

Self-attribution bias A bias in which people take too much credit for successes (*self-enhancing*) and assign responsibility to others for failures (*self-protecting*).

Self-control bias A bias in which people fail to act in pursuit of their long-term, overarching goals because of a lack of self-discipline.

Self-investment limits With respect to investment limitations applying to pension plans, restrictions on the percentage of assets that can be invested in securities issued by the pension plan sponsor.

Sell-side firm A broker/dealer that sells securities and provides independent investment research and recommendations to their clients (i.e., buy-side firms).

Semi-strong-form efficient market A market in which security prices reflect all publicly known and available information.

Semiannual bond basis yield An annual rate having a periodicity of two; also known as a *semiannual bond equivalent yield*.

Semiannual bond equivalent yield See *semiannual bond basis yield*.

Seniority ranking Priority of payment of various debt obligations.

Sensitivity analysis Analysis that shows the range of possible outcomes as specific assumptions are changed.

Separately managed account (SMA) An investment portfolio managed exclusively for the benefit of an individual or institution.

Serial maturity structure Structure for a bond issue in which the maturity dates are spread out during the bond's life; a stated number of bonds mature and are paid off each year before final maturity.

Settlement The process that occurs after a trade is completed, the securities are passed to the buyer, and payment is received by the seller.

Settlement date Date when the buyer makes cash payment and the seller delivers the security.

Settlement price The official price, designated by the clearinghouse, from which daily gains and losses will be determined and marked to market.

Share repurchase A transaction in which a company buys back its own shares. Unlike stock dividends and stock splits, share repurchases use corporate cash.

Shareholder activism Strategies used by shareholders to attempt to compel a company to act in a desired manner.

Shareholder engagement The process whereby companies engage with their shareholders.

Shareholders' equity Assets less liabilities; the residual interest in the assets after subtracting the liabilities.

Sharpe ratio The average return in excess of the risk-free rate divided by the standard deviation of return; a measure of the average excess return earned per unit of standard deviation of return.

Shelf registration A type of public offering that allows the issuer to file a single, all-encompassing offering circular that covers a series of bond issues.

Short The seller of an asset or derivative contract. Also refers to the position of being short an asset or derivative contract.

Short position A position in an asset or contract in which one has sold an asset one does not own, or in which a right under a contract can be exercised against oneself.

Short-run average total cost The curve describing average total cost when some costs are considered fixed.

Short selling A transaction in which borrowed securities are sold with the intention to repurchase them at a lower price at a later date and return them to the lender.

Shortfall risk The risk that portfolio value or portfolio return will fall below some minimum acceptable level over some time horizon.

Shutdown point The point at which average revenue is equal to the firm's average variable cost.

Side letters Side agreements created between the GP and specific LPs. These agreements exist *outside* the LPA. These agreements provide additional terms and conditions related to the investment agreement.

Simple interest The interest earned each period on the original investment; interest calculated on the principal only.

Simple linear regression (SLR) A regression that summarizes the relation between the dependent variable and a single independent variable.

Simple random sample A subset of a larger population created in such a way that each element of the population has an equal probability of being selected to the subset.

Simple random sampling The procedure of drawing a sample to satisfy the definition of a simple random sample.

Simple yield The sum of the coupon payments plus the straight-line amortized share of the gain or loss, divided by the flat price.

Simulation Computer-generated sensitivity or scenario analysis that is based on probability models for the factors that drive outcomes.

Simulation trial A complete pass through the steps of a simulation.

Single-step format With respect to the format of the income statement, a format that does not subtotal for gross profit (revenue minus cost of goods sold).

Sinking fund arrangement Provision that reduces the credit risk of a bond issue by requiring the issuer to retire a portion of the bond's principal outstanding each year.

Situational influences External factors, such as environmental or cultural elements, that shape our behavior.

Skewed Not symmetrical.

Skewness A quantitative measure of skew (lack of symmetry); a synonym of skew. It is computed as the average cubed deviation from the mean standardized by dividing by the standard deviation cubed.

Slope coefficient The coefficient of an independent variable that represents the average change in the dependent variable for a one-unit change in the independent variable.

Small country A country that is a price taker in the world market for a product and cannot influence the world market price.

Smart beta Involves the use of simple, transparent, rules-based strategies as a basis for investment decisions.

Smart contract A computer program that is designed to self-execute on the basis of pre-specified terms and conditions agreed to by parties to a contract.

Socially responsible investing An investment approach that excludes investments in companies or industries that deviate from an organization's beliefs and sometimes includes investments with favorable environmental or social profiles.

Soft-bullet covered bonds Covered bonds for which bond default and payment acceleration of bond cash flows may be delayed upon sponsor default until a new final maturity date is reached.

Solvency With respect to financial statement analysis, the ability of a company to fulfill its long-term obligations.

Solvency ratios Ratios that measure a company's ability to meet its long-term obligations.

Solvency risk The risk that an organization does not survive or succeed because it runs out of cash, even though it might otherwise be solvent.

Sovereign bond A bond issued by a national government. Also called "Sovereign."

Spearman rank correlation coefficient A measure of correlation applied to ranked data.

Special dividend A dividend paid by a company that does not pay dividends on a regular schedule, or a dividend that supplements regular cash dividends with an extra payment.

Special purpose entity A non-operating entity created to carry out a specified purpose, such as leasing assets or securitizing receivables; can be a corporation, partnership, trust, or limited liability partnership formed to facilitate a specific type of business activity. Also called *special purpose vehicle*, *special purpose company*, or *variable interest entity*.

Special purpose vehicle See *special purpose entity*.

Specific identification method An inventory accounting method that identifies which specific inventory items were sold and which remained in inventory to be carried over to later periods.

Speculative demand for money The demand to hold speculative money balances based on the potential opportunities or risks that are inherent in other financial instruments. Also called *portfolio demand for money*.

Speculative money balances Monies held in anticipation that other assets will decline in value.

Split coupon bond See *deferred coupon bond*.

Sponsored A type of depository receipt in which the foreign company whose shares are held by the depository has a direct involvement in the issuance of the receipts.

Spot curve A sequence of yields-to-maturity on zero-coupon bonds. Sometimes called *zero* or *strip curve* (because coupon payments are "stripped" off the bonds).

Spot markets Markets in which assets are traded for immediate delivery.

Spot prices The price of an asset for immediately delivery.

Spot rates A sequence of market discount rates that correspond to the cash flow dates; yields-to-maturity on zero-coupon bonds maturing at the date of each cash flow.

Spread In general, the difference in yield between different fixed-income securities. Often used to refer to the difference between the yield-to-maturity and the benchmark.

Spread over the benchmark See *required yield spread*.

Spread risk Bond price risk arising from changes in the yield spread on credit-risky bonds; reflects changes in the market's assessment and/or pricing of credit migration (or downgrade) risk and market liquidity risk.

Spurious correlation Refers to: 1) correlation between two variables that reflects chance relationships in a particular dataset; 2) correlation induced by a calculation that mixes each of two variables with a third variable; and 3) correlation between two variables arising not from a direct relation between them but from their relation to a third variable.

Stacked bar chart An alternative form for presenting the frequency distribution of two categorical variables, where bars representing the sub-groups are placed on top of each other to form a single bar. Each sub-section is shown in a different color to represent the contribution of each sub-group, and the overall height of the stacked bar represents the marginal frequency for the category.

Stackelberg model A prominent model of strategic decision making in which firms are assumed to make their decisions sequentially.

Stagflation The combination of a high inflation rate with a high level of unemployment and a slowdown of the economy.

Staggered boards Election process whereby directors are typically divided into multiple classes that are elected separately in consecutive years—that is, one class every year.

Stakeholder management The identification, prioritization, and understanding of the interests of stakeholder groups and managing the company's relationships with these groups.

Stakeholders Individuals or groups of individuals who may be affected either directly or indirectly by a decision and thus have an interest, or stake, in the decision.

Standard deviation The positive square root of the variance; a measure of dispersion in the same units as the original data.

Standard error of the estimate A measure of the fit of a regression line, calculated as the square root of the mean square error. Also known as the *standard error of the regression* and the *root mean square error*.

Standard error of the forecast A measure of the uncertainty associated with a forecasted value of the dependent variable that depends on the standard error of the estimate, the variability of the independent variable, the deviation of the forecasted independent variable from the mean in the regression, and the number of observations.

Standard error of the slope coefficient The standard error of the slope, which in a simple linear regression is the ratio of the model's standard error of the estimate (s_e) to the square root of the variation of the independent variable.

Standard normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Standardizing A transformation that involves subtracting the mean and dividing the result by the standard deviation.

Standards of conduct Behaviors required by a group; established benchmarks that clarify or enhance a group's code of ethics.

Standing limit orders A limit order at a price below market and which therefore is waiting to trade.

Stated annual interest rate A quoted interest rate that does not account for compounding within the year. Also called *quoted interest rate*.

Statement of changes in equity (statement of owners' equity) A financial statement that reconciles the beginning-of-period and end-of-period balance sheet values of

shareholders' equity; provides information about all factors affecting shareholders' equity. Also called *statement of owners' equity*.

Statement of financial condition The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of financial position The financial statement that presents an entity's current financial position by disclosing resources the entity controls (its assets) and the claims on those resources (its liabilities and equity claims), as of a particular point in time (the date of the balance sheet).

Statement of operations A financial statement that provides information about a company's profitability over a stated period of time.

Static trade-off theory of capital structure A theory pertaining to a company's optimal capital structure. The optimal level of debt is found at the point where additional debt would cause the costs of financial distress to increase by a greater amount than the benefit of the additional tax shield.

Statistic A summary measure of a sample of observations.

Statistically significant A result indicating that the null hypothesis can be rejected; with reference to an estimated regression coefficient, frequently understood to mean a result indicating that the corresponding population regression coefficient is different from zero.

Status quo bias An emotional bias in which people do nothing (i.e., maintain the status quo) instead of making a change.

Statutory voting A common method of voting where each share represents one vote.

Step-up coupon bond Bond for which the coupon, fixed or floating, increases by specified margins at specified dates.

Stochastic oscillator A momentum indicator that compares a particular closing price of a security to a range of the security's prices over a certain period of time.

Stock dividend A type of dividend in which a company distributes additional shares of its common stock to shareholders instead of cash.

Stock split An increase in the number of shares outstanding with a consequent decrease in share price, but no change to the company's underlying fundamentals.

Stop-loss order See *stop order*.

Stop order An order in which a trader has specified a stop price condition. Also called *stop-loss order*.

Store of value The quality of tending to preserve value.

Store of wealth Goods that depend on the fact that they do not perish physically over time, and on the belief that others would always value the good.

Straight-line method A depreciation method that allocates evenly the cost of a long-lived asset less its estimated residual value over the estimated useful life of the asset.

Straight voting A shareholder voting process in which shareholders receive one vote for each share owned.

Strategic analysis Analysis of the competitive environment with an emphasis on the implications of the environment for corporate strategy.

Strategic asset allocation The set of exposures to IPS-permissible asset classes that is expected to achieve the client's long-term objectives given the client's investment constraints.

Stratified random sampling A procedure that first divides a population into subpopulations (strata) based on classification criteria and then randomly draws samples from each stratum in sizes proportional to that of each stratum in the population.

Street convention A yield measure that neglects weekends and holidays; the internal rate of return on cash flows assuming payments are made on the scheduled dates, even when the scheduled date falls on a weekend or holiday.

Stress testing A specific type of scenario analysis that estimates losses in rare and extremely unfavorable combinations of events or scenarios.

Strong-form efficient market A market in which security prices reflect all public and private information.

Structural (or cyclically adjusted) budget deficit The deficit that would exist if the economy was at full employment (or full potential output).

Structural subordination Arises in a holding company structure when the debt of operating subsidiaries is serviced by the cash flow and assets of the subsidiaries before funds can be passed to the holding company to service debt at the parent level.

Structured data Data that are highly organized in a predefined manner, usually with repeating patterns.

Structured financial instrument A financial instrument that shares the common attribute of repackaging risks. Structured financial instruments include asset-backed securities, collateralized debt obligations, and other structured financial instruments such as capital protected, yield enhancement, participation, and leveraged instruments.

Subjective probability A probability drawing on personal or subjective judgment.

Subordinated debt A class of unsecured debt that ranks below a firm's senior unsecured obligations.

Subordination Form of internal credit enhancement that relies on creating more than one bond tranche and ordering the claim priorities for ownership or interest in an asset between the tranches. The ordering of the claim priorities is called a senior/subordinated structure, where the tranches of highest seniority are called senior followed by subordinated or junior tranches. Also called *credit tranching*.

Substitutes Said of two goods or services such that if the price of one increases the demand for the other tends to increase, holding all other things equal (e.g., butter and margarine).

Sum of squares error (SSE) The sum of the squared deviations of (1) the value of the dependent variable and (2) the value of the dependent variable based on the estimated regression line. Also referred to as the *residual sum of squares*.

Sum of squares regression (SSR) The sum of the squared deviations of (1) the value of the dependent variable based on the estimated regression line and (2) the mean of the dependent variable.

Sum of squares total (SST) The sum of the squared deviations of the dependent variable from its mean; the variation of the dependent variable. Also referred to as the *total sum of squares*.

Sunk cost A cost that has already been incurred.

Supervised learning A machine learning approach that makes use of labeled training data.

Supply shock A typically unexpected disturbance to supply.

Support In technical analysis, a price range in which buying activity is sufficient to stop the decline in the price of a security.

Support tranches Classes or tranches in CMOs that protect PAC tranches from prepayment risk.

Supranational bond A bond issued by a supranational agency such as the World Bank.

Surety bond Form of external credit enhancement whereby a rated and regulated insurance company guarantees to reimburse bondholders for any losses incurred up to a maximum amount if the issuer defaults.

Survey approach An estimate of the equity risk premium that is based on estimates provided by a panel of finance experts.

Survivorship bias Relates to the inclusion of only current investment funds in a database. As such, the returns of funds that are no longer available in the marketplace (have been liquidated) are excluded from the database. In addition, *backfill bias* is another problem, whereby certain surviving hedge funds may be added to databases and various hedge fund indexes only after they are initially successful and start to report their returns.

Sustainability linked loans These are any types of loan instruments and/or contingent facilities (such as bonding lines, guarantee lines, or letters of credit) that incentivize the borrower's achievement of ambitious, pre-determined sustainability performance objectives.

Sustainable growth rate The rate of dividend (and earnings) growth that can be sustained over time for a given level of return on equity, keeping the capital structure constant and without issuing additional common stock.

Sustainable investing A term used in a similar context to responsible investing, but its key focus is on factoring in sustainability issues while investing.

Sustainable rate of economic growth The rate of increase in the economy's productive capacity or potential GDP.

Swap contract An agreement between two parties to exchange a series of future cash flows.

Syndicated loan A loan from a group of lenders to a single borrower.

Syndicated offering A bond issue underwritten by a group of investment banks.

Systematic risk Risk that affects the entire market or economy; it cannot be avoided and is inherent in the overall market. Systematic risk is also known as non-diversifiable or market risk.

Systematic sampling A procedure of selecting every *k*th member until reaching a sample of the desired size. The sample that results from this procedure should be approximately random.

Tactical asset allocation The decision to deliberately deviate from the strategic asset allocation in an attempt to add value based on forecasts of the near-term relative performance of asset classes.

Tactical asset allocation (TAA) A portfolio strategy that shifts the percentages of assets held in various asset classes (or categories) to take advantage of market opportunities. Allocation shifts can occur within an asset class or across asset classes.

Tag cloud see **word cloud**.

Target capital structure A company's chosen proportions of debt and equity.

Target independent A bank's ability to determine the definition of inflation that they target, the rate of inflation that they target, and the horizon over which the target is to be achieved.

Target semideviation A measure of downside risk, calculated as the square root of the average of the squared deviations of observations below the target (also called target downside deviation).

Tariffs Taxes that a government levies on imported goods.

Tax base The amount at which an asset or liability is valued for tax purposes.

Tax expense An aggregate of an entity's income tax payable (or recoverable in the case of a tax benefit) and any changes in deferred tax assets and liabilities. It is essentially the income tax payable or recoverable if these had been determined based on accounting profit rather than taxable income.

Tax loss carry forward A taxable loss in the current period that may be used to reduce future taxable income.

Taxable income The portion of an entity's income that is subject to income taxes under the tax laws of its jurisdiction.

Taxable temporary differences Temporary differences that result in a taxable amount in a future period when determining the taxable profit as the balance sheet item is recovered or settled.

Technical analysis A form of security analysis that uses price and volume data, often displayed graphically, in decision making.

Technology The process a company uses to transform inputs into outputs.

Tender offer Corporate takeover mechanism that involves shareholders selling their interests directly to the group seeking to gain control.

Tenor The time-to-maturity for a bond or derivative contract. Also called *term to maturity*.

Term maturity structure Structure for a bond issue in which the bond's notional principal is paid off in a lump sum at maturity.

Term structure See *maturity structure*.

Term structure of credit spreads The relationship between the spreads over the "risk-free" (or benchmark) rates and times-to-maturity.

Term structure of yield volatility The relationship between the volatility of bond yields-to-maturity and times-to-maturity.

Terminal stock value The expected value of a share at the end of the investment horizon—in effect, the expected selling price. Also called *terminal value*.

Terminal value The expected value of a share at the end of the investment horizon—in effect, the expected selling price.

Terms of trade The ratio of the price of exports to the price of imports, representing those prices by export and import price indexes, respectively.

Test of the mean of the differences A statistical test for differences based on paired observations drawn from samples that are dependent on each other.

Text analytics The use of computer programs to analyze and derive meaning from typically large, unstructured text- or voice-based datasets.

Thematic investing An ESG investment approach that focuses on investing in themes or assets specifically relating to ESG factors, such as clean energy, green technology, or sustainable agriculture.

Thin-Tailed Describes a distribution that has relatively less weight in the tails than the normal distribution (also called platykurtic).

Third-degree price discrimination When the monopolist segregates customers into groups based on demographic or other characteristics and offers different pricing to each group.

Time-period bias The possibility that when we use a time-series sample, our statistical conclusion may be sensitive to the starting and ending dates of the sample.

Time-series data A sequence of observations for a single observational unit of a specific variable collected over time and at discrete and typically equally spaced intervals of time (such as daily, weekly, monthly, annually, or quarterly).

Time tranching The creation of classes or tranches in an ABS/MBS that possess different (expected) maturities.

Time value The difference between the market price of the option and its intrinsic value.

Time value decay Said of an option when, at expiration, no time value remains and the option is worth only its exercise value.

Time value of money The principles governing equivalence relationships between cash flows with different dates.

Time-weighted rate of return The compound rate of growth of one unit of currency invested in a portfolio during a stated measurement period; a measure of investment performance that is not sensitive to the timing and amount of withdrawals or additions to the portfolio.

Tokenization The process of representing ownership rights to physical assets on a blockchain or distributed ledger.

Top-down analysis An investment selection approach that begins with consideration of macroeconomic conditions and then evaluates markets and industries based upon such conditions.

Total comprehensive income The change in equity during a period resulting from transaction and other events, other than those changes resulting from transactions with owners in their capacity as owners.

Total cost The summation of all costs, for which costs are classified as fixed or variable.

Total factor productivity A scale factor that reflects the portion of growth unaccounted for by explicit factor inputs (e.g., capital and labor).

Total fixed cost The summation of all expenses that do not change as the level of production varies.

Total invested capital The sum of market value of common equity, book value of preferred equity, and face value of debt.

Total probability rule A rule explaining the unconditional probability of an event in terms of probabilities of the event conditional on mutually exclusive and exhaustive scenarios.

Total probability rule for expected value A rule explaining the expected value of a random variable in terms of expected values of the random variable conditional on mutually exclusive and exhaustive scenarios.

Total return Measures the price appreciation, or percentage change in price of the securities in an index or portfolio, plus any income received over the period.

Total return index An index that reflects the price appreciation or percentage change in price of the constituent securities plus any income received since inception.

Total return swap A swap in which one party agrees to pay the total return on a security. Often used as a credit derivative, in which the underlying is a bond.

Total variable cost The summation of all variable expenses.

Tracking error The standard deviation of the differences between a portfolio's returns and its benchmark's returns; a synonym of active risk.

Tracking risk The standard deviation of the differences between a portfolio's returns and its benchmarks returns. Also called *tracking error*.

Trade creation When regional integration results in the replacement of higher cost domestic production by lower cost imports from other members.

Trade credit A spontaneous form of credit in which a purchaser of the goods or service is financing its purchase by delaying the date on which payment is made.

Trade diversion When regional integration results in lower-cost imports from non-member countries being replaced with higher-cost imports from members.

Trade payables Amounts that a business owes to its vendors for goods and services that were purchased from them but which have not yet been paid.

Trade protection Government policies that impose restrictions on trade, such as tariffs and quotas.

Trade surplus (deficit) When the value of exports is greater (less) than the value of imports.

Trading securities Under US GAAP, a category of debt securities held by a company with the intent to trade them. Also called *held-for-trading securities*.

Traditional investment markets Markets for traditional investments, which include all publicly traded debts and equities and shares in pooled investment vehicles that hold publicly traded debts and/or equities.

Transactions money balances Money balances that are held to finance transactions.

Transfer payments Welfare payments made through the social security system that exist to provide a basic minimum level of income for low-income households.

Transparency Said of something (e.g., a market) in which information is fully disclosed to the public and/or regulators.

Treasury stock method A method for accounting for the effect of options (and warrants) on earnings per share (EPS) that specifies what EPS would have been if the options and warrants had been exercised and the company had used the proceeds to repurchase common stock.

Tree-Map Another graphical tool for displaying categorical data. It consists of a set of colored rectangles to represent distinct groups, and the area of each rectangle is proportional to the value of the corresponding group.

Trend A long-term pattern of movement in a particular direction.

Treynor ratio A measure of risk-adjusted performance that relates a portfolio's excess returns to the portfolio's beta.

Triangle pattern In technical analysis, a continuation chart pattern that forms as the range between high and low prices narrows, visually forming a triangle.

Trimmed mean A mean computed after excluding a stated small percentage of the lowest and highest observations.

Trimodal A distribution that has the three most frequently occurring values.

Triple bottom In technical analysis, a reversal pattern that results when the price forms three troughs at roughly the same price level. A triple bottom is used to predict a change from a downtrend to an uptrend.

Triple top In technical analysis, a reversal pattern that results when the price forms three peaks at roughly the same price level. A triple top is used to predict a change from an uptrend to a downtrend.

True yield The internal rate of return on cash flows using the actual calendar, including weekends and bank holidays.

Trust deed The governing legal credit agreement, typically incorporated by reference in the prospectus. Also called *bond indenture*.

Turn-of-the-year effect Calendar anomaly that stock market returns in January are significantly higher compared to the rest of the months of the year, with most of the abnormal returns reported during the first five trading days in January.

Two-dimensional rectangular array A popular form for organizing data for processing by computers or for presenting data visually. It is comprised of columns and rows to hold multiple variables and multiple observations, respectively (also called a **data table**).

Two-fund separation theorem The theory that all investors regardless of taste, risk preferences, and initial wealth will hold a combination of two portfolios or funds: a risk-free asset and an optimal portfolio of risky assets.

Two-sided hypothesis test A test in which the null hypothesis is rejected in favor of the alternative hypothesis if the evidence indicates that the population parameter is either smaller or larger than a hypothesized value; occurs when the alternative hypothesis is stated as not equal to the hypothesized population parameters.

Two-way table See *contingency table*.

Two-week repo rate The interest rate on a two-week repurchase agreement; may be used as a policy rate by a central bank.

Type I error The error of rejecting a true null hypothesis; a false positive.

Type II error The error of not rejecting a false null hypothesis; false negative.

Unanticipated (unexpected) inflation The component of inflation that is a surprise.

Uncommitted lines of credit The least reliable form of bank borrowing in which a bank offers, without formal commitment, a line of credit for an extended period of time but reserves the right to refuse any request for its use.

Unconditional probability The probability of an event *not* conditioned on another event.

Underemployed A person who has a job but has the qualifications to work a significantly higher-paying job.

Underlying An asset that trades in a market in which buyers and sellers meet, decide on a price, and the seller then delivers the asset to the buyer and receives payment. The underlying is the asset or other derivative on which a particular derivative is based. The market for the underlying is also referred to as the *spot market*.

Underwriter A firm, usually an investment bank, that takes the risk of buying the newly issued securities from the issuer and then reselling them to investors or to dealers, thus guaranteeing the sale of the securities at the offering price negotiated with the issuer.

Underwritten offering A type of securities issue mechanism in which the investment bank guarantees the sale of the securities at an offering price that is negotiated with the issuer. Also known as *firm commitment offering*.

Unearned revenue A liability account for money that has been collected for goods or services that have not yet been delivered; payment received in advance of providing a good or service. Also called *deferred revenue* or *deferred income*.

Unemployed People who are actively seeking employment but are currently without a job.

Unemployment rate The ratio of unemployed to the labor force.

Unexpected inflation The component of inflation that is a surprise.

Unimodal A distribution with a single value that is most frequently occurring.

Unit elastic An elasticity with a magnitude of negative one. Also called *unitary elastic*.

Unit labor cost The average labor cost to produce one unit of output.

Unit normal distribution The normal density with mean (μ) equal to 0 and standard deviation (σ) equal to 1.

Unitranche debt Consists of a hybrid or blended loan structure that combines different tranches of secured and unsecured debt into a single loan with a single, blended interest rate.

Units-of-production method A depreciation method that allocates the cost of a long-lived asset based on actual usage during the period.

Univariate distribution A distribution that specifies the probabilities for a single random variable.

Universal owners Long-term investors, such as pension funds, that have significant assets invested in globally diversified portfolios.

Unlimited funds An unlimited funds environment assumes that the company can raise the funds it wants for all profitable investments simply by paying the required rate of return.

Unsecured debt Debt that gives the debtholder only a general claim on an issuer's assets and cash flow.

Unsponsored A type of depository receipt in which the foreign company whose shares are held by the depository has no involvement in the issuance of the receipts.

Unstructured data Data that do not follow any conventionally organized forms.

Unsupervised learning A machine learning approach that does not make use of labeled training data.

Up transition probability The probability that an asset's value moves up.

Uptrend A pattern that occurs when the price of an asset moves higher over a period of time.

Validity instructions Instructions which indicate when the order may be filled.

Valuation allowance A reserve created against deferred tax assets, based on the likelihood of realizing the deferred tax assets in future accounting periods.

Valuation ratios Ratios that measure the quantity of an asset or flow (e.g., earnings) in relation to the price associated with a specified claim (e.g., a share or ownership of the enterprise).

Value at risk A money measure of the minimum value of losses expected during a specified time period at a given level of probability.

Value investors With reference to equity investors, investors who are focused on paying a relatively low share price in relation to earnings or assets per share.

VaR See *value at risk*.

Variable A characteristic or quantity that can be measured, counted, or categorized and that is subject to change (also called a field, an attribute, or a feature).

Variable costs Costs that fluctuate with the level of production and sales.

Variance The expected value (the probability-weighted average) of squared deviations from a random variable's expected value.

Variation margin Additional margin that must be deposited in an amount sufficient to bring the balance up to the initial margin requirement.

Veblen goods Goods that increase in desirability with increasing price.

Vega A measure of the sensitivity of an option's price to changes in the underlying's volatility.

Venture capital Investments that provide "seed" or startup capital, early-stage financing, or later-stage financing (including mezzanine-stage financing) to companies that are in early development stages and require additional capital for expansion or preparation for an initial public offering.

Venture capital fund A hedge fund that seeks to buy, optimize, and ultimately sell portfolio companies to generate profits. See *private equity fund*.

Venture capital funds Funds that seek to invest in, optimize, and eventually exit portfolio companies to generate profits. See *private equity funds.*

Vertical analysis Common-size analysis using only one reporting period or one base financial statement; for example, an income statement in which all items are stated as percentages of sales.

Vertical demand schedule Implies that some fixed quantity is demanded, regardless of price.

Visual technique The most common and readily available method of initial data assessment. Experts in pattern recognition maintain that the visual (or "eyeball") technique is still the most effective way of searching for recognizable patterns.

Visualization The presentation of data in a pictorial or graphical format for the purpose of increasing understanding and for gaining insights into the data.

Volatility The standard deviation of the continuously compounded returns on the underlying asset.

Voluntarily unemployed A person voluntarily outside the labor force, such as a jobless worker refusing an available vacancy.

Voluntary export restraint A trade barrier under which the exporting country agrees to limit its exports of the good to its trading partners to a specific number of units.

Vote by proxy A mechanism that allows a designated party—such as another shareholder, a shareholder representative, or management—to vote on the shareholder's behalf.

Warrant Attached option that gives its holder the right to buy the underlying stock of the issuing company at a fixed exercise price until the expiration date.

Waterfall Represents the distribution method that defines the order in which allocations are made to LPs and GPs. There are two major types of waterfall: *deal by deal* (or *American*) and *whole of fund* (or *European*).

Weak-form efficient market hypothesis The belief that security prices fully reflect all past market data, which refers to all historical price and volume trading information.

Wealth effect An increase (decrease) in household wealth increases (decreases) consumer spending out of a given level of current income.

Web-based lender A lender that operates primarily on the internet, offering loans in relatively small amounts, typically to small businesses in need of cash.

Weighted average cost method An inventory accounting method that averages the total cost of available inventory items over the total units available for sale.

Weighted average cost of capital A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average cost of capital (WACC) A weighted average of the after-tax required rates of return on a company's common stock, preferred stock, and long-term debt, where the weights are the fraction of each source of financing in the company's target capital structure.

Weighted average coupon rate Weighting the mortgage rate of each mortgage loan in the pool by the percentage of the mortgage outstanding relative to the outstanding amount of all the mortgages in the pool.

Weighted average life A measure that gives investors an indication of how long they can expect to hold the MBS before it is paid off; the convention-based average time to receipt of all principal repayments. Also called *average life*.

Weighted average maturity Weighting the remaining number of months to maturity for each mortgage loan in the pool by the amount of the outstanding mortgage balance.

Weighted mean An average in which each observation is weighted by an index of its relative importance.

Winsorized mean A mean computed after assigning a stated percentage of the lowest values equal to one specified low value and a stated percentage of the highest values equal to one specified high value.

Word cloud A visual device for representing textual data, which consists of words extracted from a source of textual data. The size of each distinct word is proportional to the frequency with which it appears in the given text (also known as **tag cloud**).

Working capital The difference between current assets and current liabilities.

Working capital management The management of a company's short-term assets (such as inventory) and short-term liabilities (such as money owed to suppliers).

World price The price prevailing in the world market.

Yield duration The sensitivity of the bond price with respect to the bond's own yield-to-maturity.

Yield to maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield to redemption* or *redemption yield*.

Yield-to-maturity Annual return that an investor earns on a bond if the investor purchases the bond today and holds it until maturity. It is the discount rate that equates the present value of the bond's expected cash flows until maturity with the bond's price. Also called *yield-to-redemption* or *redemption yield*.

Yield-to-redemption See *yield-to-maturity*.

Yield-to-worst The lowest of the sequence of yields-to-call and the yield-to-maturity.

Zero-coupon bond A bond that does not pay interest during its life. It is issued at a discount to par value and redeemed at par. Also called *pure discount bond*.

Zero-volatility spread (Z-spread) Calculates a constant yield spread over a government (or interest rate swap) spot curve.

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.